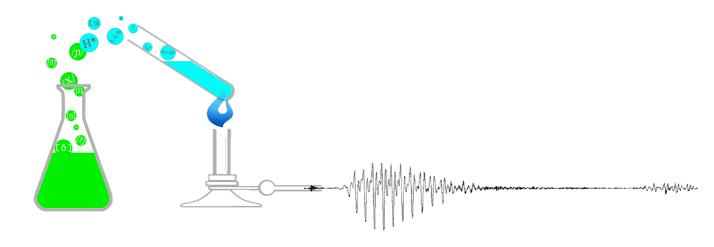
Perspectivas actuales en el

análisis fónico del habla

Tradición y avances en la fonética experimental



Editado por **Adrián Cabedo Nebot**

Perspectivas actuales en el análisis fónico del habla.

Tradición y avances en la fonética experimental

Perspectivas actuales en el análisis fónico del habla.

Tradición y avances en la fonética experimental

Editado por:

Adrián Cabedo Nebot

PERSPECTIVAS ACTUALES EN EL ANÁLISIS FÓNICO DEL HABLA: TRADICIÓN Y AVANCES EN LA FONÉTICA EXPERIMENTAL

Editor: Adrián Cabedo Nebot ISBN: 978-84-606-8969-0

Anejo núm. 7 de Normas. Revista de Estudios Lingüísticos Hispánicos [Recibido: 31 de enero de 2015. Evaluado: marzo de 2015. Aceptado: abril de 2015.]

© del texto: los autores

www.uv.es/normas (ISSN: 2174-7245) Departamento de Filología Española Universitat de València

Libro electrónico de acceso gratuito

Texto en línea:

https://ojs.uv.es/index.php/normas/pages/view/anejos.

Monografías científicas

Maquetación: Tecnolingüística S. L.

Contenido

Autores	11
Prefacio	17
Introducción	19
Parte I. Análisis suprasegmentales	
Melodías para preguntar en portugués y en castellano. Un análisis contrastivo ARAÚJO LEITE, MONIQUE; FONT-ROTCHÉS, DOLORS; HUELVA, ENRIQUE	23
Intonation of Yes/No question in Portuguese: analysis of the relationship between intonation and lexical stress in the AMPER-POR corpus CRUZ, REGINA; SEARA, IZABEL; MOUTINHO, LURDES	35
La entonación cubana en zonas rurales: La Habana y Santa Clara Dorta, Josefa; Díaz Cabrera, Chaxiraxi; Hernández Díaz, Beatriz	45
Diferencias acústicas entre los acentos prenucleares en catalán Fernández Planas, Ana Ma.; Roseano, Paolo; Elvira García, Wendy; Cerdà Massó, Ramon; Martínez Celdrán, Eugenio	57
Prosodia e interpretación: La calidad de las interpretaciones desde una perspectiva prosódica FIGUEROA LÓPEZ, HELENA	67
Rasgos melódicos de enunciados neutros del español hablado por brasileños FONSECA DE OLIVERIA, ALINE; MATEO RUIZ, MIGUEL	79
Realización fonética del acento español en entonación ascendente Kimura, Takuya; Sensui, Hirotaka; Takasawa, Miyuki; Toyomaru, Atsuko; Atria, José Joaquín	91
Positional effects of prosodic prominence in spoken Italian LOMBARDI VALLAURI, EDOARDO	99
Sobre los indicadores acústico-melódicos de la ironía verbal en la conversación espontánea MARTÍNEZ HERNÁNDEZ, DIANA	113
"Es un hombre famoso" o "¿Es un hombre famoso?" Rasgos melódicos de las interrogativas absolutas del español hablado por suecos MARTORELL MORALES, LAURA; FONT-ROTCHÉS, DOLORS	127
Aproximaciones a la caracterización prosódica de los subactos, la unidad discursiva mínima del sistema Val.Es.Co. PASCUAL ALIAGA, ELENA	137

Aspectos prosódicos en una zona de transición lingüística: la Franja de Aragón Romera Barrios, Lourdes; Fernández Planas, Ana Ma.; Roseano, Paolo; Elvira-García, Wendy; Carrera Sabaté, Josefina; Martínez Celdrán, Eugenio	151
Diferencias perceptivas entre los acentos tonales prenucleares en catalán ROSEANO, PAOLO; FERNÁNDEZ PLANAS, ANA MA.; ELVIRA-GARCÍA, WENDY; CERDÀ MASSÓ, RAMÓN; MARTÍNEZ CELDRÁN, EUGENIO	163
La relación de las características del lenguaje corporal y la voz en el discurso mendaz SÁNCHEZ CONDE, CRISTINA	175
Variación fonética de los tonos estándar del chino mandarín. Un estudio de corpus YAO, JUNMING; GARRIDO ALMIÑANA, JUAN MARÍA	187
Parte II. Enseñanza de segundas lenguas	
Propostes didàctiques d'entonació de cortesia atenuant per al nivell b1 de català com a llengua estrangera Devís Herraiz, Empar; Bartolí Rigol, Marta	199
Percepción auditiva y visual como estrategias de adquisición de sonidos nuevos del español por estudiantes rusos DMÍTRIEVA, ANNA	211
La presencia de la lengua escrita en los procesos de aprendizaje y adquisición de la pronunciación del español como lengua extranjera: ¿podríamos hablar de interferencia? GIRALT LORENZ, MARTA; CANTERO SERENA, FRANCISCO JOSÉ	223
Análisis descriptivo de las estrategias de aprendizaje y creencias en los procesos de adquisición y aprendizaje de la pronunciación en E/LE GIRALT LORENZ, MARTA	233
L1-L2map: una base de datos fónica para la enseñanza de la pronunciación de segundas lenguas MARTÍNEZ-PARICIO, VIOLETA; KOREMAN, JACQUES; HUSBY, OLAF	241
Acquisition of L2 Portuguese stops by Chinese-speaking learners - Considerations NUNES, ANA MARGARIDA BELÉM	249
Enseñanza de E/LE a alumnos de edades tempranas en una escuela de Irlanda del Norte PIQUER VIVES, INMA	259
La percepción de oposiciones fonético/fonológicas en L2 ROMITO, LUCIANO; FRONTERA, MANUELA	269
Aspectos prosódicos de las interrogativas en aprendientes italianos de E/LE SAVY, RENATA; LUQUE MOYA, JOSE ANTONIO	283

Parte III. Fonética clínica

Velocidad de articulación y estructuración del espacio vocálico en función del acento en la afasia: un estudio preliminar BAQUÉ, LORRAINE	297
La lingüística clínica a la vista de la fonética: adquisición y pérdidas fonéticas Tordera Yllescas, Juan Carlos	309
Parte IV. Fonética forense	
Análisis semiautomático y automático de vocales con fines forenses en habla de micro y de móvil Dorta, Josefa; Jorge Trujillo, Carolina; Martín Gómez, José Antonio	32 3
El ritmo lingüístico en la caracterización del locutor. Percepción y aplicaciones judiciales Marrero Aguiar, Victoria; López Bascuas, Luis Enrique; Martín Fernández, José Luis	335
Efectos del cambio de lengua en la comparación de voces mediante LTAS ROSEANO, PAOLO; FERNÁNDEZ PLANAS, ANA MA.; ELVIRA-GARCÍA, WENDY; MARTÍNEZ CELDRÁN, EUGENIO	347
Parte V. Fonética normativa	
La fonètica en la gramàtica d'una institució normativitzadora (GNV, 2006) Saragossà, Abelard	359
Parte VI. Fonética vocálica y consonántica	
Las estrategias compensatorias de los hablantes chilenos: una ampliación acústica de la investigación de /s/ implosiva BOLYANATZ, MARIŠKA A.	373
Sons, cervells i usos lingüístics. Caracterització diafàsica de vocals tòniques del lleidatà en contextos multilingües Carrera-Sabaté, Josefina	387
Effects of place of articulation on Spanish coda /s/ perception DEL SAZ, MARÍA	399
Estudio del cambio lingüístico en el sistema vocálico del habla pasiega MORENO CARRAL, MANUEL	411
Normalització de les vocals mitjanes anteriors del català en parla espontània RIUS-ESCUDÉ, AGNÈS; TORRAS COMPTE, FRANCINA	421
Tratamiento perceptivo de las categorías /p,t,k/ /b,d,g/. Esbozo del espacio acústico de distribución categorial a partir de los parámetros VOT, Explosión y Transición ROSAS SALVANS, ALBA	431

Resilabificación de la aspiración de /-s/ ante oclusiva dental sorda. Parámetros acústicos y variación social VIDA-CASTRO, MATILDE	441
Parte VII. Tecnologías del habla	
Una herramienta para la transcripción prosódica automática con etiquetas Sp_ToBI en Praat ELVIRA-GARCÍA, WENDY; ROSEANO, PAOLO; FERNÁNDEZ PLANAS, ANA MA.; MARTÍNEZ CELDRÁN, EUGENIO	455
Estudio comparativo de la información que aportan los rasgos articulatorios y acústicos de los fonemas al reconocimiento automático de habla SIMÓN, JAVIER; LLOMBART GIL, JORGE; MIGUEL ARTIAGA, ANTONIO; LLEIDA SOLANO, EDUARDO	465

Autores

Araújo Leite, Monique

Universidad de Brasilia

Atria, José Joaquín

University College London

Baqué, Lorraine

Universitat Autònoma de Barcelona

Bartolí Rigol, Marta

Universitat de Barcelona

Bolyanatz, Mariška A.

Universidad de California, Los Ángeles

Cantero Serena, Francisco José

Universitat de Barcelona

Carrera-Sabaté, Josefina

Universitat de Barcelona

Cerdà Massó, Ramon

Universitat de Barcelona

Cruz, Regina

Universidade Federal do Pará

Del Saz, María

Universidad de Sevilla

Devís Herraiz, Empar

Universitat de Barcelona

Díaz Cabrera, Chaxiraxi

S.E.G.A.I. Laboratorio de Fonética. Universidad de La Laguna

Dmítrieva, Anna

Universidad Pablo de Olavide

Dorta, Josefa

S.E.G.A.I. Laboratorio de Fonética. Universidad de La Laguna

Elvira-García, Wendy

Universitat de Barcelona

Fernández Planas, Ana M.ª

Universitat de Barcelona

Figueroa López, Helena

Universitat de València

Fonseca De Oliveria, Aline

Universitat de Barcelona

Font-Rotchés, Dolors

Universitat de Barcelona

Frontera, Manuela

Università della Calabria

Garrido Almiñana, Juan María

Universitat Pompeu Fabra

Giralt Lorenz, Marta

Universitat de Barcelona

Hernández Díaz, Beatriz

S.E.G.A.I. Laboratorio de Fonética. Universidad de La Laguna

Huelva, Enrique

Universidad de Brasilia

Husby, Olaf

Norwegian University of Science and Technology

Jorge Trujillo, Carolina

S.E.G.A.I. Laboratorio de Fonética. Universidad de La Laguna

Kimura, Takuya

Universidad Seisen/Universidad Sofia

Koreman, Jacques

Norwegian University of Science and Technology

Lleida Solano, Eduardo

Universidad de Zaragoza, ViVoLab; Instituto Universitario de Investigación en Ingeniería de Aragón

Llombart Gil, Jorge

Universidad de Zaragoza, ViVoLab; Instituto Universitario de Investigación en Ingeniería de Aragón

Lombardi Vallauri, Edoardo

Università Roma Tre

López Bascuas, Luis Enrique

Universidad Complutense de Madrid

Luque Moya, Jose Antonio

Universidad de Málaga

Marrero Aguiar, Victoria

UNED

Martín Fernández, José Luis

Universidad Complutense de Madrid

Martín Gómez, José Antonio

S.E.G.A.I. Laboratorio de Fonética. Universidad de La Laguna

Martínez Celdrán, Eugenio

Universitat de Barcelona

Martínez Hernández, Diana

Universitat de València

Martínez-Paricio, Violeta

Norwegian University of Science and Technology

Martorell Morales, Laura

Universitat de Barcelona

Mateo Ruiz, Miguel

Universitat de Barcelona

Miguel Artiaga, Antonio

Universidad de Zaragoza, ViVoLab; Instituto Universitario de Investigación en Ingeniería de Aragón

Moreno Carral, Manuel

Universitat Pompeu Fabra

Moutinho, Lurdes

Universidade de Aveiro

Nunes, Ana Margarida Belém

University of Macau

Pascual Aliaga, Elena

Universitat de València

Piquer Vives, Inma

Universitat de Barcelona

Rius-Escudé, Agnès

Universitat de Barcelona

Romera Barrios, Lourdes

Universitat de Barcelona

Romito, Luciano

Università della Calabria

Rosas Salvans, Alba

Universitat Autònoma de Barcelona/Laboratori FLexSem

Roseano, Paolo

Universitat de Barcelona

Sánchez Conde, Cristina

UIMP-CSIC

Saragossà, Abelard

Universitat de València

Savy, Renata

Università degli studi di Salerno

Seara, Izabel

Universidade Federal de Santa Catarina

Sensui, Hirotaka

Universidad Nanzan

Simón, Javier

Universidad de Zaragoza, ViVoLab

Takasawa, Miyuki

Universidad Asia

Tordera Yllescas, Juan Carlos

Universitat de València

Torras Compte, Francina

Universitat de Barcelona

Toyomaru, Atsuko

Universidad Takushoku

Vida-Castro, Matilde

Universidad de Málaga

Yao, Junming

Universidad Internacional Menéndez Pelayo

Prefacio

El ámbito internacional recoge en las últimas dos décadas importantes estudios sobre la fonética experimental. La idea de este volumen ha sido recoger artículos que pongan de manifiesto ese interés creciente por el análisis experimental y que reflejen un amplio bagaje de líneas temáticas. Con ello se ofrece al futuro investigador (y también al actual) una motivación para la gesta de nuevos estudios.

En este volumen 64 investigadores de renombre han participado en los 39 artículos que componen las 7 secciones temáticas establecidas para cubrir el ámbito de la fonética experimental. En estas se han mantenido líneas tradicionales de investigación, como el análisis suprasegmental, la enseñanza de lenguas o los análisis vocálicos y consonánticos, junto a otros más innovadores, como la fonética forense, las tecnología del habla o la fonética clínica.

Así mismo, los investigadores participantes proceden de importantes universidades internacionales, como la Universitat de València, la Universitat de Barcelona, la Universidad Takushoku, la Università degli studi di Salerno, la Università Roma Tre, la UNED, etc.

Quiero dar las gracias a Antonio Hidalgo y a Manuel José Aguilar Ruiz que, de manera desinteresada, han participado en la revisión editorial de este trabajo.

Adrián Cabedo Nebot Editor

Introducción

Siguiendo con su espíritu plural e innovador la colección Anejos de *Normas, Revista de Estudios Lingüísticos Hispánicos* da paso en esta ocasión a la publicación del volumen monográfico *Perspectivas actuales en el análisis fónico del habla. Tradición y avances en la fonética experimental.* Y en mi modesta opinión hay que felicitarse por la oportunidad de este libro.

Para los especialistas en el estudio fónico representará sin duda un compendio de trabajos punteros relacionados con el estudio experimental e interdisciplinar del componente fónico de la lengua, aspectos muy novedosos en el ámbito de interés de la Fonética y Fonología, tanto desde el punto de vista del español (de ahí la presencia de este volumen en la colección Anejos de *Normas*) como desde una perspectiva más general, como demuestran los diversos trabajos que tienen como objeto de estudio lenguas diferentes al español. Por otra parte, aquellos que se acerquen iniciáticamente al volumen hallarán con toda probabilidad algún trabajo que satisfaga su curiosidad científica.

Y es que una de las primeras virtudes de este trabajo que sin duda apreciará el lector es la riqueza y diversidad de sus enfoques, desarrollados por algunos de los investigadores más señalados a nivel nacional e internacional.

Siguiendo el orden de aparición de los estudios que se presentan, se encontrarán inicialmente investigaciones relacionadas con los aspectos suprasegmentales, considerados en ciertos casos desde una perspectiva contrastiva y, en su mayor parte, vinculados con algunos de los modelos de investigación más avanzados por lo que respecta a los estudios entonativos. Se aborda por igual el problema de la diversidad entonativa dialectal y el análisis interdisciplinar de la entonación a través de sus relaciones con la pragmática, la kinésica o la proxémica.

En la misma medida en que en la actualidad resulta de suma relevancia el estudio aplicado de la fonética desde el punto de vista de su utilidad para la enseñanza de segundas lenguas, también constituyen un bloque relevante en este volumen los estudios dedicados a aspectos didácticos del componente fónico y aspectos metodológicos (herramientas y estrategias) útiles para la adquisición y aprendizaje de la pronunciación de segundas lenguas, lo que incluye la consideración de factores que pueden determinar dicha adquisición/aprendizaje (como la edad) o un interés específico por el componente prosódico en este ámbito didáctico.

Estudios más representativos de la aplicabilidad de la fonética a otras disciplinas se hallan, respectivamente, en los bloques dedicados a la fonética clínica (donde se abordan aspectos de la cuestión más generales, y otros más particulares como la afasia) y a la fonética forense (donde se asume el reconocimiento forense de la voz mediante el análisis de sonidos vocálicos, del ritmo lingüístico o la consideración del cambio de lenguas en el corpus analizado).

Existen también estudios que abordan la descripción del componente fónico desde una perspectiva más "clásica", como no podía ser de otra manera en un estudio de la cuestión que se pretenda global. A este ámbito corresponden los trabajos que se dedican al estudio normativo de la fonética o a la descripción de asuntos específicos relativos a problemas de índole acústica, articulatoria o perceptiva.

No faltan, en fin, algunos estudios, proyectados hacia el futuro de la fonética experimental, dedicados a las posibles aplicaciones tecnológicas del componente fónico.

El volumen que ahora se ofrece, pues, representa un ejemplo muy claro de innovación científica y voluntad de avance en el conocimiento de una de las disciplinas que, dentro de las llamadas ciencias humanas, constituye hoy por hoy, uno de ámbitos donde más posibilidades de desarrollo se aventuran en el futuro más próximo. El turno para opinar sobre estas ideas corresponde ahora al lector, a quien invitamos a acercarse a este *Análisis fónico del habla* sin prejuicios y con intención de actualizar sus conocimientos de la materia.

Antonio Hidalgo Navarro Co-director de Normas. Revista de Estudios Lingüísticos Hispánicos.

Parte I. Análisis suprasegmentales

Melodías para preguntar en portugués y en castellano. Un análisis contrastivo

ARAÚJO LEITE, MONIQUE; FONT-ROTCHÉS, DOLORS; HUELVA, ENRIQUE

1 Introducción

El creciente interés por los estudios sobre la entonación como un fenómeno lingüístico importante nos lleva a profundizar cada vez más nuestras investigaciones acerca de su función en la adquisición de lenguas.

Cantero (2002: 18) caracteriza la entonación como "las variaciones de frecuencia fundamental (F₀) que cumplen una función lingüística a lo largo de la emisión de voz". Es un fenómeno que actúa en tres niveles lingüísticos (prelingüístico, lingüístico y paralingüístico) a través de un proceso dinámico que integra y estructura el discurso (Cantero & Mateo, 2011).

El presente trabajo forma parte de una investigación más amplia, que tiene por objetivos establecer los patrones melódicos de todas las regiones del portugués de Brasil y desarrollar propuestas didácticas dirigidas a la enseñanza y al aprendizaje de lenguas¹. Para llevarla a cabo, se está utilizando el método *Análisis Melódico del Habla* (Cantero, 2002), el cual se ha demostrado muy adecuado para obtener datos precisos de la entonación para ser aplicados a la enseñanza de lenguas. Referente al portugués de Brasil y con esta metodología, hasta el momento ya se han descrito los patrones /+interrogativos/ de São Paulo (Mendes, 2013), de Goiás (Cantero & Font-Rotchés, 2013), de Río de Janeiro (Paixão, 2014) y de Minas Gerais (Araújo, 2014), los cuales son

¹Esta investigación se enmarca dentro de los trabajos del Grup de Recerca en Entonació i Parla (2014 SGR169) y del proyecto I+D Análisis del habla y modelos didácticos (FF12013-41915-P) financiado por el Ministerio de Economía y Competitividad.

coincidentes, pero con algunos rasgos melódicos propios de cada variedad estudiada, así como también se ha caracterizado la interlengua del español hablado por portugueses de Brasil (Fonseca, 2013).

Ahondando en aspectos de la entonación, en este artículo presentaremos un análisis comparativo entre los patrones /+interrogativos/de las preguntas absolutas del portugués de Brasil (PB) hablado en la provincia de Minas Gerais, obtenidos en nuestras investigaciones previas (Araújo, 2014), con los establecidos para el castellano peninsular (Cantero & Font-Rotchés, 2007; Font-Rotchés & Mateo, 2011, 2013) y evidenciaremos las semejanzas y diferencias entre los patrones de ambas lenguas, que pueden favorecer o dificultar la comunicación a los aprendientes brasileños de español y a los aprendientes españoles de portugués de Brasil.

2 Corpus y metodología

En una investigación previa del portugués de Brasil en habla espontánea de la provincia de Minas Gerais (Araújo, 2014), hemos establecido cuatro patrones melódicos /+interrogativos/, descritos a partir de 60 preguntas absolutas producidas por un total de 32 informantes (22 hombres y 10 mujeres). Estos enunciados han sido extraídos de grabaciones audiovisuales de diversos programas de la televisión brasileña, que corresponden a los géneros de *Talk show y Reality Show*, así como también de grabaciones ocultas realizadas a nativos brasileños en una situación de habla espontánea.

El método de análisis empleado en nuestra investigación es el método *Análisis Melódico del Habla* (AMH), propuesto en Cantero (2002), aplicado y ampliado en Font-Rotchés (2007) y expuesto en forma de protocolo en Cantero & Font-Rotchés (2009). Se trata de un método que permite el análisis de habla espontánea y que, frente a otros métodos, ofrece un criterio de segmentación de las melodías del habla exclusivamente fónico, independiente de otros niveles de análisis.

Los instrumentos de extracción y análisis de los datos que hemos utilizado para nuestra invetigación fueron el software Praat (Boersma & Weenink, 1992-2012) y el Script de Praat para la semiautomatización del proceso, elaborado y propuesto por Mateo (2010).

Para el análisis de los enunciados, se extraen los datos acústicos en Praat – valores absolutos en hercios– se estandarizan (para un proceso manual, véase todos los pasos expuestos de forma minuciosa en el protocolo de Cantero & Font-Rotchés, 2009, y para el de semiautomatización, véase Mateo, 2010). Los valores relativos de los enunciados, aptos para poder realizar comparaciones entre voces de informantes de distinto sexo y edad, se presentan en gráficos para poder clasificarlos, compararlos y preparar pruebas perceptivas que nos permitan establecer generalizaciones lingüísticas.

Al compartir una misma metodología, la investigación sobre el portugués de Brasil y la del español peninsular, nos ha permitido comparar los rasgos melódicos de los contornos entonativos de cada corpus. Para ello, nos hemos basado en la estructura básica del contorno de la figura 1.

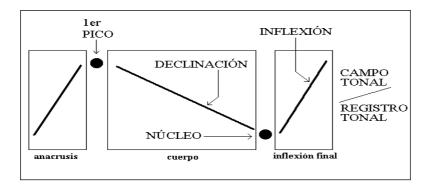


Figura 1. Estructura del contorno entonativo (Cantero & Font-Rotchés, 2007)

El contorno entonativo está constituido por tres partes separadas por sílabas tónicas de gran relevancia entonativa: *anacrusis*, se trata de las sílabas átonas anteriores a la primera sílaba tónica del contorno o *primer pico*; *cuerpo*, va desde el primer pico hasta la última vocal tónica o *núcleo*; e *inflexión final*, se sitúa en las sílabas posteriores al núcleo y llega hasta el final, y aporta la carga significativa más importante del contorno.

3 Los patrones /+interrogativos/

En este apartado, describiremos los cuatro patrones /+interrogativos/ que hemos obtenido para el portugués de Minas Gerais y los cuatro del español peninsular, además de sus variedades, su presencia en los corpus y algunos ejemplos obtenidos.

3.1 Los patrones interrogativos del portugués de Brasil

El Patrón A. Inflexión final ascendente (30 % \sim 60 %). Este patrón, caracterizado por una inflexión final ascendente que se sitúa entre un 30 % y un 60 %, posee dos variaciones: en A.1 no hay presencia de primer pico y el cuerpo se mantiene plano; en A.2, hay un primer pico prominente, seguido de una leve declinación (véase en la figura 2).

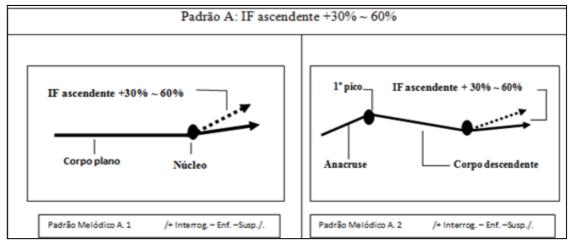


Figura 2. El patrón melódico interrogativo A del portugués de Brasil (Araújo, 2014)

En nuestro corpus hemos hallado un total de 20 contornos (33,3 %) de los cuales 9 siguen el patrón melódico A.1, y 6, el A.2. Además, hay 5 enunciados que solo presentan inflexión final.

En el contorno de la figura 3, que sigue el patrón A.1, podemos observar que el contorno melódico carece de primer pico, porque el ascenso es inferior a un 10 %; en cuanto al cuerpo, es casi plano –solo desciende un 9 %–, mientras que la inflexión final tiene un ascenso de un 59,5 %.

En el contorno de la figura 4, en cambio, sí hay presencia de primer pico, con un ascenso muy leve, de un 10,2 %, en *então*, seguido de un descenso significativo en el cuerpo y una inflexión final de 54 %.

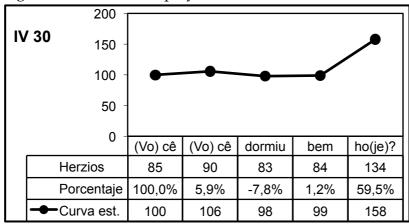


Figura 3. Gráfico del enunciado Você dormiu bem hoje? '¿Has dormido bien hoy?'

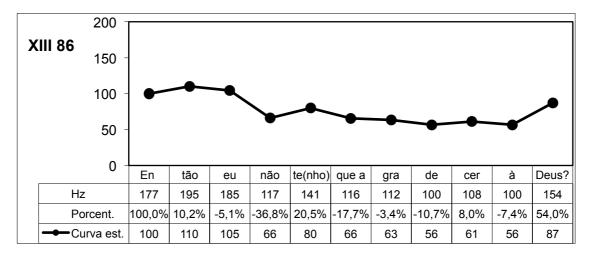


Figura 4. Gráfico del enunciado Então eu não tenho que agradecer à Deus? 'Así pues, ¿no tengo que agradecérselo a Dios?'

El Patrón B. Inflexión final (ascendente + 30 % ~ descendente -30 %). También este patrón, con una inflexión final con una horquilla de +30 % a -30 %, presenta dos variedades condicionadas por los rasgos del primer pico y del cuerpo: en B.1 no hay presencia de primer pico y el cuerpo se mantiene plano; en B.2 hay primer pico, seguido de una leve declinación (véase en la figura 5).

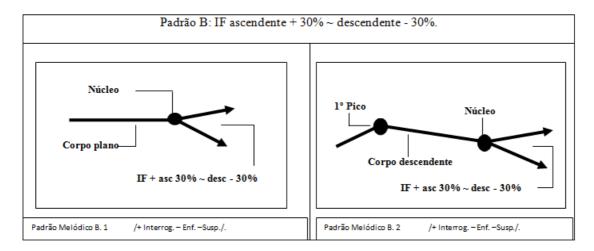


Figura 5. El patrón melódico interrogativo B del portugués de Brasil (Araújo, 2014)

En el corpus hay un total de 21 contornos, que equivalen a un 35 %: 11 siguen el patrón B.1; 7, el B.2, y 3 solo presentan inflexión final.

El contorno de la figura 6 contiene los rasgos del patrón B.1: ausencia de primer pico, cuerpo plano e inflexión final con un descenso de un 26,1 %.

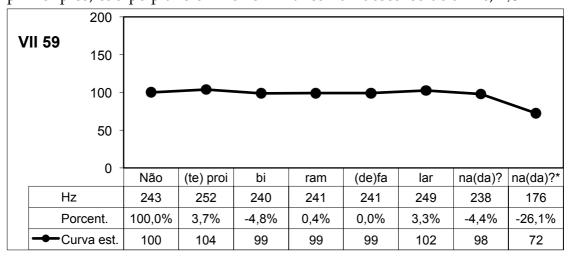


Figura 6. Gráfico del enunciado Não te proibiram de falar nada? '¿No te prohibieron de decir nada?'

El Patrón C. Inflexión final con núcleo elevado (véase figura 7). Se trata de un patrón que no presenta primer pico, el cuerpo se mantiene plano y la inflexión final posee el núcleo o la sílaba tónica final, que se encuentra en la cúspide de un ascenso y que varía de un 10 % hasta un 40 %, seguida de un descenso hasta llegar al final. En nuestro corpus hemos hallado un total de 12 contornos (18,5 %).

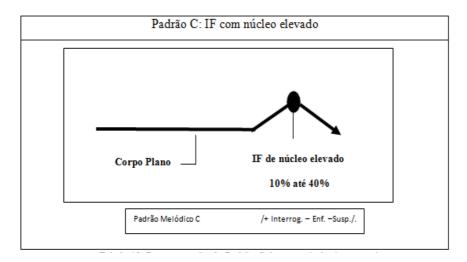


Figura 7. El patrón melódico interrogativo C del portugués de Brasil (Araújo, 2014)

En la figura 8, se encuentra el contorno de la pregunta ¿Ah, o Tatá voltou pra novela? '¿Ah, el Tatá ha regresado a la telenovela?', que contiene los rasgos melódicos del patrón C: no hay primer pico, la declinación del cuerpo permanece prácticamente toda plana y la inflexión final se destaca por su núcleo elevado que culmina en -ve-, en la cima de un ascenso de un 17,3 %, y que le sigue un descenso de un 21,6 % hasta el final.

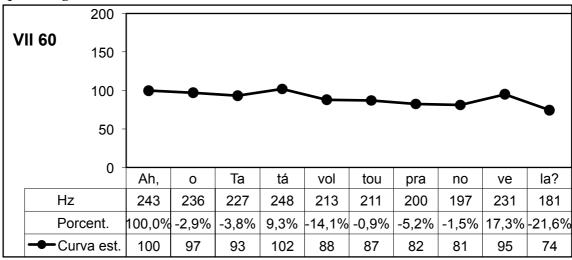


Figura 8. Gráfico del enunciado ¿Ah, o Tatá voltou pra novela? 'Ah, El Tatá ha regresado a la telenovela?'

El Patrón D. Inflexión final ascendente-descendente (véase figura 9). Se trata de un patrón que no presenta primer pico, el cuerpo se mantiene plano y la inflexión final es circunfleja, es decir, tiene dos direcciones, un ascenso y un descenso, y tres segmentos tonales. El ascenso de la inflexión final suele estar entre un 10 % y un 45 %. En nuestro corpus hemos hallado un total de 8 contornos (13,3 %).

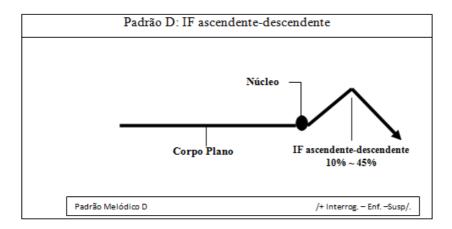


Figura 9. El patrón melódico interrogativo D del portugués de Brasil (Araújo, 2014)

En la figura 10, vemos un contorno con primer pico y después de un leve ascenso, $mas\ \acute{e}\ s\acute{o}$, que posteriormente desciende en el cuerpo abruptamente, un 35 %, para dar inicio a la inflexión final, que empieza en la sílaba se-. Esta sílaba tónica se alarga en un ascenso de un 45,8 % y seguidamente desciendeun 37,9 % hasta -xo.

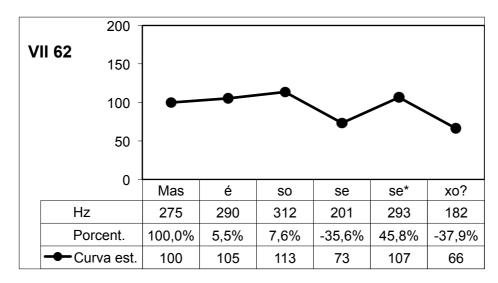


Figura 10. Gráfico del enunciado Mas é só sexo? ('¿Pero es solo sexo?')

3.2 Los patrones interrogativos del español peninsular

Acerca del corpus de las interrogativas del castellano peninsular, se describieron los patrones II, III y IV en Cantero & Font-Rotchés (2007). Posteriormente, en Font-Rotchés & Mateo (2011), se presentó el patrón XIII y en Font-Rotchés & Mateo (2013), se demuestra la presencia y la distribución de todos los patrones en la zona meridional y Canarias. Se trata de un corpus con 186 enunciados-pregunta, emitidos por 110 informantes, 44 hombres y 66 mujeres, obtenidos a partir de grabaciones de programas televisivos, básicamente de programas producidos en canales de las zonas dialectales estudiadas, y que fueron emitidos en contextos de debates, concursos, entrevistas, reportajes temáticos y programas similares.

El Patrón melódico II. Inflexión final ascendente (+70%). Este patrón se caracteriza por un primer pico en la cima de un ascenso que puede llegar a ser

de hasta un 40 % y que suele tener lugar en la primera sílaba tónica, un cuerpo con un descenso suave y una inflexión final con un ascenso igual o superior a un 70 %, que supera de forma visible el primer pico. Se encuentra en 30 enunciados, un 27,5 % (véase en la figura 11).

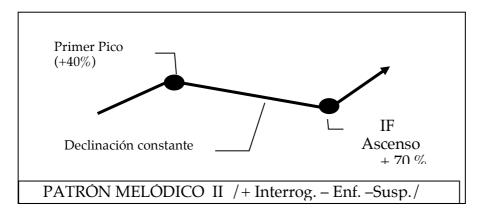


Figura 11. Patrón II del español peninsular (Font-Rotchés & Mateo, 2013)

En la figura 12, vemos un contorno que sigue el patrón II, con un primer pico que culmina un ascenso de un 22,7 %, un cuerpo con un descenso suave y una inflexión final superior a un 70 %.

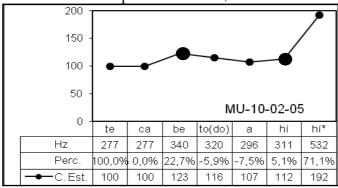


Figura 12. Gráfico del enunciado ¿Te cabe todo ahí?

El Patrón melódico III. Inflexión final ascendente ($40\% \sim 70\%$). Este patrón, caracterizado por una inflexión final ascendente que se sitúa entre un 40% y un 70%, presenta un primer pico desplazado a una sílaba átona posterior y un cuerpo con un descenso suave. Siguen este esquema 31 preguntas absolutas del corpus, un 28,5% (véase en la figura 13).

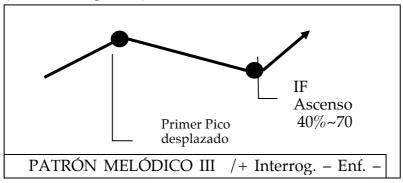


Figura 13. Patrón III del español peninsular (Font-Rotchés & Mateo, 2013)

En la figura 14 encontramos un ejemplo de contorno, que sigue el patrón III, en el que el primer pico se encuentra en *per*-, sílaba átona posterior a la primera tónica. Además presenta un cuerpo con un descenso muy suave y una inflexión final con un ascenso de un 53,4 %.

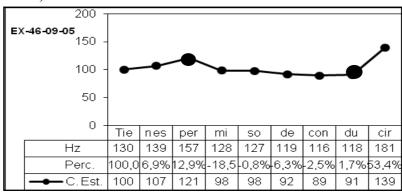


Figura 14. Gráfico del enunciado ¿Tienes permiso de conducir?

El Patrón melódico IV. Inflexión final ascendente-descendente. Este patrón presenta dos variantes (véase figura 15), según las características del primer pico y del cuerpo: IVa, primer pico en un punto elevado y desplazado a una átona posterior y cuerpo descendente; IVb, cuerpo en declinación plana y no presenta primer pico. En la inflexión final, con un movimiento ascendente-descendente, el primer valor debe empezar en la última sílaba tónica y presentar, desde este punto en adelante, un total de tres valores y dos direcciones, un ascenso y un descenso. Los contornos que siguen este patrón son menos frecuentes, 16, que el resto, un 14,7 %, y, entre ambas variedades predomina el tipo IVa.

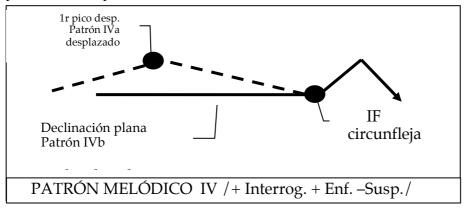


Figura 15. Patrón IV del español peninsular (Font-Rotchés & Mateo, 2013)

En la figura 16 tenemos un ejemplo de contorno que sigue el patrón IVa: primer pico en la cúspide de un ascenso de un 16,7 %, un descenso suave y una inflexión final, que empieza en la última sílaba tónica, viu-, asciende un 54,3 % hasta -da, sílaba que se alarga y desciende hasta el final, un 30,2 %.

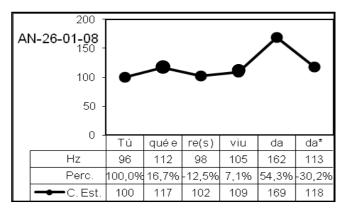


Figura 16. Gráfico del enunciado ¿Tú qué eres, viuda?

El Patrón melódico XIII. Cuerpo e Inflexión final ascendentes. Este patrón se caracteriza por un cuerpo ascendente, en general, sin primer pico, y una inflexión final ascendente. El ascenso se calcula global y se estableció en un 90 % para percibir los contornos como enfáticos (Font-Rotchés & Mateo, 2011). Las preguntas del corpus que siguen este patrón son 32, un 29,3 %, una presencia parecida al de los patrones II y III. Véase en la figura 17.

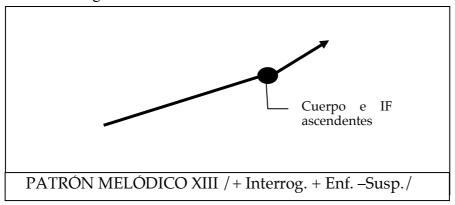


Figura 17. Patrón XIII del español peninsular (Font-Rotchés & Mateo, 2013)

El contorno melódico de la figura 18 coincide con la estructura del patrón XIII: empieza con el primer pico en la sílaba *Pe*-, le sigue un descenso, 13,3 %, y continúa con un cuerpo en dirección ascendente, que culmina en la inflexión final, *allí*, también ascendente. En este caso, el ascenso total es de 124 %.

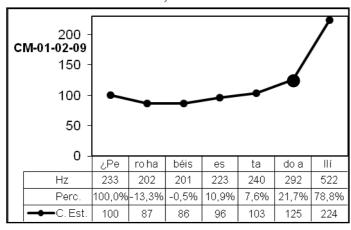


Figura 18. Gráfico del enunciado ¿Pero habéis estado allí?

4 Análisis contrastivo y conclusiones

A partir de los resultados de las investigaciones de Araújo (2014) y Cantero & Font-Rotchés (2007) y Font-Rotchés & Mateo (2013) sobre las preguntas absolutas del portugués de Brasil (hablado en Minas Gerais) y las del español peninsular, respectivamente, hemos podido observar que hay diferencias y aproximaciones significativas entre los patrones de ambas lenguas.

En el patrón A de las interrogativas del portugués de Brasil, la inflexión final es ascendente con una horquilla de 30 % a 60 % y predominan contornos sin primer pico y cuerpo plano (variante A.1). Los patrones II y III del español peninsular también poseen la inflexión final ascendente, igual o superior a un 70 %, el patrón II, y entre un 40 % y un 70%, el III, pero sus contornos se caracterizan, de forma general, por un cuerpo en un descenso leve y un primer pico, que tiene lugar en la primera sílaba tónica, en el II, y en una átona posterior, en el III. Estos patrones con final ascendente son los más frecuentes en ambas lenguas.

El patrón B de las interrogativas del portugués presenta una inflexión final ascendente +30 % o descendente -30 %. Este patrón probablemente dejará de ser /+interrogativo/ cuando se exponga a pruebas de percepción, por el mero hecho que se trata de contornos que tienen lugar en un contexto favorecedor para que se interpreten como interrogativos, pero que no se reconocen cuando se emiten aislados. Esto ocurre en contornos de preguntas absolutas del español con un ascenso inferior a un 40 % o con un descenso.

El patrón D del portugués y el patrón IV poseen una inflexión final circunfleja ascendente-descendente y coinciden especialmente el patrón D y la variante IVb por la ausencia de primer pico y la presencia de cuerpo plano; sin embargo, la variante IVa del castellano, que es la más utilizada, presenta primer pico y un cuerpo en descenso suave.

Finalmente, el patrón C del portugués, con inflexión final de núcleo elevado, y el patrón XIII del español, con cuerpo e inflexión final ascendentes, no tienen un patrón equivalente en la otra lengua.

Así, pues, aunque haya coincidencias en la dirección de la inflexión final, cabe decir que los contornos del portugués de Brasil del patrón A presentan ascensos menos marcados, entre un 30 % y un 60 %, que los del el patrón II del castellano que son iguales o superiores a un 70 %. Estos contornos del patrón A del portugués son más parecidos a los del patrón III del castellano. Referente a los patrones D (del portugués) y IV (del castellano, con inflexión final circunfleja), la horquilla de ascenso es parecida.

El rasgo de la inflexión final no es el único que distancia algunos patrones de ambas lenguas, sino que también los hay en el primer pico y en el cuerpo. De forma general, los patrones del portugués de Brasil tienen una tendencia muy pronunciada a realizar los contornos interrogativos sin primer pico y cuerpo plano; en cambio, en castellano, suelen tener primer pico, con un ascenso de hasta un 40 %, y un cuerpo con un descenso suave. Se trata pues de rasgos distintivos relevantes entre ambas lenguas a los que tenemos que añadir las diferencias en la inflexión final y la existencia de patrones, el C del portugués y el XIII del castellano, que no tienen equivalente en la otra lengua.

Tras realizar el análisis contrastivo de los rasgos melódicos de estos patrones, nos hemos dado cuenta de que las interrogativas de ambas lenguas, aunque *a priori* parecía que compartían bastantes semejanzas, no es así. Solo los patrones D y IVa tienen la misma estructura, el resto presentan diferencias relevantes en las diversas partes del contorno.

Tales diferencias entre los patrones presentados anteriormente demuestran que los hablantes nativos del portugués de Brasil y del castellano peninsular pueden sufrir dificultades en el proceso de adquisición de esas lenguas. Si un aprendiente utiliza los patrones melódicos que está acostumbrado en su lengua materna al realizar una pregunta en la lengua que está adquiriendo, corre el riesgo de que su interlocutor no lo comprenda o cometa una mala interpretación de su enunciado. De esta manera, surgen la confusión o los malentendidos en la comunicación, dado el uso de patrones melódicos inadecuados, que afectan la eficacia y la fluidez del discurso.

Por consiguiente, es imprescindible dar continuidad a estas investigaciones puesto que pueden contribuir a estimular el desarrollo de posibles aplicaciones didácticas que beneficien la adquisición de la entonación del portugués de Brasil y del español peninsular.

5 Bibliografía

- Araújo, M. L. (2014). Entonação das interrogativas e das declarativas do português brasileiro falado em Minas Gerais: Modelos para o Ensino de Línguas. Dissertação de Mestrado. Departamento de Línguas Estrangeiras e Tradução, Universidade de Brasília.
- Boersma, P. & Weenink, D. (1992-2012): Praat: doing Phonetics by Computer. Institute of Phonetic Sciences, University of Amsterdam. En http://www.praat.org
- Cantero, F. J. (2002). Teoría y análisis de la entonación. Barcelona: Ed. de la Universitat de Barcelona.
- Cantero, F. J. & Font-Rotchés, D. (2007). Entonación del español peninsular en habla espontánea: patrones melódicos y márgenes de dispersión. Moenia, 13, 69-92.
- Cantero, F. J. & Font-Rotchés, D. (2009). Protocolo para Análisis Melódico del Habla. Estudios de Fonética Experimental, XVIII, 17-32.
- Cantero, F. J. & Font-Rotchés, D. (2013). The Intonation of absolute questions of Brazilian Portuguese. Linguistics and Literature Studies, 1 (3), 142-149.
- Cantero, F. J. & Mateo, M. (2011). Análisis Melódico del habla: complejidad y entonación en el discurso. Oralia. Análisis del discurso oral, 14, 105-128.
- Fonseca, A. (2013). Caracterización del español hablado por brasileños. Tesis Doctoral. Universitat de Barcelona. Recuperado el 14 de 1 de 2015, de http://diposit.ub.edu/dspace/bitstr

- eam/2445/54688/5/03.AFdO_3de3. pdf.
- Font-Rotchés, D. (2007). L'entonació del català. Barcelona: Publicacions de l'Abadia de Montserrat.
- Font-Rotchés, D.; Mateo, M. (2011). Absolute interrogatives in Spanish, a new melodic pattern. Anais do VII Congresso Internacional da Abralin, Curitiba (Brasil), 1111-1126.
- Font-Rotchés, D.; Mateo, M. (2013). Entonación de las interrogativas absolutas del español peninsular del sur en habla espontánea. Onomázein, 28, 256-275.
- Mateo, M. (2010): Protocolo para la extracción de datos tonales y curva estándar en Análisis Melódico del Habla (AMH)". *Phonica*, 6, 49-90. Recuperado el 14 de 1 de 2015, http://www.publicacions.ub.edu/re vistes/phonica6/documentos/702.p df
- Mendes, R. S. (2013). A entonação no processo de ensino-aprendizagem de PLE: proposta didática para o ensino de modelos de entonação interrogativa do português do Brasil Estado de São Paulo. Dissertação de Mestrado. Departamento de Línguas Estrangeiras e Tradução. Universidade de Brasilia.
- Paixão, V. B. (2014). A prosódia das interrogativas totais na fala carioca: Fala espontânea versus leitura. Dissertação de Mestrado. Faculdade de Letras. Universidade Federal do Rio de Janeiro.

Intonation of Yes/No question in Portuguese: analysis of the relationship between intonation and lexical stress in the AMPER-POR corpus

CRUZ, REGINA; SEARA, IZABEL; MOUTINHO, LURDES

1 Introduction

This work confronts three varieties of Portuguese, taking into account the relationship between intonation and lexical stress in yes/no question from the AMPER-POR corpus. The Portuguese varieties here analyzed are two spoken in Brazil (BP), one dialect from North of Brazil (Belem) and other from South (Florianopolis). Both dialects have a common base in their origin, the Portuguese spoken in Azores Island. For this reason, we compared the data of these two BP varieties with the Portuguese spoken in the Azores (EP) Island. To carry out the analysis itself, four sentences representing each kind of Portuguese lexical stress Portuguese were chosen - oxytone (cv.cv.CV), paroxytone (cv.CV.cv) and proparoxytone (CV.cv.cv) - a total of 12 sentences analyzed from each speaker were used, each sentence repeated three times to

the present analysis that focused on yes/no question of AMPER-POR corpus. We analyzed data from two speakers of each dialect -one male and one female-with the lowest school level. Therefore, the final corpus analyzed for this study consists of 72 utterances (2 speakers x 3 dialects x 4 sentences x 3 lexical stresses) of the type Subject Verb Complement (SVO). More precisely the movements of F0 variation in the last element of the sentence were verified. For this, we took averages in semitones (ST) of F0 movements on the vowels as determined by AMPER project.

We organized this paper in four parts. Firstly, we will present information on similar historical and sociolinguistic aspects of the Portuguese dialects analyzed. After that, we will describe the methodology used to data analysis and how it is structured the Portuguese corpus for the AMPER project. Finally, we will present the results of this analysis.

2 An Intralinguistic analysis

We analyze three Portuguese dialects in this paper. Two BP dialects, one from North, spoken in Belem, and other from South, spoken in Florianopolis. We confronted these two BP dialects with an EP dialect, the EP spoken in one Azores Island, in Sao Miguel Island. As these dialects have similar historical aspects in their origin, we compare their data in order to verify if they have any common prosodic characteristics.

2.1 Historical motivation

Both Florianopolis (1726) and Belem (1616) are Brazilian cities founded in the Brazilian colonial period (1500-1882) and these cities received for a long time Portuguese speakers from Azores as immigrants during the colonial period (Furlan, 1989), Florianopolis (1748 - 1756) and Belem (1645-1876). Florianopolis received 6071 Azores immigrants, a number which caused the population increase of 144,6 %. From 1855 until 1856, 7309 Azores people moved to Brazil, whose 3.371 were from Sao Miguel Island. Several historians registered Amazonia as a final destination of the Azores immigrants. Timothy Coates (1998) and Arthur Reis (1993 *apud* Chambouleyron, 2008), for example, note several and different periods of the arrival of Azores people to Belem: in 1675, 50 couples or 234 people arrived to Para; in 1677 other 50 couples or 235 people and in 1752, 430 Azores people arrived to Belem.

According to Furlan (1989: 31), the Azores migration to Northern and Southern Brazil had a politic and military strategy to keep Brazil under Portuguese Kingdom domain.

"Foi para diminuir a sobrecarga demográfica dos Açores e para tomar posse dos extremos Norte e Sul do Brasil, expostos à invasão espanhola, que a Coroa portuguesa estimulou e dirigiu de 1748/1756 a vinda de açorianos e madeirenses ao Norte (Maranhão e Grão Pará) e ao Sul (hoje Rio Grande do Sul e Santa Catarina)"

Chambouleyron (2008: 190) confirms this fact:

"O fato é que, como o próprio Timothy Coates indica, durante o século XVII, as capitanias do Maranhão e do Pará "mantiveram uma ligação com os Açores", caracterizada pela migração de indivíduos decorrente de "sobrepopulação e

uma série de terramotos". Para esse autor, foram enviados moradores das ilhas nos anos de 1619, 1622, 1649, 1667, 1673, 1674 e 1677".

Both Brazilian Portuguese varieties have the same Portuguese dialect in their origins. For this reason, we decided to compare these three dialects in order to find any similarities at the prosodic level.

2.2 Sociolinguistic cues

The BP spoken in Belem and Florianopolis has same sociolinguistic similarities, such as a similar personal pronominal pattern and a strong palatalization of alveolar fricative in coda position.

The majority of Brazilian Portuguese dialects do not use more 'TU' (you) as a second person pronoun, as these dialects prefere 'VOCÊ' (Lopes, 2009). The use of TU as second person is still productive in extreme north and extreme south of Brazil, more precisely in Belem (Soares & Leal, 1993) and Florianopolis dialects (Loregian-Penkal, 2012).

According Loregian-Penkal (2004:12) "a presença de flexão verbal de segunda pessoa constitui-se como marca de identidade do falar do ilhéu de Florianopolis".

Scherre (2007) shows that in relation to Florianopolis:

"100% na fala espontânea de Florianópolis – região Sul –, (BONFÁ, PINTO; LUIZ, 1997, p.10-11), área de alternância do pronome tu/você com predomínio do uso de tu - 76%, com 43% com morfologia verbal explícita, ou seja, com concordância: dados da década de 90 (Loregian-Penkal, 2004, p.133,167)"

The same author analyzed data from Soares & Leal (1993) in relation to Belem and states that "para a região Norte, estado do Pará, Soares e Leal (1993, p.51) evidenciam que a cidade de Belém exibe 77% de uso global do tu com concordância variável" (Scherre, 2007: 203).

The use of 'TU' as a second person is a conservative variant in BP. The BP dialects spoken in Belem and Florianopolis have another conservative variant, a strong palatalization of the alveolar fricative in coda position (Brescancini, 2002, 1996; Šmaiclová, 2010).

Brescancini (2002) showed that palatalization of alveolar in coda position is a strong process in BP spoken in Florianopolis, she found 83% of palatalization cases of alveolar in coda position in her data (Haupt & Berri, 2009).

Carvalho (2000) found 69% of palatalization cases of the alveolar sound in coda position in her data from BP spoken in Belem.

Jesus (2013) analyzed BP data from 25 Brazilian capitals and found the same distribution between alveolar and palatal fricative in coda position for Belem and Florianopolis. In both cities there is a categorical palatalization of /S/ in the words like 'prostituta', but there is a strong variation in /S/ in the words such as 'estilingue', in this case BP spoken in both cities has two possibilities an alveolar or a palatal for pronunciation of the /S/ in the coda position.

One question motivated this work:

What do these dialects (EP spoken in Azores and BP spoken in Florianopolis and Belem) have in common with respect to prosodic/intonation level?

3 Methodology

Some informations about AMPER corpus is important for understanding the comparative analysis here stablished.

3.1 Features of AMPER-POR¹ corpus

The AMPER project coordination has determined common guidelines for corpus formation:

- there is a unique corpus for Portuguese, with some semantic adaptations;
- the corpus is composed by 66 sentences;
- the speaker doesn't supposed to have contact with written form of the sentences during fieldwork;
- at least two native speakers are recorded, socially stratified by sex and school level:
- only speaker above 30 years old are chosen for participating to fieldwork

All methodological procedures, encompassing 5 steps, are determined by AMPER project coordination, in order to keep a homogeneity for the corpora formed.

The Portuguese corpus is controlled syntactic, phonetic and prosodically.

3.1.1 Syntactic Features

Syntactically the corpus has only SVO - Subject, Verb and Complement - sentences. The nominal phrases (NP) can be placed whether at the subject or complement position. There are two kind of NP: 1) simples, formed only by a determinant plus a noun or; 2) expanded, formed by a determinant plus a noun plus an adjective (AP) or an adverb (PP).

3.1.2 Phonetic Features

The sentences have the same size, same number of syllables. The Portuguese corpus has the sentence with 10, 13 or 14 syllables. The noun placed at the intonation core of phrases represents the three kinds of lexical stress in Portuguese: a) oxytone stress (cv.cv.CV); b) paroxytone stress (cv.CV.cv) or c) proparoxytone stress (CV.cv.cv).

3.1.3 Prosodic Features

At prosodic level, each sentence is pronounced in two intonation modalities, declarative and yes/no question. This total allowed us the comparison established for this work.

3.2 Data from AMPER-POR Corpus

In order to verify our working hypothesis, we chose only data of speakers with lowest school level for this work. We analyzed data from one female and one male of each Portuguese dialect. You can see each dialect analyzed at the table 1 and its speaker identified by a code.

¹ http://pfonetica.web.ua.pt/AMPER-POR.htm.

city	dialect	Female speaker code	Male speaker code
Belem	BP	BE01	BE02
Florianopolis	BP	BB01	BB02
Azores	EP	01c1	01c2

Table 1. The Portuguese speaker's profile

We decided to analyze only yes/no question of AMPER-POR corpus. We have chosen four yes/no questions of each Portuguese lexical stress. The stress in Portuguese can fall on one of three last syllables of the word as you can note at the Table 2 below:

Lexical Stress	stress pattern
Proparoxytone	CV.cv.cv
Paroxytone	cv.CV.cv
Oxytone	cv.cv.CV

Table 2. Portuguese lexical stress

If the stress is on the final syllable of the word, this word has is oxytone lexical stress. This kind of stress is more productive in the borrowing words in Portuguese, like words from French, African and Amerindian languages found in Portuguese lexical inventory. The AMPER corpus has the following oxytone words: for the BP, 'bisavô' (grandfather), 'nadador' (swimmer) and 'Salvador' (name of a Brazilian city); for the EP, 'capataz' (foreman), 'popular' and 'Canadá' (Canada country). Only the oxytone words placed at the end of sentence are analyzed here. You can see the selected sentences for each variety and their oxytone words with their morphological and syntactic features at the Table 3.

code	phrase	grammar classification	BP	EP
pwki	NP)	Noun	O pássaro gosta do bisavô?	A música fala do capataz?
pwgi	AP	Adjective	O pássaro gosta do bisavô nadador?	A música fala do capataz popular ?
pydi	PP	Adverb	O pássaro gosta do Renato de Salvador ?	A música fala do fadista do Canadá?

Table 3. The AMPER-POR corpus' sentences with oxytone word in the final NP

If the stress falls on the penultimate syllable of the word, this word has a paroxytone lexical stress. The AMPER corpus has the following paroxytone words: for BP, 'Renato' (a personal name), 'pateta' (goofy) and 'Veneza' (the Italian city Venice); for the EP, 'fadista' (the traditional Portuguese singer), 'castiço' (meaning "of good greed", no reference was found in English) and 'Capelas' (chapels). This kind of stress is the common stress pattern in Portuguese. This is the same Latin metrical structure. The majority of the Portuguese words have the stress falling on the penultimate syllable. You can see the selected sentences for each variety and their paroxytone words with their morphological and syntactic features at the Table 4.

code	phrase	grammar classification	ВР	EP
pwti	NP	Noun	O pássaro gosta do Renato?	A música fala do fadista ?
pwsi	AP	Adjective	O pássaro gosta do Renato pateta ?	A música fala do fadista castiço?
pysi	PP	Adverb	O pássaro gosta do Renato de Veneza ?	A música fala do fadista das Capelas?

Table 4. The AMPER-POR corpus' sentences with paroxytone word in the final NP

If the stress falls on the antepenultimate syllable of the word, this word has a proparoxytone lexical stress. The AMPER corpus has the following proparoxytone words: for the BP, 'pássaro' (bird), 'bêbado' (drunk) and 'Mônaco' (Monaco); for the EP, 'música' (music), 'típica' (tipical) and 'México' (Mexico). The proparoxytone words in Portuguese have a lower frequency of occurrence in relation to paroxytone words and even in relation to oxytone words. Its low frequency is justified by the fact the proparoxytone words were introduced in the Portuguese system quite late, after the Portuguese language had its linguistic autonomy from the Latin language. You can see the selected sentences for each variety and their proparoxytone words with their morphological and syntactic features at the Table 5.

code	phrase	grammar classification	BP	EP
twpi	NP (SN)	Noun	O Renato gosta do pássaro ?	O fadista fala da música?
twvi	AP (SAdj)	Adjective	O Renato gosta do pássaro bêbado ?	O fadista fala da música típica?
pyzi	PP (SPrep)	Adverb	O pássaro gosta do Renato de Mônaco ?	A música fala do fadista do México?

Table 5. The AMPER-POR corpus' sentences with proparoxytone word in the final NP

Besides these words, we have chosen three more yes/no questions from each variety and these yes/no questions are available at the Table 6.

code	stress	BP	EP
pwpi	CV.cv.cv	O pássaro gosta do pássaro?	A música fala da música?
twti	cv.CV.cv	O Renato gosta do Renato?	O fadista fala do fadista?
kwki	cv.cv.CV	O bisavô gosta do bisavô?	O capataz fala do capataz?

Table 6. The AMPER-POR corpus' sentences with the same word in the both phrase positions

The nouns 'pássaro', 'Renato' and 'bisavô' from BP and 'capataz', 'fadista' and 'música' from EP are placed at the both syntactic positions, subject and object, in the same sentence. We start the presentation of our analysis by these yes/no question.

4 Resultats

The Figure 1 presents the results of the sentence with the oxytone nouns - bisavô from Brazilian Portuguese and capataz from European Portuguese - in both initial and final position in the sentence. In order to establish a comparison the data, we convert the F0 measures of Hertz in semitones (ST). The white line

illustrates the data from EP. The black line illustrates data from the BP. The full line illustrates the data from the male speaker. The dotted line illustrates the data from the female speaker. The different markers represent the different lexical stresses: a sphere represents the oxytone stress (cv.cv.CV); a square marker indicates the paroxytone stress; and a diamond signals the proparoxytone stress. The Figure 1a shows the F0 variations overall sentence. The Figure 1b shows the F0 variations only on the end of sentence.

In according to our hypothesis, this position - final NP – supposedly has the main F0 variations for discriminate the dialects. You can note that there is a convergent movement of F0 at the stressed syllable for every dialect, we can note a small difference of F0 movements at the unstressed syllables.

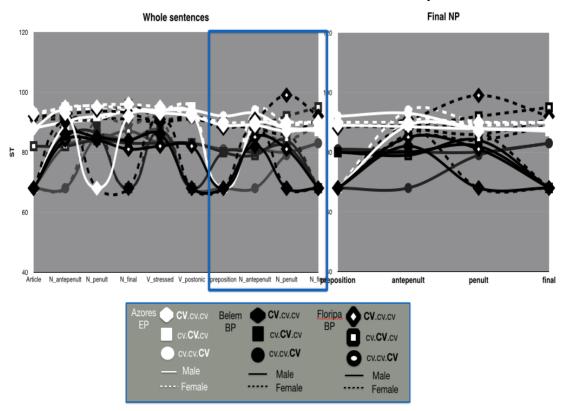


Figure 1. The F0 variation (ST) at the kwki, twti and pwpi sentences with the same oxytone (cv.cv.CV), paroxytone (cv.CV.cv) and proparoxitone (CV.cv.cv) word in the core of both initial and final NPs

From this point on, we will show only the F0 movements on the word placed on the final position of the sentence. We will see the F0 movements in each set of word according to lexical stress.

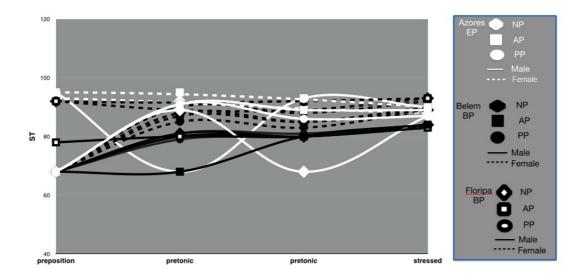


Figure 2. The F0 variation (ST) at the final of sentence with the set of oxytone words (cv.cv.CV)

The Figure 2 shows the F0 curves on final NP with the oxytone words placed at the end of the sentences. We can state that the F0 curves at the end of the sentence with the oxytone word in all dialects converge to stressed syllable, the final syllable of the word. The variations are registered at The unstressed syllables of the phrase, registering the relevant F0 variation.

The same situation of the oxytone word is observed in the yes/no question with the proparoxytone word (CV.cv.cv) placed at the final position.

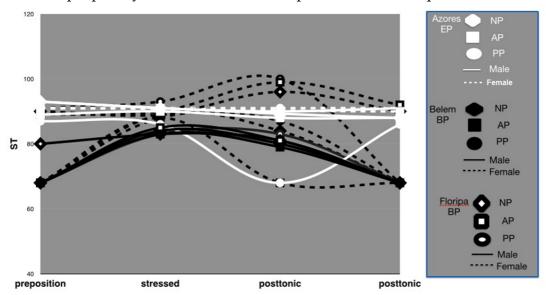


Figure 3. The F0 variation (ST) at the final of sentence with the set of proparoxytone words

The F0 curve pattern of each dialect is more evident when the proparoxytone word is place at the final of sentence. In this case, yes/no questions of Belem have more F0 regular curves than Florianopolis and Azores.

We can find a well-formed pattern for the set of paroxytone words, the most common kind of lexical stress in Portuguese language. Every Portuguese variety has a F0 movement pattern for paroxytone words placed at the final position of yes/no question. The Azores data gives us the impression of a mixed pattern of both of Brazilian Portuguese varieties. We can note a circumflex movement of the F0 curve at the core region for all of paroxytone

words in each Portuguese dialect. However, it appears that there is a difference of alignment of the F0 peak.

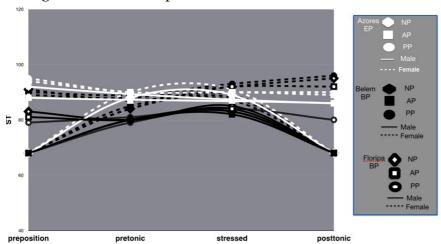


Figure 4. The F0 variation (ST) at the final of sentence with the set of paroxytone words

The first results confirm a possibility of an intralinguistic analysis with the Portuguese varieties analyzed here. The paroxytone words seem to show less variation in relation to other kind of lexical stress. The Portuguese dialects spoken in Florianopolis and in Azores seem to have more variability in relation to the Portuguese spoken in Belem which seems to have a more regular behavior of the F0 curve for each lexical word and for each kind of phrase. This is an early comparative study, but its results gives us cues that we are on the right way. In order to confirm the result on the relationship between stress and F0 variation in the characterization of intonational modalities of Portuguese, we intend to amplify this analysis for completely set of AMPER data for these three Portuguese varieties analyzed here. The Portuguese data of the AMPER corpus can give us cues that there is a great prosodic variation between Portuguese spoken in Azores, Belem and Florianopolis, but this variation can be accounted for intralinguistic.

5 Conclusion

This study shows it is possible to carry out a comparative analysis between data from tree Portuguese dialects, Azores, Belem and Florianopolis. It seems to have a well-formed pattern for yes/no question which in placed at the last position of the sentence the paroxytone words.

Overall, Florianopolis and Azores appear to have more variability in relation to their intonation and lexical stress, on the other hand Belem seems more regular in relation to F0 curves on yes/no question and lexical stress.

6 References

Brescancini, C. (2002). *A fricativa palatoalveolar e sua complexidade*. Tese (Doutorado em Letras), Pontificia Universidade Católica do Rio Grande do Sul, Porto Alegre.

Brescancini, C. (1996). A palatalização da fricativa alveolar não-morfêmica em posição de coda no português de origem açoriana do município de Florianopolis: uma abordagem nãolinear. Dissertação (Mestrado em Letras), Universidade Federal de Santa Catarina, Florianópolis.

Carvalho, R. (2000). A variação do /S/ pósvocálico na fala de Belém. Dissertação (Mestrado em Letras), Universidade Federal do Para, Belém.

Chambouleyron, R. (2008). A Amazônia colonial e as ilhas atlânticas. *Canoa do*

- Tempo. Revista do Programa de Pósgraduação em História, Manaus, v.2, n.1, jan./dez. 2008, pp. 187-204.
- Coates, T. (1998). *Degredados e órfãs*. Lisboa: CNCDP.
- Furlan, O. (1989). Influência Açoriana no Português do Brasil em Santa Catarina. Florianópolis: Editora da UFSC.
- Haupt, C. & Berri, A. (2009). O processo de palatalização na fala de florianopolitanos nativos em *corpus* de fala espontâneo e controlado. *Letrônica*, Porto Alegre, v.2, n.2, pp. 1-21.
- Jesus, C. (2013). Palatalização do /S/ em capitais brasileiras, com base em dados do ALiB: o caso de estilingue e de prostituta. Revista Inventário, Salvador, 13a. edição, jul./dez. 2013, 1-12. En http://www.inventario.ufba.br/13/palatalizacao-s-capitais-brasileiras-base-dados-alib-claudia-jesus.pdf
- Lopes, C. (2009). Retratos da mudança no sistema pronominal: o tratamento carioca nas primeiras décadas do século XX. In: Arnaldo Cortina; Silvia Maria Gomes da Conceição Nasser. (Org.). Sujeito e Linguagem: Séries Trilhas Linguísticas. Araraquara: Cultura Acadêmica, v.17, pp. 47-74.
- Loregian-Penkal, L. (2004). (Re)análise da referência de segunda pessoa na fala da Região Sul. Tese (Doutorado em Letras), Universidade Federal do Parana, Curitiba. Orientadora: Odete Menon. 260 p.
- Reis, A. (1993). A política de Portugal no vale Amazônico. Belém: Secult.
- Scherre, M. (2007). Aspectos sincrônicos e diacrônicos do imperativo no Português Brasileiro. *Alfa*, São Paulo, 51(1): pp. 189-222.
- Smaiclova, G. (2010). Palatalização do /S/ em coda silábica o português falado na Costa da Lagoa em Florianópolis. *Work. Pap. Linguist.*, Florianópolis, n. esp.: pp. 33-44.
- Soares, I. & Leal, M. G. (1993). Do senhor ao tu: uma conjugação em mudança. *Moara*. Revista do curso de mestrado (UFPA), Belém, n.1, pp. 27-64.

La entonación cubana en zonas rurales: La Habana y Santa Clara

DORTA, JOSEFA; DÍAZ CABRERA, CHAXIRAXI; HERNÁNDEZ DÍAZ, BEATRIZ

1 Introducción

El presente trabajo se enmarca en el proyecto de investigación *Estudio comparativo de la entonación y delacento en zonas fronterizas del español (FFI2014-52716-P)¹* vinculado al proyecto internacional AMPER (*Atlas Multimedia de Prosodia del Espacio Románico*), nacido en Francia en el año 2002² con el objetivo de realizar un atlas de prosodia de las lenguas románicas. Nuestro interés por la relación histórica entre las variedades atlánticas de estas zonas, de una parte, y los fines comparativos del proyecto AMPER, por otra, dieron origen al reciente *Estudio comparativo preliminar de la entonación de Canarias, Cuba y Venezuela* (Dorta ed., 2013), que constituye el precedente más inmediato del trabajo que aquí presentamos. En esta obra abordamos la prosodia de estas tres variedades a partir de un corpus de habla urbana; en el presente trabajo continuamos el estudio de una de ellas, la cubana, teniendo en cuenta un corpus de habla rural.

La entonación del español de Cuba ha sido estudiada partiendo de diferentes modelos de análisis. Después del trabajo de Haden y Matluck (1973) donde quedan descritos, siguiendo el esquema tradicional de niveles tonales de la Escuela Americana, diversos patrones típicos de esta variedad, la trilogía Aspectos de la entonación hispánica (1996-1998), de García Riverón, es la obra que dedica mayor atención al estudio prosódico del español cubano tratando de

¹Proyecto de I+D del programa estatal de fomento de la investigación científica y técnica de Excelencia, subprograma estatal de generación del conocimiento (convocatoria 2014).

² Coordinado por Michel Contini (Universidad Stendhal-Grenoble 3) y Antonio Romano (Universidad de Turín).

determinar el sistema de entonación de la lengua mediante el establecimientode entonemas y variantes de entonema que son descritos acústica y funcionalmente¹.

En el marco del modelo AM, Sosa (1999) hace una interpretación fonológica de los contornos melódicos de La Habana caracterizando las declarativas con un pretonema L*+H en el que destaca la superioridad del segundo pico, y un tonema L* L%. La interrogativa absoluta, según este autor, culmina con el tonema H+H* L% en final claramente circunflejo. Toledo (2004), por su parte, basándose en un corpus de habla semiespontánea y seleccionando los acentos tonales paroxítonos, observa una gran variabilidad tonal en el pretonema H* en combinaciones monotonales y bitonales, indicando con ello el pico tonal en la sílaba acentuada.

En el marco del proyecto AMPER, la entonación del español de Cuba ha sido objeto de diferentes trabajos comparativos, como el de Fernández Pérez-Terán, Dorta, Ramos & García Riverón (2007) o Dorta & Díaz (2013), donde se considera la relación entre la entonación canaria y la de La Habana desde el punto de vista perceptivo, con un altísimo reconocimiento de interrogativas de las dos variedades². En el trabajo más amplio, ya citado, Dorta ed. (2013) se analiza la entonación cubana de zonas urbanas y se compara con la de Canarias y Venezuela. El corpus elegido es el formal (SVO) del proyecto AMPER, al igual que en Dorta & Martín Gómez (2014) donde se analiza un hablante de La Habana. En ambos trabajos se evidencian las semejanzas entre los patrones interrogativos y declarativos cubanos y canarios, caracterizados los primeros por un final circunflejo³ y por una estructura bicumbre con un pico al final del SN y otro al final del SV los segundos. En otros trabajos (Dorta & Martín Gómez, 2012 o Dorta, Díaz & Martín Gómez, 2012) se había analizado, además, el habla espontánea comprobando -mediante la aplicación del Sp_ToBI- las grandes semejanzas, pero también las diferencias, que presentan las interrogativas de las dos variedades.

2 Metodología

El corpus que hemos analizado es el experimental fijo del proyecto AMPER emitido en las modalidades declarativa e interrogativa por dos mujeres sin estudios superiores de zonas rurales de Cuba: occidental (Surgidero, La Habana) y central (Placetas, Santa Clara); está integrado por 9 frases simples sin expansión (SVO), con todas las combinaciones entre agudas, llanas y esdrújulas en los SN y SP. Las grabaciones se realizaron *in situ* con el Handy Recorder H4n profesional y fueron segmentadas posteriormente con el Goldwave Digital Audio Editor (versión 4.25). El análisis acústico se hizo con rutinas desarrolladas en Matlab por la Universidad de Oviedo (Brezmes Alonso, 2007)⁴ a partir de las originales creadas por Antonio Romano (1995). En el presente estudio se ha analizado solo la F0 (en Hz) y su relativización (en St). Para considerar las diferencias perceptivamente relevantes se ha elegido el umbral diferencial de 1,5 St (Rietveld & Gussenhoven 1985, aplicado al español por Toledo 2000, Pamies *et al.* 2002 o Dorta ed., 2013). Se ha tenido en cuenta,

³ En Cuba, este final circunflejo ya había sido señalado por García Riverón, Bermúdez, Pedrosa y Marrero (2010) para la interrogación neutral. En Canarias, ha sido ampliamente estudiado en los numerosos trabajos realizados en el marco de

¹ En los últimos años, el Centro de Lingüística Aplicada de Santiago de Cuba está prestando atención a los estudios de entonación, como lo evidencia la tesis doctoral de Alex Muñoz Alvarado Sistema y norma entonativos en el habla de Santiago de Cuba defendida en 2012.

² En el primer estudio de 2007 los auditores fueron cubanos; en Dorta y Díaz (2013) fueron canarios.

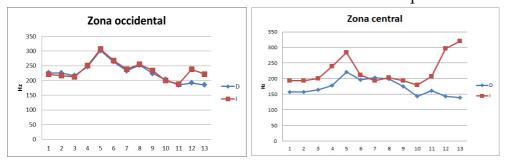
⁴ La licencia de Matlab en el Laboratorio de Fonética de la ULL es la n.º 256105. Véase el replanteamiento de las nuevas rutinas en López Bobo et al. (2007).

además, el tono medio (TM) de las informantes como valor de referencia. El etiquetaje fonético-fonológico que se realiza en el presente trabajo sigue las pautas propuestas en Dorta ed. (2013).

3 Resultados

3.1 Descripción general

Según Sosa (1999: 198), las interrogativas de La Habana "mantienen una altura global más alta" que las declarativas. Esta característica, sin embargo, se hace más evidente en la zona central (Placetas, Sta. Clara, figura 2) que en la occidental (Surgidero, La Habana, figura 1) donde prácticamente se superponen los contornos de las dos modalidades hasta el núcleo en que se distancian.



Figuras 1-2. F0 en oraciones SVO con inicio y final llano

Por otra parte, las dos modalidades analizadas sitúan sus inicios (I) por debajo del tono medio (TM) de las informantes (figura 3), si bien solo en la zona central la distancia respecto de este supera el umbral perceptivo de 1,5 St. La diferencia entre las dos zonas, en cambio, no llega a dicho umbral en ninguna de las dos modalidades.

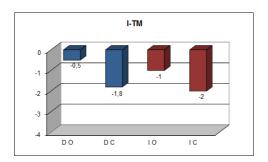
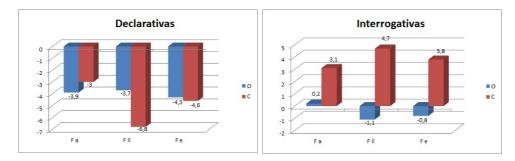


Figura 3. Diferencia en St entre el inicio y el TM

Si consideramos el tipo de acento con que terminan las oraciones analizadas se observa que en las declarativas (figura 4) la única diferencia significativa entre las dos zonas aparece en los finales llanos (3,1 St a favor de la zona occidental, más próxima al TM). En las interrogativas (figura 5), la diferencia, a favor de la zona central, supera ampliamente el umbral en los tres acentos, sobre todo en llanos y esdrújulos, debido al marcado ascenso que experimenta la F0 en esta zona (véase la figura 2)⁵.

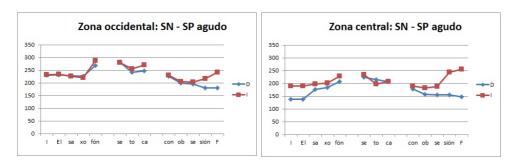
⁵ La representatividad de este tonema final deberá ser corroborada a partir de corpus más amplios de la misma o de otras zonas rurales del centro de Cuba puesto que su aparición, en lugar del final circunflejo, podría estar motivada por el tipo de corpus en que se basa el trabajo, tal como se ha comprobado en las interrogativas de algunas islas canarias como La Gomera (v. gr. Díaz, 2013).



Figuras 4 y 5. Diferencia en St entre el tono medio y el final agudo, llano y esdrújulo de las declarativas e interrogativas

3.2 Comportamiento de la F0 según el tipo de acento léxico que abre y cierra las oraciones

Tomando como referencia la vocal tónica, tanto en el sintagma nominal (SN) como en el preposicional (SP), el acento léxico determina localmente un salto de la F0, por lo general significativo. En efecto, en los SN agudos de la zona occidental (figura 6), la F0 se mantiene desde el inicio y asciende significativamente en la última tónica (3 St en declarativas; 4,6 St en interrogativas) superando el TM (3 y 3,2 St, respectivamente). En la zona central (figura 7), en cambio, las dos modalidades inician el ascenso de la F0 en la segunda vocal del sintagma inicial⁶ aunque, en este movimiento ascendente, seguimos encontrando un salto significativo de la pretónica a la tónica, si bien menos importante que en la zona occidental (2,1 y 2,2 St en declarativas e interrogativas, respectivamente). Como en Surgidero, en Placetas la vocal tónica sobrepasa el TM con valores que superan o se aproximan al umbral (2,9 St en declarativas y 1,1 St en interrogativas).



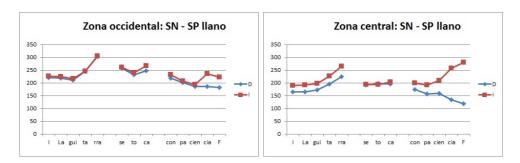
Figuras 6-7. Declarativas vs interrogativas con inicio y final agudo en Surgidero y Placetas

Como puede apreciarse en los gráficos precedentes, en el SP la F0 desciende significativamente ya desde la segunda vocal⁷ y se mantiene hasta la pretónica, si bien en la última tónica vuelve a darse un movimiento importante que acelera el descenso hacia el final en las declarativas de la zona occidental (-1,4 St) e inflexiona la F0 en las interrogativas de ambas zonas, determinando un repentino ascenso que contrasta con la dirección anterior, mucho más significativo en la zona central que en la occidental (+4,5 y +1,1 St,

⁶ +4,2 St en declarativas. Obsérvese que este primer ascenso significativo se produce en el paso del artículo al sustantivo. ⁷-2,2 y -1,9 St en declarativas e interrogativas de la zona occidental; -2 St en declarativas de la zona central. Obsérvese que este descenso significativo se produce en el paso de la preposición al sustantivo.

respectivamente)⁸. En definitiva, el movimiento de la tónica nuclear es determinante para diferenciar las dos modalidades⁹ e, incluso, entre las dos variedades cubanas.

En los SN llanos (figuras 8-9), la presencia de la tónica en las dos zonas rurales analizadas se evidencia a través del ascenso significativo de la F0¹⁰ hacia el primer pico máximo que recae en la postónica. En todos los casos, este ascenso vuelve a situar la F0 por encima del TM de las informantes, si bien solo en las declarativas lo hace de manera significativa (1,5 St en la zona occidental y 2 St en la zona central).



Figuras 8-9. Declarativas vs. interrogativas con inicio y final llano en Surgidero y Placetas

La diferencia entre las dos zonas y las dos modalidades también vuelve a ser patente en el núcleo entonativo. Así, en la zona occidental la F0 decae de manera relativamente brusca hasta la tónica en declarativas e interrogativas pero es justo a partir de este punto donde las dos modalidades se separan: en la primera, sigue descendiendo de manera menos brusca (-0,4 St); en las interrogativas, en cambio, asciende en la postónica (3,4 St)¹¹. En la zona central es también el movimiento tonal a partir de la tónica el que marca claramente la distancia entre las dos modalidades: ascenso pronunciado en las interrogativas¹² y descenso brusco en las declarativas (-5,2 St).

En los SN esdrújulos (figuras 10-11), se repite con pocas variantes el mismo esquema tonal para las dos modalidades. Así, como en los SN agudos, en las dos variedades cubanas el acento se manifiesta a través de un acusado ascenso tonal que sitúa la F0, ya desde la segunda vocal, significativamente por encima del TM (1,8 y 3,4 St en las declarativas de la zona occidental y central, respectivamente; 2,2 St en las interrogativas de las dos zonas)¹³.

⁸ En la zona central, este movimiento significativo sitúa la Fo a 2,2 St por encima del TM.

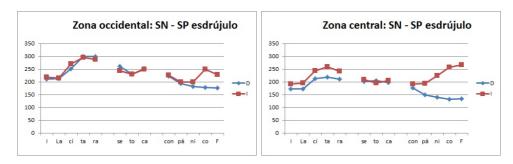
⁹Con una diferencia de altura tonal de 3,1 St en la zona occidental y de 7,8 St en la zona central, a favor de las interrogativas.

¹⁰ En declarativas, 2,6 St en la zona occidental y 2,3 St en la central; en interrogativas, 2,2 y 2,4 St, respectivamente.

¹¹ Al final de la postónica la Fo desciende ligeramente (-0,9 St) por lo que no podemos considerar que se trata del final circunflejo al que aludimos en la Introducción donde el descenso se produce de manera significativa a partir de un pico tonal que se da en la tónica y no en la postónica como sucede aquí.

¹² Realmente, el movimiento ascendente en esta modalidad comienza a partir de la pretónica pero es a partir de la tónica donde el ascenso se comienza a marcar de manera brusca (5,1 St desde la tónica hasta el final).

¹³Los valores de este ascenso en el paso de la pretónica a la tónica inicial son: en declarativas, 2,9 St en la zona occidental y 3,8 St en la central; en interrogativas, 3,9 y 3,8 St, respectivamente.

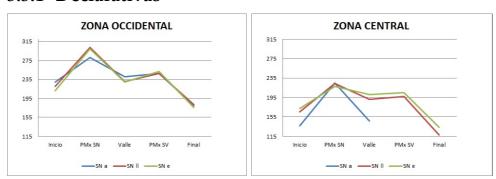


Figuras 10-11. Declarativas vs. interrogativas con inicio y final esdrújulo en Surgidero y Placetas

En los SP, asimismo, las declarativas de ambas zonas tienen una trayectoria similar, con descenso tonal significativo en la tónica nuclear (-2,5 y -2,8 St en zona occidental y central, respectivamente)¹⁴. Las interrogativas, sin embargo, presentan patrones finales diferentes: en la zona occidental, la tónica desciende significativamente respecto de la pretónica (-2,1 St), se mantiene en la postónica y asciende de manera repentina en la última vocal, con un valor que supera ampliamente el umbral perceptivo (3,9 St)¹⁵; en la zona central la vocal tónica no registra movimiento alguno respecto de la anterior¹⁶, si bien en la postónica se produce un ascenso significativo de 2,5 St que culmina en el final absoluto.

3.3 El rango tonal y el escalonamiento de los picos

3.3.1 Declarativas



Figuras 12 y 13. Picos tonales de las declarativas

Los gráficos 12-13 que preceden muestran una estilización de las curvas a partir de los diferentes intervalos de F0. Si nos centramos en el PMx1 (véanse los valores de la tabla 1), cuando el trisílabo inicial es llano o esdrújulo, los picos de la zona occidental son más prominentes que los de la central, con campos tonales más amplios, tanto por los valores del ascenso inicial (I-P1) como por los del descenso hacia el valle posterior (P1-V), donde las diferencias entre las dos zonas son siempre significativas. En cambio, cuando el trisílabo inicial es agudo, la zona central el PMx1 tiene mayor prominencia con valores diferenciales que superan ampliamente el umbral en relación con la zona occidental.

 $^{^{14}}$ Este descenso vuelve a situar la Fo por debajo del TM, con valores significativos de $^{-2,7}$ St en la zona occidental y $^{-2,8}$ St en la zona central.

 $^{^{15}}$ En los esdrújulos de la zona occidental se da un pequeño descenso (-0,7 St) en la última vocal átona que, como en los llanos (nota 13) no podemos considerar como exponente de un final circunflejo.

¹⁶ Si bien en la zona occidental el descenso en la tónica sitúa la Fo significativamente por debajo del TM (-3 St), en la zona central la pretónica ya se encuentra a -2 St respecto de este, de ahí que la tónica se mantenga sin descender pues ya se encuentra significativamente por debajo del TM.

	Zona Occi	idental		Zona Central				
	I-P1	P1-V	V-P2	P2-F	I-P1	P1-V ¹⁷	V-P2	P2-F
a	3,3	-2,5	0,2	-5,2	8,6	-7,4		
11	5,2	-4,5	1,1	-5,2	5,1	-2,8	0,5	-8,9
e	5,9	-4,5	1,3	-5,9	4,1	-1,4	0,3	-7,4

Tabla 1. Valores relativos en St del rango tonal de los picos del SN y SV de las declarativas

Comparando los tres acentos léxicos observamos que, si en la zona central los rangos tonales del primer pico son, en las estructuras agudas, significativamente más amplios que en las llanas y esdrújulas, en la zona occidental ocurre lo contrario, siendo los picos de las agudas los menos relevantes desde el punto de vista perceptivo, con valores por encima del umbral respecto de las otras dos.

Al considerar el rango anterior del PMx₂(V-P₂), observamos que los valores son irrelevantes pues en ninguna de las dos zonas cubanas se supera el umbral diferencial. En cambio, el descenso desde este pico hasta el final (P2-F) es importante en todos los casos, aunque los finales de la zona central son siempre más abruptos que los de la zona occidental con diferencias que llegan al umbral perceptivo (1,5 St en los esdrújulos) o lo superan (2,2 St en agudos y 3,7 St en los llanos). En todo caso, este último descenso es muy significativo y similar en los tres acentos de la zona occidental; en la central, si bien es bastante relevante desde el punto de vista perceptivo en los tres acentos, sobresalen los llanos por encima de los agudos y esdrújulos.

Por último, si consideramos el escalonamiento de los picos en esta modalidad (tabla 2) encontramos que en la zona occidental cubana el último pico se sitúa siempre por debajo del primero, superando en todos los casos el umbral perceptivo de 1,5 St, por lo que el escalonamiento tonal es descendente. En la central, excluidas las configuraciones agudas puesto que la estructura de la oración es monocumbre y el descenso final se adelanta al segundo acento tonal, la pendiente descendente solo supera el umbral en los llanos.

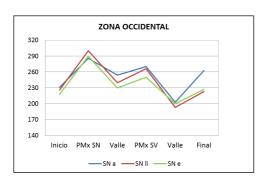
	Zona O	ccidental		Zona Central			
	PMx1	PMx2	Diferencia PMx1-PMx2	PMx1	PMx2	Diferencia PMx1-PMx2	
a	278	243	-2,3	225			
11	298	245	-3,4	224	197	-2,2	
e	297	247	-3,2	218	205	-1,1	

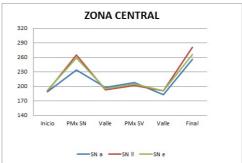
Tabla 2. Valores absolutos y diferencia relativa en St de los picos de las declarativas

3.3.2 Interrogativas

La estilización de las curvas en esta modalidad tiene los perfiles de las figuras 14-15. Como puede verse, los contornos son bicumbres en el prenúcleo (frontera SN/SV y SV/SP) como en las declarativas pero se diferencian de estas por la entonación ascendente del núcleo.

¹⁷ En las combinaciones agudas de la zona central se produce el descenso final desde el pico 1 (P1-F).





Figuras 14 y 15. Picos tonales de las interrogativas

Ahora bien, la diferencia fundamental entre las dos zonas analizadas es que en la occidental el rango tonal posterior del pico situado en la frontera SV/SP (P2-V) es significativamente mayor que en la zona central (2,5 St); en cambio, el tonema ascendente es mucho más abrupto en los tres finales de la zona central. Salvando estas diferencias, los valores del ascenso hacia el primer pico (I-P1) son, tal como se puede observar en la tabla 3, similares en ambas zonas coincidiendo en el hecho de que, de las tres estructuras acentuales, este ascenso es inferior en los agudos (como en las declarativas de la zona occidental); las diferencias rozan o incluso superan el umbral en relación con los otros dos acentos (1,2 y 1,2 St respecto de llanos y esdrújulos en la zona occidental; 2,1 y 1,5 St, respectivamente, en la zona central).

	Zona Occidental						Zona Central				
	I-P1	P1-V	V-P2	P2-V	V-F	I-P1	P1-V	V-P2	P2-V	V-F	
a	3,7	-2,0	0,6	-4,7	3,3	3,7	-2,9	0,8	-2,2	5,8	
11	4,9	-3,9	1,4	-4,8	2,3	5,8	-5,5	0,8	-1	6,7	
e	4,9	-4,2	1,3	-3,7	3,1	5,2	-4,9	0,9	-1,2	5,7	

Tabla 3. Valores relativos en St del rango tonal de los picos del SN y SP de las interrogativas

En cuanto al núcleo entonativo, las dos zonas tienen, como decíamos, patrones bien diferenciados: en la central, el ascenso final (V-F), significativo en todos los casos, no deja ver diferencias relevantes entre los tres tipos de final desde el punto de vista perceptivo. Este último ascenso es, por otro lado, el de mayor rango de toda la oración aunque solo en los agudos la diferencia respecto del rango del pico inicial es significativa (2,1 St). En la zona occidental, asimismo, el último ascenso muestra valores sobre el umbral en las tres estructuras acentuales próximos entre sí.

En relación con el escalonamiento de los picos, observamos en esta modalidad (tabla 4) que en las dos zonas estudiadas la frecuencia más elevada se da siempre en el pico inicial, con diferencias que superan el umbral en los agudos y sobre todo en los llanos y esdrújulos por lo que el escalonamiento es descendente como en las declarativas.

	Zona C	Occidental		Zona Central			
	PMx1	PMx2	Diferencia	PMx1	PMx2	Diferencia	
			PMx1-PMx3			PMx1-PMx2	
a	286	270	<i>-</i> 1,5	234	208	-2	
11	300	266	-5,1	265	202	-4,7	
е	290	250	-3,5	259	205	-4	

Tabla 4. Valores absolutos y diferencia relativa en St de los picos de las interrogativas

4 Caracterización fonético-fonológica: acentos tonales y tonos de frontera final

4.1 Los precedentes

En Dorta ed. (2013) el acento tonal inicial de las declarativas e interrogativas de las zonas urbanas de La Habana y de Santa Clara es bitonal /L+H*/. En el acento nuclear, las declarativas se caracterizan por un acento bajo seguido de un tono de frontera también bajo /L* L%/; las interrogativas tienen un final circunflejo, con alineación del pico con la última tónica y un tono de frontera por debajo del tono medio /H* L%/.

4.2 Etiquetaje prosódico en las zonas rurales

De acuerdo con las pautas de etiquetaje propuestas en Dorta ed. (2013), en las tablas 5 y 6 mostramos las variantes fonéticas y las invariantes fonológicas de los acentos tonales y tonos de frontera de las declarativas e interrogativas de las zonas rurales de La Habana y de Santa Clara establecidas a partir de los resultados expuestos en este trabajo.

Desde el punto de vista fonético (tabla 5) lo más destacado en el acento inicial es que en las zonas rurales el pico tonal aparece siempre desplazado, con la única excepción de los agudos de Surgidero.

	Acento 1 (SN)			Ace	Acento 2 (SV)			Acento 3 (SP)			Tono de frontera final %		
	a	11	e	a	11	e	a	11	e	a	11	e	
Surgidero	L+> H*	L+>H*	L+>H*	L*	L*	L*+! H	L*	L*	L*	L%	L%	L%	
Placetas	L+> H*	L+>H*	L+>H*	H*	L*	H*	L*	L*	L*	L%	L%	L%	
Surgidero	L+ H*	L+>H*	L+>H*	L*	L*+!H	L*	L+H*	L*	L*	MH %	MH %	MH %	
Placetas	L+> H*	L+>H*	L+>H*	L*	L*	L*	L+H*	L*+H	L*	Н%	Н%	Н%	

Tabla 5. Interpretación fonética de los acentos tonales y tonos de frontera de las declarativas e interrogativas de zona rural

No obstante, como en las zonas urbanas, fonológicamente (tabla 6) no hay discrepancias puesto que, en las dos modalidades analizadas, el acento tonal inicial es el mismo /L+H*/.

		Acento 1 (SN)	Acent	Acento 2 (SV)		Acento 3 (SP)	Tono de
			a	11	е		frontera final %
D	Surgidero	L+H*	L*	L*	L*+H	L*	L%
	Placetas	L+H*	H*	L*	H*	L*	L%
I	Surgidero	L+H*	L*	L*+H	L*	H* (a)/L*(ll,e)	H%(a, ll,e)
	Placetas	L+H*	L*	L*	L*	H*(a)/L* (ll,e)	H%(a,ll.e)

Tabla 6. Interpretación fonológica de los acentos tonales y tonos de frontera de las declarativas e interrogativas de zona rural

También coincide fonética y fonológicamente el habla rural con la urbana en el acento nuclear y en el tono de frontera de las declarativas (L*L%). Las

interrogativas, en cambio, tienen una diferencia importante: en las zonas urbanas estudiadas con anterioridad, la acentuada nuclear coincide con un pico tonal a partir del cual se da el descenso típico del final circunflejo (H* L%). En las zonas rurales, en cambio, no se da ni fonética ni fonológicamente este acento nuclear: lo más relevante es que fonológicamente la tónica tiene un tono alto en agudos /H*/ y bajo /L*/ en llanos y esdrújulos con tono de frontera alto /H%/ en todos los casos. Ahora bien, como ya se ha dicho, la consideración de este tipo de final ascendente como rasgo particular del habla rural de Cuba debería ser corroborada a partir del análisis de corpus más amplios y/o más espontáneos.

5 Conclusiones

Del análisis realizado, las conclusiones más importantes son las siguientes:

- 1) En el SN de ambas modalidades destaca la realización tardía de los picos tonales, generalmente en la postónica. Así, más que con el máximo tonal, el acento léxico se relaciona con el TM, pues es en la tónica donde este se supera significativamente mediante una inflexión o un salto significativo de la F0 que acelera el ascenso hacia el PMx1. El desplazamiento del pico tonal es, según nuestra interpretación, un hecho fonético por lo que desde el punto de vista fonológico se confirma en las dos modalidades de las zonas rurales, como en las urbanas, la existencia de un acento bitonal inicial /L+H*/ con la tónica inicial asociada a un tono alto y no bajo (L*+H) como había señalado Sosa (1999).
- 2) En el SP se encuentran las principales diferencias entre las dos modalidades consideradas. En las declarativas, el acento léxico marca un cambio en la progresión del descenso de la F0: o bien se da un tono sostenido en la pretónica y un descenso brusco en la tónica (SP agudos de Surgidero y SP llanos de Placetas), o bien un descenso brusco en la tónica seguido de un tono sostenido en la postónica (SP llanos de Surgidero y SP esdrújulos de las dos zonas). En cualquier caso, fonética y fonológicamente se confirma en las dos zonas la existencia de un tono bajo en la tónica y en la frontera final (L* L%) característicos del español general. Por el contrario, en las interrogativas la inflexión tonal en el SP hace que la F0, significativamente por debajo del TM ya desde la primera sílaba, vuelva a situarse por encima de este. A diferencia de las zonas urbanas no se registró el final circunflejo /H* L% / (Dorta *et al.*, 2013), ni tampoco /H+H* L% / (Sosa 1999), sino /H* H% / en agudos y /L* H% / en llanos y esdrújulos de las dos zonas.
- 3) Los inicios de las dos modalidades se sitúan por debajo del TM, aunque solo de manera significativa en la zona central. En cuanto a los finales, en las declarativas siempre se colocan significativamente por debajo del TM; en las interrogativas, ambas zonas se diferencian: el final se coloca en torno al tono medio en Surgidero y significativamente por encima de él en Placetas.
- 4) El escalonamiento tonal descendente de los picos caracteriza las declarativas e interrogativas rurales estudiadas; las diferencias entre los dos picos máximos superan siempre el umbral de 1,5 St salvo en los esdrújulos de las declarativas de Placetas.

6 Bibliografía

Brezmes Alonso, D. (2007). Desarrollo de una aplicación software para el análisis de características fundamentales de la voz (Proyecto de fin de carrera). Universidad de Oviedo.

Díaz, Ch. (2013). Contribución al Atlas Prosódico de Canarias (AMPERCan):

- Declarativas e interrogativas de La Gomera (Islas Canarias) (Tesis doctoral). Universidad de La Laguna.
- Dorta, J. (ed.). (2013). Estudio comparativo preliminar de la entonación de Canarias, Cuba y Venezuela. Madrid-Santa Cruz de Tenerife, España: La Página Ediciones, Colección Universidad. Participan: J. Dorta, E. Mora, B. Hernández Díaz, Ch. Díaz Cabrera, N. Rojas, J. A. Martín Gómez & C. Jorge.
- Dorta, J., & Díaz, Ch. (2013). Proximidad perceptivo-entonativa en dos variedades atlánticas: el caso canario-cubano. *Lengua y Habla*, 17, 34 54. Revista electrónica del Centro de Investigación y Atención Lingüística (C.I.A.L.).
- Dorta, J., Díaz, Ch., & Martín Gómez, J. A. (2012). Continuidad prosódica y variación en habla experimental y espontánea de Canarias y Cuba: el caso de las interrogativas con patrón nuclear ascendente-descendente. Coloquio Las variaciones diasistémicas y sus interdependencias (ΔΙΑΙΙ), 19-21 de noviembre de 2012, Institut for Engelsk, Germansk og Romansk, Copenhague, Dinamarca.
- Dorta, J., & Martín Gómez, J. A (2012). Análisis comparativo de la entonación: estudio preliminar de las interrogativas no pronominales y pronominales canario-cubanas en habla espontánea. *LEA*, 34/2, 197-222.
- Dorta, J., & Martín Gómez, J. A. (2014). Estudio preliminar de la fonética y fonología de la entonación de La Habana en el marco de Amper-Cuba. Fonética Experimental, Espacio Europeo de Educación Superior e Investigación, 3, 189-210. Madrid, España: Editorial Arco Libros.
- Fernández Pérez-Terán, F., Dorta Luis, J., Ramos, D., & García Riverón, R. (2007). La interrogativa absoluta en el español de Canarias y Cuba: estudio perceptivo. En J. Dorta (Ed.), La prosodia en el ámbito lingüístico románico, 371-387. Madrid-Santa Cruz de Tenerife, España: La Página Ediciones, Colección Universidad.
- García Riverón, R. (1996): Aspectos de la entonación hispánica. I Metodología. II Análisis acústico de muestras del español de Cuba. III Las funciones de la entonación en el español de Cuba. Cáceres, España: Servicio de

- Publicaciones de la Universidad de Extremadura.
- García Riverón, R., Bermúdez, M., Pedrosa, A., & Marrero, A. (2010). El sistema de entonación del español de Cuba a la luz del modelo de análisis melódico del habla. *Phonica*, 6, 3-25.
- Haden, E., & Matluck, J. (1973). El habla culta de La Habana: análisis fonológico preliminar. *Anuario de Letras*, 11, 5-33.
- López Bobo, M. J., Muñiz Cachón, C., Díaz Gómez, L., Corral Blanco, N., Brezmes Alonso, D., & Alvarellos Pedrero, M. (2007). Análisis y representación de la entonación. Replanteamiento metodológico en el marco del proyecto AMPER. En J. Dorta (ed.), La prosodia en el ámbito lingüístico románico, 17-34. Madrid-Santa Cruz de Tenerife, España: La Página Ediciones, Colección Universidad.
- Pamies, A., Fernández Planas, A. M., Martínez Celdrán, E., Ortega Escandell, A., & Amorós Céspedes, M. C. (2002). Umbrales tonales en español peninsular. En J. Díaz García (ed.), Actas del II Congreso Nacional de Fonética Experimental, 272-278. Sevilla, España: Universidad de Sevilla.
- Rietveld, A., & Gussenhoven, C. (1985). On the relation between pitch excursion size and prominence. *Journal of Phonetics*, New York, Academic Press, 13, 299-308.
- Romano, A. (1995). Développement d'un environnement de travail pour l'étude des structures sonores et intonatives de la parole (DEA en Sciences du Langage). ICP, Univ. Stendhal-Grenoble 3.
- Sosa, J. M. (1999). La Entonación del español. Su estructura fónica, variabilidad y dialectología. Madrid: Cátedra.
- Toledo, G. A. (2000). Taxonomía tonal en español. *Language Design*, 3, 1-20.
- Toledo, G. A. (2004). Prominencia H*: una muestra de español de Cuba. *Estudios de Fonética Experimental*, 13I, 181-202.

Diferencias acústicas entre los acentos prenucleares en catalán

FERNÁNDEZ PLANAS, ANA MA.; ROSEANO, PAOLO; ELVIRA GARCÍA, WENDY; CERDÀ MASSÓ, RAMON; MARTÍNEZ CELDRÁN, EUGENIO

1 Introducción

La caracterización de los acentos prenucleares en catalán ha sido objeto de debate científico en los últimos años (Prieto et al., 2009; Martínez Celdrán & Fernández Planas, 2010; Crespo, 2010). Los estudios publicados sobre el tema se pueden clasificar a partir de su foco de atención (rango de los movimientos tonales, alineación de la diana tonal alta, alineación de la diana tonal baja) y de las modalidades oracionales que tratan (declarativa neutra, interrogativa neutra, o ambas). Estos estudios van desde una perspectiva fonológica en la que se describen cualitativamente las características de los acentos considerados (p. ej. Prieto et al., 2009) hasta descripciones fonéticas que presentan grados de detalle diferentes (p. ej. Martínez Celdrán, 2007 o Crespo, 2010).

El tema que cuenta con más publicaciones es, sin duda, la alineación de la diana tonal alta en el pretonema de las declarativas neutras (p. ej. Prieto et al., 2010; Carrera-Sabaté et al. 2010; Martínez Celdrán et al. 2005, entre otros muchos). Las autoras de dicho estudio concluyen que la diana tonal alta no está asociada a la frontera de palabra. Sin embargo, la frontera de palabra tiene un efecto sobre la alineación de la diana tonal misma ya que actúa como una barrera que trata de frenar el desplazamiento del pico a la palabra siguiente, efecto que se ha descrito también para otras lenguas (Roseano, 2012).

La alineación de la diana tonal alta en las interrogativas ha sido mucho menos tratada que en las enunciativas (Fernández Planas et al., 2004; Martínez Celdrán, 2007; Carrera Sabaté et al., 2010) y los estudios que la tratan

mencionan el hecho de que el primer pico se encuentra más a la derecha en las interrogativas que en las declarativas.

La diana tonal baja no ha gozado de demasiada atención (Prieto, 2001, 2005, 2014; Prieto et al., 2009; Crespo, 2010). Aunque estas publicaciones señalan que este parámetro se alinea con el comienzo de la sílaba tónica en las frases declarativas (L+>H*) y con el final de la misma sílaba (L*+H) en las interrogativas, no basan sus afirmaciones en un análisis acústico exhaustivo.

El rango de los movimientos de los acentos tonales ascendentes prenucleares es el parámetro menos recurrente (Martínez Celdrán et al., 2005; Martínez Celdrán, 2007; Carrera Sabaté et al., 2010). Se ha destacado que el primer pico del prenúcleo de las interrogativas es más agudo que el de las declarativas y que la diferencia suele superar los 2 semitonos.

En suma, aunque de la bibliografía disponible se deduce que hay diferencias acústicas entre los acentos prenucleares de declarativas e interrogativas absolutas en catalán central, los trabajos citados no abarcan todas las diferencias posibles, ni las analizan de forma rigurosa y no coinciden en señalar cuál es el parámetro más relevante. El presente trabajo viene a llenar este vacío puesto que pretende comparar sistemática y exhaustivamente todas las características acústicas de los acentos prenucleares de las frases declarativas neutras y de las interrogativas absolutas informativas, a saber: duración del sujeto (o el verbo), valor inicial de F0 de la frase, posición del primer valle en relación con la sílaba tónica, posición del primer pico en relación con la frontera de palabra, rango del movimiento tonal dentro de la sílaba tónica, rango del movimiento tonal entre el principio de la sílaba tónica y el pico y, finalmente, rango del movimiento tonal entre el valle y el pico.

2 Metodología

El corpus analizado está integrado por 330 frases (3 repeticiones de 5 frases – con estructura sintáctica SVO o VO– en 2 modalidades –declarativa neutra e interrogativa absoluta neutra– producidas por 11 hablantes nativos de catalán oriental central –7 mujeres y 4 hombres–). Para neutralizar los posibles efectos de la longitud de los constituyentes sobre el fraseo prosódico, los sujetos o los verbos analizados del pretonema están constituidos por una sola palabra prosódica (incluyendo palabras funcionales, si es el caso). Las cinco frases en cuestión son las siguientes: La Bàrbara menjava mandarines./?, La Marina mesurava pantalons./?, El Germà maniobrava maquinària./?, Plantàveu albergínies./? y Menjarà merengues./?

Las frases se obtuvieron con una variante del Discourse Completion Task (Blum-Kulka et al., 1989; Billmyer & Varghese, 2000; Félix-Brasdefer, 2010) que consiste en pedir a los informantes que utilicen, en el contexto que se les presenta, una frase que han leído previamente. Se grabaron en el Laboratorio de Fonética de la Universidad de Barcelona con una grabadora digital Marantz modelo PMD620 y un micrófono direccional Shure SM58. El material acústico se analizó mediante un conjunto de scripts de Praat (Boersma & Weenink 2014, Elvira-García, 2014) que, a partir de un .txt anotado y un archivo de sonido, extraen y guardan en una base de datos los valores de duración y de F0 en puntos específicos escogidos para poder estudiar los objetivos propuestos.

El análisis estadístico se efectuó con el paquete SPSS 20. Se llevaron a cabo análisis bivariados con una variable cualitativa (modalidad oracional) y otra cuantitativa (p. ej., rango en st del movimiento tonal). Tras comprobar la normalidad de los datos, se estudió la significatividad de los resultados

mediante el test paramétrico de la t de Student si la distribución era normal y, si no lo era, mediante el test no paramétrico de la U de Mann-Whitney al comparar dos categorías (p. ej. modalidad declarativa vs. modalidad interrogativa), o la prueba de Krusal-Wallis al contrastar los resultados de tres categorías (p. ej., el tipo acentual –llano, agudo y esdrújulo–).

3 Resultados

Por lo que respecta a las oraciones con estructura sintáctica SVO, los resultados son los que aparecen en los apartados siguientes.

3.1 Duración del sujeto

Con el objeto de comprobar si hay diferencias significativas de duración entre el prenúcleo de las declarativas y de las interrogativas, como se ha postulado para otras lenguas (Cangemi y D'Imperio, en prensa), se ha medido este parámetro.

Los resultados medios señalan que la diferencia entre la duración entre ambas modalidades es de +18 ms. Porcentualmente, en general, el sujeto de las declarativas es un 4 % más largo que el de las interrogativas. La t de Student sugiere que esta pequeña diferencia es estadísticamente significativa (t = 3,531, p <0,05); sin embargo, una diferencia del 4 % parece perceptivamente muy poco relevante si se compara con el umbral de perceptibilidad de diferencias de duración (33 % en Pamies & Fernández Planas, 2004); y, además, no hay congruencia entre los informantes puesto que mientras para algunos las interrogativas son más rápidas, para otros lo son las declarativas.

3.2 Valor inicial de F0 de la frase

Según Navarro Tomás (1947), en español el valor inicial de F0 es superior en las interrogativas que en las declarativas. En este trabajo, para evaluar la incidencia de este parámetro hemos calculado el valor inicial de F0 en st de cada enunciado en comparación con el valor medio de F0 de todas las frases producidas por el mismo informante. Los resultados señalan que las frases declarativas comienzan en un nivel de F0 ligeramente más alto, aunque la diferencia no pasa el umbral psicoacústico de 1,5 st (Pamies et al., 2002) y la prueba de la U de Mann-Whitney indica que no es estadísticamente significativa (U = 4686, p> 0,05). Así pues, nuestros datos para el catalán no confirman la afirmación de Navarro Tomás.

3.3 Posición del primer valle

Como se ha indicado en la introducción, la diferencia acústica entre los acentos prenucleares L*+H y L+>H* del catalán se fudamenta en la alineación del valle tonal: mientras que en la estructura L*+H –que caracteriza el prenúcleo de las interrogativas– el valor mínimo de F0 se coloca al final de la sílaba tónica, en la estructura L+>H*, que aparece en el prenúcleo de las declarativas, el valle se sitúa al comienzo de la tónica (p. ej. Prieto et al., 2009); sin embargo, para el catalán la bibliografía disponible no ofrece datos cuantitativos que establezcan la distinción.

En este estudio para calcular con precisión y de manera normalizada la posición del valle en relación con la sílaba tónica, se ha adoptado el siguiente

sistema: se ha atribuido un valor de 100 a la duración de cada sílaba y se ha establecido que el punto cero de la variable tiempo coincida con el principio de la tónica. Así, si el valle cae en medio de la tónica su posición sería de +50; si cayera en la mitad de la primera postónica, de +150; y si cayera en medio de la pretónica, de -50 (el mismo tipo de estandarización se utilizará para evaluar los parámetros relacionados con la posición del pico).

El análisis de los datos del corpus permite obtener los resultados de la tabla 1 en este parámetro. Como se observa, el valle de las declarativas se encuentra más a la derecha que el valle de las interrogativas; en las declarativas, el valle se sitúa próximo al principio de la tónica o al final de la pretónica mientras que en las interrogativas se sitúa en la segunda mitad de la tónica. Además, el tipo acentual parece influir en el sentido de que el valle está más a la izquierda en las agudas que en las llanas y en las llanas que en las esdrújulas.

Medias	Sujeto esdrújulo	Sujeto llano	Sujeto agudo
Frases declarativas	+2	-15	-25
Frases interrogativas	+60	+73	+60

Tabla 1. Medias de los valores normalizados de la posición del valle en relación con el principio de la sílaba tónica

La t de Student indica que la diferencia de alineación del valle entre declarativas e interrogativas es estadísticamente significativa (t = -13.611, p <0,01), de acuerdo con lo que habían señalado estudios anteriores (Prieto, 2001, 2005, 2014, Prieto et al., 2009, Crespo 2010). Por lo que respecta al tipo acentual, la prueba de Kruskal-Wallis confirma que no es significativo en las interrogativas y sí en las declarativas (H=19,621, p <0,01).

3.4 Posición del primer pico en relación con el inicio de la sílaba tónica

Con el mismo sistema expuesto en el apartado anterior se ha evaluado este parámetro, especialmente relevante por cuanto permitirá dilucidar la pertinencia o no desde el punto de vista acústico de una diferencia de alineación muy poco tratada en la bibliografía, la diferencia entre L*+H y L+>H*. Los resultados indican, como se ve en la tabla 2, que el pico de las declarativas siempre está más a la izquierda que el de las interrogativas y que en ambas modalidades el pico está más a la izquierda según la siguiente progresión respecto al tipo acentual: esdrújulas>llanas>agudas. La prueba de la U de Mann-Whitney indica que la diferencia de alineación entre declarativas e interrogativas es estadísticamente significativa (U = 1169, p <0,01). Y respecto al efecto del tipo acentual, la prueba de Kruskal-Wallis confirma que es significativo tanto en las declarativas (H = 50,046, p <0,01) como las interrogativas (H = 34,037, p <0,01).

Medias	Sujeto esdrújulo	Sujeto llano	Sujeto agudo
Frases declarativas	+234	+209	+166
Frases interrogativas	+298	+280	+239

Tabla 2. Medias de los valores normalizados de la posición del pico en relación con el principio de la sílaba tónica

3.5 Posición del primer pico en relación con la frontera de palabra

Se ha empleado el mismo sistema que en el estudio de los parámetros anteriores, pero en este caso se ha situado el cero al final de la palabra. Este parámetro es interesante porque muchos estudios se han ocupado de él (Prieto, 1997; Estebas, 2002, 2003; Fernández Planas et al., 2004; Prieto, 2006; Martínez Celdrán, 2007; Estebas & Prieto, 2007; Feldhausen, 2008; Prieto et al., 2010; Carrera Sabaté et al., 2010).

Los resultados del análisis muestran que el pico tonal de las declarativas se encuentra más a la izquierda que el pico de las interrogativas y que en ambas modalidades el pico está más a la derecha según la siguiente progresión: agudas>llanas>esdrújulas, como se ve en la tabla 3. La prueba de la U de Mann-Whitney indica que esta diferencia de alineación entre modalidades es estadísticamente significativa (U = 2,490, p < 0,01). Por otra parte, la prueba de Kruskal-Wallis confirma que el efecto del tipo acentual es significativo tanto en las declarativas (H = 79,555, p < 0,01) como las interrogativas (H = 68,329, p < 0,01).

Medias	Sujeto esdrújulo	Sujeto llano	Sujeto agudo
Frases declarativas	-65	+9	+66
Frases			
interrogativas	-48	+43	+105

Tabla 3. Medias de los valores normalizados de la posición del pico tonal en relación con la frontera de palabra

3.6 Rango del movimiento tonal dentro de la sílaba tónica

El interés de este parámetro reside en comprobar y cuantificar una afirmación de Prieto et al. (2009) según la cual una diferencia entre las declarativas y las interrogativas consiste en el hecho de que las primeras presentan movimientos tonales ascendentes en el interior de las tónicas prenucleares, mientras que en las segundas el ascenso tonal tiene lugar en las postónicas. Por ello se ha calculado en semitonos el rango entre el punto inicial y el punto final de la sílaba tónica. Efectivamente los datos señalan sin lugar a dudas que las declarativas muestran un ascenso que supera el umbral psicoacústico de 1,5 st (Pamies et al., 2002) a diferencia de lo que ocurre en las tónicas de las interrogativas. Los resultados medios aparecen en la tabla 4. El test de la t de Student indica que esta diferencia es estadísticamente significativa (t = -9,292, p<0,01). Sin embargo, el efecto del tipo acentual según la prueba de Kruskal-Wallis no es significativo en ninguna de las dos modalidades oracionales.

Medias	Sujeto esdrújulo	Sujeto llano	Sujeto agudo	TOTAL
Frases			+3,5	
declarativas	+4,0	+2,7		+3,4
Frases				
interrogativas	-0,1	-0,9	+0,1	-0,3

Tabla 4. Medias de los valores normalizados en st del rango del movimiento tonal dentro de la sílaba tónica

3.7 Rango del movimiento tonal entre el principio de la sílaba tónica y el pico

Entre las diferencias de rango, muy poco tratadas en la bibliografía, en este parámetro se ha buscado cuantificar la diferencia de rango en semitonos entre el inicio de la tónica y el pico en los acentos prenucleares de las declarativas neutras y de las interrogativas absolutas. De forma obvia, el ascenso es más acusado en las interrogativas que en las declarativas aunque en ambos casos se sobrepasa el umbral psicoacústico de 1,5 st, como se ve en la tabla 5. La diferencia entre ambas modalidades es estadísticamente significativa según la t de Student (t = -8,677, p < 0,01), pero no hay significación estadística respecto al efecto del tipo acentual según la prueba de Kruskal-Wallis.

Medias	Sujeto esdrújulo	Sujeto llano	Sujeto agudo	TOTAL
Frases			+5,9	
declarativas	+7,0	+6,1		+6,3
Frases				
interrogativas	+8,5	+8,8	+8,0	+8,4

Tabla 5. Medias de los valores en st del rango del movimiento tonal entre el principio de la sílaba tónica y el pico

3.8 Rango del movimiento tonal entre el valle y el pico

Respecto a este parámetro, los resultados obtenidos a partir del análisis de nuestros datos confirman que el ascenso de F0 es mayor en las interrogativas que en las declarativas y que en ambos casos supera el umbral de 1,5 st (véase la tabla 6). Lógicamente, se trata de la diferencia mayor de rango entre todos los parámetros estudiados dedicados a él. La diferencia entre las dos modalidades respecto a este punto es estadísticamente significativa según la t de Student (t = -14,042, p < 0,01); pero el efecto del tipo acentual, en cambio, según la prueba de Kruskal-Wallis no es significativo, ni en las interrogativas ni en las declarativas.

Medias	Sujeto esdrújulo	Sujeto llano	Sujeto agudo	TOTAL
Frases			+6,7	
declarativas	+7,2	+6,5		+6,8
Frases				
interrogativas	+10,7	+10,2	+9,7	+10,2

Tabla 6. Medias de los valores normalizados en st del rango del movimiento tonal entre el valle y el pico tonal

3.9 Resultados de las frases VO

En los apartados anteriores han quedado reflejados los resultados obtenidos a partir del análisis acústico efectuado en el pretonema de las 198 frases con estructura SVO y también su significatividad estadística. Las 132 frases con estructura VO dan resultados absolutamente similares en todos los parámetros estudiados. No se han encontrado diferencias significativas (ni desde el punto de vista psicoacústico ni desde el punto de vista estadístico) en cuanto a la duración y al valor inicial de F0, mientras que se han encontrado para todos los demás parámetros.

En estas frases de estructura VO se observa, como en las frases de estructura SVO, que el primer valle tonal se sitúa más a la derecha en las interrogativas que en las declarativas (t = -10.2, df = 130, p < 0.01). Respecto a la posición del

primer pico en relación con el principio de la sílaba tónica, también el pico se encuentra significativamente más a la derecha en las interrogativas que a las declarativas; y más a la derecha en frases con verbo llano que en las frases con verbo agudo (U = 1469, p <0,01).

Estas frases VO también presentan diferencias estadísticamente significativas (U = 1074, p < 0.01) con respecto a la alineación del pico tonal en comparación con la frontera de palabra. El pico tonal de las declarativas se encuentra más a la izquierda que el pico de las interrogativas y, tanto en las declarativas como las interrogativas, el pico está más a la derecha en las agudas que en las llanas y más en las llanas que en las esdrújulas.

Sobre el movimiento tonal en el interior de la sílaba tónica, se observa de nuevo que las declarativas contienen una subida psicoacústicamente significativa, mientras que en las interrogativas no hay ningún movimiento tonal significativo. Estas diferencias quedan reflejadas estadísticamente (U = 1.023, p < 0.01).

Respecto al rango del movimiento tonal entre el principio de la tónica y el pico, en estas frases también el ascenso de F0 pasa el umbral de 1,5 st tanto en las declarativas como en las interrogativas y estas últimas presentan un movimiento más acusado que aquellas de forma estadísticamente significativa (U = 783, p < 0.01).

Finalmente, sobre el rango del movimiento tonal entre el primer valle y el pico siguiente, se confirma de nuevo que el ascenso de F0 es más acusado en las interrogativas que en las declarativas y que la diferencia es estadísticamente significativa (U = 511, p < 0.01).

4 Discusión

Entre los resultados obtenidos hemos descrito los efectos de dos variables independientes (modalidad oracional y tipo acentual) sobre algunas variables dependientes (duración del sujeto, valor inicial de F0, alineación del valle y del pico de F0 y aspectos relativos al rango del movimiento tonal). Ninguna de las variables independientes tiene efectos significativos sobre la duración del sujeto y el valor de F0 de inicio de las frases. Sobre el rango de los movimientos, se ha podido confirmar el resultado de Martínez Celdrán et al. (2005) según el cual el ascenso de F0 es más acusado en las interrogativas que en las declarativas y no está condicionado por el tipo acentual de palabra. La imagen de la figura 1 ilustra qué sucede en un espacio virtual cuando se cruzan los valores de movimiento de rango en el interior de la tónica y de movimiento de rango entre el valle y el pico. Como se ve, aunque existe una zona de solapamiento, existe una tendencia muy clara a definir dos grandes grupos, uno para las enunciativas y otro para las interrogativas en función de las diferencias de rango en los movimientos tonales.

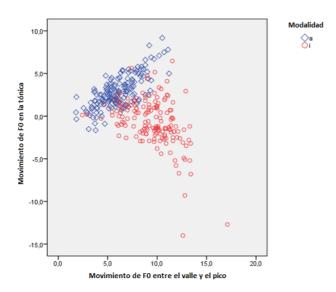


Figura 1. Clasificación de las oraciones por modalidades a partir de los rangos de los movimientos

Por lo que respecta a la alineación de la diana tonal baja o valle, se confirma experimentalmente la hipótesis de Prieto (2001, 2005, 2014), según la cual el valor mínimo de F0 se encuentra en el inicio de la tónica en las declarativas y próximo a su fin en las interrogativas.

Sobre la alineación de las dianas tonales altas, cuestión sobre la que existen abundantes descripciones entonativas del catalán para las oraciones declarativas (Prieto, 1997, 2006; Estebas, 2002, 2003; Estebas & Prieto, 2007; Feldhausen, 2008; Prieto et al., 2010) pero no para las interrogativas, nuestros resultados confirman que las interrogativas presentan el pico más a la derecha que las enunciativas tanto en relación con el inicio de la tónica como con el final de la palabra y en todos los casos existen efectos significativos del tipo acentual de palabra.

Vistas estas diferencias acústicas, se impone ahora ver qué sucede desde el punto de vista perceptivo (Roseano et al., 2014) para cada parámetro, tal como propusieron Martínez Celdrán & Fernández Planas (2010) y qué implicaciones tiene todo ello sobre el etiquetaje de los acentos pretonemáticos o prenucleares.

5 Conclusiones

Este trabajo pretendía comprobar si existían diferencias psicoacústicamente y estadísticamente significativas entre los acentos prenuclears de enunciados declarativos neutros e interrogativos absolutos neutros en catalán central.

Los resultados obtenidos con nuestros datos señalan que no hay ninguna diferencia significativa en cuanto a la duración del sujeto y el valor inicial de F0, pero sí las hay en cuanto a la alineación de las dianas tonales y el rango de movimientos. Sobre la alineación, se puede destacar que tanto el valle como el pico se alinean más a la derecha en las interrogativas que en las declarativas. En cuanto al movimiento de subida de F0, se observa que es más amplio en las interrogativas que en las declarativas y, además, que en el prenúcleo de las declarativas se da un movimiento significativo de F0 en la sílaba tónica que no aparece en las preguntas.

6 Bibliografía

- Billmyer, K., & Varghese, M. (2000). Investigating instrument-based pragmatic variability: Effects of enhancing discourse completion tests. Applied Linguistics, 21 (4), 517-552.
- Blum-Kulka, S., House, J., & Kasper, G. (1989). Investigating cross-cultural pragmatics: An Introductory overview. En S. Blum-Kulka, J. House & G. Kasper (eds.), Cross-cultural pragmatics: Requests and apologies. Norwood: Ablex, 1-34.
- Boersma, P., & Weenink, D. (2014). Praat: Doing phonetics by computer [Computer program]. Versión 5.3.75. Recuperado el 30 de 4 de 2014, de http://www.praat.org/
- Cangemi, F., & D'Imperio, M. (en prensa).

 Beyond f0: Sentence modality and speech rate. En J. Romero y M. Riera (eds.), Selected papers from the 5th Phonetics and Phonology in Iberia Conference. Amsterdam: John Benjamins.
- Carrera-Sabaté, J., Fernández Planas, A. M., & Martínez Celdrán, E. (2010).

 Declaratives i interrogatives absolutes del català en el marc del projecte internacional Atles Multimèdia de Prosòdia de l'Espai Romànic. Caplletra, 49, 133-167.
- Crespo-Sendra, V. (2010). L'entonació de les oracions interrogatives absolutes: un estudi de contrast dialectal. Lengua y Literatura, 21, 199-226.
- Elvira-García, W. (2014). Extraction of prosodic data and alignment v. 1.0. [praat script], Recuperado el 30 de 4 de 2014, de http://stel.ub.edu/labfon/
- Estebas Vilaplana, E. (2002). The modelling of prenuclear accents in Central Catalan declaratives. Catalan Journal of Linguistics, 2, 97-114.
- Estebas Vilaplana, E. (2003). Catalán prenuclear accents: Evidence for word-edge tones. En M. J. Solé, D. Recasens & J. Romero (eds.), Proceedings of the 15th International Congress of the Phonetic Sciences. Barcelona: Causal Productions, 1779-1782.
- Estebas, E., & Prieto, P. (2007). Production and perception of word-edge tones in Catalan and Spanish. En M.

- González, E. Fernández & B. González (eds.), Actas del III Congreso de Fonética experimental. Santiago de Compostela: Universidad de Santiago de Compostela, 279-290.
- Feldhausen, E. (2008). The Prosody-Syntax Interface in Catalan. Tesis doctoral, Humanwissenschaftlichen Fakultät der Universität Potsdam.
- Felix-Brasdefer, J. C. (2010). Data collection methods in speech act performance: DCTs, role plays, and verbal reports. En A. Martínez-Flor & E. Usó-Juan (eds.), Speech Act Performance: Theoretical, Empirical, and Methodological Issues. Amsterdam/Philadelphia: John Benjamins, 41-56.
- Fernández Planas, A. M., Martínez Celdrán, E., Carrera Sabaté, J., van Oosterzee, C., Salcioli Guidi, V., Castellví Vives, J, & Szmidt, D. (2004). Interrogatives absolutes al barceloní i al tarragoní (estudi contrastiu). Estudios de fonética experimental, XIII, 129-155.
- Martínez Celdrán, E., Fernández Planas, A. M., Salcioli Guidi, V., Carrera Sabaté, J., & Espuny Monserrat, J. (2005). Approche de la prosodie du dialecte de Barcelona. Géolinguistique, Hors serie 3, 153-175.
- Martínez Celdrán, E. (2007). Los dialectos catalanes y su prosodia. En J. Dorta (ed), Temas de dialectología, La Laguna (Tenerife): Instituto de Estudios Canarios, 123-140.
- Martínez Celdrán, E., & Fernández Planas, A. M. (2010). ¿Son diferentes las estructuras L*+H y L+>H* perceptivamente? Presentación en el II Congrés sobre entonació del català i Cat_ToBI, Barcelona, Universitat Pompeu Fabra, 9 de 7 de 2010.
- Navarro Tomás, T. (1974). Manual de entonación española. Madrid: Guadarrama.
- Pamies Bertrán, A., & Fernández Planas, A. M. (2004). La percepción de la duración vocálica en español. En J. D. Luque Durán (ed.), Actas del V Congreso Andaluz de Lingüística General. Granada: Granada Linguistica, tomo I, 501-512.
- Pamies Bertrán, A., Fernández Planas, A. M., Martínez Celdrán, E., Ortega Escandell, A., & Amorós Céspedes, M. C. (2002). Umbrales tonales en

- español peninsular. En J. Díaz García (ed.), Actas del II Congreso de Fonética Experimental. Sevilla: Universidad de Sevilla, 272-278.
- Prieto, P. (1997). Prosodic manifestation of syntactic structure in Catalan. En F. Martínez Gil & A. Morales Frente (eds.), Issues in the Phonology of the Major Iberian Languages. Washington DC: Georgetown University Press, 179-199.
- Prieto, P. (2001). L'entonació dialectal del català: el cas de les frases interrogatives absolutes. En A. Bover, M. R. Lloret & M. Vidal-Tibbits (eds.), Actas del Noveno Coloquio de la North American Catalan Society. Barcelona: Publicacions de l'Abadia de Montserrat, 347-377.
- Prieto, P. (2005). En torno a la asociación tonal en el modelo métrico-autosegmental. Puntos controvertidos en su aplicación al catalán. Revista Internacional de Lingüística Iberoamericana, 6, 3, 9-28.
- Prieto, P. (2006). Word-edge tones in Catalan. Italian Journal of Linguistics, 17, 39-71.

- Prieto, P. (2014). The Intonational Phonology of Catalan. En S. A. Jun (ed.), Prosodic Typology 2. Oxford: Oxford University Press, 43-80.
- Prieto, P., Aguilar, L., Mascaró, I., Torres Tamarit, F. J., & Vanrell, M. M. (2009). L'etiquetatge prosòdic Cat_ToBI. Estudios de Fonética Experimental, 18, 287-309.
- Prieto, P., Estebas, E., & Vanrell, M. M. (2010). The relevance of prosodic structure in tonal articulation. Edge effects at the prosodic word level in Catalan and Spanish. Journal of Phonetics 38, 4, 688-707.
- Roseano, P. (2012). La prosòdia del friülà en el marc de l'Atles Multimèdia de Prosòdia de l'Espai Romànic. Tesis doctoral, Universitat de Barcelona.
- Roseano, P., Fernández Planas, A. M., Elvira-García, W., Cerdà Massó, R., & Martínez Celdrán, E. (2014): Diferencias perceptivas entre los acentos tonales prenucleares en catalán. Comunicación presentada en el VI Congreso Internacional de Fonética Experimental. Valencia, noviembre de 2014.

Agradecimientos: Este trabajo se ha realizado con una ayuda del MCI, ref FFI2012-35998.

Prosodia e interpretación: La calidad de las interpretaciones desde una perspectiva prosódica

FIGUEROA LÓPEZ, HELENA

1 Introducción

En el desempeño de su profesión, los intérpretes prestan un servicio a terceros, por lo que de su destreza depende mantener y aumentar el número de encargos. Por tanto, es de vital importancia cuidar la calidad de las interpretaciones, aspecto que preocupa enormemente a los intérpretes. Es por eso por lo que, desde los años ochenta (década en la que se inicia la investigación teórica y empírica en torno a la interpretación), se han realizado numerosos estudios con el objetivo de determinar cuáles son los parámetros que inciden en la valoración de una interpretación.

Dada su similitud en lo que al encargo se refiere, se ha estudiado tanto en el ámbito de la interpretación simultánea como en el de la consecutiva. No obstante, en el presente estudio se trabaja con interpretaciones simultáneas y, por tanto, solo se reseñarán los trabajos relativos a esta modalidad. Con todo, gran parte de las conclusiones extraídas de la bibliografía de este campo pueden aplicarse a la modalidad consecutiva.

El primero que se interesó por el aspecto de la calidad en la modalidad simultánea fue Bühler (1986), para lo cual elaboró una lista de dieciséis criterios que presentó a un total de cuarenta y siete intérpretes profesionales, a quienes pidió que puntuasen esos criterios según su importancia. A partir de los resultados, Bühler buscaba inferir la opinión de los usuarios. Los criterios señalados como más relevantes fueron (en este orden): correspondencia con el sentido del mensaje original, cohesión lógica de la versión, corrección del uso terminológico, interpretación completa, fluidez de la prestación, corrección gramatical, acento nativo y voz agradable.

Posteriormente, Kurz (1989) quiso comprobar la validez de esta correlación directa entre las expectativas de los intérpretes y las de los usuarios. Con este objetivo, realizó una encuesta empleando los criterios de este autor, pero esta vez dirigida a los usuarios. De este modo, concluyó que los criterios puntuados como más importantes coincidían entre ambos grupos, aunque con un porcentaje inferior.

No obstante, estos estudios se centraban únicamente en valoraciones previas a la actuación de un intérprete, pero todavía no se había realizado ningún estudio acerca de las evaluaciones tras la escucha de una interpretación. Con el objetivo de subsanar esta falta de referencias, Miriam Shlesinger (1994) y Franz Pöchhacker (1994) publicaron sendos trabajos de los que se extrae, principalmente, que los aspectos prosódicos juegan un papel crucial en la comprensión del mensaje emitido y en la satisfacción del usuario.

Posteriormente, se han realizado múltiples estudios que buscan dilucidar cuáles son los efectos de la prosodia de los intérpretes sobre los usuarios, y cómo estos efectos determinan una valoración positiva o negativa por parte de estos últimos. Destacan los trabajos de Garzone (2003), Palkowska y Wolańska (2008) y Holub (2010). En lo que se refiere al español, no podemos olvidar la contribución de Ángela Collados (1998), cuyo estudio se enmarca bajo el proyecto de investigación ECIS (Evaluación de la Calidad en Interpretación Simultánea), de vigencia actualmente, en el que además de español se trabaja con otras lenguas.

Dado que estos estudios se asemejan experimentalmente al de Shlesinger (1994), comenzaremos explicando qué hizo su autora. Shlesinger (1994) escogió diez interpretaciones simultáneas (seis hebreo-inglés y tres inglés-hebrero) realizadas por ocho intérpretes de conferencias profesionales, las transcribió y pidió a los propios intérpretes que las rehablasen. Las grabaciones fueron mostradas a dos grupos de sujetos, que no eran profesionales de la interpretación: el grupo A se encargó de la versión interpretada y el grupo B de los rehablados. A continuación, cada participante respondió a un cuestionario con el que se pretendía determinar cuáles habían sido las diferencias de comprensión y retención de la información entre un grupo y otro. El cuestionario demostró que los niveles de comprensión y retención de la información de los oyentes eran inferiores en el grupo de las interpretaciones que en el de los rehablados.

Por su parte, Collados (1998) trató de establecer cuál es la relación entre las expectativas y los juicios reales de valor. En primer lugar, solicitó a cuarenta y dos usuarios especialista en interpretación y a quince intérpretes profesionales que determinasen, por orden de relevancia, cuáles son los criterios de calidad que debe cumplir una buena interpretación. Después, les pidió que evaluasen una serie de simulaciones de interpretaciones simultáneas (inglés-español), que se diferenciaban por su entonación, bien monótona, bien viva; y por correspondencia con el sentido original: unas eran totalmente fieles y otras perdían o tergiversaban la información. De este modo, se descubrió que las interpretaciones con entonación viva, aunque con errores, recibían una puntuación mayor que las versiones monótonas con correspondencia total del sentido con el texto original.

En esta misma línea, Garzone (2003) pidió a dieciséis participantes que puntuasen sus criterios para determinar una interpretación como buena; posteriormente, comparó estas expectativas con juicios de valor de interpretaciones reales (al italiano), las cuales presentaban o no prosodia irregular. De nuevo, las interpretaciones con una prosodia anómala recibieron

una puntación inferior, no solo en cuanto a las características de la producción, sino también en cuanto a calidad de voz, fidelidad y coherencia.

Por su parte, Palkowska y Wolańska (2008) también presentaron cuestionarios de expectativas de calidad a quince participantes. Paralelamente, tomaron un texto de Derecho en italiano que se declamó para que pudiese ser interpretado al polaco. De esta interpretación, que incluía tanto el original como la traducción, se obtuvieron tres videos distintos: uno con entonación monótona pero correspondencia total con el discurso original; otro con entonación viva, pero sin correspondencia total con el original; y el último con entonación viva y correspondencia total con el discurso original. Los errores del segundo video (a saber, cifras, omisiones, cambios de gradación, incoherencias, adiciones y tergiversaciones del significado) fueron introducidos en laboratorio. El discurso original fue, en todos los casos, pronunciado con entonación monótona, para que los sujetos del estudio no percibieran el discurso como excelente y eso empañara su valoración de las interpretaciones.

Los quince participantes, expertos legales, fueron divididos en tres grupos, y cada grupo vio un vídeo. Las autoras constataron que los sujetos fueron incapaces de detectar los errores de la segunda versión; por tanto, concluyeron que la falta de correspondencia no afecta a la valoración, puesto que los usuarios no pueden evaluarla. Además, comprobaron que los juicios de valor están claramente relacionados con la entonación: a mejor entonación, evaluación más favorable.

Por último, cabe reseñar el trabajo de Holub (2010), en el que se pretende verificar la hipótesis de que, como apunta Shlesinger (1994), la entonación monótona afecta negativamente a la comprensión del mensaje por parte de la audiencia. Holub (2010) grabó, en condiciones de laboratorio, una interpretación al alemán con entonación viva a partir de un vídeo de una ponencia en inglés. A continuación, esta pista de audio fue modificada mediante el programa de análisis fonético del habla Praati para obtener otra versión con entonación monótona. Un total de sesenta y tres expertos en la materia de la interpretación (marketing postmoderno) fueron divididos en dos grupos y cada uno de ellos escuchó una de las interpretaciones. En ambos casos, se simuló que la interpretación se estaba realizando en directo. Posteriormente, los participantes respondieron a un cuestionario en el que se medía, por un lado, su grado de comprensión y, por otro lado, su grado de satisfacción. El estudio concluyó que, efectivamente, la monotonía afecta negativamente no solo al entendimiento del mensaje, sino también a la valoración total de la praxis profesional del intérprete.

Gracias a todo este trabajo de campo, observamos que, aunque las encuestas previas a la actuación del intérprete premian la correspondencia total con el sentido original, las encuestas posteriores demuestran que el usuario prefiere una interpretación con pérdidas informativas, pero clara e inteligible, frente a una totalmente fiel al mensaje, pero confusa e incomprensible.

Es innegable que se espera que el intérprete dé cuenta del sentido total de lo que el orador está diciendo en la otra lengua, puesto que su misión consiste en ser puente entre dos partes. Sin embargo, en la mayoría de ocasiones el usuario no puede comprobar la *fidelidad* del intérprete. Por tanto, lo valora en tanto que orador y centra su atención en la producción discursiva, donde la prosodia desempeña un importante papel.

¹Praat es una herramienta de análisis fonético del habla desarrollada por Paul Boersma y David Weenink en el Instituto de Ciencias Fonéticas de la Universidad de Ámsterdam.

Dentro de la prosodia, uno de los aspectos con mayor relevancia es la entonación, puesto que desempeña un papel esencial en el proceso comunicativo: cumple una función demarcativa, una función social representativa y una función expresiva (Prieto, 2002). Así pues, una entonación adecuada es crucial en interpretación simultánea. De un lado, el ponente original habrá empleado una entonación determinada, con las consiguientes intenciones comunicativas asociadas, y el intérprete ha de saber dar cuenta de ello. De otro lado, el intérprete también modulará su discurso de acuerdo a las convenciones de la lengua de llegada y buscará, a partir de la entonación, una complicidad con sus interlocutores, es decir, los clientes.

Como hemos visto a lo largo de este apartado sobre el estado de la cuestión, todos los estudios en relación a la calidad de las interpretaciones se apoyan principalmente en encuestas, tanto de corpus reales como de muestras creadas *ad hoc*. En cambio, en un estudio previo al que aquí presentamos (pendiente de publicación) se planteó un enfoque distinto: primero, describir prosódicamente el habla de los intérpretes y, después, realizar las encuestas de valoración. En el presente artículo se presenta la segunda fase; sin embargo, se hará referencia al estudio previo, cuyos principales resultados también se incluirán en el apartado correspondiente. El objetivo conjunto de ambas fases es, en última instancia, contribuir a la línea de investigación que analiza los factores que afectan a la calidad de las interpretaciones, haciendo especial hincapié en el papel preeminente de la prosodia.

La descripción prosódica se realizó en base a un corpus real de interpretaciones simultáneas emitidas por televisión, para así poder retratar lo más fielmente posible el ámbito laboral de los intérpretes. En segundo lugar, y dada la función social representativa de la entonación, las interpretaciones escogidas tenían como lengua meta únicamente el español, y eran traducciones del inglés.

En la caracterización prosódica realizada en el estudio previo, se compararon los valores medios obtenidos del análisis de los intérpretes con valores medios que caracterizan el español coloquial. En este sentido, este estudio también fue pionero: se partió de la idea de que la interpretación se caracteriza por su inmediatez (Pöchhacker, 2004: 11), huyendo así de la concepción tradicional que define a la interpretación como traducción oral, en oposición a la traducción escrita. Así pues, asumimos que las marcas prosódicas ideales de una interpretación simultánea deberían asemejarse a las del registro coloquial que, como bien apunta Briz (1996: 32), se caracteriza especialmente por la ausencia de preparación.

Se analizaron diez intérpretes (cinco masculinos y cinco femeninas; una de las cuales era no nativa) y, de cada uno de ellos, se extrajeron veinte muestras. Cada muestra constituía un grupo fónico, es decir, una secuencia melódica situada entre pausas de más de 0,2 segundos (Cabedo, 2010). Las doscientas muestras resultantes fueron analizadas prosódicamente mediante el programa de análisis fónico Praat, que permite obtener de manera clara y sencilla los valores relativos a los parámetros que se decidieron estudiar: la intensidad, la velocidad de habla y la frecuencia fundamental (F0).

Se concluyó que, tal y como apuntaba Shlesinger (1994), efectivamente podemos hablar de una entonación propia o extraña de los intérpretes. En nuestro caso, las particularidades venían dadas por su comparación con los rasgos característicos del español coloquial. De este modo, determinamos que la prosodia de los intérpretes posee una intensidad media-baja; una velocidad de elocución media-baja o baja; y una F0 variable, con gamas tonales muy diversas.

Como se indica más arriba, y dado el objetivo del estudio, tras la descripción prosódica, creímos necesario completar la investigación con una encuesta, que nos permitiese inferir qué características prosódicas son mejor valoradas. Tras la presentación del marco teórico, explicitamos a continuación la metodología empleada en esas encuestas.

2 Metodología

Para realizar la encuesta de calidad, se ha utilizado el corpus recopilado para el estudio previo (pendiente de publicación): interpretaciones simultáneas correspondientes a diez intérpretes: cinco masculinos y cinco femeninas, siendo una de ellas no nativa.

El objetivo de la encuesta es doble. En primer lugar, comprobar la correspondencia o no entre las expectativas del usuario sobre una interpretación y la valoración que posteriormente realiza tras haber escuchado un ejemplo, para lo cual retomamos los estudios de Bühler (1986) y Kurz (1989). En segundo lugar, establecer relaciones entre los parámetros prosódicos de intensidad, velocidad de elocución y F0, y las valoraciones de los encuestados. Es por eso por lo que la encuesta se divide en cuatro secciones: (a) un breve cuestionario para conocer el nombre, edad y nivel de estudios del participante, (b) una pregunta sobre sus expectativas, (c) una petición de valoración de muestras de interpretaciones simultáneas y (d) una comparación entre pares de intérpretes.

Para establecer las expectativas de los encuestados, se ha replicado el estudio de Kurz (1989). Se presenta a los encuestados con una lista (en orden aleatorio) de los ocho criterios señalados por Bühler (1986) como más importantes en una interpretación de calidad. Los encuestados deben ordenar estos criterios según su importancia, de mayor (1) a menor (8), teniendo en mente lo que para ellos constituye una interpretación simultánea de óptima calidad.

Con las secciones (c) y (d) se busca obtener valoraciones de los intérpretes que, posteriormente, puedan ser relacionadas con los parámetros prosódicos descritos en el anterior estudio. Debido a que los valores de F0 presentaban una enorme variabilidad intralocutor e interlocutor, se decidió separar su evaluación y realizarla por comparación entre pares de interpretaciones. Así pues, en la sección (c) el encuestado ha de evaluar una muestra de audio de cada uno de los diez intérpretes. En lugar de emplear los grupos fónicos, se decidió mostrar fragmentos de entre 40 y 60 segundos, para así facilitar la valoración gracias a una mayor duración. Los encuestados deben pulsar para escuchar la interpretación y, posteriormente, puntuarla del 1 (pésima) al 5 (excelente). Después, de manera opcional, pueden justificar su respuesta en un cuadro de texto. Por último, deben indicar del 1 (muy lenta) al 5 (muy rápida) cómo perciben la velocidad de habla del intérprete.

Por su parte, en la sección (d) se presentan dos pares de intérpretes, uno caracterizado por una entonación más vívida y otro, por una entonación más monótona. El encuestado debe escoger la que más le guste. En este caso, se han empleado los grupos fónicos, ya que, por un lado, habían sido etiquetados previamente en INTSINT y, por otro lado, aun siendo breves, permiten percibir claramente los patrones melódicos de un hablante.

Las encuestas se han realizado vía Internet por medio del servicio gratuito que ofrece e-encuesta², puesto que permite subir audios y resulta muy intuitivo. Además,

² Para más información: http://e-encuesta.com/.

el hecho de que sea una plataforma virtual de libre acceso permite una mayor difusión y la llegada, por tanto, a un mayor número de encuestados. Por último, e-encuesta realiza automáticamente un análisis estadístico preliminar, lo que permite obtener estadísticos básicos de un gran número de datos.

La encuesta ha sido contestada por un total de 33 encuestados de entre 18 y 35 años, de los cuales el 93,94~% ha realizado estudios de Grado, Diplomatura o Licenciatura y un 6,06 % de ciclo formativo de grado superior, de modo que obtenemos un nivel de estudios suficientemente homogéneo.

Cabe señalar que tres de estos encuestados solo completaron la encuesta hasta la sección (b) y no prosiguieron con las valoraciones. No obstante, sus apreciaciones sobre las expectativas se han tenido en cuenta, ya que contribuyen al esbozo de lo que se entiende a priori por "interpretación simultánea de alta calidad". Además, gracias a la estadística es posible contabilizar sus respuestas sin alterar los resultados.

A continuación, se presentan los resultados obtenidos siguiendo esta metodología.

3 Resultados

3.1 Expectativas de los usuarios

Los ocho criterios que debían ordenarse de más a menos importante, quedaron jerarquizados tal y como se muestra en la tabla 1.

Presente estudio	Bühler (1986)	Kurz (1989)
Correspondencia con el sentido del mensaje original	Correspondencia con el sentido del mensaje original	Correspondencia con el sentido del mensaje original
Cohesión lógica de la versión	Cohesión lógica de la versión	Cohesión lógica de la versión
Fluidez del discurso	Corrección del uso terminológico	Corrección del uso terminológico
Interpretación completa	Fluidez del discurso	Interpretación completa
Corrección gramatical	Corrección gramatical	Fluidez del discurso
Corrección del uso terminológico	Interpretación completa	Voz agradable
Voz agradable	Voz agradable	Acento nativo
Acento nativo	Acento nativo	Corrección gramatical

Tabla 1. Comparativa de las encuestas

Si comparamos la jerarquía del presente estudio con las obtenidas por Bühler (1986) y Kurz (1989), se observa que son muy similares: en los tres casos, el criterio correspondencia con el sentido del mensaje original es siempre el más importante, mientras que el de voz agradable queda relegado a los puestos más inferiores.

3.2 Intensidad y velocidad de elocución

Los intérpretes mejor y peor valorados, junto con sus puntuaciones, se presentan por orden en la tabla 2. Se reflejan únicamente los porcentajes de la puntuación en la que más votos han obtenido, aunque para el orden que aquí se presenta se han tenido en cuenta todas las puntuaciones obtenidas.

Mejor valorados		Peor valorados		
Intérprete 8	70 % en puntuación excelente (5)	Intérprete 7	53,3 % en puntuación mala (2)	
Intérprete 1	36,7 % en puntuación excelente (5)	Intérprete 3	43,3 % en puntuación mala (2)	
Intérprete 9	60 % en puntuación buena (4)	Intérprete 4	33,3 % en puntuación mala (2)	
Intérprete 5	50 % en puntuación buena (4)			
Intérprete 10	50 % en puntuación buena (4)			

Tabla 2. Sección (c) - Intérpretes mejor valorados y peor valorados

Comparemos ahora estas valoraciones con los resultados sobre intensidad obtenidos en el estudio previo. En la tabla 3, se pueden consultar los valores medios extraídos para los intérpretes y los correspondientes al español coloquial, descritos por Cabedo (2007). Los valores medios correspondientes a cada uno de los intérpretes pueden consultarse en la tabla 4. Para facilitar las comparaciones, se han marcado en verde los intérpretes mejor valorados (resaltando aquellos considerados como excelentes) y en rojo los peor valorados (resaltando aquel considerado como pésimo). La no nativa aparece subrayada.

	Intensidad máxima (dB)	Intensidad media (dB)
Media del intérprete	77,144	66,35
Media conversación	82,946	77,50

Tabla 3. Valores medios generales de intensidad (estudio previo y Cabedo, 2007)

	Intensidad máxima (dB)	Intensidad media (dB)
Intérprete 1	75,336	66,08
Intérprete 2	76,290	70,00
Intérprete 3	82,021	71,68
Intérprete 4	81,238	70,83
Intérprete 5	77,061	66,07
Intérprete 6	74,009	61,01
Intérprete 7	79,764	65,89
Intérprete 8	75,942	64,61
Intérprete 9	76,491	63,97
Intérprete 10	78,166	68,67

Tabla 4. Valores medios de intensidad de cada intérprete

Contrariamente a lo que cabría esperar, puesto que se asemejan a las medidas prototípicas de una conversación, las intensidades más bajas han sido las correspondientes a interpretaciones mejor valoradas. En cambio, aquellas interpretaciones cuya intensidad se acerca más a la de la conversación, referente que tomamos en el anterior estudio y que mantenemos aquí como ideal, son las que peor puntuación han recibido.

Con el objetivo de saber si estas valoraciones estaban influenciadas por la acción de la velocidad de elocución, las comparamos con las medidas obtenidas para ese parámetro en el estudio previo. Asimismo, también prestamos atención a las evaluaciones lenta-rápida dadas para cada muestra por parte de los encuestados. En la tabla 5 se observa la media obtenida para el intérprete en el estudio anterior, acompañada de la media descrita para el español coloquial por Cabedo (2007). A su vez, en la tabla 6 se muestra la media de cada intérprete, junto con las valoraciones dadas por los encuestados. Utilizamos el mismo sistema de colores expuesto anteriormente.

Se puede observar que el intérprete mejor valorado (8) y el peor valorado (7) coinciden, respectivamente, con la mayor y menor velocidad registradas. De hecho, el intérprete 7 es el que recibe la categorización como *más lenta* según los usuarios de la encuesta (*vid.* tabla 6): el 63,3 % lo puntúan como de *velocidad baja* (2).

	sílabas/s
Media del intérprete	5,67
Media conversación	6,53

Tabla 5. Valores medios de velocidad de elocución (estudio previo y Cabedo, 2007)

	sílabas/s	Valoraciones de la velocidad
Intérprete 1	5,69	Velocidad correcta (3) → 53,3 %
Intérprete 2	5,60	Velocidad correcta (3) → 40 %
Intérprete 3	5,37	Velocidad rápida (4) → 46,7 %
Intérprete 4	6,17	Velocidad correcta (3) → 43,3 %
Intérprete 5	5,33	Velocidad correcta (3) → 66,7 %
Intérprete 6	5,26	Velocidad correcta (3) → 53,3 %
Intérprete 7	4,49	Velocidad baja (2) → 63,3 %
Intérprete 8	7,18	Velocidad correcta (3) → 50 %
Intérprete 9	5,58	Velocidad correcta (3) → 46,7 %
Intérprete 10	5,71	Velocidad correcta (3) → 60 %

Tabla 6. Valores medios de velocidad de elocución de cada intérprete

Sin embargo, esta correspondencia no se refleja de manera tan clara en otras valoraciones de la velocidad. Observemos que, por ejemplo, el intérprete 4 es el tercer peor valorado y, sin embargo, su habla es más rápida que la del resto de intérpretes que han recibido mejores puntuaciones. Por otra parte, en el caso de la intérprete 3, no nativa, observamos que su ratio es bastante bajo y, sin embargo, es percibida como la que más rápido habla. Así pues, tras una primera aproximación, no podemos establecer relaciones entre habla más rápida e interpretación de mayor calidad, ni viceversa.

Por último, revisamos los comentarios en busca de otros motivos que respalden las valoraciones anteriores. En la mayoría de los casos, los argumentos se apoyan en la percepción de la calidad de voz del intérprete. En el caso de los intérpretes mejor valorados (8 y 1) los comentarios son similares y objetan que estos intérpretes "suenan naturales", "transmiten seguridad", "emplean un tono de voz agradable" o "poseen una voz muy convincente". Por lo que respecta a los intérpretes peor valorados, en el caso de 7 y 4, los encuestados arguyen que los intérpretes "no parecen seguros de lo que están diciendo, "no transmiten seguridad", "el discurso no resulta natural" o "la voz es bastante apagada". En el caso de la intérprete 3, todas las justificaciones de las valoraciones negativas giran en torno a su condición de no nativa; así, podemos leer: "cuesta de entender", "las frases tienen poco sentido", "no pronuncia correctamente" o "el acento es demasiado marcado".

3.3 Frecuencia fundamental

En la sección (d), los encuestados debían responder a la pregunta "¿Qué intérprete le gusta más?", escogiendo uno de los dos grupos fónicos que se les presentaba. Este ejercicio se repetía dos veces y, en ambos casos, un audio correspondía a una entonación monótona, mientras que el otro era ejemplo de una entonación más viva.

En la tabla 7, que muestra los porcentajes de votación para uno u otro grupo fónico en cada uno de los dos bloques, queda patente la preferencia por una entonación más enérgica.

	Entonación monótona	Entonación viva
	Intérprete 4	Intérprete 6
Bloque 1	30 %	70 %
	Intérprete 7	Intérprete 8
Bloque 2	3,3 %	96,7 %

Tabla 7. Sección (d) - Preferencias de entonación

A continuación, mostramos los pares de grupos fónicos del bloque 2, para que sirvan como ejemplo paradigmático de las diferencias de entonación encontradas en el estudio anterior. Asimismo, dado que la inmensa mayoría de los encuestados coincide a la hora de escoger la entonación viva como la preferida, servirá como modelo de a qué debe asemejarse una interpretación de óptima calidad desde el punto de vista entonativo. En las figuras 1 y 2 (intérprete 7 y 8, respectivamente), se presentan los grupos fónicos tal y como se muestran en Praat tras la estilización de la curva melódica y su codificación mediante el sistema de transcripción INTSINT.

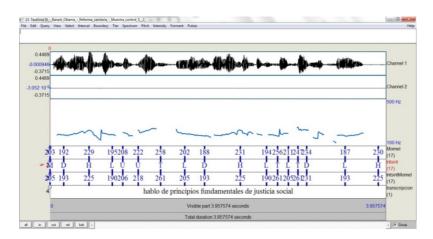


Figura 1. Sección (b) - Bloque 2 - Intérprete 7

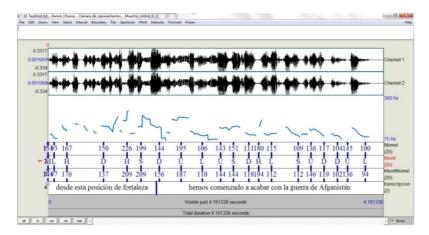


Figura 2. Sección (d) - Bloque 2 - Intérprete 8

4 Conclusiones

Recordemos que el objetivo del presente estudio es establecer relaciones entre la caracterización prosódica de un grupo de intérpretes, en base a los parámetros de intensidad, velocidad de elocución y F0, y las valoraciones emitidas por un grupo de encuestados tras escucharlos interpretar. De este modo, se persigue dilucidar qué características prosódicas debería tener una interpretación de óptima calidad. Mediante este estudio aproximativo se pretende contribuir a la línea de investigación en torno a la cuestión de qué factores afectan a la calidad de las interpretaciones, concretamente en la modalidad simultánea.

En primer lugar, queda patente la ausencia de coherencia entre las expectativas de los usuarios sobre cuáles son los parámetros que caracterizan una interpretación ideal y las valoraciones que realizan tras escuchar una interpretación, como ya habían demostrado Shlesinger (1994) y Pöchhacker (1994). En la sección (b), la voz queda relegada a los últimos puestos de la clasificación; sin embargo, en la sección (c) y (d) se juzga a los intérpretes mayoritariamente según el timbre de su voz y su entonación, premiando a aquellos que resultan agradables al oído y que muestran una entonación enérgica. De hecho, observando los comentarios de los encuestados, vemos que la prosodia afecta a la valoración de otros parámetros que, en principio, nada tienen que ver con ella, como la fidelidad o la coherencia.

En segundo lugar, no podemos establecer relación alguna entre una valoración positiva y una velocidad de elocución alta o baja. Sin embargo, por lo que respecta a la intensidad, parece que los encuestados se decantan por los intérpretes que presentan una intensidad baja. En el estudio previo, donde se describía prosódicamente el habla de los intérpretes (pendiente de publicación), se afirmaba que el habla de los intérpretes simultáneos debería asemejarse prosódicamente al registro coloquial para considerarse de óptima calidad. No obstante, tras las encuestas, concluimos que o bien este enfoque es erróneo y, por lo tanto, hay que buscar un nuevo registro con el que comparar el habla de los intérpretes; o bien, tal y como afirmaba Shlesinger (1994), los intérpretes poseen una prosodia que los caracteriza.

Con todo, no cabe duda de la existencia de una relación directa entre valoración positiva y entonación viva. En la sección (d) comprobamos que los encuestados se inclinan por las interpretaciones moduladas. Combinando estos resultados con algunos de los argumentos aducidos en la sección (c), observamos que los intérpretes son a veces percibidos como menos profesionales cuando hablan monótonamente, hayan o no transmitido correctamente el contenido del mensaje original. Así pues, nos atrevemos a aventurar que la curva melódica es un elemento decisivo a la hora de valorar a un intérprete como bueno o malo.

Por supuesto, somos conscientes de las limitaciones de este estudio y de la necesidad de ampliarlo. En primer lugar, se podrían estudiar los distintos parámetros de manera aislada, para así poder determinar cuál es el papel exacto que desempeña la entonación en las valoraciones. Por otra parte, sería interesante realizar un estudio perceptivo en torno a la cuestión de qué aspectos prosódicos llevan a un usuario a determinar el nivel de seguridad del intérprete, ya que el tema de la confianza transmitida ha sido ampliamente mencionado en las justificaciones. Por último, resultaría enriquecedor incluir el estudio prosódico de las pausas ya que, junto con la intensidad, la velocidad de elocución y la F0, es un parámetro crucial en la caracterización prosódica.

Además, en interpretación se suelen encontrar colocaciones anómalas de las pausas.

Con todo, parece claro que la prosodia desempeña un papel crucial en las evaluaciones que los usuarios hacen de los intérpretes. Así pues, además del conocimiento lingüístico e enciclopédico, creemos que es esencial que los intérpretes ejerciten la voz y dominen las distintas funciones que la prosodia ejerce en el proceso comunicativo. De esta forma, incluso, podrían utilizarla como mecanismo de compensación para aumentar su credibilidad: cuanto más convincente resulte, mejor será valorado.

5 Bibliografía

- Briz, A. (1996). El español coloquial: situación y uso. Madrid: Arco Libros.
- Cabedo, A. (2007). Marcas prosódicas del registro coloquial. *CAUCE*, 30, 41-56. Recuperado el 20 de 4 de 2014, de http://cvc.cervantes.es/literatura/cauce/pdf/cauce30/cauce30_03.pdf
- Cabedo, A. (2010). *La segmentación prosódica en español coloquial*. Valencia: Quaderns de Filologia.
- Collados, Á. (1998). Quality assessment in simultaneous interpreting: The importance of nonverbal communication. En F. Pöchhacker & M. Shlesinger (eds.), *The Interpreting Studies Reader*. London/New York: Routledge, 327-336.
- Garzone, G. (2003). Reliability of quality criteria evaluation in survey research. En Á. Collados, M. M. Fernández & D. Gile (eds.), La evaluación de la calidad en interpretación. Granada: Comares, 23-30.
- Holub, E. (2010). Does Intonation Matter?
 The impact of monotony on listener comprehension. *The Interpreters'*Newsletter, 15, 117-126. Recuperado el 4 de 4 de 2014, de http://www.openstarts.units.it/dspace/bitstream/10077/4753/1/Holub IN15.pdf
- Palkowska, K. & Wolańska, M. (2008). Importance of nonverbal communication in simultaneous interpreting. *Scientific Bulletin of "Politehnica" University of Timisoara Transactions on Modern Languages*. 7/1-2; 37-40. Recuperado el 5 de 4 de 2014, de http://www.cls.upt.ro/files/buletin/2008/palkowska wolanska.pdf
- Pöchhacker, F. (1994). Quality assurance in simultaneous interpreting. En C. Dollerup y A. Lindegaard (eds.), Teaching Translation and Interpreting 2: Insights, aims and visions. Papers from the Second

- Languages International Conference Elsinore, 1993. 8, 233.
- Pöchhacker, F. (2004). *Introducing Interpreting Studies*. London-New York: Routledge.
- Prieto, P. (2002). *Entonació: Models, teoria, mètodes.* Barcelona: Ariel.
- Shlesinger, M. (1994). Intonation in the production and perception of simultaneous interpretation. En S. Lambert y B. Moser-Mercer (eds.), Bridging the Gap: Empirical research in interpretation. simultaneous John **Benjamins** Amsterdam: Publishing Company, 225-236. Recuperado el 4 de 4 de 2014, de http://interpreters.free.fr/simultane ous/schlesinger.htm

Rasgos melódicos de enunciados neutros del español hablado por brasileños

FONSECA DE OLIVERIA, ALINE; MATEO RUIZ, MIGUEL

1 Introducción

Si entendemos la competencia comunicativa, a partir del pensamiento complejo, como un todo más allá de la suma de sus partes, con una perspectiva que contempla el carácter multilingüe de cualquier hablante, esto es, como la interrelación de códigos que conforman su competencia: códigos verbales y no verbales, pero también las variedades lingüísticas que maneja (dialectales, de registro, etc.); la competencia fónica no puede ser entendida como algo añadido a posteriori sino como un elemento inherente a la propia competencia comunicativa (Cantero, 2008). En ella juega un papel fundmental la prosodia, el primer componente lingüístico que adquirimos en nuestra L1; la entonación ejerce un papel primordial en la comunicación, pues contribuye a la fragmentación del discurso en unidades semánticas menores, con el fin de facilitarle al receptor el procesamiento del mensaje (Cortés, 2002).

Para Cantero (2002), al hablar no emitimos sonidos encadenados unos tras otros, sino que estructuramos nuestro discurso siguiendo una jerarquía fónica en la que las vocales son el eje principal, puesto que son los segmentos emitidos de forma consciente por el hablante. Por lo tanto, los sonidos son integrados y jerarquizados en el discurso. Por esto se puede decir que los fenómenos entonativos (acento, ritmo y entonación) cumplen un papel determinante, tanto en la producción fónica como en la percepción de la lengua oral.

El trabajo que aquí presentamos forma parte de una investigación que aborda el análisis y comparación de los diversos tipos de enunciados del español emitidos por hablantes brasileños, caracterizar el acento extranjero de estos hablantes y tiene como finalidad última el desarrollo de aplicaciones didácticas para el aprendizaje del español por hablantes brasileños, que adquieran una completa competencia fónica.

Así, si en Mateo & Fonseca de Oliveira (en prensa) nos ocupamos de los patrones interrogativos, aquí nos hemos centrado en el análisis de los rasgos melódicos característicos de los enunciados neutros producidos por brasileños cuando hablan en español en habla espontánea. Nos hemos focalizado en la entonación prelingüística y contrastamos sus características con las descritas para el español hablado por nativos peninsulares.

2 Metodología

La presente investigación se basa en los presupuestos de la teoría del Análisis Melódico del Habla (AMH), propuestos en Cantero (2002) y Font-Rotchés (2007), que ofrece un método de análisis formal, *Análisis Melódico del Habla* (véase el protocolo de actuación en Cantero & Font-Rotchés, 2009), basado en el análisis acústico y perceptivo del discurso con medios instrumentales, en el proceso se utilizó el programa de análisis y síntesis de voz *Praat* (Boersma & Weenink, 1992-2014), con la finalidad de describir el fenómeno de la entonación desde un punto de vista fonético y fonológico. El modelo establece la existencia de tres tipos de entonación: la prelingüística, la lingüística y la paralingüística.

La entonación *prelingüística* engloba los fenómenos de acento, ritmo y entonación que funcionan solidariamente como elementos fónicos que estructuran el discurso, al margen de cualquier otra dimensión significativa. En este nivel, Cantero (2002) considera que la entonación actúa como un "contenedor" lingüístico que integra y delimita las unidades del discurso. Es lo que Quilis (1993) llamaba "función delimitadora" y "función integradora" de la entonación. Sus principales manifestaciones son el acento *extranjero* y el acento *dialectal*.

La entonación *lingüística* comprende las características melódicas cuyo rendimiento fonológico permite caracterizar y distinguir las unidades funcionales de la entonación, sus "signos lingüísticos". En el español, fueron identificados ocho tonemas (resultantes de la combinación binaria de las características fonológicas: /±interrogativo/, /±suspendido/, /±enfático/), que normalmente son llamados de entonación interrogativa, suspendida, enfática o neutra –ausencia de todos los rasgos– (Cantero, 2002). Los diferentes patrones melódicos que los hablantes producen y perciben a partir de la combinación de dichos rasgos constituyen las melodías típicas del español y fueron descritas por Cantero & Font-Rotchés (2007): un total de doce patrones diferentes, con sus variantes y márgenes de dispersión; posteriormente, Font-Rotchés & Mateo (2011, 2013) describen un nuevo patrón, el XIII, al analizar un corpus más amplio, de variedades del español (Mateo, 2014).

Por último, la entonación *paralingüística* comprende las variantes melódicas de los referidos tonemas, dentro de los amplios márgenes de dispersión de cada uno de ellos. Tales melodías permiten expresar emociones particulares, características discursivas idiolectales o de la personalidad del hablante, algunos usos pueden estar variablemente codificadas, en comunidades delimitadas de habla.

Con el presente estudio se busca seguir caracterizando el perfil melódico del español hablado por brasileños (las características de su "acento extranjero", en este caso, de los enunciados neutros) a partir de la descripción de los rasgos que configuran la entonación prelingüística, presentada de forma detallada en Cantero & Mateo (2011), y cuya principal función es, como hemos señalado anteriormente, integrar el discurso.

2.1 Selección de enunciados

El presente estudio se ha realizado a partir de un corpus de habla espontánea genuina, con 50 enunciados declarativos de entonación neutra, emitidos por 12 informantes procedentes de diferentes regiones de Brasil. Dicho corpus fue extraído de un corpus más amplio, de 511 enunciados, elaborado para fundamentar la investigación sobre las características de la entonación del español hablado por brasileños (Fonseca de Olivera, 2013).

A partir de dicho trabajo, nos hemos centrado en el análisis e interpretación de los contornos de cada enunciado desde el punto de vista de la comprensión de la comentada entonación prelingüística, para caracterizar la del español hablado por brasileños; pues desde su conocimiento se puede entender el *perfil melódico* (Cantero & Devís, 2011) de dicha "variedad" del español.

Este análisis permitirá identificar los procesos de transferencia e interferencia presentes en la interlengua de los brasileños durante el desarrollo de la adquisición de su competencia fónica.

3 Resultados

La entonación neutra es la menos marcada desde el punto de vista fonético y fonológico, pues se caracteriza por la ausencia de cualquiera de los rasgos: /+interrogativo/, /+enfático/ y /+suspendido/.

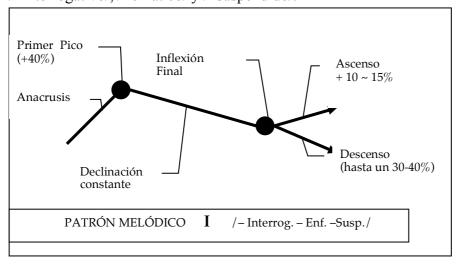


Figura 1. Patrón melódico neutro (Cantero & Font, 2007)

El patrón melódico de la entonación neutra en español estándar, cuyo perfil podemos observar en la figura 1, presenta la inflexión final descendente (hasta un 40% de descenso) o ascendente (hasta un 15% de ascenso), más allá de estos valores, la melodía deja de ser neutra.

En el habla espontánea la entonación neutra no es tan frecuente como las demás (Cantero & Font-Rotchés, 2007), en el caso de nuestro corpus encontramos 359 enunciados con entonación marcada: /+interrogativa/,

/+suspendida/ y /+ enfática/ y 152 enunciados con entonación neutra (Fonseca de Olivera, 2013). De ellos, seleccionamos 50 enunciados neutros para profundizar, en este estudio, en su análisis; trabajo que ha revelado diferencias considerables en los rasgos melódicos constituyentes de los contornos considerados de entonación neutra, si los comparamos con los establecidos para el español estándar

Presentamos a continuación las prinicipales características de los enunciados neutros del español hablado por brasileños para cada una de los tres elementos que constituyen el contorno melódico: el primer pico, la declinación y la inflexión final.

3.1 Primer pico

A través del análisis del corpus se verificó que ,de los 50 enunciados analizados, 8 no presentan primer pico, es decir, el 16 % de los contornos no tienen este rasgo melódico. Se identificaron 33 contornos con picos múltiples, con tendencia a alcanzar la misma altura tonal del primero, es decir, el 66 % de los enunciados presentan esta característica. En tan solo 9 enunciados se identifica la presencia de un supuesto primer pico, es decir, un 18 % del total; de estos, en un 64,3 % de los casos recae en vocal átona y en un 35,7 %, en vocal tónica.

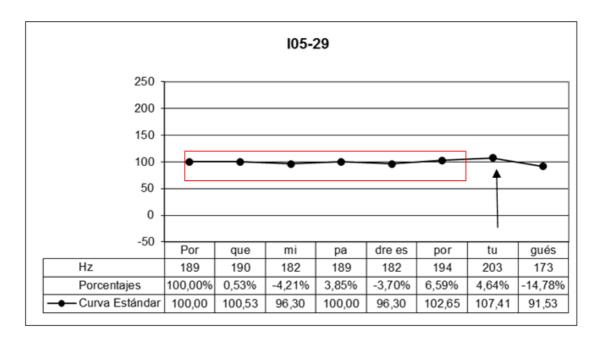


Figura 2. Contorno I05-29: "Porque mi padre es portugués"

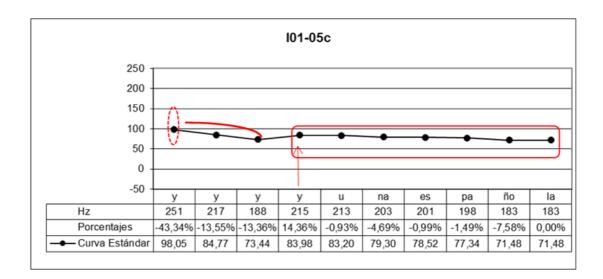


Figura 3. Contorno I01-05c: "Y una española"

Puesto que se considera como criterio para la existencia de primer pico la presencia de declinación, se percibe que al no haber declinación en el cuerpo de los contornos entonativos no se puede considerar que haya primer pico en los enunciados neutros analizados, como podemos observar en el ejmplo de la figura 2, ejemplo de contorno entonativo sin primer pico y con cuerpo completamente plano, que presenta apenas un movimiento tonal final descendente, de un -14,76 %.

En la figura 3, observamos un contorno entonativo iniciado con un primer pico en vocal átona con inflexión descendente, seguida de cuerpo plano, que mantiene la misma altura tonal hasta el final y ausencia de inflexión final.

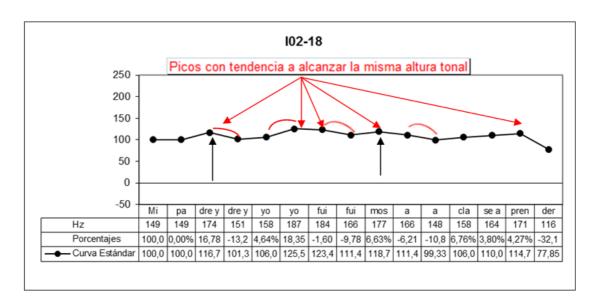


Figura 4. Contorno I02-18: "Mi padre y yo fuimos a clase aprender"

En cambio, el contorno de la figura 4 presenta un primer pico en vocal átona, seguido de múltiples picos con tendencia a alcanzar la misma altura tonal del primer pico. Cuerpo en zigzag causado por los movimientos tonales recurrentes de ligaduras y prominencias tonales. La inflexión final presenta un movimiento tonal final marcadamente descendente (-32,16 %)

3.2 Declinación

El cuerpo de los contornos puede ser agrupado en dos tipologías: planos, es decir, prácticamente sin ningún tipo de movimiento tonal, y ondulados o en zigzag, debido a los movimientos causados por prominencias y ligaduras internas en las vocales, o sea, discretas inflexiones tonales que ocurren a lo largo de los enunciados dentro de los segmentos vocálicos, constituidos por movimientos tonales comprendidos, en general con variaciones de entre el 10 % y el 20 %, aunque en algunos casos pueden darse movimientos más acusados.

El primer caso, cuerpos casi planos, se produce en 22 casos (44 %) de los enunciados del corpus analizado; el resto, 28 (56 %), presentan una declinación ondulada.

En la figura 5, se presenta un ejemplo de contorno entonativo del primer tipo comentado, con cuerpo plano, se mntiene la misma altura tonal en todos los segmentos, en contraste con la declinación descrita para el español estándar (figura 1). No hay inflexión final, es de apenas un -3,65 %, no perceptible.

En la segunda tipología, las variaciones tonales entre las vocales forman picos o movimientos ondulatorios, como podemos observar en el ejemplo de la figura 6, un contorno en el que el cuerpo presenta forma de zigzag, causado por ligaduras y prominencias tonales, formando picos con tendencia a alcanzar la misma altura tonal a del primer pico.

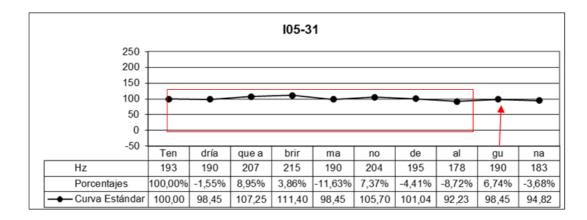


Figura 5. Contorno I05-31: "Tendría que abrir mano de alguna"

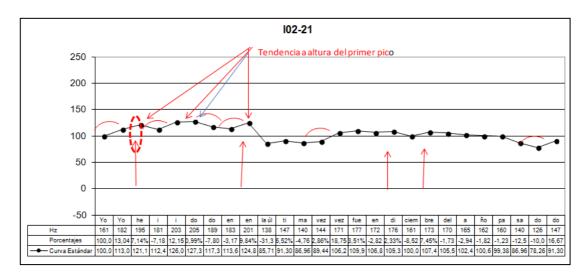


Figura 6. Contorno I02-21: "Yo he ido en la última vez fue en diciembre del año pasado sí"

Contornos similares, con variaciones no tan regulares, han sido identificadas en las variedades meridionales del español peninsular (Mateo, 2014).

3.3 Inflexión final

Por último, por lo que respecta a la inflexión final, la parte del contorno más relevante en el caso de la entonación lingüística; los resultados muestran que 27 de los enunciados analizados (54 %) presentan inflexión final y los otros 23 (46 %) enunciados no, pero la mayoría de ellos muestran movimientos tonales finales en su mayoría descendentes, 80 % del total, con descenso comprendido entre -1 % y -53 %; el 20 % restante se divide en un 10 % con inflexión ascendente, de entre +1 % y +16 %, con una única incidencia de un movimiento marcadamente ascendente, con subida final de +72 %. Finalmente, un 10 % presenta movimientos finales circunflejos de ascenso-descenso y descenso-ascenso.

En la figura 7, observamos el contorno de un enunciado con inflexión final iniciado en la última vocal pretónica del contorno con movimiento descendente, de casi un 23 %. En el caso del contorno de la figura 8, podemos ver un perfil en el que ha ocurrido un pronunciado movimiento tonal final descendente, con una variación de un -52,15 %, superior a la del patrón descrito para el español estándar, en el que, como en casos anteriores, los picos tienden a alcanzar la misma altura tonal del primer pico, marcados con las flechas superiores en el gráfico.

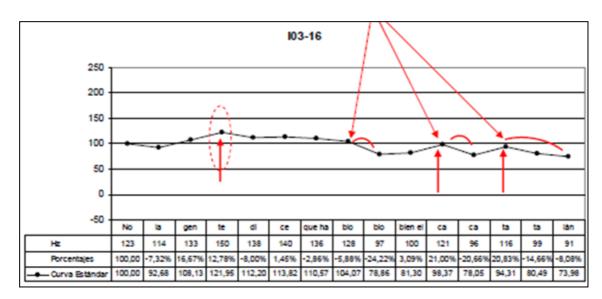


Figura 7. Contorno I03-16: "No la gente dice que hablo bien el catalán"

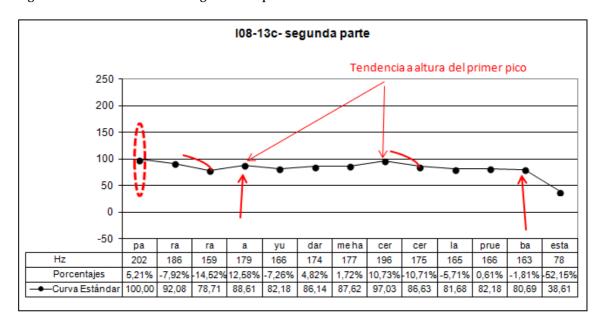


Figura 8. Contorno I08-13c-segunda parte: "para ayudarme a hacer la prueba esta"

En la figura 8, podemos ver un ejemplo de enunciado con inflexión final iniciado en la última vocal pretónica del contorno con movimiento descendente (-52,15 %) en contraste con el contorno del enunciado de la figura 9, de un enunciado con perfil plano, sin inflexión final.

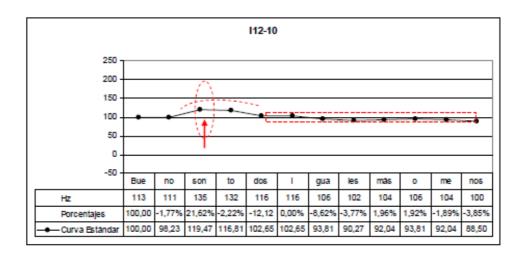


Figura 9. Contorno I12-10: "Bueno son todos iguales más o menos"

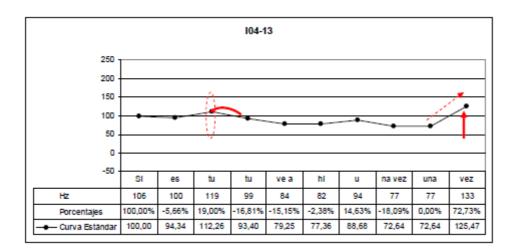


Figura 10. Contorno I04-13: "Sí estuve ahí una vez una vez"

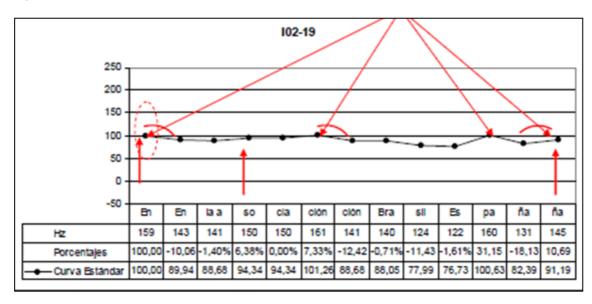


Figura 11. Contorno I02-19: "En la asociación Brasil España"

Para finalizar, ejemplos de un contorno con un ascenso de +72,73 % (figura 10), caso excepcional, y un enunciado que presenta un contorno con

movimiento tonal final circunflejo de descenso-ascenso, -18,13 % / +10,69 % (figura 11).

4 Conclusiones

El análisis del corpus de enunciados neutros permite concluir que los contornos entonativos presentan una entonación prelingüística con rasgos melódicos distintos de los definidos para el español, lo cual indica que la forma de estructuración del discurso sigue otro tipo de jerarquía. Los fenómenos más marcados son la presencia de picos múltiples con tendencia a alcanzar la misma altura tonal del primer pico. Indicios de primer pico, en general, en vocal átona, y en otras situaciones ausencia total primer pico, ausencia de declinación, cuerpos planos (44 %), o bien con movimientos ondulados o en zigzag (56 %), debido a la presencia de prominencias tonales y masivas ligaduras internas que conforman diversos picos con tendencia a alcanzar la misma altura tonal del primero.

La inflexión final es, generalmente, descendente o bien, no existe inflexión final, pero con movimiento tonal con tendencia descendente (80 % de los casos analizados). El porcentaje de descenso se sitúa entre -1 % y -53 %. Hemos encontrado también contornos con final ascendentes con subidas comprendidas, excepto en un caso, entre +1 % y +16 % (un 10 % de los casos). Finalmente, el 10 % de los enunciados presentan movimientos tonales finales circunflejos de ascenso-descenso y de descenso-ascenso.

Estas diferencias con el perfil melódico descrito en la figura 1, junto con las diferencias en la entonación lingüística, como por ejemplo una menor inflexión final (Fonseca de Oliveira, 2013), son las que permiten hablar de ese "acento extranjero" del español hablado por brasileños. Su trabajo en el aula, y fuera de ella, permitirá alcanzar la adquisición plena de la competencia fónica a la que nos referíamos en la introducción de este trabajo y evitará los malentendidos que se pueden derivar deluso de perfiles melódicos con diferentes significados en la L1 (español) y L2 (español hablado por brasileños) con contornos de su L1 (el portugués).

5 Bibliografía

- Boersma, P., & Weenink, D. (2014). *PRAAT*. *Doing phonetics by computer*.

 Amsterdam: Institute of Phonetic
 Sciences, University of Amsterdam.
 En http://www.praat.org
- Cantero, F. J. (2002). *Teoría y análisis de la entonación*. Barcelona: Edicions de la Universitat de Barcelona.
- Cantero, F. J. (2008). Complejidad y competencia comunicativa. *Revista Horizaontes de Lingüística Aplicada*, 7, 71-87.
- Cantero, F. J., & Devís, E. (2011). Análisis melódico de la interlengua, La investigación de la entonación española. En A. Hidalgo, Y. Congosto & M. Quilis, El estudio de la prosodia en España en el siglo XXI: perspectivas y ámbitos. Valencia: Quaderns de Filologia, anejo 75, 285-300

- Cantero, F. J., & Font-Rotchés, D. (2007). Entonación del español peninsular en habla espontánea: patrones melódicos y márgenes de dispersión. *Moenia*, 13, 69-92.
- Cantero, F. J., & Font-Rotchés, D. (2009). Protocolo para el análisis melódico del habla. *Estudios de Fonética Experimental*, XVIII, 17-32.
- Cantero, F. J., & Mateo, M. (2011). Análisis Melódico del Habla: complejidad y entonación en el discurso. *Oralia*, 14, 105-127.
- Cortés, M. (2002). Didáctica de la prosodia del español: la acentuación y la entonación. Madrid: Edinumen.
- Fonseca de Oliveira, A. (2013). Caracterización de la entonación del español hablado por brasileños. Tesis doctoral. Departamento de Didàctica

- de la Llengua i la Literatura. Universitat de Barcelona.
- Fonseca de Oliveira, A., & Mateo , M. (en prensa). A entonação pré-linguística do español e do español falado por brasileiros: análise contrastiva. *Anais do VIII Congresso International ABRALIN*. Ed. Abralin. Associação Brasileira de Lingüística. Brasil
- Font-Rotchés, D. (2007). *L'entonació del català*. Barcelona: Publicacions de l'Abadia de Montserrat.
- Font-Rotchés, D., & Mateo, M. (2011).

 Absolute interrogatives in Spanish, a new melodic pattern. Anais do VII Congresso International ABRALIN. Ed. Abralin. Associação Brasileira de Lingüística. Brasil, 1111-1125. Recuperado el 12 de enero de 2015, de https://docs.google.com/file/d/0B0 SxBl5IJQMEZTc4MDBlZGUtOTk1Yi 00MWI5LWEwZDMtMTJjOGQ3MW MwMWQ5/edit?pli=1
- Font-Rotchés, D., & Mateo, M. (2013).

 Entonación de las interrogativas absolutas del español peninsular del sur en habla espontánea. *Onomazein*, 28, 256-275. Recuperado el 16 de enero de 2015, de http://www.onomazein.net/Articulos/N28/28-20_Font_FINAL.pdf
- Mateo, M. (2014). *La entonación del español meridional*. Tesis doctoral. Departamento de Didàctica de la Llengua i la Literatura. Universitat de Barcelona.
- Mateo, M., & Fonseca de Oliveira, A. (en prensa). ¿Estás enfadado o me lo dices? Rasgos melódicos del español hablado por brasileños. Horizontes científicos y planificación académica en la didáctica de lenguas y literaturas. Braga: Universidade do Minho.
- Navarro Tomás, T. (1944 [1974]). Manual de Entonación Española. Nueva York: Hispanic Institute. Madrid: Guadarrama.
- Quilis, A. (1981). Fonética acústica de la lengua española. Madrid: Gredos.

Realización fonética del acento español en entonación ascendente

KIMURA, TAKUYA; SENSUI, HIROTAKA; TAKASAWA, MIYUKI; TOYOMARU, ATSUKO; ATRIA, JOSÉ JOAQUÍN

1 Introducción

El presente trabajo informa sobre el resultado de un reanálisis de los datos obtenidos en el experimento perceptivo que realizamos en España y en Japón en el año 2010. Para los detalles de su procedimiento y de sus resultados, remitimos a Kimura, Sensui, Takasawa, Toyomaru y Atria (2012). Aquí nos fijamos en la fase de producción, tema no tratado en el trabajo citado, más concretamente en los rasgos prosódicos observados en los estímulos utilizados en el experimento.

Este trabajo es una versión escrita de la comunicación presentada en el VI Congreso Internacional de Fonética Experimental, celebrado los días 5-7 de noviembre de 2014 en la Universitat de València, con el título "Realización fonética del acento español en diferentes contextos entonativos y su percepción por los estudiantes japoneses", aunque la parte que trata la percepción por los estudiantes japoneses está aquí reducida, puesto que ya está explicado detalladamente en el trabajo arriba citado.

2 Resumen del experimento perceptivo y de sus resultados

Aquí nos limitamos a presentar someramente solo la parte esencial del experimento en cuestión. Para más detalles, véase Kimura *et al.* (2012).

Para estudiar la influencia de la entonación oracional sobre la percepción de la posición acentual del español por los hablantes nativos y por los estudiantes japoneses, elegimos como "palabras objetivo" a tres palabras españolas y tres logatomos, diferenciados únicamente en la posición acentual (p. ej. número/numero/numeró; núlido/nulido/nulidó), y las enmarcamos en seis contextos diferentes. Los contextos elegidos fueron los indicados a continuación, en donde la letra X representa la posición de la palabra objetivo en las oraciones.

- (1) Aislado (*Aisl* en lo que sigue): "X".
- (2) En posición final de una oración afirmativa (*Fin-afir*): "Dijo la palabra X".
- (3) En posición no final de una oración afirmativa (Nofin-afir): "Dijo la palabra X la semana pasada".
- (4) En posición no final de una oración negativa (Nofin-neg): "No dijo la palabra X la semana pasada. La dijo ayer".
- (5) En posición final de una oración interrogativa absoluta (Fin-Q): "¿Dijo la palabra X?"
- (6) En posición no final de una oración interrogativa absoluta (Nofin-Q): "¿Dijo la palabra X la semana pasada?"

Se prepararon 270 enunciados como los estímulos en total, 120 estímulos con el contexto Aisl y 30 con cada uno de los otros cinco contextos, a partir de enunciados grabados de un hablante masculino, nativo de Cuenca, España.

Participaron en el experimento un total de 51 estudiantes universitarios japoneses especializados en español y 43 hablantes nativos del español peninsular, quienes contestaron en cuál de las tres sílabas se percibió el acento en cada uno de los 270 estímulos.

La figura 1 muestra los resultados. Se ve que mientras los hablantes españoles perciben los acentos con gran certeza en casi todos los contextos estudiados, los estudiantes japoneses tienden a no acertar la posición del acento cuando la palabra objetivo se pronuncia con entonación ascendente (los contextos Nofin-afir, Nofin-neg y Fin-Q).

El contexto Fin-Q es el contexto más típico de la entonación ascendente, porque es la posición final de una oración interrogativa absoluta. En los contextos Nofin-afir y Nofin-neg también son favorables a la aparición del patrón entonativo ascendente, porque la palabra objetivo se sitúa al final de un grupo entonativo que constituye la primera mitad de una oración enunciativa, que casi obligatoriamente termina con un patrón entonativo ascendente (p. ej., Dijo la palabra número ∕ la semana pasada 🔻).

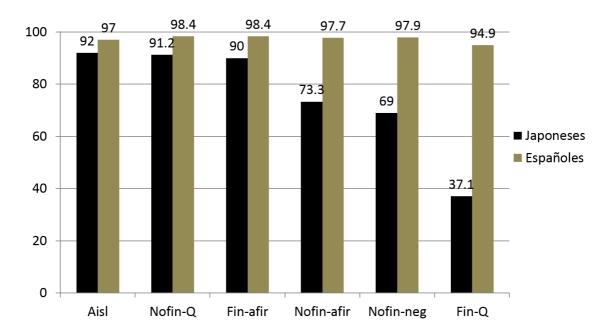


Figura 1. Proporción (%) de percepciones correctas del acento

La relativamente baja tasa de aciertos de los japoneses en la entonación ascendente puede atribuirse a la diferencia de la manera de realizarse el acento léxico en los dos idiomas. El acento léxico japonés se caracteriza por la existencia de una bajada tonal desde la mora acentuada a la siguiente. El tono siempre desciende inmediatamente después de la mora acentuada, aun en un contexto entonativo ascendente (Saito, 2006: 125-126; Vance, 2008: 142-143). Este patrón tonal es parecido al que muestran los enunciados del español con entonación descendente, típicamente cuando se pronuncia una palabra aislada (Real Academia Española/Asociación de Academias de la Lengua Española, 2011: 366-368).

Por otra parte, el acento léxico español muestra comportamientos diferentes cuando se sitúa en un contorno entonativo ascendente. En estas circunstancias ocurre lo que se conoce como "desplazamiento del pico tonal", y como indica la RAE *et al.* (2011: 366 y 451), el acento léxico no siempre es marcado por un aumento de la frecuencia, debido a las restricciones impuestas por la entonación, y el pico tonal puede coincidir con el acento nuclear u ocupar posiciones pretónicas o postónicas.

En las secciones que siguen, veremos en detalle las curvas del F0 de los estímulos con entonación ascendente utilizados en el exprimento arriba mencionado, para ver sus características, sobre todo las relaciones entre la posición de la sílaba acentuada y el movimiento tonal.

3 Curva melódica de las palabras objetivo en contexto de la entonación ascendente

A partir de aquí nos ocuparemos exclusivamente de los 90 enunciados con los contextos *Nofin-afir*, *Nofin-neg* y *Fin-Q*, en los que la palabra objetivo está situada en la entonación ascendente, contextos desfavorables para la percepción acentual de los estudiantes japoneses.

La figura 2 muestra la curva del F0 (frecuencia fundamental) de uno de los estímulos utilizados, "Dijo la palabra *numeró* la semana pasada". Se observa que el F0 desciende en la segunda mitad de la palabra "palabra", que la palabra

objetivo comienza manteniendo el tono bajo, y que el tono vuelve a subir en la segunda mitad de la palabra objetivo "numeró". Este patrón descendente-ascendente (una curva convexa hacia abajo) es común para todos los 90 enunciados en consideración. Dicho de otra manera, existe un punto mínimo local del F0 dentro del marco que abarca desde el punto de comienzo de la sílaba -bra hasta el fin de la palabra objetivo.

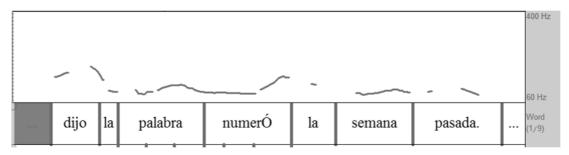


Figura 2. Curva F0 del enunciado "Dijo la palabra numeró la semana pasada"

En lo que sigue, siguiendo a Yanagida, Trípodi, Toledo y Gurlekian (2009: 622), utilizaremos los *log z-scores*, calculados con la siguiente ecuación, en vez de los valores crudos en Hz:

$$\log z$$
-score = $\log (f0 / mg) / \log \deg$

donde *mg* indica la media geométrica y *deg* la desviación estándar geométrica. El objetivo de la utilización de este cálculo es anular los fenómenos microprosódicos innecesarios y estandarizar las diferencias entre los enunciados.

Hicimos medición del F0 en los siguientes 12 puntos de cada enunciado: los puntos de 1/6, 3/6 y 5/6 a partir del comienzo de cada una de las cuatro sílabas, que son la sílaba *-bra* y las tres sílabas que constituyen la palabra objetivo. Llamamos m_n al punto de medición n de la sílaba m, y notamos $z(m_n)$ al log z-score (del F0) del punto m_n . Llamamos para comodidad "la sílaba 0" a la sílaba *-bra* (m = 0, 1, 2, 3; n = 1, 2, 3) (véase la figura 3).

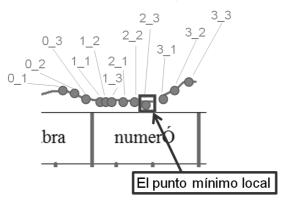


Figura 3. Los 12 puntos de medición y el punto mínimo local del F0

3.1 Relación entre la sílaba acentuada intencionada y la curva melódica

La tabla 1 muestra la relación entre la sílaba acentuada intencionada y el punto de medición con el mínimo local del (log z-score del) F0, según el contexto. Las casillas vacías indican un cero.

Se observa que el F0 toma su valor mínimo local inmediatamente antes del comienzo de la sílaba acentuada, salvo en casi una mitad de los casos con el acento en la primera sílaba, donde el punto mínimo local se atrasa hasta el 1_1. En el contexto *Fin-Q*, hay una tendencia global de atrasarse el punto mínimo local.

Punto mínimo local F0→		0_3	1_1	1_2	1_3	2_1	2_2	2_3	3_1	Total
Síl. acentuada	Contexto									
	Fin-Q	2	6	2						10
1ª.	Nofin-afir	4	6							10
	Nofin-neg	6	4							10
	Fin-Q			3	2	3	2			10
2ª.	Nofin-afir			7	2	1				10
	Nofin-neg			7	3					10
	Fin-Q							9	1	10
3 ^a .	Nofin-afir					2	2	6		10
	Nofin-neg						3	7		10
Tota	1	12	16	19	7	6	7	22	1	90

Tabla 1. El punto mínimo local del F0 según el contexto y la sílaba acentuada

La tabla 2 enseña la sílaba donde el F0 muestra el mayor ascenso entre las tres sílabas de la palabra objetivo (por ser más exacto, el m con el mayor $z(m_3)$ $-z(m_1)$).

En los contextos *Nofin-afir* y *Nofin-neg*, el mayor ascenso ocurre en la sílaba acentuada sin ninguna excepción. En el contexto *Fin-Q*, el mayor ascenso ocurre en la segunda sílaba cuando la primera es acentuada, y en la tercera cuando la primera o la segunda es acentuada.

Síl. acentuada	Contexto	Síl. cor	Total		
Sii. accittada	Contexto	1ª.	2ª.	3ª.	Total
	Fin-Q	1	8	1	10
1^a .	Nofin-afir	10			10
	Nofin-neg	10			10
	Fin-Q			10	10
2 ^a .	Nofin-afir		10		10
	Nofin-neg		10		10
3ª.	Fin-Q			10	10
	Nofin-afir			10	10
	Nofin-neg			10	10
Tota	21	28	41	90	

Tabla 2. La sílaba con mayor ascenso del F0 según el contexto y la sílaba acentuada

3.2 Relación entre la sílaba con el mayor ascenso tonal y la tendencia de la reacción de los sujetos

La tabla 3 muestra la relación entre la sílaba con mayor ascenso y el juicio de los sujetos que escucharon los estímulos.

Síl. con mayor	Contexto (núm. de estímulos	Reacción de los españoles (%)			Reacción de los japoneses (%)		
ascenso	correspondientes)	1.a	2.ª	3.ª	1.a	2.ª	3.ª
	Fin-Q (1)	95.4	4.6	0.0	29.4	21.6	49.0
1ª.	Nofin-afir (10)	99.1	0.9	0.0	70.0	16.5	13.5
	Nofin-neg (10)	98.6	1.4	0.0	60.6	23.3	16.1
	Fin-Q (8)	99.4	0.3	0.3	35.8	26.7	37.5
2 ^a .	Nofin-afir (10)	4.4	94.4	1.2	8.8	54.5	36.7
	Nofin-neg (10)	2.1	95.8	2.1	10.0	54.7	35.3
	Fin-Q (21)	9.0	41.5	49.5	21.9	26.8	51.4
3ª.	Nofin-afir (10)	0.0	0.2	99.8	1.6	2.7	95.7
	Nofin-neg (10)	0.0	0.5	99.5	3.3	4.7	92.0

Tabla 3. Reacción de los sujetos según la sílaba con mayor ascenso y el contexto

Mientras que las reacciones de los japoneses carecen de una tendencia clara, los españoles juzgan generalmente que la sílaba con el mayor ascenso tonal es la acentuada. Hay, sin embargo, dos excepciones notables: primero, los españoles contestan que la primera sílaba está acentuada a los enunciados del contexto Fin-Q con el mayor ascenso en su segunda sílaba; y segundo, para los enunciados del contexto Fin-Q con el mayor ascenso en su tercera sílaba, la mitad de las contestaciones de los sujetos españoles juzgan la segunda sílaba como la acentuada y la otra mitad juzga la tercera como la acentuada (las figuras en **negritas**). Este último hecho, sin embargo, no significa que los españoles contesten al azar, como lo indica la tabla 4, que muestra las relaciones entre la sílaba acentuada intencionada y los juicios de los sujetos españoles, exclusivamente de los 21 enunciados indicados en **negritas** en la tabla 3.

Síl. acentuada	Reacción de los españoles (%)					
(núm. de estímulos correspondientes)	1ª.	2ª.	3ª.			
1 ^a . (1)	97.7	2.3	0.0			
2 ^a . (10)	8.1	86.7	5.1			
3 ^a . (10)	0.9	0.2	98.8			

Tabla 4. Reacción de los sujetos españoles a los estímulos del contexto Fin-Q con el mayor ascenso en la tercera sílaba

Este alto índice de respuestas correctas parece indicar que los sujetos españoles utilizan otras claves aparte del ascenso tonal por lo menos ante algunos enunciados. En lo que sigue estudiaremos la duración sílabica.

4 El acento y la duración silábica

La tabla 5 muestra la relación entre la posición acentual y la duración sílabica observadas en los estímulos utilizados. Las duraciones están mostradas en valores relativos, siendo el 1 la duración de la sílaba *-bra* inmediatamente antes de la palabra objetivo. Se observa una clara tendencia de prolongar la sílaba acentuada. Esta observación hace pensar que los sujetos españoles, también en la fase de percepción, juzgan la posición del acento combinando el movimiento tonal y la duración silábica.

	_	Duración relativa de la					
Síl. acentuada	Contexto	1ª. síl.	2ª. síl.	3ª. síl.			
	Fin-Q	1.10 (±0.13)	0.81 (±0.08)	0.79 (±0.14)			
1 ^a .	Nofin-afir	1.14 (±0.14)	0.82 (±0.13)	0.87 (±0.13)			
	Nofin-neg	1.17 (±0.14)	0.88 (±0.12)	$0.80~(\pm 0.16)$			
	Fin-Q	$0.92 (\pm 0.07)$	1.10 (±0.15)	$0.91 (\pm 0.14)$			
2 ^a .	Nofin-afir	0.97 (±0.11)	1.09 (±0.11)	0.97 (±0.14)			
	Nofin-neg	0.94 (±0.09)	1.07 (±0.15)	0.98 (±0.12)			
	Fin-Q	$0.86 (\pm 0.10)$	0.93 (±0.09)	1.21 (±0.23)			
3ª.	Nofin-afir	0.90 (±0.18)	0.88 (±0.19)	1.43 (±0.09)			
	Nofin-neg	$0.87 (\pm 0.16)$	0.99 (±0.18)	1.50 (±0.15)			

Tabla 5. Duración relativa de cada sílaba de la palabra objetivo (promedio (\pm desviación estándar) de la proporción de la duración sobre la de la sílaba -bra), según la sílaba acentuada y el contexto

5 A modo de conclusión

Las observaciones que hemos mostrado parecen indicar los siguientes hechos prosódicos que ocurren en la entonación ascendente:

- 1) El F0 toma su valor mínimo local inmediatamente antes del comienzo de la sílaba acentuada, salvo en el contexto *Fin-Q*, en cuyo caso se retrasa la ocurrencia del F0 mínimo local.
- 2) El F0 asciende violentamente dentro de la sílaba acentuada. Este ascenso también se retrasa en el contexto *Fin-Q*.
- 3) La sílaba acentuada tiene más duración que las no acentuadas colindantes.

Estos hechos observados sugieren fuertemente la importancia de percibir en qué momento empieza el ascenso tonal y en qué sílaba sube más el tono para una percepción correcta del acento léxico español, por lo menos en contextos entonativos ascendentes. Esto supone una situación inoportuna para los estudiantes japoneses, puesto que el acento léxico japonés se percibe casi exclusivamente por un descenso tonal y los hablantes del japonés son a menudo poco sensibles al comienzo de subida.

Para una mejor comprensión de la producción y la percepción del acento en entonación ascendente, y para aplicar los conocimientos obtenidos a la enseñanza del E/LE a los estudiantes japoneses, quedan por investigar los siguientes temas, entre otros:

- 1) ¿Cómo combinan los españoles (o, más generalmente, los hablantes nativos del español) la información tonal y la duración sílabica para percibir el acento?
- 2) ¿Cómo perciben, o tratan de percibir, el acento español los estudiantes japoneses del E/LE?

Con el objetivo de investigar estos temas, estamos planeando un nuevo experimento de percepción con dos grupos de sujetos, uno de los españoles y otro de los estudiantes japoneses que aprenden E/LE, en el que se pide que los sujetos indiquen la sílaba acentuada como en nuestro experimento anterior, pero esta vez utilizando estímulos con F0 y duración modificados digitalmente, para ver cómo utilizan los sujetos de cada grupo para juzgar la posición del acento.

6 Agradecimientos

Agradecemos profundamente a los asistentes al VI Congreso Internacional de Fonética Experimental por su atención. Nuestro agradecimiento va también, por supuesto, a todos los participantes en nuestros experimentos, en España y en Japón, y a todos los que nos ayudaron para hacer posibles los experimentos.

7 Bibliografía

- Kimura, T., Sensui, H., Takasawa, M., Toyomaru, A., & Atria, J. J. (2012). Influencia de la entonación española en la percepción del acento por parte de estudiantes japoneses. Estudios de Fonética Experimental, XXI, Barcelona, 11-42. En http://stel.ub.edu/labfon/sites/default/files/XXI-05-TKimura.pdf
- Real Academia Española & Asociación de Academias de la Lengua Española (2011). Nueva gramática de la lengua española: fonética y fonología, Madrid: Espasa Libros.
- Saito, Y. (2006). Nihongo Onseigaku Nyuumon [Kaiteiban] (Introducción a la

- fonética japonesa [Edición revisada]), Tokio: Sanseido.
- Vance, T. J. (2008). *The Sounds of Japanese*. Cambridge: Cambridge University Press.
- Yanagida, R., Trípodi, M., Toledo, G. & Gurlekian, J. A. (2009). Análisis comparativo de la entonación en declarativas absolutas sin expansión, *Actas del IV Coloquio Argentino de la IADA*, La Plata, Argentina, 617-632. En http://www.iada-web.org/actas-del-iv-coloquio-argentino-de-la-iada/

Positional effects of prosodic prominence in spoken Italian

LOMBARDI VALLAURI, EDOARDO

1 Information Structure and prosody

In spoken language, utterances express Information Structure (IS) mainly through acoustic patterns. For the main categories of IS that are related to prosody, we adopt the definitions proposed by Cresti (1992; 2000) and Lombardi Vallauri (2001; 2009b), based on the presence/absence of illocutionary force. The Focus is the part of an utterance which carries illocutionary force and realizes the informational purpose of the utterance itself. The Topic is the part of an utterance that has no illocutionary force, whose function is to allow the comprehension of the Focus with respect to the discourse.

Similar definitions for Topic and Focus (Theme-Rheme, Topic-Comment) are more or less explicitly adopted by many other studies concerned with the acoustic correlates of IS (e.g. Halliday, 1989; Ladd, 1978; 1996; Pierrehumbert, 1980; Selkirk, 1984; and, closer to the object of our analysis, Avesani, 2000; Avesani & Vayra, 2004; Avesani *et al.*, 2007; Breen *et al.*, 2010; D'Imperio, 2002; Féry & Krifka, 2008; Frascarelli, 2000; 2004; Frascarelli & Hinterölzl, 2007).

Classical studies (Ladd, 1996) and more recent investigations (Breen et al. 2010; Lee & Xu, 2010) maintain that a focused item might involve a complex combination of different acoustic cues, namely duration, pitch and intensity. Accordingly, the correlation between focused items and prosody must be investigated by considering prosodic prominence as a complex set of acoustic features, whose automatic identification is highly complex.

2 Definition and Detection of Prominence

Following e.g. Couper-Kuhlen, 1986; Jensen, 2004; Kohler, 2006; Mertens, 1991; Terken, 1991, we can define prosodic prominence as a perceptual phenomenon, continuous in its nature, emphasizing segmental units with respect to their surrounding context, and supported by a complex interaction of prosodic and phonetic/acoustic parameters.

In this view, the most relevant role in supporting sentence prominence is played by (i) *pitch accent* (Bolinger, 1958) i.e. movements in F0 profile; (ii) *force accent*, i.e. mainly intensity and segmental durations, as shown in Table 1:

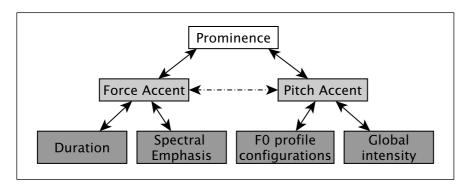


Table 1: Relations between perceived prominence, linguistic-prosodic and acoustic parameters.

Tamburini's (2005; 2006; 2009) experiments, to which we refer for more details, confirmed such relations for different languages, and introduced a prominence function which assigns a continuous prominence level to each syllabic nucleus using only acoustic information:

$$Prom^{i} = W_{FA} \cdot \left[SpEmph_{SPLH-SPL}^{i} \cdot dur^{i} \right] + W_{PA} \cdot \left[en_{ov}^{i} \cdot \left(A_{event}^{i}(at_{M}, at_{m}) \cdot D_{event}^{i}(at_{M}, at_{m}) \right) \right]$$

 $(SpEmph_{Splh-Spl})$ = the spectral emphasis; dur = the nucleus duration; en_{ov} = the overall energy in the nucleus; A_{event} , D_{event} = the parameters derived from the TILT model (Taylor, 2000) as a function of the maxima alignment type – at_{m} – and the minima alignment type – at_{m} . All parameters are referred to the generic syllable nucleus i.)

3 Experimental data

We present here two experiments aimed at comparing IS resulting from expert annotators' judgements to the corresponding position and level of the Main Prominence, identified through Tamburini's algorithm. Portions of linguistic material in utterances from two corpora of spoken Italian were labeled by three linguists as Topic or Focus according to the following criteria:

1.- the perceptive impression (based on acoustic patterns) that a certain part of the utterance conveys illocutionary force, thus being also responsible for the linguistic act carried out by the utterance itself, i.e. for its being an

¹The experiments were carried out in collaboration with Gloria Gagliardi and Fabio Tamburini.

assertion, a question, a request, a command or any other pragmatically relevant act (Cresti, 2000);

- 2.- the confirmation of this information status through illocutionary change tests (Lombardi Vallauri, 1996);
- 3.- the evaluation of the preceding context, aimed at establishing which information may be considered as active (Chafe, 1987; 1992) at the utterance time, i.e. Given, and consequently less likely to be in Focus, and which information may be considered inactive, i.e. New, and consequently more likely to be in Focus.

Only utterances obtaining the agreement of all annotators were considered.

The data for the first experiment were extracted from the "Bonvino" corpus:² 47 utterances were selected from three conversations; the corresponding waveforms were then extracted, and a reference transcription was manually added to mark the syllabic nuclei needed for prominence identification. The data for the second experiment were selected from the map-task sections of *CLIPS*.³ 184 utterances were selected, respectively 64 for Italian spoken in Rome, 59 for Florence and 61 for Naples.

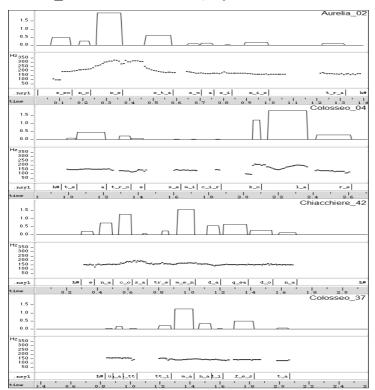
In both experiments, the annotators identified the unit of Focus and possible units of Topic or Appendix if present. They also estimated Focus breadth and possible contrastiveness. The three main classes of utterances were considered:4 (a) Topic-Focus; (b) Broad Focus (extending to the whole utterance); (c) Focus-Appendix (i.e. Narrow and possibly Contrastive Focus located to the left and followed by a Topic⁵). Then the identification of Prominence was made through the automatic algorithm, leading to representations such as those in Table 2:

² A section of Ar.Co.Dip. (Bonvino, 2005), consisting of 12 conversations by young, educated speakers from Rome. 3 Cf. Albano Leoni, 2003: informants are males and females between 20 and 30 born and raised in different Italian cities, with high education degrees.

⁴ This means that utterances containing hesitations, repetitions and parenthetical intonation units were discarded.

⁵ For the nature of Appendixes as postposed Topics, cf. Lombardi Vallauri, 2009a, 2014.

Table 2: The prominence function profiles – Prom – and pitch profiles for some utterances. Aurelia_02: "Secondo me $_{\rm T}$ | stava sulla sinistra $_{\rm F}$ ". Colosseo_04: "Il teatro è semicircolare $_{\rm F}$ ". Chiacchiere_42: "E' una cosa tremenda $_{\rm F}$ | quella donna $_{\rm A}$ ". Colosseo_37: "Una settimana $_{\rm F}$ | di festa $_{\rm A}$ ".



The results of both experiments, summarized in Table 3, show relevant correlations between IS and the position of the Main Prominence, with no significant differences between the considered Italian varieties:

TOPIC - FOCUS							
Syllable marked by	y the m	ain promine	nce			No main	
Internal of Topic	Last c	of Topic	Internal of Foo	cus	Last of Focus	prominence	
3	62		3		5	15	
FOCUS - APPENE	DIX (Na	rrow and/o	Contrastive Foci	ıs)			
Syllable marked by	y the m	ain promine	nce			No main	
Internal of Focus	Internal of Focus Last of Fo		Gocus Appendix Last of		of Appendix	prominence	
11	75		2			4	
BROAD FOCUS						•	
Syllable marked by the main prominence					No main		
Internal of Focus			Last of Focus			prominence	
16		15				20	

Table 3: Position of the Main Prominence with different IS patterns

4 The function of Prominence: demarcation comes before culmination

The positions occupied by automatically measured Prominence in association with the information patterns in the corpus (Table 3) can be summarized as follows:

Topic-Focus

- the majority of utterances have the Main Prominence at the Right end of the Topic;
- a minority seems not to distinguish between the two units, with comparable Prominences

Narrow Focus (at the Left, followed by Appendix)

always marked by the Main Prominence at the Right of the Focus.

Broad Focus

 about half of the utterances have the Main Prominence at the Right, the other half have no Main Prominence, but several minor/equivalent Prominences.

In sum, only constituents located at the left of the utterance (Topic or Narrow Focus), seem to be steadily associated to the Main Prominence. More precisely, what bears the Prominence is the *right end* of such constituents. This suggests that the primary function of the Main Prominence may be *demarcation*, rather than culmination. In other words, its first, immediate effect may be that of drawing a boundary between two information units, no matter if also characterizing one of them by means of a specific, recognizable contour.

This doesn't mean that different intonation patterns cannot express different kinds of Focuses and Topics, effecting different types of illocutions and pragmatic functions. But the *bare presence and position* of the Main Prominence (as it results from our measurements) may suffice to signal addressees if the utterance contains a boundary between Information Units, and where. Then, once the Main Prominence has signaled a boundary between two units, the specific contour of the unit located to the right signals whether it is a Focus or an Appendix. Which, in turn, determines if the one to the left is a Topic or a Narrow Focus.

This may not be a complete account of the features involved in the encoding of IS, but it may be useful, at least as long as there is little agreement in the literature as to what should be the supposedly different qualitative features of different Topics and Focuses.¹ As long as the discussion among scholars is open about what should be regarded as the specific intonational contours of IS categories (and if there are any), a purely structural hypothesis may shed some light on how speakers and hearers can exchange information. In strictly structural terms, the recognition of IS categories can happen even just on the basis of the minimal cues exposed in Table 4.

¹ Cf. Avesani (2000), Avesani - Vayra (2004), Avesani et al. (2005, 2007), Bocci, Avesani (2006), Bonvino (2004, 2005), D'Imperio (1997, 2002b), Face, D'Imperio (2005), Frascarelli (2000, 2004), Frascarelli, Hinterhölzl (2007), Féry, Krifka (2008), Breen et al. (2010), etc.

IS unit	left margin marked by:	right margin marked by:	
Topic	beginning of the utterance and	Main Prominence on the last stressed	
	of the intonational contour	syllable of the Topic	
Right Focus	Main Prominence on the last stressed	end of the utterance and	
after Topic	syllable of the Topic	of the intonational contour	
Broad Focus	beginning of the utterance and	end of the utterance and	
	of the intonational contour	of the intonational contour	
Narrow	beginning of the utterance and	Main Prominence on last stressed syllable	
Focus (at the	of the intonational contour	of Focus and beginning of Appendix flat	
Left)		contour	
Appendix	Main Prominence on last stressed	end of utterance	
	syllable of the Focus and beginning of		
	Appendix flat contour		

Table 4: Minimal perceptual cues for the recognition of IS units.

This hypothesis may be seen as giving the simplest possible explanation of:

- (a) why Topics are marked more strongly than both Broad Focuses and Right Focuses following a Topic, though the communicative import of Focuses is greater than that of Topics. This may be due to the fact that Topics, unlike Broad or Right Focuses, being followed by another major Information Unit within the same utterance, need that the boundary between the two is signaled;
- (b) why Narrow Focuses located at the Left are also strongly marked. This may be due to the same reason, since also Left Focuses are followed by a boundary between Information Units within the utterance.

This is an *exquisitely structural*, *positional explanation* of how the Main Prominence may allow recognition of Information Units in Italian (and possibly elsewhere); in other words, *presence and position of the Prominence*, not its qualitative characteristics, are proposed as the essential features allowing speakers to recognise the IS of utterances. This is summarized in Hypothesis (1):

(1) A Positional Hypothesis on Main Prominence The Main Prominence marks the boundary between Information Units within the utterance.

It can be further observed that the only *qualitative* difference strictly needed in order to recognize the IS of an utterance is that between the marking of a Topic and the marking of a Left, Narrow Focus, because both are followed by another unit. That difference can be signaled by the different intonation contours of the following units (respectively a Right Focus or an Appendix). Then, marking it also by means of specific different intonational contours for the Topic and the Left Focus themselves already constitutes a certain degree of redundancy.

The absence of a Main Prominence, or its being located on the last stressed syllable of the utterance, both signal a Broad Focus (not preceded by a Topic), whose boundaries match the boundaries of the whole utterance, and consequently do not need to be signaled by a Main Prominence.

This view is summarized in Table 5, containing the minimal steps by which the addressee can "compute" the Information Structure of an utterance:

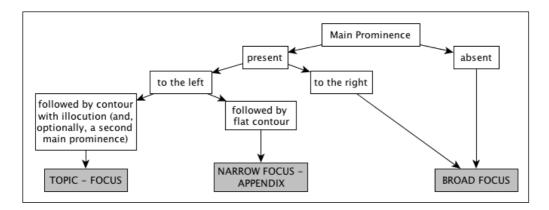


Table 5: Minimal steps for the recognition of IS units

In this interpretation, the expression of IS may simply obey the *law of least effort*. The only elements strictly needed are:

- (i) on a structural level: a Main Prominence per utterance, whose position (in the middle or, optionally, at the end) signals if the utterance contains a boundary between the Information Units or not; and
- (ii) on a qualitative level: the difference between a Focus contour expressing illocution, and the contour of an Appendix, devoid of illocution.

Now, since the different Focus contours are independently needed to express the different illocutions of utterances (i.e. the different linguistic acts: Cresti, 2000), the specific cost strictly required for expressing IS ends up being very low. Marking all information units with culminative, "descriptive" Prominences would cost more effort than simply marking the boundaries, because:

- (a) distinguishing Topic from Focus would require two qualitatively different Prominences (one for each) instead of just one (at the boundary);
- (b) distinguishing Broad Focus from Narrow Focus would require two recognizably different Prominences, because not only Narrow, but also Broad Focuses would need a distinct Prominence.

It should not be a surprise if language prefers to work in a more economic way, namely *marking* (*compulsorily*) *only the marked element*, i.e. not Broad, but Narrow Focus. Table 5 shows that this situation is well represented in the corpus.

201748.			
	Utterances corresponding to Hypothesis (1)	Utterances not corresponding to Hypothesis (1)	
Rome – Bonvino	40 (85.10%)	7 (14.90%)	
Rome – Clips	46 (71.88%)	18 (28.12%)	
Florence – Clips	42 (71.19%)	17 (29.81%)	
Naples – Clips	43 (70.49%)	18 (29.50%)	
TOTAL	170 (73.59%)	61 (26.41%)	

Table 5: Foreseen vs. unforeseen results for IS acoustic realization in the corpus.

This means that the algorithm located Prominences in positions that confirm Hypothesis (1) in almost 74% of utterances. Considering the average convergence between human evaluations and automatic algorithms for the

detection of Prominence attested in literature, this would be already a fairly good result. But we will see right away that Hypothesis (1) is more strongly confirmed by our data.

5 The Topic - Focus continuum

A minority of the utterances perceived as Topic-Focus bear no Main Prominence (Table 3). Conversely, a minority of those evaluated as Broad Focuses have an internal Main Prominence, in a position similar to that of Topic-Focus structures. In other words, utterances acoustically measurable as Broad Focuses can be perceived as Topic-Focus, and vice versa. This can be explained with Topic-Focus and Broad Focus being not separate and reciprocally exclusive structures, rather the extremes of a continuum. The middle of the continuum is occupied by utterances where the boundary between the units is not neatly marked, and the distinction between the two possible Information Structures remains under- or unspecified.

This amounts to saying that, prosodically, the speaker is not bound to decide between Topic-Focus and Broad Focus. This should not surprise if we consider that in each actual situation the speaker and the addressee can evaluate prosodic cues differently, and they are aware thereof. As a consequence, both know that a certain amount of fuzziness must go with the perception of IS.

But it is not only a matter of correct understanding on the part of the addressee for an Information status (Topic or Focus) which is univocally established on the part of the source. It is reasonable that a content can be focused at an unlimited variety of degrees (Daneš, 1967, 1974; Firbas, 1966, 1987, 1989; Sgall 1975; Sgall *et al.* 1973): focusing must not necessarily be either 100% or null. And, more radically, parts of utterances can remain underspecified as for their Information status. Many utterances can be something intermediate between Topic-Focus and Broad Focus, with uncertain Information status of a certain amount of information, typically "in the middle", and the Main Prominence not clearly recognizable. In other words, Topic vs. Focus cannot be a black & white story, rather one in a grey scale. This is the case for the utterances in Figure 3 from our corpus.

 $^{1\,\}mathrm{The}$ matching between perception and measurement usually reveals no more than strong tendencies. The role of discrepancies is evident from Breen et al. (2010), just to cite one recent and very careful study.

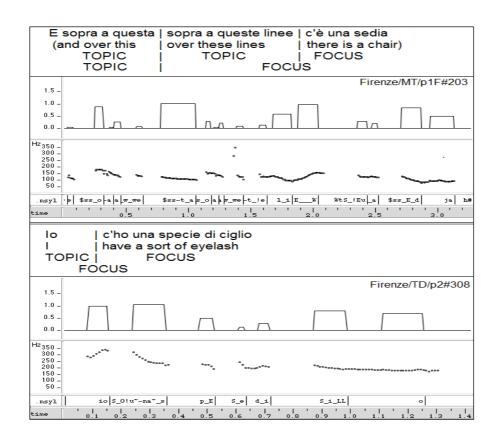


Figure 3: utterances underspecified between Topic-Focus and Broad Focus

From a textual perspective, the absence of a clear-cut distinction between Topic-Focus and Broad Focus reflects their being structures often possible in the same contexts, and their often not influencing subsequent discourse in a decisively different way, which does not hold for Focus-Appendix structures. Then, more in general, underspecification of some features is quite common in many aspects of linguistic structure, not only in IS. For instance, an expression like "Jay's book" is open to different, freely enriched interpretations, such as the book written by Jay, which Jay has read, published, talked about recently, etc. Information less pragmatic in nature is also frequently unspecified. For example, verbal tense or agreement with the subject can remain not overtly expressed in many languages, as shown in (2):

(2) Aisukuriimu-wo taberu (Japanese) icecream-ACC eat "I/you/(s)he/we/they eat/will eat the/an icecream

If we consider all cases in our corpus where Information Structure remains underspecified between Topic-Focus and Broad Focus as consistent with the model, we obtain the new figures depicted in Table 6.

		Utterances not corresponding to Hypothesis (1)
Rome – Bonvino	43 (91.49%)	4 (8.51%)
Rome – Clips	55 (85.94%)	9 (14.06%)
Florence – Clips	53 (89.83%)	6 (10.17%)
Naples – Clips	53 (86.89%)	8 (13.11%)
TOTAL	170 (87.88%)	28 (12.12%)

Table 6: Foreseen vs. unforeseen results for IS acoustic realization in the corpus, including the utterances located on the continuum between Topic-Focus and Broad Focus

In other words, in roughly 90% of the utterances one of the following matchings between their perceptive evaluation and the results of measurement was the case:

- (a) structures evaluated as Topic-Focus, with Main Prominence at the right end of the Topic;
- (b) structures evaluated as Focus-Appendix, with Main Prominence at the right end of the Focus;
- (c) structures evaluated as Broad Focus, either with no Main Prominence or with Main Prominence at the right end;
- (d) structures evaluated either as Topic-Focus or as Broad Focus, with no evident Main Prominence.

Only in 10% of the cases, automatic measurement put the Main Prominence elsewhere. A certain amount of "noise" in the procedure is expected, because (i) human errors while assessing the distribution of Information Units cannot have been completely avoided; (ii) not all data can be exempt from the typical "flaws" of speech, such as imperfect production, changes of intention, etc.; and (iii) no algorithm automatically assigning prominence levels to syllables can be 100% error free.

6 Conclusions

At least for the examined Italian varieties, the following conclusions can be drawn:

- 1. It makes sense to consider a merely "positional" level for the working of Prominence which ensures demarcation between units, more essentially than their culmination and "description". This already provides speakers with much of the cues they need to interpret the IS of utterances in discourse. Further features may help, though with some redundancy. This is the case for the specific inonational contours of the different Information Units, which are independently aimed at expressing illocution.
- 2. The distribution of Topic and Focus need not be expressed clearly. Many utterances remain prosodically (and acoustically) underspecified in this respect. This is especially true for the distinction between Topic-Focus and Broad Focus, with no serious effects on the progression of communicative dynamism in the subsequent discourse.
- 3. If such results and the interpretation we proposed are correct, speakers seem, not surprisingly, to obey the law of least effort in minimizing the formal features they put at work to ensure the expression of relevant information.

4. The very high percent of matching between perceptual evaluations and automatic measurement seem to confirm the efficiency of Tamburini's algorithm for the automatic detection of Prominence.

7 References

- Albano Leoni, F. (2003). Tre progetti per l'italiano parlato. In N. Maraschio & T. Poggi Salani (eds.), Italia linguistica anno Mille. Italia linguistica anno Duemila, 675-683. Roma: Bulzoni.
- Avesani, C. (2000). Costruzioni marcate e non marcate in italiano. Il ruolo dell'intonazione. In D. Locchi, A. Giannini & M. Pettorino (eds.), Il parlante e la sua lingua, 1-14. Napoli: Il Torcoliere.
- Avesani, C., & Vayra, M. (2004). Focus ristretto e focus contrastivo in italiano. In F. Albano Leoni, F. Cutugno, M. Pettorino & R. Savy (eds.), Il Parlato Italiano. Atti del Convegno Nazionale, 1-20. Napoli: D'Auria.
- Avesani, C., Vayra, M., Banelli, D., & Lovisolo, M.G. (2005). Accentazione, deaccentazione e struttura informativa. In P. Cosi (ed.), Misura di Parametri. Aspetti tecnologici ed implicazioni nei modelli linguistici, 287-312. Brescia: EDK.
- Avesani, C., Vayra, M., Zmarich, Paggiaro, R., & Sperandio, D. (2007). Le basi articolatorie della prominenza accentuale in italiano. In V. Giordani, V. Bruseghini, & P. Cosi (eds.), Metodologie di valutazione e Risorse linguistiche, 1-22. CD-ROM.
- Bocci, G., & Avesani, C. (2006). Focus Contrastivo nella periferia sinistra della frase: un solo accento, ma non solo un accento. In R. Savy & C. Crocco (eds.), Analisi Prosodica. Teorie. Modelli e sistemi di annotazione, 111-141. Rimini: EDK.
- Bolinger, D. (1958). A theory of pitch-accent in English. Word, 14, 109-149.
- Bonvino, E. (2004). Fenomeni prosodici in relazione al soggetto postverbale. In F. Albano Leoni, F. Cutugno, M. Pettorino & R. Savy (eds.), Il Parlato Italiano. Atti del Convegno Nazionale (B06). Napoli: D'Auria (CD ROM).
- Bonvino, E. (2005). Le sujet postverbal. Une étude sur l'italien parlé. Paris: Ophrys.
- Breen, M., Fedorenko, E. Wagner, M., & Gibson, E. (2010). Acoustic correlates

- of information structure. Language and Cognitive Processes, 25(7/8/9), 1044-1098.
- Chafe, W. (1987). Cognitive Constraints on Information Flow. In R. S. Tomlin (ed.), Coherence and Grounding in Discourse, 21-51. Amsterdam: Benjamins.
- Chafe, W. (1992). Information Flow in Speaking and Writing. In P. Downing, S.D. Lima, & M. Noonan (eds.), The Linguistics of Literacy, 17-29. Amsterdam: Benjamins.
- Couper-Kuhlen, E. (1986). English prosody. Londn: Arnold.
- Cresti, E. (1992). Le unità d'informazione e la teoria degli atti linguistici. In G. Gobber (Ed.), Linguistica e pragmatica, 501-529. Roma: Bulzoni.
- Cresti, E. (2000). Corpus di italiano parlato. Firenze: Accademia della Crusca.
- Daneš, F. (1967). Order of Elements and Sentence Intonation. In Studies to Honor Roman Jakobson, 499-512. The Hague-Paris: Mouton.
- Daneš, F. (1974). Functional Sentence Perspective and the Organization of the Text. In F. Daneš (ed.), Papers on Functional Sentence Persepctive, 106-128. Prague: Academia/The Hague: Mouton.
- D'Imperio, M. (1997). Breadth of Focus, modality and prominence perception in Neapolitan Italian. In M. Ainsworth-Darnell, & M. D'Imperio (Eds.), The Ohio State University UWP in Linguistics – Papers from the Linguistic Laboratory, 50, 19-39.
- D'Imperio, M. (2002). Italian Intonation: An overview and some questions. Probus, 14(1), 37-69.
- Face, T. L., & D'Imperio, M. (2005). Reconsidering a focal typology: Evidence from Spanish and Italian. Italian Journal of Linguistics, 17, 271-289.
- Féry, C., & Krifka., M. (2008). Information structure. Notional distinctions, ways of expression. In P. van Sterkenburg (Ed.), Unity and diversity of languages, 123-136. Amsterdam: Benjamins.

- Firbas, J. (1966). On Defining the Theme in Functional Sentence Analysis. Travaux Linguistiques de Prague, 1, 267-280.
- Firbas, J. (1987). On the Delimitation of the Theme in Functional Sentence Perspective. In R. Dirven, & V. Fried (Eds.), Functionalism in Linguistics, 137-156. Amsterdam: Benjamins.
- Firbas, J. (1989). Degrees of communicative dynamism and degrees of prosodic prominence (weight). Brno Studies In English, 18, 21-66.
- Frascarelli, M. (2000). The Syntax-Phonology Interface in Focus and Topic Constructions in Italian. Dordrecht: Kluwer.
- Frascarelli, M. (2004). L'interpretazione del Focus e la portata degli operatori sintattici. In F. Albano Leoni, F. Cutugno, M. Pettorino & R. Savy (eds.), Il Parlato Italiano. Atti del Convegno Nazionale (B06). Napoli: D'Auria (CD ROM).
- Frascarelli, M. & Hinterhölzl, R. (2007). Types of Topics in German and Italian. In S. Winkler, & K. Schwabe (Eds.), On Information Structure, Meaning and Form, 87-116. Amsterdam: Benjamins.
- Gagliardi, G., Lombardi Vallauri, E., & Tamburini, F. (2012a). La prominenza in italiano: demarcazione più che culminazione. In M. Falcone, & A. Paoloni (eds.), La voce nelle applicazioni, 255-270. Roma: Bulzoni
- Gagliardi, G., Lombardi Vallauri, E., & Tamburini, F. (2012b). A Topologic view of Topic and Focus marking in Italian. In N. Calzolari, K. Choukri, T. Declerck, M. U. Doğan, B. Maegaard, J. Mariani, J. Odijk, & S. Piperidis (eds.), Proceedings 8th International Conference on Language Resources and Evaluation LREC 2012, 948-955. Istanbul: LREC.
- Halliday, M.A.K. (1989). Spoken and Written Language. Oxford: Oxford University Press.
- Jensen, C. (2004). Stress and Accent. Phd thesis, University of Copenhagen.
- Kohler, K. J. (2006). What is emphasis and how is it coded? In R. Hoffmann & H. Mixdorff (eds.), Proceedings of Speech Prosody, 3, 748-751. Dresden: TUD.
- Ladd, D. R. (1978). The Structure of Intonational Meaning. Bloomington: Indiana University Press.

- Ladd, D. R. (1996). Intonational Phonology.

 Cambridge: Cambridge University

 Press.
- Lee, Y., & Xu, Y. (2010). Phonetic Realization of Contrastive Focus in Korean. In Proceedings of Speech Prosody, 5, Chicago. In http://speechprosody2010.illinois.ed u/papers/100975.pdf
- Lombardi Vallauri, E. (1996). A simple test for theme and rheme in the clause complex. Language Sciences, 17(4), 357-378.
- Lombardi Vallauri, E. (2001). La teoria come separatrice di fatti di livello diverso. L'esempio della struttura informativa dell'enunciato. In F. Albano Leoni, E. Stenta Krosbakken, R. Sornicola & C. Stromboli (eds.), Dati empirici e teorie linguistiche, 151-173. Roma: Bulzoni.
- Lombardi Vallauri, E. (2009a). Appendix or Postposed Topic: where does the difference lie? In L. Mereu (ed.), Information Structure and its Interfaces, 387-411. Berlin: Mouton De Gruyter.
- Lombardi Vallauri, E. (2009b). La struttura informativa. Forma e funzione negli enunciati linguistici. Roma: Carocci.
- Lombardi Vallauri, E. (2014). What can Japanese -wa tell us about the function of Appendixes. Faits de Langues, 43, 61-86.
- Lombardi Vallauri, E., & Tamburini, F. (2012). Topic and Focus marking in an Italian corpus: some results of algorithmic measurement and structural interpretation. In H. Mello, M. Pettorino, & T. Raso (eds.), Proceedings of the VII GSCP International Conference 2012 (Belo Horizonte): Speech and Corpora, 191-196. Firenze: Firenze University Press.
- Mertens, P. (1991). Local prominence of acoustic and psychoacoustic functions and perceived stress in french. In Proceedings of the 12th International Congress of Phonetic Sciences (ICPhS), Aix-en-Provence, vol. 3, 218-221. In http://bach.arts.kuleuven.be/pmert ens/papers/icphsc1991.pdf
- Pierrehumbert, J. (1980). The Phonology and Phonetics of English Intonation. (PhD thesis, MIT. Distributed 1988, Indiana University Linguistics Club.
- Selkirk, E. (1984). Phonology and Syntax: The Relation between Sound and

- Structure. Cambridge, MA: MIT Press.
- Sgall, P. (1975). Conditions of the Use of Sentences and a Semantic Representation of Topic and Focus. In E. Keenan (Ed.), Formal Semantics of Natural Language, 297-312. Cambridge: Cambridge University Press.
- Sgall, P., Hajicová, E., & Benesová, E. (1973). Topic, Focus and Generative Semantics. Kronberg Taunus: Scriptor.
- Tamburini, F. (2005). Automatic Prominence Identification and Prosodic Typology. In Proceedings InterSpeech 2005, Lisbon, 1813-1816. In https://www.academia.edu/573109 3/Automatic_prominence_identificat ion_and_prosodic_typology
- Tamburini, F. (2006). Reliable Prominence Identification in English Spontaneous Speech. In Proceedings of Speech Prosody, 2006, Dresden (PS1-9-19). In https://www.academia.edu/573110 1/Reliable_Prominence_Identificatio n_in_English_Spontaneous_Speech
- Tamburini, F. (2009). Prominenza frasale e tipologia prosodica: un approccio acustico. In G. Ferrari, R. Benatti, & M. Mosca (eds.), Linguistica e modelli tecnologici di ricerca, 437-455. Roma: Bulzoni.
- Taylor, P. A. (2000). Analysis and Synthesis of Intonation using the Tilt Model. Journal of the Acoustical Society of America, 107, 1697-1714.
- Terken, J. (1991). Fundamental Frequency and perceived prominence parameters. Journal of the Acoustical Society of America, 87, 1768-1776.

Sobre los indicadores acústico-melódicos de la ironía verbal en la conversación espontánea

MARTÍNEZ HERNÁNDEZ, DIANA

1 Introducción

Prosodic contrasts are argued to be an important mechanism for communicating implicit emotional and intentional information in speech—and a means to understanding traditional notions of an ironic tone (Bryant, 2010: 545).

Tradicionalmente, numerosos autores han hecho referencia a un tipo de tono especial propio en la realización de enunciados irónicos como expresión de un determinado matiz emocional o mental (Haverkate, 1985; Alvarado Ortega, 2006; Ruiz Gurillo, 2010). Sin embargo, teniendo en cuenta que las apariciones de la ironía verbal son considerablemente más frecuentes entre hablantes con una relación de familiaridad estrecha (Gibbs, 2000), son pocos los estudios de tipo experimental que han tenido como objeto de investigación el reconocimiento de rasgos prosódicos atribuidos a los enunciados irónicos en el marco de la conversación espontánea (Attardo, Eisterhold, Hay & Poggi, 2003; Padilla García, 2004, 2011; Cheang & Pell, 2008; Becerra Valderrama, 2011, 2012). Así, el principal objetivo de nuestro estudio será el de ofrecer una revisión teórica, a modo de estado de la cuestión, sobre una de las principales cargas pragmáticas en la caracterización de las expresiones irónicas; esto es, la prosodia. De este modo, retomaremos los estudios que han abordado la interfaz prosodia e ironía con el fin de observar el significado que ha podido darse al denominado "tono irónico" (Padilla García, 2004).

Facilitaremos, pues, el terreno para dar paso a estudios futuros que tomen en consideración estos hechos ya sea directa o indirectamente.

En este sentido, pese a que no aludamos en nuestras aclaraciones de un modo preciso a información relativa a los indicadores lingüísticos (uso de evidenciales, superlativos, hipérboles, lítotes) y kinésicos (movimientos corporales y faciales), hemos de advertir que estos acompañan, obviamente, a los enunciados expuestos y se unen en su función pragmática al componente melódico como índices decisivos en la identificación y comprensión de la ironía. Lo mismo ocurre con la naturaleza contextual (condicionantes pragmáticos) de este tipo de enunciados que, junto a las señales prosódicas, empujan a la óptima comprensión de juicios e intenciones irónicas (Gibbs, 1994) y constituyen necesariamente una forma de desambiguar el mensaje dentro del discurso espontáneo:

The production and processing of disambiguation cues involves certain costs, but without the cues, interlocutors risk misunderstanding. To minimize the potential costs and maximize relevance, verbal irony should contain features that assist listeners in reaching correct interpretations, particularly in contexts with reduced available information (Bryant & Fox Tree, 2002: 100).

2 La ironía verbal en interacción

Diversos autores han tratado de definir la ironía desde las diferentes teorías pragmáticas (Clark & Gerrig, 1984; Ducrot, 1984; Horn, [1984] 1998; Grice, 1989; Wilson & Sperber, 1992; Levinson, 2000; Utsumi, 2000), sin embargo, se han producido definiciones notablemente dispares en la tarea descriptiva de aclarar el concepto y percibirlo desde un punto de vista pragmático. En este sentido, Ruiz Gurillo (2010: 96) concluye que "tales acercamientos a la ironía no son fácilmente aplicables a la amplia idiosincrasia de este hecho pragmático". En definitiva, parece que estos modelos teóricos pretenden ofrecer un marco único para capturar la esencia de la ironía, y han aportado enfoques (no mutuamente excluyentes) demasiado estrechos para revelar qué es la ironía y demasiado amplios para expresar lo que hace y, por consiguiente, difícilmente han podido contribuir al desarrollo de una teoría global del uso del lenguaje irónico. De este modo, consideramos que una caracterización adecuada de la ironía lingüística puede examinarse mejor a través de sus formas y funciones, más claramente reveladas en la conversación (Clift, 1999).

Grosso modo, nuestra propuesta investigadora se justifica en los llamados análisis interaccionistas integrados en el Análisis Conversacional americano, en la Etnografía de la Comunicación de Gumperz y en la Sociolingüística Interaccional de Goffman. Entre los seguidores de estas teorías, el Grupo GRIALE¹ opta por el modelo propuesto por Levinson (2000) y propone una explicación objetiva de la ironía, la cual se concibe como un fenómeno pragmático que se apoya en indicadores y marcas para su interpretación inferencial más allá de los contextos particulares en los que se produce (indicadores lingüísticos, kinésicos, paralingüísticos y acústico-melódicos). Aunque los llamados análisis interaccionistas se integran en un ámbito todavía novedoso para el estudio de la ironía, han comenzado a desarrollarse algunos trabajos (Jorgensen, 1996; Clift, 1999; Anolli, Ciceri & Infantino, 2000; Gibbs & Colston, 2001; Giora & Gur, 2003;

¹ Nos referimos a los estudios correspondientes al Grupo GRIALE (Grupo de Investigación sobre ironía y humor en español) de la Universidad de Alicante. Sus estudios están centrados en el análisis pragmático de la ironía verbal en español, así como en su tipología y aplicaciones.

Kotthof, 2003; o Eisterhold, Attardo & Boxer, 2006) centrados en factores no menos importantes dentro de la progresión conversacional, derivados de las relaciones interpersonales (edad, género, relaciones de poder y solidaridad...) y de diversos aspectos psicosociales (ironía, humor, cortesía, descortesía, etc.) (Padilla García, 2005; Hidalgo Navarro, 2006, 2009; Padilla García y Alvarado Ortega, 2005, 2008). Así pues, los usos lingüísticos del discurso irónico adquieren un largo alcance, como ya proponían Gibbs (1994), Roberts & Kreuz (1994), Toplak & Katz (2000), o Bryant (2012), hasta el punto de llegar a preguntarnos "What can't irony do?" (Bryant, 2012: 675).

En especial, de estos estudios podemos deducir que, (1) el significado generalmente negativo comunicado por una declaración irónica se silencia (Dews, Kaplan & Winner, 1995) y parece menos grosero (especialmente cuando se manifestan críticas triviales) en el discurso espontáneo (Jorgensen, 1996), y (2) que las diferentes formas de ironía a menudo tienen distintas funciones comunicativas (Roberts & Kreuz, 1994; Lee & Katz, 1998) y evocan diferentes respuestas emocionales en los oyentes (Leggitt y Gibbs, 2000).

En cierta medida, los análisis interaccionistas desmienten ciertas convicciones sobre el empleo de la ironía en interacción: nos referimos al carácter negativo que ha caracterizado el estudio de la ironía o a la ausencia de tales enunciaciones en el discurso femenino. En un análisis puramente interaccionista basado en ejemplos reales, parece que el proceso inferencial se complica o al menos las experiencias sociales y contextuales se multiplican. El empleo de la ironía en interacción, entre otros muchos factores, puede relacionarse directamente con el fenómeno de la cortesía e implicar solidaridad en situaciones de igualdad social, así como presentar una cohesión estrecha con el fenómeno del humor y ser capaz de generar situaciones humorísticas y/o divertidas.

Por lo que respecta a nuestra labor investigadora, pues, partimos del hecho de que la conversación se organiza sobre unos principios pragmáticos y, por consiguiente, la interpretación de los enunciados irónicos está condicionada por factores contextuales (propios del contorno físico inmediato) y por principios sociopragmáticos (el entorno sociocultural y la adecuación a las diferentes situaciones comunicativas). De este modo, las distintas estrategias fónicas que se desarrollan en situaciones discursivas coloquiales expresan sentidos comunicativos que contribuyen a la modalidad funcional del enunciado irónico: es el caso, por ejemplo, de la llamada (des)cortesía intralingüística o estratégica, que puede implicar crítica, burla, broma, negociación, desacuerdo, confirmación... Sin embargo, la investigación de las funciones pragmáticas del componente fónico en expresiones irónicas, como veremos, tiene escasa tradición en el ámbito hispánico (Haverkate, 1994; Padilla García, 2004, 2008; Becerra Valderrama, 2011; Grupo GRIALE).

3 El papel de la prosodia en la marcación de efectos pragmáticos

Diferentes investigadores (Rockwell, 2000; Attardo et al., 2003; Padilla García, 2004; Cheang & Pell, 2008; Ruiz Gurillo, 2008) han mostrado la importancia de la prosodia en el reconocimiento de enunciados irónicos como uno de los recursos lingüísticos más usados en su comprensión, junto con otro tipo de señales, como las relacionadas con la expresión facial. Si bien, la mayoría de los estudios relacionados con la entonación española (Navarro Tomás, 1918, 1944; Pierrehumbert, 1980; Pierrehumbert & Beckman, 1988; Ladd, 1996) se han ocupado de patrones de entonación interrogativos y declarativos

en el discurso de laboratorio, en gran medida, a través de técnicas relacionadas con la lectura, ignorando así, los factores que influyen en la entonación del habla espontánea. En general, la relación entre entonación y significado pragmático ha experimentado lagunas importantes por lo que a su investigación se refiere. En este sentido, Face (2003) ya trató de comparar ciertas propiedades fonéticas en contornos declarativos de habla espontánea y habla de laboratorio, concluyendo que:

- En el habla espontánea, es común encontrar ejemplos de palabras desacentuadas.
- Así, Face (2001a, 2001b, 2002), Hualde (2002) y Nibert (2000) muestran que en las oraciones declarativas de laboratorio, la alineación de F0 se realiza en la sílaba tónica, a diferencia del habla espontánea, en la que se realiza, normalmente, en la sílaba postónica.
- Comúnmente, la F0 se eleva al final de las oraciones declarativas en el habla espontánea.
- Además, aunque los tonemas descendentes son característicos en el habla de laboratorio, los datos de su estudio (Face, 2003) muestran que dichos tonemas también se exhiben en el habla espontánea. Este resultado parece indicar que tales marcas se utilizan en las expresiones que comunican información sobre la edad.

Del mismo modo, Hansson (2003) se atreve a explorar sobre la prosodia en habla espontánea y en el discurso leído concluyendo que hay diferencias notables por lo que se refiere a la velocidad de elocución y a los límites prosódicos caracterizados por señales de tono elevadas. Tampoco podemos olvidar los estudios que Hidalgo (1998, 2001) ha realizado para el caso del español peninsular sobre la entonación del habla espontánea. De sus estudios se infiere que los descensos y ascensos del tono sirven, particularme, para enfatizar ciertas funciones pragmáticas (énfasis de palabra, expresión del humor, etc.).

Ahora bien, dichas propiedades fonéticas y verbales, propias del habla espontánea y ausentes en el discurso de laboratorio, pueden variar de acuerdo con el sentido o la función pragmática, ya sea a través de la carga emocional del enunciado (Prieto, 2004), de la actitud del hablante hacia un tema, del énfasis o de la importancia comunicativa que quiera darse a un elemento del discurso, etc. En este sentido, cada vez nuestro aprendizaje sobre la fonética y la pragmática de la entonación del habla espontánea, aunque se encuentre aún en sus etapas más iniciales, parece encaminarse hacia un análisis fonológico de esta variedad del habla (Rajiv, 2006).

4 Marcas acústico-melódicas de la ironía

La incidencia del sistema verbal y de los sistemas no verbales en determinados fenómenos pragmático-discursivos implica y favorece la constitución del mensaje irónico a través del empleo de indicadores lingüísticos y no lingüísticos; no obstante, se ha considerado que las marcas verbales y no verbales funcionan de forma secundaria en la especificación del sentido irónico. De este modo se refleja en los estudios de Muecke (1978), Schoentjes (2003) o Utsumi (2000), los cuales afirman que estas evidencias textuales no funcionan como guías determinantes en la detección del sentido irónico del enunciado. Alertan de que los indicadores verbales y no verbales acompañan al enunciado solamente para expresar actitudes negativas.

En este sentido, existe un debate sobre si el lenguaje irónico incluye necesariamente un tono de voz especial como señal prosódica esencial para inferir el significado de los enunciados irónicos (Cutler, 1974; Grice, 1989), o,

contrariamente, se trata de una propiedad prescindible para la comprensión de tales inferencias (Gibbs, 1986a, 1986b, 1994; Kreuz & Roberts, 1995; Kreuz, 1996). Pues bien, a fin de que el oyente capte la intención irónica y con ello la información lingüística del enunciado irónico, a la información estilística le deben acompañar otras marcas lingüísticas y paralingüísticas que dirijan el procesamiento léxico, gramatical, semántico y pragmático de tales enunciados. Esto no quiere decir que la interpretación del enunciado irónico venga determinada únicamente por indicadores lingüísticos y no lingüísticos, dado que los indicios contextuales adquieren también una relevancia ineludible en la presencia de este fenómeno pragmático-discursivo.

Así pues, en los estudios sobre ironía (Warning, 1976; Myers Roy, 1978, 1981; Roster, 1978; Eggs, 1979), existen varios centros de interés fundamentales, entre los que destacan: su estructura lingüística, su definición conceptual y el análisis del llamado tono irónico. Encontramos autores que afirman que la ironía se realiza mediante determinados medios lingüísticos (Myers Roy, 1978) y autores que niegan la existencia de estos (Berg, 1978). Distinguen los primeros dos niveles gramaticales: el prosódico y el sintáctico. Algunos autores (Warning, 1976; Fraser, 1979) han atribuido un papel primordial a la curva entonacional (acento de intensidad, alargamiento de la cantidad silábica, articulación nasal), no obstante, la valoración de los medios prosódicos ha originado ciertas controversias. La mayoría de investigadores (Haverkate, 1985; Cutler, 1974) han vinculado el análisis del componente fonético-acústico del acto verbal irónico con factores puramente sintácticos, ciñéndose a la manifestación de la ironía en el lenguaje escrito:

[...] cualquiera que sea la influencia de factores prosódicos, estos no pueden superar el nivel de una condición necesaria; en ningún caso se les puede calificar de condición suficiente para que se verifique la ironía verbal (Haverkate, 1985: 347).

Asimismo, algunos investigadores sugieren una vinculación estrecha entre características vocales y características léxico-semánticas. Kreuz & Roberts (1995) examinan el fenómeno de la hipérbole en las interpretaciones de la ironía verbal y afirman que el tono de voz puede acompañar también al uso de esta figura.

Un estudio decisivo en el tratamiento acústico de la ironía verbal en habla espontánea ha sido el trabajo de Bryant & Fox Tree (2002). Sobre la base de la Teoría de la Relevancia predicen que ciertas señales acústicas, junto a diversas informaciones contextuales, pueden funcionar como marcas desambiguadoras en situaciones que carecen de otras fuentes de información (canal visual, por ejemplo) para inferir la intención irónica en el habla espontánea. Para estos autores (2002: 99), "speech should contain semantic disambiguation cues as a function of the richness of contextual information". Observan la ironía verbal como un hecho que contiene información acústica y contextual, no solo en situaciones que carecen de un canal visual, sino también debido a la falta de familiaridad entre los interlocutores, lo cual supone un esfuerzo extra para no ser malinterpretados. De este modo, tal limitación de procesamiento fundamentada en diferentes rasgos prosódicos se asemeja, por ejemplo, a situaciones en las que los padres hablan a los bebés utilizando contornos exagerados dada la inmadurez de estos para los sistemas preceptivos, así como su falta de capacidad lingüística (Fernald, 1992). En suma, "the benefits of effective communication in these situations are worth the costs of extra speech production efforts" (Bryant & Fox Tree, 2002: 101). De este modo, la incongruencia existente entre las características prosódicas y el contenido mencionado ayudan a transmitir tanto la intención (enunciado eco) como la actitud del hablante. Los resultados vertidos de su análisis muestran que la ironía verbal, en definitiva, contiene características prosódicas que ayudan a los oyentes en la diferenciación del discurso espontáneo.

Entre los estudios que hacen referencia a la identificación y descripción de un tono de voz irónico, teniendo en cuenta variables de tipo contextual e intencional, en el ámbito hispánico, destacan los trabajos de Padilla García (2004, 2011) y Becerra Valderrama (2011, 2012). En un trabajo sobre la identificación prosódica de la ironía, Padilla García (2011) trata de presetnar ciertos rasgos prosódicos mediante un análisis de producción e interpretación considerando fonéticamente dos corpus de enunciados orales: un corpus inducido y un corpus espontáneo (por medio de ejemplos extraídos de programas de radio y de televisión). Por medio de este estudio concluye que:

- El elemento fundamental para contrastar un enunciado irónico y un enunciado neutro es la duración, tanto a nivel local como a nivel global. Además, tal parámetro se ve reforzado por la presencia de pausas y procesos de silabeo. En suma, el aumento de la duración manifestada mediante alargamientos y separación por pausas constituye un elemento "casi obligatorio" en la producción de la ironía.
- El aumento de F0 e intensidad, por otro lado, se corresponden con elementos menos relevantes o determinantes en la identificación de ironía. El aumento de F0 se ve reflejado en el pico inicial y en la inflexión final, así como en sílabas remarcables o focos acentuales. Por lo que corresponde a la intensidad, el autor hace referencia al enunciado de forma global y a las palabras enfatizadas.

En la prueba interpretativa de su estudio, un 50 % de los entrevistados opina que tanto los factores fónicos como contextuales presentan la misma importancia, un 48 % opina que los factores fónicos son los más relevantes, y un 2 % que lo es el contexto. En definitiva, de su estudio se deriva que "el contexto es una ayuda fundamental, pero no la única, ni, por supuesto, la más importante" (Padilla García, 2011: 223).

En un estudio anterior (Padilla García, 2009), este mismo autor define subtipos de enunciaciones irónicas a propósito de una serie de rasgos acústicos, pragmáticos y kinésicos concretos. Para ello, se aleja de los conocidos rasgos binarios y simbólicos que propone la fonología y, desde postulados psicofonéticos, se centra en la denominada codificación expresiva o emotiva en el análisis de los elementos acústico-melódicos. Afirma que el tono irónico es una marca pragmática con valor procedimental que pertenece a la función expresiva y que comporta rasgos acústicos de manera gradual, "moviéndose entre los márgenes de dispersión que proporcionan los elementos fónicos codificados en la función representativa o simbólica" (Padilla García, 2009: 149). Por medio de su estudio, este autor plantea cuatro tipos de enunciaciones irónicas vinculadas con situaciones comunicativas concretas:

- Aumento del volumen de la voz: enunciación sarcástica con actitud descortés hacia el oyente e insertada en contextos conflictivos como riñas o discusiones generalizadas.
- Exageración de rasgos kinésicos (gestos faciales, sonrisa): enunciación falsa o exageradamente sincera disfrazada de una falsa cortesía y oculta bajo actitudes combativas o conflictivas.
- Aumento de F0 en la inflexión final de la curva del grupo fónico y velocidad de emisión lenta y relajada: enunciación irónica propiamente dicha con carácter jocoso o divertido.

 Ralentización en la velocidad de habla y falta de manifestación kinésica: enunciación irónica neutra o plana referida a aquellos enunciados ambiguos, "más sutiles e inteligentes" (Padilla García, 2009: 154) sobre los que el contenido irónico está al servicio del contexto sociocultural (Padilla García, 2005).

_

En suma, este autor trata de caracterizar el tono irónico en relación con la tipología enunciativa utilizada: desde la enunciación más marcada pragmática y acústicamente (enunciación sarcástica), hasta la ironía menos exagerada (enunciación plana) con un número menor de rasgos acústico-melódicos. Así, del análisis pragmático de Padilla (2009) se deduce que, la prosodia no se revela en la totalidad del corpus y que sus componentes acústicos (velocidad, intensidad y F0) se manifiestan de forma gradual.

Por su parte, por medio de un análisis acústico, Becerra Valderrama (2012) trata de examinar los rasgos prosódicos en dos formas de ironía en español (hipérboles y preguntas retóricas) asociadas a la burla, la crítica, la ira y la tristeza. Para su estudio toma como base los valores promedio de algunos parámetros acústicos (F0, intensidad y duración) y concluye que los enunciados irónicos con cierta intención irónica (burla, crítica, ira y tristeza) presentan particularidades acústicas diferentes (aumento de F0, intensidad y duración) a los enunciados con entonación neutra. Para el caso de la hipérbole, el estudio muestra una diferencia significativa en la duración, la frecuencia fundamental y la intensidad, esto es, presenta valores más bajos en los enunciados de burla y crítica en comparación con la enunciación neutra. La F0 e intensidad son mayores en la entonación de crítica en relación con la burla, y la duración asciende para el caso de la burla. En las preguntas retóricas, la diferencia se encuentra en los valores de F0, duración e intensidad de la crítica, con valores superiores respecto a las demás actitudes (ira, tristeza, burla y neutralidad). La autora constata la existencia de un tono de voz irónico, ahora bien, afirma que "no parece haber un patrón prosódico consistente que permita expresar actitudes irónicas distintas, como la burla y la crítica, por medio de la entonación" (Becerra Valderrama, 2012: 204). Así pues, un estudio más reciente de la autora (Becerra Valderrama & Igoa González, 2013) concluye con resultados parecidos: el protagonista indiscutible en la identificación de intenciones irónicas es el contexto y afirman que "...la entonación no es una propiedad suficientemente fiable a la hora de enjuiciar la intención irónica del hablante" (2013: 111). En este sentido, la prosodia parece ser un elemento más bien dependiente de la información contextual, aunque esto no significa que no se trate de un apoyo sustancial para el reconocimiento de burlas o críticas irónicas.

Pues bien, estudios sobre el inglés, asimismo, han analizado la existencia de rasgos prosódicos otorgados a los enunciados irónicos (Cutler, 1974; Haiman, 1998; Milosky & Ford, 1997; Anolli *et al.*, 2000; Rockwell, 2000, 2007; Attardo *et al.*, 2003; Bryant & Fox Tree, 2005; Cheang & Pell; 2008). La mayoría de estos estudios parte de la base de la Teoría de la Relevancia para explicar el reconocimiento de la ironía verbal en habla espontánea; una visión que, en definitiva, hace coincidir la forma proposicional de un enunciado con su expresión inferencial. En cierta medida, estas marcas o señales de desambiguación (información acústica y contextual) reducen al mínimo el coste de procesamiento y maximizan la relevancia con el fin de alcanzar interpretaciones correctas (Bryant & Fox Tree, 2002).

A título ilustrativo, son diversas las marcas que se han establecido para el caso de la ironía, sobre todo, por lo que se refiere a las actitudes o intenciones

comunicativas expresadas en el discurso televisivo entendido como una forma de habla espontánea (Attardo, 2000):

- Reducción de la altura tonal en sílabas acentuadas (Haiman, 1998).
- Nasalización (Cutler, 1974; Myers Roy, 1977; Muecke, 1978; Schaffer, 1982a; Chen, 1990; Haiman, 1998).
- Acento fuertemente exagerado (Cutler, 1974; Myers Roy, 1977; Schaffer, 1982a; Barbe, 1995).
- Sucesión de sílabas prominentes ("beat clash") (Uhmann, 1996; Haiman, 1998; Attardo, 2000).
- Variaciones en el ritmo (Uhmann, 1996).
- Discurso lento (Cutler, 1974; Fónagy, 1976).
- Alargamiento de sílabas (Myers Roy, 1977; Schaffer, 1982a; Barbe, 1995; Adachi 1996; Haiman, 1998; Anolli et al., 2000; Rockwell 2000).
- Pausas largas (Schaffer, 1982a; Haiman, 1998).
- Risas (Schaffer, 1982a; Haiman, 1998).
- F0 elevada (Schaffer, 1982b; Adachi, 1996; Rockwell, 2000).
- Reajuste tonal significativo ("Strong within-statement contrast") (Attardo, 2000).
- "Softened voice" (Muecke; 1978).

Además, los resultados obtenidos sobre los distintos esquemas melódicos asociados a este tipo de enunciados han sido harto heterogéneos:

- Contorno plano (ni ascenso ni descenso) ("Compressed pitch pattern") (Fónagy 1976; Myers Roy 1978; Shapely 1987; Milosky & Wrosbleski, 1994; Barbe 1995; Haiman 1998; Attardo, 2000).
- Contornos ascenso-descenso en expresiones irónicas del tipo "is that so", y tonos bajos con afirmaciones del tipo "a likely story" (Muecke, 1978; Bolinger, 1985, 1989).
- Entonación relativamente monótona (Haiman, 1998).
- Entonación de pregunta (ascenso tonal) (Schaffer, 1982a; Attardo et al., 2003).
- Patrones de entonación exagerados (Muecke, 1978; Haiman, 1998; Attardo, 2000).

Así pues, parece ser que, una vez revisados los resultados de los trabajos mencionados, existe cierto acuerdo por lo que se refiere al estudio de los rasgos acústicos: F0 variable, mayor duración y mayor intensidad, así como en sus respectivos correlatos perceptivos, tono más bajo, ritmo lento y aumento del volumen de la voz.

5 Conclusiones

Este apartado comparativo sobre la capacidad multimodal de la prosodia nos obliga, en cierta medida, a mencionar algunos factores que deberían estar presentes en trabajos experimentales de esta índole.

Por una parte, habría que considerar las características contextuales en estudios de este ámbito, puesto que resultaría un factor clave para una correcta interpretación de estas expresiones (Kreuz & Glucksberg, 1989; Gibbs, 1994; Utsumi, 2000). De este modo, partiendo de una correcta caracterización sobre los tipos de enunciados irónicos (ciertamente, hoy día resulta un tema escasamente tratado) se podría demostrar experimentalmente cuán arraigada está la información contextual a estos subtipos irónicos. En este sentido, existirán enunciados irónicos en los que las señales acústicas podrán ser suficientes para la detección de la intención irónica (Bryant & Fox Tree, 2005),

otros en los que las señales contextuales y el tono de voz influirán perceptivamente (Woodland & Voyer, 2011) y otras situaciones en las que la información contextual cobrará mayor peso frente a la información prosódica (Bryant & Fox Tree, 2002; Becerra Valderrama & Igoa González, 2013). Cabe preguntarse, además, si estos marcadores trabajan de manera jerarquizada. En opinión de algunos autores (Attardo, 2000), parece que estamos frente a una jerarquía en la que las señales de comportamiento preceden a las pistas de entonación, y estas lo hacen sobre las pistas semánticas. Si bien, ante estas afirmaciones sobre las que se afirma que las pistas de entonación anulan a las semánticas (Mehrabian & Wiener, 1967; Argyle & Trower 1979), algunas investigaciones (Beattie, 1980; Krauss, Apple, Morency, Wenzel & Winton, 1981; Noller 1986) han demostrado que la semántica de las expresiones anula las pistas visuales y entonativas.

Por otra parte, estudiar la jocosidad, el sarcasmo, la hipérbole, las preguntas retóricas, la burla o la crítica como interpretaciones de la ironía verbal, ciertamente podrá suponer un tono de voz característico como una marca definitoria en su comprensión, si bien, esta será diferente según la forma y la información emocional e intencional expresada (Leggitt & Gibbs, 2000). Tal y como señala Gibbs (2000: 23), "irony is not a single category of figurative language, but includes a variety of types, each of which is motivated by slightly different cognitive, linguistic, and social factors and conveys somewhat different pragmatic meanings". No olvidemos que en la conversación espontánea se utiliza una amplia gama de estrategias de comunicación con la finalidad de transmitir significados diversos, de ahí que la evidencia lingüística y paralingüística cobre una importancia máxima, sin embargo, actualmente existen pocas investigaciones sistemáticas y cuantitativas que den cuenta de la diversidad de formas en que la ironía se utiliza y la frecuencia con la que cada tipo de ironía aparece en la conversación espontánea (Kreuz & Roberts, 1995; Kreuz, 1996; Kreuz, Roberts, Johnson & Bertus, 1996; Colston, 1997; Colston & Keller, 1998; Padilla García, 2009). Así,

For instance, some forms of irony are affiliative, whereas others are sources of estrangement between individuals (Coser, 1959; Seckman & Couch, 1989). Irony is routinely used in the ongoing flow of conversation between group members to affirm their solidarity by directing comments at individuals who are not group members and not deemed worthy of group membership (Gibbs, 2000: 12).

Hemos de tener presente que estas diferentes formas de discurso irónico pueden comunicar diferentes mensajes (Leggitt & Gibbs, 2000; Bryant, 2012) a través de diferentes señales prosódicas que están estrechamente relacionadas con señales emocionales (Attardo *et al.*, 2003; Bryant & Fox Tree, 2005; Bryant, 2010, 2011). Los subtipos de ironía, pues, se entremezclan de acuerdo con las diferentes combinaciones emocionales e intencionales. De este modo, se observarán marcas prosódicas diversas según los tipos de comunicación irónica, dado que "(...) the significant variation in affect and intentions in ironic language use should drive variable prosodic forms" (Bryant, 2010: 547). Por lo tanto, es posible que no haya una entonación irónica en particular, sino que podrán existir diferentes patrones de entonación que señalen que algo inusual está ocurriendo (Attardo, 2000). Afirma Cruttenden (1986: 114) que la naturaleza contrastiva de los indicadores irónicos, esto es, "a mismatch between tone and context may reinforce an ironical effect already present in the co-occurring grammar and lexis", elimina la idea de que existe una entonación

irónica específica. Asimismo, examinar la prosodia de la ironía a nivel local (información lingüística) y a nivel global (información emocional) nos proporcionaría, además, una investigación rígida sobre el uso de ambas señales prosódicas (locales y globales) afectadas por diversos contextos comunicativos (Eady & Cooper, 1986; McRoberts, Studdert-Kennedy & Shankweiler, 1995; Pell, 2001). Los resultados vertidos en estudios como los de Bryant & Fox Tree (2005) demuestran que la ironía verbal implica características vocales a nivel global y a nivel local que se superponen a medida que el comportamiento comunicativo varía.

Lo cierto es que, hablar del tono como un marcador contrastivo no es un error, pero no debe hacerse de forma aislada o separada de la pragmática (Attardo, 2000), ni puede hablarse de una sola señal de entonación irónica. Hasta aquí la ironía, observada como un hecho que contiene información acústica y contextual, nos permite expresar ciertas intenciones y actitudes presentes en los procesos de producción y de percepción del habla. De este modo, podemos concluir que un enfoque basado en la forma-función podría descubrir un vínculo entre las múltiples señales de procesamiento prosódico y sus respectivas funciones comunicativas (en relación con los subtipos de ironía).

6 Bibliografía

- Adachi, T. (1996). Sarcasm in Japanese. *Studies in Language*, 20, 1, 1-36.
- Alvarado Ortega, B. (2006). Las marcas de la ironía. *Interlingüística*, 16, 1134-8941
- Anolli, L., Ciceri, R., & Infantino, M. G. (2000). Irony as a game of implicitness: Acoustic profiles of ironic communication. *Journal of Psycholinguistic Research*, 29, 3, 275-311.
- Argyle, M., & Trower, P. (1979). *Person to Person: Ways of Communicating*. Nueva York: Harper & Row.
- Attardo, S. (2000). Irony markers and functions: Towards a goal-oriented theory of irony and its processing. Rask. Internationalt Tidsskrift for Sprog og Kommunikation, 12, 3-20.
- Attardo, S., Eisterhold, J., Hay, J., & Poggi, I. (2003). Multimodal markers of irony and sarcasm. Humor: International Journal of Humor Research, 16, 2, 243-260.
- Barbe, K. (1995). *Irony in Context*. Ámsterdam: Benjamins.
- Beattie, G. W. (1980). Language and nonverbal communication the essential synthesis? *Linguistics*, 19, 11/12, 1165-1183.
- Becerra Valderrama, M. I. (2011). El papel de la prosodia y el contexto en la comprensión y producción de dos formas de ironía en español (Tesis Doctoral Inédita). Madrid: Universidad Autónoma de Madrid.

- Becerra Valderrama, M. I. (2012). Rasgos prosódicos en la producción de dos formas de ironía en español. *Lingüística*, 28, 191-205.
- Becerra Valderrama, M. I., & Igoa González, J. M. (2013). La intención irónica en las hipérboles y en las preguntas retóricas en español. *Forma y Función*, 26, 2, 99-124.
- Berg, W. (1978). Uneigentliches Sprechen. Zur Pragmatik und Semantik von Metapher, Metonymie, Ironie, Litotes und rhetorische Frage. Tubinga: Narr.
- Bolinger, D. (1985). Intonation and its Parts: Melody in Spoken English. Londres: Edward Arnold.
- Bolinger, D. (1989). Intonation and its Uses: Melody in grammar and discourse. Londres: Edward Arnold.
- Bryant, G. A. (2010). Prosodic contrasts in ironic speech. *Discourse Processes*, 47, 7, 545-566.
- Bryant, G. A. (2011). Verbal irony in the wild. *Pragmatics and Cognition*, 19, 2, 291-309.
- Bryant, G. A. (2012). Is verbal irony special. Language and Linguistics Compass, 6, 11, 673-685.
- Bryant, G. A., & Fox Tree, J. E. (2002). Recognizing verbal irony in spontaneous speech. *Metaphor and Symbolic Activity*, 17, 2, 99-117.
- Bryant, G. A., & Fox Tree, J. E. (2005). Is there an ironic tone of voice?. *Language and Speech*, 48, 3, 257-277.

- Cheang, H. S., & Pell, M. D. (2008). The sound of sarcasm. *Speech Communication*, 50, 366-381.
- Chen, R. (1990). *Verbal Irony as Implicature* (Tesis Doctoral Inédita). Ball State University, Muncie, Indiana.
- Clark, H. H., & Gerrig. R. J. (1984). On the Pretense Theory of Irony. *Journal of Experimental Psychology: General*, 113, 1, 121-126.
- Clift, R. (1999). Irony in conversation. *Language in Society*, 28, 4, 523-553.
- Colston, H. (1997). "I've never seen anything like it": Overstatement, understatement, and irony. *Metaphor and Symbol*, 12, 43-58.
- Colston, H. L., y Keller, S. B. (1998): You'll never believe this: Irony and hyperbole in expressing surprise. *Journal of Psycholinguistic Research*, 27, 4, 499-513.
- Cruttenden, A. (1986). *Entonación*. Barcelona: Teide.
- Cutler, A. (1974). On saying what you mean without meaning what you say. En W. La Galy, A. R. Fox, & A. Bruck (eds.), Papers from the Tenth Regional Meeting of the Chicago Linguistic Society, 117-127. Chicago: Chicago Linguistic Society
- Dews, S., Kaplan, J., & Winner, E. (1995). Why not say it directly? The social functions of irony. *Discourse Processes*, 19, 347-367.
- Ducrot, O. (1984). *El decir y lo dicho*. Buenos Aires: Hachette.
- Eady, S. J., & Cooper, W. E. (1986). Speech intonation and focus location in matched statements and questions. *Journal of the Acoustical Society of America*, 80, 402-415.
- Eggs, E. (1979). Eine Form des uneigentlichen Sprechens: die Ironie. *Folia Linguistica*, 13.3, 4, 413-435.
- Eisterhold, J., Attardo, S., & Boxer, D. (2006). Reactions to irony in discourse: evidence for the least disruption principle. *Journal of Pragmatics*, 38, 1239-1256.
- Face, T. (2001a). Focus and early peak alignment in Spanish intonation. *Probus*, 13, 223-246.
- Face, T. (2001b). *Intonational marking of contrastive focus in Madrid Spanish*. Columbus: The Ohio State University dissertation.
- Face, T. (2002). Local intonational marking of Spanish contrastive focus. *Probus*, 14, 71-92.

- Face, T. (2003). Intonation in Spanish declaratives: differences between lab speech and spontaneous speech. *Catalan Journal of Linguistics*, 2, 115-131
- Fernald, A. (1992). Human maternal vocalizations as biologically relevant signals: An evolutionary perspective. En J.H. Barkow, L. Cosmides, & J. Tooby (eds.), *The adapted mind: Evolutionary psychology and the generation of culture,* 391-428. Nueva York: Oxford University Press.
- Fónagy, I. (1976). Radiological aspects of emotive speech. *Phonetic*, 33, 1, 31-44.
- Fraser, B. (1979). The interpretation of novel metaphors. En A. Ortony (ed.), *Metaphor and thought*, 172-186. Cambridge: Cambridge University Press.
- Gibbs, R. (1986a). Comprehension and memory for nonliteral utterances: The problem of sarcastic indirect requests. *Acta Psychologica*, 62, 41-57.
- Gibbs, R. (1986b). On the psycholinguistics of sarcasm. *Journal of Experimental Psychology: General*, 115, 3-15.
- Gibbs, R. (1994). The poetics of mind: Figurative thought, language, and understanding. Nueva York: Cambridge University Press.
- Gibbs, R. W. (2000). Irony in Talk among Friends. *Metaphor and Symbol*, 15, 5-27.
- Gibbs, R. W., & Colston, H. L. (2001). The Risks and Rewards of Ironic Communication. En L. Anolli, R. Ciceri & G. RIVA (ed.), Say Not Say:

 New Perspectives on Miscommunication, 187-200.

 Ámsterdam: IOS Press.
- Giora, R., & Gur, I. (2003). Irony in conversation: salience, role, and context effects. En B. Nerlich, Z. Todd, V. Herman, & D. Clarke (eds.), Polysemy: Flexible Patterns of Meaning in Mind and Language, 297-315. Berlin-New York: Mouton de Gruyter.
- Grice, H. (1989). Studies in the way of words. Cambridge, MA: Harvard University Press.
- Haiman, J. (1998). Talks Is Cheap: Sarcasm, Alienation, and the Evolution of Language. Oxford: Oxford University Press.
- Hansson, P. A. (2003). *Prosodic phrasing in spontaneous Swedish*. Lund, Sweden: Lunds Universitet dissertation.

- Haverkate, H. (1985). La ironía verbal: análisis pragmalingüístico. *Revista Española de Lingüística*, 15, 2, 343-391.
- Haverkate, H. (1994). Cortesía verbal: estudio pragmalingüístico. Madrid: Gredos.
- Hidalgo Navarro, A. (1998). Expresividad y función pragmática de la entonación en la conversación coloquial. Algunos usos frecuentes. *Oralia*, 1, 62-92.
- Hidalgo Navarro, A. (2001). Modalidad oracional y entonación: notas sobre el funcionamiento pragmático de los rasgos suprasegmentales en la conversación. *Moenia*, 7, 271-292.
- Hidalgo Navarro, A. (2006). La expresión de cortesía (atenuación) en español hablado: marcas y recursos prosódicos para su reconocimiento en la conversación coloquial. En M. Villayandre, Actas del XXXV Simposio Internacional de la Sociedad Española de Lingüística, 957-979. León: Universidad de León.
- Hidalgo Navarro, A. (2009): Modalización (des)cortés y prosodia: estado de la cuestión en el ámbito hispánico. Boletín de Filología de la Universidad de Chile, 44, 1, 161-195.
- Horn, L.H. (1998 [1984]). Towards a new taxonomy for Pragmatic Inference: Q-based and R-based Implicature. En A. Kasher (ed.), *Pragmatics*, 4, 383-418. Londres/ Nueva York: Routledge.
- Hualde, J. I. (2002). Intonation in Spanish and the other Ibero-Romance languages. En C. Wiltshire & J. Camps (ed.), Romance phonology and variation, 101-115. Philadelphia: John Benjamins.
- Jorgensen, J. (1996). The functions of sarcastic irony in speech. *Journal of Pragmatics*, 26, 613-634.
- Kotthoff, H. (2003). Responding to irony in different contexts: on cognition in conversation. *Journal of Pragmatics*, 35, 9, 1387-1411.
- Krauss, R. M., Apple, W., Morency, N., Wenzel, C., & Winton, W. (1981). Verbal, vocal, and visible factors in judgments of another's affect. *Journal of Personality and Social Psychology*, 40, 2, 312-320.
- Kreuz, R. (1996). The use of verbal irony: Cues and constraints. En J. Mio & A. Katz (eds.), *Metaphor: Implications* and applications, 23-38. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

- Kreuz, R., & Glucksberg, S. (1989). How to be sarcastic: The echoic reminder theory of verbal irony. *Journal of Experimental Psychology: General*, 118, 374-386.
- Kreuz, R. J. & Roberts, R. M. (1994). Why do people use figurative language? American Psychological Society, 5, 3, 158-163.
- Kreuz, R. J., & Roberts, R. M. (1995). Two cues for verbal irony: Hyperbole and the ironic tone of voice. *Metaphor and symbolic activity*, 10, 1, 21-31.
- Kreuz, R., Roberts, R., Johnson, B., & Bertus, E. (1996). Figurative language occurrence and co-occurrence in contemporary literature. En R. Kreuz & M. MacNealy (eds.), Empirical approaches to literature and aesthetics, 83-97. Norwood, NJ: Ablex.
- Ladd, R. (1996). *Intonational phonology*. Cambridge, UK: Cambridge University Press.
- Lee, C., & Katz, A. (1998). The differential role of ridicule in sarcasm and irony. *Metaphor and Symbol*, 13, 1-15.
- Leggitt, J. y Gibbs, R. (2000). Emotional reactions to verbal irony. *Discourse Processes*, 29, 1-24.
- Levinson, S. C. (2000). Presumptive meaning. The theory of Generalized Conversational Implicature. Cambridge: Massachusetts.
- McRoberts, G. W., Studdert-Kennedy, M., & Shankweiler, D. P. (1995). The role of fundamental frequency in signaling linguistic stress and affect: Evidence for a dissociation. *Perception & Psychophysics*, 57, 159-174.
- Mehrabian, A., & Morton, W. (1967).

 Decoding of inconsistent communications. *Journal of Personality and Social Psychology*, 6, 1, 109-114.
- Milosky, L., & Wrobleski, C. A. (1994). The Prosody of irony. Comunicación presentada en la *International Society* for Humor Studies Conference. Ithaca, NY.
- Milosky, L. M. & Ford, J. A. (1997). The role of prosody in children's inferences of ironic intent. Discourse Processes, 23, 47-61.
- Muecke, D. C. (1978). Irony markers. *Poetics*, 7, 363-375.
- Myers Roy, A. (1977). Towards a definition of irony. En R. W. Fasold, & R. Shuy (eds.), Studies in Language Variation,

- Washington, C.: 171-183. Georgetown University Press.
- Myers Roy, A. (1978). Irony in conversation. Ann Arbor: University Microfims International.
- Myers roy, A. (1981). The function of irony in discourse. *Text*, 1, 4, 407-423.
- Navarro Tomás, T. (1918). Manual de española. pronunciación Madrid: Centro de Estudios Históricos.
- Navarro Tomás, T. (1944). Manual de entonación española. Madrid: CSIC.
- Η. J. (2000).Phonetic and Nibert, phonological evidence intermediate phrasing in Spanish Urbana-Champaign: intonation. University of Illinois dissertation.
- Noller, P. (1986). Video primacy A further look. Journal of Nonverbal Behavior, 9, 1, 28-47.
- Padilla García, X. A. (2004). El tono irónico: estudio fonopragmático. Actual, 81, 85-98.
- Padilla García, X. A. (2005). Pragmática del orden de palabras. Alicante: Servicio de Publicaciones de la Universidad de Alicante.
- Padilla García, X. A. (2009). Marcas acústico-melódicas: el tono irónico. En L. Ruiz Gurillo, & X. A. Padilla García (eds.), Dime cómo ironizas y te diré quién eres, 155-166. Frankfurt am Main: Peter Lang.
- Padilla García, X. A. (2011). ¿Existen rasgos prosódicos objetivos en los enunciados irónicos? Oralia, 14, 203-226.
- Padilla García, X. A. & Alvarado Ortega, M. B. (2005). La ironía y la cortesía: una aproximación desde sus efectos. ELUA, 19, 33-45.
- Padilla García, X. A. & Alvarado Ortega, M. B. (2008). La ironía o cómo enmascarar un acto supuestamente amenazante. Actas del III Coloquio Internacional del Programa EDÍCE, 419-435.
- Pell, M. D. (2001). Influence of emotion and focus location on prosody in matched statements and questions. Journal of the Acoustical Society of America, 109, 1668-1680.
- Pierrehumbert, J. (1980). The phonology and phonetics of English intonation. Cambridge, MA: MIT dissertation.
- Pierrehumbert, J. & Beckman. M. (1988). Japanese tone structure. Cambridge, MA: MIT Press.

- Prieto, P. (2004). Phonological phrasing in Spanish. Barcelona: Universitat Autònoma de Barcelona.
- (2006).On Intonation's Rajiv, Relationship with Pragmatic Meaning in Spanish. En T. Face & C. A. Klee (ed.), Selected Proceedings of the 8th Hispanic Symposium, 103-115. Linguistics Somerville, MA: Cascadilla Proceedings Project.
- Roberts, R., & Kreuz, R. (1994). Why do people use figurative language? Psychological Science, 5, 159-163.
- Rockwell, P. (2000). Lower, slower, louder: Vocal cues of sarcasm. Journal of Psycholinguistic Research, 29, 5, 483-
- Rockwell, P. (2007). Vocal Features of Conversational Sarcasm: Comparison of Methods. Journal of Psycholinguistic Research, 36, 361-369.
- Roster, P. J. (1978): La ironía como método de análisis literario: la poesía de Salvador Novo. Madrid: Gredos.
- L. (2008).Ruiz Gurillo, metarrepresentaciones en el español hablado. Spanish in Context, 5, 1, 40-
- Ruiz Gurillo, L. (2009). ¿Cómo se gestiona la ironía en la conversación? RILCE, 25, 2, 363-377.
- Ruiz Gurillo, L. (2010). Las 'marcas discursivas' de la ironía. En J. L. Cifuentes Honrubia, A. Gómez González-Jover, A. Lillo & F. Yus (coord.), Los caminos de la lengua. Estudios en homenaje a Enrique Alcáraz Varó, 871-886. Alicante: Servicio de Publicaciones de la Universidad de Alicante.
- Schaffer, R. (1982a). Vocal clues for irony in English (Tesis Doctoral Inédita). Ohio State University.
- Schaffer, R. (1982b). Are there consistent vocal clues for irony? En C. S. Masek, R. A. Hendrick & M. F. Miller (eds.), Papers from the Parasession on Language and Behavior, 204-210. Chicago, IL: Chicago Linguistic Society.
- Schontjes, P. (2003). La poética de la ironía. Madrid: Cátedra.
- Shapely, M. (1987). Prosodic variation and audience response. IPrA: Papers in *Pragmatics*, 1, 2, 66-79.
- Toplak, M., & Katz., A. N. (2000). On the uses of sarcastic irony. Journal of Pragmatics, 32, 1467-88.

- Uhmann, S. (1996). On rhythm in everyday German conversation: Beat clashes in assessment utterances. En E. Couper-Kuhlen, & M. Selting (eds.), *Prosody in Conversation: Interactional Studies*, 303-365. Cambridge: Cambridge University Press.
- Utsumi, A. (2000). Verbal Irony as implicit display of ironic environment: Distinguishing ironic utterances from nonirony. *Journal of Pragmatics*, 32, 1777-1806.
- Warning, R. (1976). Ironiesignale und ironische Solidarisierung. En W. Preisendanz & R. Warning (ed.), *Das Komische*, 416-423. München: Fink.
- Wilson, D., & Sperber, D. (1992). On verbal irony. *Lingua*, 87, 53-76.
- Woodland, J., & Voyer, D. (2011). Context and Intonation in the Perception of Sarcasm. *Metaphor and Symbol*, 26, 227-239.

"Es un hombre famoso" o "¿Es un hombre famoso?" Rasgos melódicos de las interrogativas absolutas del español hablado por suecos

MARTORELL MORALES, LAURA; FONT-ROTCHÉS, DOLORS

1 Introducción

El objetivo de este trabajo es describir la entonación de las interrogativas absolutas del español hablado por suecos para determinar los rasgos melódicos de los patrones que utilizan y determinar hasta qué punto se corresponden con los cuatro patrones /+interrogativos/ del español, tres descritos en Cantero y Font-Rotchés (2007) y el cuarto en Font-Rotchés y Mateo (2011, 2013).

El estudio forma parte de una investigación más amplia dedicada a la caracterización de la entonación del español hablado por suecos¹, motivada por la observación de dificultades comunicativas en el aula debidas a una entonación inadecuada, por la falta de un estudio objetivo que permita elaborar propuestas didácticas y por el escaso o inapropiado tratamiento de la

¹ Esta investigación se enmarca dentro de los trabajos del Grup de Recerca en Entonació i Parla (2014 SGR169) y del proyecto I+D Análisis del habla y modelos didácticos (FF12013-41915-P) financiado por el Ministerio de Economía y Competitividad.

entonación en los manuales de enseñanza de español para extranjeros (Baditzné, 2012; Baditzné & Font-Rotchés, 2013).

Además, existe un creciente interés por el aprendizaje del español en Suecia, país que figura en la lista de los 10 países del mundo con más estudiantes de español (Instituto Cervantes, 2014).

Para llevarlo a cabo, se ha seguido el método Análisis Melódico del Habla, propuesto en Cantero (2002), revisado y ampliado en Font-Rotchés (2007) y expuesto en forma de protocolo en Cantero y Font-Rotchés (2009). Esta elección viene motivada porque anteriormente ya se han realizado investigaciones sobre distintas interlenguas –español hablado por chinos (Liu, 2003), italianos (Devís, 2011), húngaros (Baditzné, 2012) o portugueses (Fonseca, 2013)– y se ha demostrado muy apto para describir los rasgos melódicos que caracterizan el acento extranjero.

Las interrogativas absolutas o totales, caracterizadas por comprender el contenido de todo el enunciado y responderse con "sí" o "no", como sería el caso de "¿es un hombre famoso?", son el tipo de enunciados en los que se han detectado mayores problemas comunicativos en el aula, causados por la inadecuación entonativa. El malentendido o confusión más habitual se genera cuando un estudiante produce un enunciado con la intención de preguntar, pero sin entonación de pregunta, y el profesor lo interpreta como otro tipo de enunciado –generalmente declarativo o, a veces también, suspendido– que no requiere respuesta. De ahí el interés inicial por analizar este tipo de enunciados y obtener resultados para su posible aplicación didáctica.

2 Metodología

2.1 Corpus e informantes

Los datos del corpus se han obtenido de diversas sesiones de grabación con estudiantes de una escuela sueca de Barcelona especializada en la enseñanza del español a suecos. Los informantes aceptan o se ofrecen voluntariamente a participar en ellas sin conocer el objeto de estudio de la investigación. Las grabaciones, registradas con una grabadora digital, incluyen clases del programa habitual de los cursos que se imparten en la escuela y reuniones fuera del horario lectivo en las que se proponen juegos y conversaciones sobre diversos temas.

Para este estudio, se han extraído y transcrito 50 enunciados interrogativos absolutos, representativos de todos los informantes y de todas las sesiones, que suman un total de 15 horas de audio. A continuación aportamos ejemplos de las preguntas que hemos obtenido:

ES-05-05: ¿Los trabajadores lo pagan?

ES-07-01: ¿Puedes poner todas las conjugaciones de "tener"?

ES-10-02: ¿Pero hay, ehm, futuro de subjuntivo?

ES-16-02: ; Tiene otro nombre chino entonces también?

ES-20-04: ¿Eres de Cataluña?

Los informantes son 20 hombres y mujeres de entre 19 y 32 años, hablantes nativos de sueco, y tienen el español como lengua extranjera con un nivel mínimo B1 –el inglés es su segunda lengua–. Se excluyó a los informantes que habían aprendido el español en otros países con una variedad dialectal distinta a la peninsular, ya que el estudio se basa en esta variedad (Martorell, 2010).

2.2 Metodología

Como ya hemos comentado en la introducción, para el análisis de los datos, se utiliza la metodología empírica y experimental *Análisis Melódico del Habla*, basada en el modelo teórico propuesto en Cantero (2002), revisado y ampliado en Font-Rotchés (2007), y expuesto en forma de protocolo en Cantero y Font-Rotchés (2009). También en Cantero y Mateo (2011) se presentan los pasos de aplicación del método utilizando un *script* de semiautomatización para la obtención y procesamiento de los datos, cuyas características se exponen en Mateo (2010).

Se trata de identificar las unidades melódicas y extraer los datos acústicos en hercios, utilizando la aplicación de análisis y síntesis de voz Praat (Boersma & Weenink, 1992-2011), y relativizarlos, calculando la distancia tonal entre un valor y el siguiente en porcentaje. Con los valores relativos de cada contorno se elaboran los gráficos, los cuales ya están listos para poderse clasificar y comparar con otros contornos, ya que al relativizar los valores se han eliminado las diferencias melódicas por razón de sexo o de edad.

En nuestro caso, una vez obtenidos los rasgos melódicos y los gráficos de los enunciados, se compararon con los cuatro patrones entonativos /+interrogativos/ del español peninsular.

2.3 Los patrones interrogativos del español peninsular

Para la entonación interrogativa del español peninsular se han establecido cuatro patrones (Cantero & Font-Rotchés, 2007, los patrones II, III y IV; Font-Rotchés & Mateo, 2011, 2013, el patrón XIII), los cuales presentamos en este apartado caracterizados por sus rasgos melódicos más relevantes. Estos patrones constituirán la base con la que se van a comparar los rasgos melódicos de los contornos que configuran las preguntas absolutas de la interlengua del español hablado por suecos, con la finalidad de poder analizar hasta qué punto se asemejan y hasta qué punto difieren de la lengua meta, el español.

El patrón II (véase el gráfico 1) es neutro y el más habitual en la entonación interrogativa del español peninsular. Se caracteriza por una inflexión final ascendente superior al 70 % –en algunos casos llega a superar el 100 %–, un primer pico en la primera vocal tónica del contorno y un cuerpo en dirección descendente.

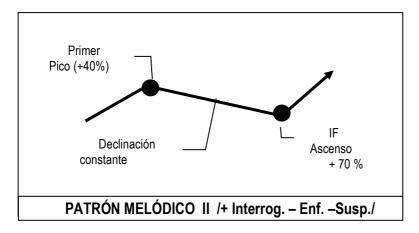


Gráfico 1. Patrón II del español peninsular (Font-Rotchés & Mateo, 2013)

El patrón III (véase el gráfico 2), también neutro, se caracteriza por una inflexión final inferior a la del patrón II, de entre un 40 y un 60 %, aunque puede llegar al 70 %, y un primer pico desplazado a una vocal átona posterior a la primera tónica del contorno. También en este caso el cuerpo tiene una dirección descendente.

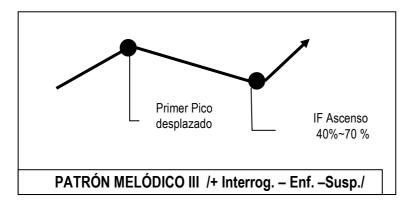


Gráfico 2. Patrón III del español peninsular (Font-Rotchés & Mateo, 2013)

El patrón IV (veáse gráfico 3), de carácter enfático, presenta una inflexión final circunfleja ascendente-descendente y se han identificado dos variantes: el patrón IVa, con un primer pico elevado desplazado a una vocal átona posterior a la primera tónica y un cuerpo descendente; y el patrón IVb, sin primer pico y un cuerpo en declinación plana.

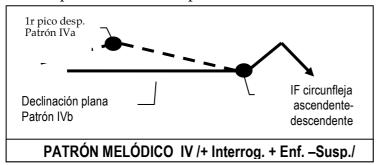


Gráfico 3. Patrón IV del español peninsular (Font-Rotchés & Mateo, 2013)

El patrón XIII (véase el gráfico 4), de carácter enfático, presenta un cuerpo ascendente, generalmente, sin primer pico, y con una inflexión final también ascendente, que suele ser de un 15 % o superior, aunque el rasgo melódico que parece determinante es el ascenso total que tiene lugar en el contorno, que se ha establecido en pruebas de percepción como igual o superior a un 90 % de ascenso.

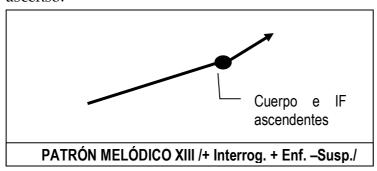


Gráfico 4. Patrón XIII del español peninsular (Font-Rotchés & Mateo, 2013)

3 Resultados

De los 50 enunciados-pregunta analizados (véase el gráfico 5), 13 (26 %) corresponden a la entonación interrogativa, 23 (46 %) a la entonación suspendida, 10 (20 %) a la entonación declarativa y 4 (8 %) a la entonación enfática, según los patrones del español peninsular descritos en Cantero & Font-Rotchés (2007).



Gráfico 5. Resultados

Por tanto, solo 13 de los 50 enunciados, un 26 %, presentan patrones interrogativos, frente a 37, un 74 %, en los que se usan otro tipo de patrones.

3.1 Contornos melódicos de la interlengua interrogativos

De los 13 patrones interrogativos del corpus, los informantes han utilizado mayoritariamente, en 10 casos, el patrón típico II de las interrogativas neutras, caracterizado por una inflexión final ascendente de +70 %, como vemos en el gráfico 6.

En este contorno, la inflexión final, que tiene lugar en *también*, presenta un 72,6 % de ascenso, lo que garantiza que va a ser interpretada como una pregunta por los hablantes nativos. Referente al primer pico, está situado en una sílaba tónica posterior *-lum-* a la primera tónica *-*Tie-, después de un leve ascenso de un 16 %, y el cuerpo presenta un descenso suave hasta la inflexión final. Aunque el primer pico presente este rasgo distinto que le aleja del patrón II, el rasgo de la inflexión final es suficiente para reconocerse el enunciado como pregunta.

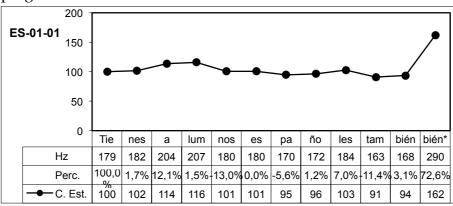


Gráfico 6. Patrón II: "¿Tienes alumnos españoles también?"

Los aprendientes usan, en menor grado, otros patrones interrogativos, como el patrón III de las interrogativas neutras, en 1 caso (véase el gráfico 7), caracterizado por un ascenso final menos pronunciado, de entre 40 y 70 % –la inflexión final tiene lugar en la palabra *Dylan*, que presenta un 47,5 % de ascenso— compensado con el desplazamiento del primer pico a una vocal átona posterior –en la sílaba *-ta* de *No te gusta*—, con un ascenso muy marcado de un 127 % (un 100 % equivale a una octava de la escala musical).

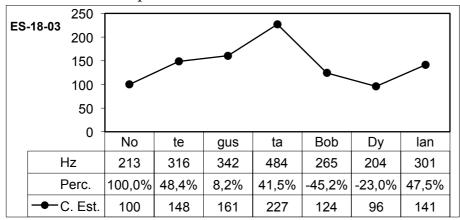


Gráfico 7. Patrón III: "¿No te gusta Bob Dylan?"

Y el patrón IV de las interrogativas enfáticas, en dos casos, caracterizado por una inflexión final circunfleja ascendente-descendente, como vemos en el gráfico 8.

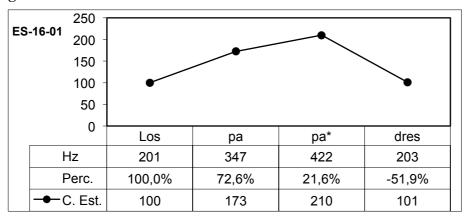


Gráfico 8. Patrón IV: "¿Los padres?"

No se ha encontrado ningún enunciado producido con el patrón XIII de las interrogativas enfáticas, de cuerpo e inflexión final ascendentes.

3.2 Contornos melódicos de la interlengua no interrogativos (74 %)

De los 37 (74 %) contornos que no siguen un patrón interrogativo, la mayoría se caracterizan por presentar una terminación más propia de patrones suspendidos, un 46 %, o de declarativos, un 20 %, y en menor grado, de enfáticos, un 8 %.

3.2.1 Suspendidos

Los informantes han producido 23 preguntas con patrones suspendidos, de inflexión final no terminativa, correspondientes en 19 casos a los patrones VIa y VIb (Cantero & Font-Rotchés, 2007), caracterizados por un ascenso final inferior al 70 %, o al 40 % con el primer pico desplazado, respectivamente. En los 4 casos restantes, se trata de contornos parecidos al patrón XIII de las interrogativas enfáticas, de cuerpo e inflexión final ascendente, pero que no alcanzan un ascenso total del 90 % para ser percibidos como preguntas (Font-Rotchés & Mateo, 2013). Los informantes utilizan patrones suspendidos posiblemente porque no llegan a producir el ascenso necesario que los convertiría en preguntas.

En el gráfico 9 vemos un ejemplo de pregunta producida con inflexión final no terminativa, correspondiente al patrón suspendido VIa, lo que provoca que se perciba como inacabada. En este caso, la inflexión final cuenta con un ascenso de un 24 % y además el contorno carece de primer pico. Realmente, se aleja mucho de cualquiera de los patrones interrogativos de la lengua meta.

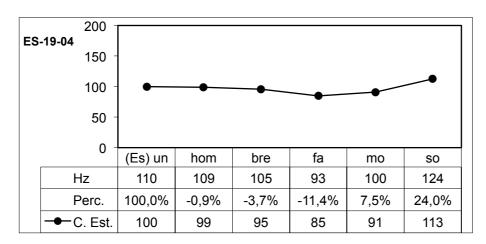


Gráfico 9. Patrón VIa: "¿Es un hombre famoso?"

Esta pregunta se percibe como si el informante tuviera que continuar hablando para ampliarla, por ejemplo: "¿es un hombre famoso... o solo es conocido aquí?", pero realmente su intervención ha terminado y espera respuesta. Se trata de un malentendido que suele darse en el aula, cuando un estudiante hace una pregunta y se produce un largo silencio hasta que el profesor responde porque este esperaba más información al no interpretar la pregunta como terminada.

En el gráfico 10 vemos un ejemplo de pregunta producida con un contorno similar al patrón XIII de las interrogativas enfáticas, con un cuerpo progresivamente ascendente, pero con un ascenso total de 34 %, muy por debajo del 90 % establecido para que sea percibido como pregunta.

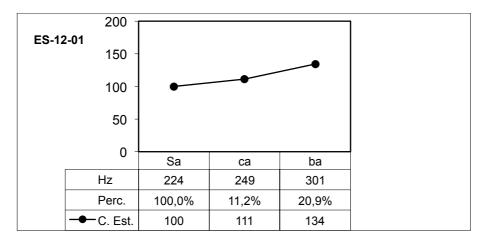


Gráfico 10. Patrón XIII con ascenso total insuficiente: "¿Sacaba?"

Este enunciado no se percibe como una pregunta, sino más bien como un comentario inacabado, lo que también se ha observado como causa habitual de confusión o malentendido en el aula, ya que en este caso el profesor no responde, porque no lo interpreta como pregunta, o duda de la intención comunicativa del mensaje del estudiante.

3.2.2 Declarativos

Los aprendientes han utilizado en 10 casos patrones declarativos para producir preguntas, correspondientes al patrón I de la entonación neutra o no marcada (Cantero & Font-Rotchés, 2007), con una inflexión final de ascenso inferior a 15 % o descenso de hasta 40 %. Probablemente porque les cuesta producir el ascenso final adecuado y, al no ser necesario en su lengua, no son conscientes de su rendimiento lingüístico.

La mayoría de estos patrones en nuestro corpus presentan un ascenso o descenso muy poco pronunciado, como se muestra en el siguiente ejemplo del gráfico 11, con un descenso de -5,4 % en la inflexión final, lo que provoca que se perciban como enunciados declarativos y no como preguntas, y por tanto, que se generen también problemas comunicativos.

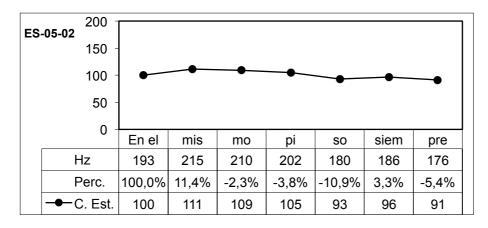


Gráfico 11. Patrón I: "¿En el mismo piso siempre?"

En el contexto de este enunciado no queda claro si se trata de una pregunta o una afirmación, lo que también suele darse en el aula cuando un estudiante formula una pregunta pero el profesor no responde porque interpreta que se trata de un comentario o confirmación y no de una duda.

3.2.3 Enfáticos

Y finalmente, los informantes han utilizado patrones enfáticos, en 4 casos, concretamente el patrón VII (Cantero & Font-Rotchés, 2007). Probablemente porque hay aspectos expresivos como la incredulidad o la sorpresa que dominan a la intención de preguntar, como ocurre en el ejemplo del gráfico 12.

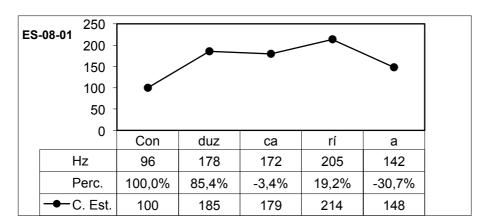


Gráfico 12. Patrón VII: "¿Conduzcaría?"

4 Conclusiones

Los informantes, hablantes nativos de sueco y de español como lengua extranjera, aún teniendo un nivel mínimo de B1, presentan dificultades para producir patrones interrogativos, y suelen emitir los enunciados interrogativos como declarativos o suspendidos, lo que causa deficiencias comunicativas.

Sería conveniente enseñarles, en primer lugar, a producir contornos con el ascenso necesario para preguntar y, posteriormente, otros tipos de melodías propias de las preguntas más complejas.

Valoramos positivamente la utilidad de proseguir con la investigación para obtener más datos y contextos, profundizar en el análisis de la interlengua, y poder elaborar propuestas didácticas que cubran estas carencias.

5 Bibliografía

Baditzné, K. (2012). Spanish Intonation of Hungarian learners of Spanish: yes or no questions. Tesis doctoral. Università Eötovös Loránd, Budapest. Biblioteca Phonica, 15. Recuperado el 14 de 1 de 2015, de http://www.publicacions.ub.edu/revistes/phonica-biblioteca/

Baditzné, K., & Font-Rotchés, D. (2013). The intonation of Spanish yes-no questions in spontaneus speech and in a didactic material. *Beszédkutatás*, 2013, 76-93.

Boersma, P. & Weenink, D. (1992-2011).

PRAAT. Doing phonetics by computer.

Institute of Phonetic Sciences,
University of Amsterdam. En

http://www.praat.org

Cantero, F. J. (2002). *Teoría y análisis de la entonación*. Barcelona: Ed. Universitat de Barcelona.

Cantero, F. J. & Font-Rotchés, D. (2007). Entonación del español peninsular en habla espontánea: patrones melódicos y márgenes de dispersión. *Moenia*, 13, 69-92.

Cantero, F. J. & Font-Rotchés, D. (2009). Protocolo para el análisis melódico del habla. *Estudios de Fonética Experimental*, XVIII, 17-32.

Cantero, F. J. & Mateo, M. (2011). Análisis melódico del habla: complejidad y entonación en el discurso. *Oralia*, 14, 105-127.

Devís, E. (2011). La entonación del español hablado por italianos. *Didáctica* (*Lengua y Literatura*), 23, 35-58.

- Fonseca, A. (2013). Caracterización de la entonación del español hablado por brasileños. Tesis doctoral. Dep. Didáctica de la Lengua y la Literatura. Universidad de Barcelona.
- Font-Rotchés, D. (2007): L'entonació del català. Barcelona: Publicacions de l'Abadia de Montserrat.
- Font-Rotchés, D. & Mateo, M. (2011). Absolute interrogatives in Spanish, a new melodic pattern. *Actas del VII Congresso Internacional da ABRALIN, Curitiba, Brasil*.
- Font-Rotchés, D. & Mateo, M. (2013). Entonación de las interrogativas absolutas del español meridional en habla espontánea. *Onomázein*, 28, 256-275.
- Instituto Cervantes (2014). El español: una lengua viva. Informe 2014.
- Liu, Y. H. (2003). *La entonación del español hablado por taiwaneses*. Tesis doctoral. Universitat de Barcelona. Biblioteca Phonica, 2 (2005). Recuperado el 14 de 1 de 2015, de http://www.publicacions.ub.edu/revistes/phonica-biblioteca/
- Martorell, L. (2010). Les interrogatives absolutes de l'espanyol parlat per suecs.
 Trabajo de fin de máster. Dep.
 Didáctica de la Lengua y la
 Literatura. Universidad de
 Barcelona. Dipòsit Digital UB.
 Recuperado el 14 de 1 de 2015, de
 http://diposit.ub.edu/dspace/handl
 e/2445/32302?mode=full.
- Mateo, M. (2010). Protocolo para la extracción de datos tonales y curva estándar en análisis melódico del habla (AMH). *Phonica*, 6, 49-90. Recuperado el 14 de 1 de 2015, de http://www.ub.edu/lfa/phonica.ht m

Aproximaciones a la caracterización prosódica de los subactos, la unidad discursiva mínima del sistema Val.Es.Co.

PASCUAL ALIAGA, ELENA

1 Introducción

La prosodia es considerada un criterio fundamental para segmentar el discurso oral en unidades discursivas. Numerosos modelos de segmentación (Degand & Simon, Morel & Boileau, 1998; Moneglia & Cresti, 2006, Lombardi, 2014, entre otros) adoptan criterios prosódicos para dividir la cadena hablada en diferentes segmentos con un sentido informativo determinado. En el caso de la propuesta del sistema de unidades del grupo Val.Es.Co. (Valencia Español Coloquial), la prosodia adquiere especial relevancia a la hora de delimitar las unidades menores del discurso conversacional coloquial, entre las cuales se encuentra el subacto.

El subacto se define desde el modelo Val.Es.Co. como la unidad informativa mínima del discurso conversacional, identificable a partir de criterios semántico-informativos y prosódicos (Briz y Grupo Val.Es.Co., 2003, 2014). El criterio prosódico fundamental que se emplea para reconocer subactos toma como base la unidad prosódica grupo entonativo (Quilis et al., 2003), y establece que todo subacto se corresponde con un grupo de entonación.

Sin embargo, como han demostrado diversos estudios (Cabedo 2009, 2014), el criterio prosódico sobre el que se apoya la segmentación en subactos presenta

una serie de conflictos: en primer lugar, la segmentación de la cadena fónica en grupos entonativos plantea el inconveniente de estar sometida a una percepción auditiva subjetiva ¹; y en segundo lugar, la premisa teórica de la correspondencia entre subactos y grupos de entonación no siempre se cumple, ya que en muchas ocasiones las fronteras de la unidad discursiva no coinciden con las de la unidad prosódica.

Ante todo ello, y dada la necesidad de utilizar el criterio semánticoinformativo para segmentar la conversación en subactos, surge la pregunta de hasta qué punto la unidad grupo entonativo –y por extensión, el criterio prosódico en sí- es útil para llevar a cabo la segmentación de una conversación en unidades informativas mínimas, esto es, en subactos.

Este interrogante se ha considerado desde una perspectiva teórica (Cabedo, 2009, 2014); este artículo presenta un estudio descriptivo con el fin de validar empíricamente las consideraciones teóricas previas. Este trabajo se articula como una aproximación inicial a la cuestión, y parte de los siguientes objetivos:

El estudio se propone, en primer lugar, contrastar la posibilidad de emplear otra unidad prosódica, el grupo fónico, para segmentar subactos. El grupo fónico es una unidad prosódica definida por Navarro Tomás (1918: 29-30) como «la porción de discurso comprendida entre dos pausas sucesivas de la articulación». Según Cantero (2002: 77), esta suele coincidir con un sintagma gramatical, si se tienen en cuenta fenómenos como el acento, con lo cual, podría ser operativa a la hora de segmentar el discurso en subactos.

En segundo lugar, se pretende obtener datos en relación a la correspondencia entre subactos y grupos entonativos, con el fin de poder establecer de forma objetiva la eficacia del criterio prosódico del modelo Val.Es.Co.

En tercer lugar, se lleva a cabo una descripción de rasgos prosódicos relativos a la duración y a la velocidad de habla de los grupos entonativos que se corresponden con subactos.

2 Algunas anotaciones sobre la relación entre informatividad y prosodia

2.1 El papel de la prosodia en la segmentación del discurso oral

La prosodia es un factor determinante para la construcción de un sentido dado en el lenguaje hablado (Hidalgo Navarro, 2006: 12). La distribución de la información en el discurso oral está determinada por los elementos suprasegmentales y no siempre obedece a las necesidades fisiológicas de la articulación (Chafe, 1994: 57):

[S]peech is not produced in a continuous, uninterrupted flow but in spurts. [...] The need to breath would alone produce the spurtlike quality of speech, but if one examines the linguistic and psychological nature of the spurts, it becomes clear that more is involved. [...] [T]his

_

Las últimas investigaciones (Cabedo, 2009) han establecido nuevos parámetros para delimitar segmentos fónicos de forma más rigurosa: por un lado, se ha redefinido el concepto de grupo entonativo, de forma que se propone una nueva unidad ("grupo prosódico") en base a criterios más objetivos; y por otro lado, se ha confeccionado el algoritmo MESTEL (Modelo Estadístico para la Selección de Términos Entonativos Ligados), que permite calcular la probabilidad de que una frontera fónica constituya un grupo prosódico. No obstante, dado el carácter preliminar de este estudio, se ha optado por realizar el análisis de los grupos entonativos según los criterios aportados por Quilis et al. (1993).

physiological requirement operates in happy synchrony with some basic functional segmentations of discourse.

La articulación prosódica del flujo de información en segmentos o unidades responde a la forma en la que un hablante desea presentar la información (Brown y Yule, 1983: 157); el oyente, por otro lado, también identifica unidades lingüísticas a partir de la división del discurso en bloques fónicos (Cantero, 2002:85)². En este sentido, la prosodia puede considerarse un criterio que proporciona elementos objetivos y mesurables (pausas, unidades entonativas, cambios melódicos o de la F0) para dividir muestras de habla (Pons, 2014). Es lógico, pues, que muchos sistemas de segmentación discursiva tomen el componente fónico como criterio principal para llevar a cabo la división del discurso en unidades (véanse las propuestas de Degand & Simon, 2009; Morel & Boileau, 1998; Moneglia & Cresti, 2006; o Lombardi, 2014, entre otras).

Sin embargo, tal y como expresa Pons (2014), cabe preguntarse hasta qué punto la prosodia debe considerarse el factor primordial para la delimitación de unidades discursivas.

2.2 El caso del subacto, la unidad mínima del sistema Val.Es.Co.

El sistema de unidades del grupo Val.Es.Co. se define como un modelo funcional pensado "por y para la conversación coloquial española" (Briz y Grupo Val.Es.Co. 2003, 2014) que presenta las siguientes características: se estructura en tres niveles (el dialógico, que se vincula con el orden social o externo de la conversación, y el monológico, que se relaciona con estructura interna del discurso); es jerárquico (las unidades menores son los constituyentes inmediatos de las superiores); es recursivo (las unidades menores contienen unidades de su mismo nivel o de un orden superior); y se agrupa por dimensiones (la social, la estructural y la informativa). El modelo presenta un total de ocho unidades: el discurso, el diálogo, el intercambio, la alternancia de turnos, la intervención, el turno, el acto y el subacto.

Los distintos niveles, dimensiones y unidades se interrelacionan, como puede observarse en la tabla 1. El subacto es la unidad mínima del sistema. Pertenece al plano informativo del discurso, ya que a través de ella se puede observar la articulación y distribución del contenido informativo, y a su vez, se sitúa en el nivel monológico:

² Cantero (2002: 213) afirma, además, que «la entonación, como fenómeno fónico y, solidariamente con el acento, [es] previ[a] a la propia estructuración lingüística del discurso». Expresa que esta posee una función prelingüística previa, por tanto, a la organización informativa del discurso.

NIVEL	DIMENSIONES		
	ESTRUCTURAL	SOCIAL	INFORMATIVA
Dialógico	discurso diálogo intercambio	alternancia de turnos	
Monológico	intervención acto	turno	subacto

Tabla 1. Niveles, dimensiones y unidades del sistema Val.Es.Co. (Briz y Grupo Val.Es.Co., 2014: 14)

El subacto se define en el sistema Val.Es.Co. como una «unidad monológica estructural, constituyente inmediato del acto, caracterizada por constituir un segmento informativo identificable, habitualmente, mediante marcas semánticas y prosódicas» (Briz y Grupo Val.Es.Co., 2014: 53). La siguiente figura sintetiza los diferentes tipos de subacto que, hasta el momento, se han planteado en la propuesta de Val.Es.Co.3:

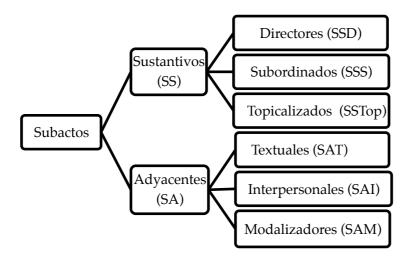


Figura 1. Tipos de subactos

Los subactos se definen e identifican por medio de marcas semánticas y prosódicas. En función del tipo de subacto, prevalecen unas marcas u otras para identificarlos: el criterio semántico-informativo tiene más relevancia para la delimitación de subactos sustantivos —que presentan un contenido proposicional-, y el fónico, para el reconocimiento de subactos adyacentes — cuyo contenido no es proposicional-. A continuación se describe qué elementos componen los criterios prosódico e informativo de reconocimiento de subactos.

2.2.1 El criterio prosódico

El grupo Val.Es.Co. considera el criterio prosódico un "factor determinante y absolutamente necesario para la identificación y reconocimiento de subactos" (Hidalgo Navarro, 2006: 138). El criterio prosódico toma como base la unidad grupo entonativo, según la definen Quilis *et. al.* (1993: 56-57):

³ De ahora en adelante se denominará a los diferentes tipos de subacto según las abreviaturas que aparecen junto a cada nombre.; p. ej., se empleará SSD para hacer referencia a los subactos sustantivos directores.

Porción de discurso comprendida entre dos pausas, entre pausa e inflexión del fundamental, entre inflexión del fundamental y pausa o entre dos inflexiones del fundamental, que configuran una unidad sintáctica más o menos larga o compleja (sintagma, cláusula, oración).

Teóricamente, los subactos presentan autonomía fónica y conforman grupos entonativos independientes; sin embargo, Cabedo (2009, 2014) advierte que no siempre hay una correspondencia unívoca entre la unidad prosódica y la discursiva: por ejemplo, las rupturas de la cadena fónica, la presencia de subactos suspendidos (que son entonativamente truncos, pero informativamente completos, puesto que el interlocutor los reconstruye) o la velocidad de habla son fenómenos que manifiestan la no correspondencia entre subactos y grupos entonativos.

Además de la correspondencia con la unidad fónica grupo entonativo, hay otras marcas prosódicas que permiten reconocer tipos de subactos:

- La altura tonal de la sílaba final de un grupo entonativo denota dependencia (si esta presenta un tono bajo ↓) –función integradora de la entonación– o independencia (si su tono es alto ↑) –función demarcativa de la entonación– de dicho grupo de entonación frente al segmento entonativo que le sigue.
- Una realización prosódica prominente que presente una elevada F0 puede ayudar a distinguir conectores pragmáticos (SAT) de conjunciones.
- Los SSSTop también pueden identificarse a partir de rasgos prosódicos (Hidalgo y Padilla, 2006): no suele haber una pausa entre el segmento previo o posterior al SSSTop; en los segmentos topicalizados a la izquierda puede percibirse un énfasis prosódico del tonema final del grupo entonativo del subacto –la F0 final alcanza o supera el promedio tonal del hablante–; y los segmentos topicalizados a la derecha se integran en la curva melódica principal.

2.2.2 El criterio semántico-informativo

Un subacto es un segmento informativo que puede presentar tres tipos de contenido proposicional (Briz y Grupo Val.Es.Co., 2014: 57):

- Informaciones proposicionales primarias (narrativas, descriptivas, argumentativas, factitivas, etc.); es el caso de los SSD, que son elementos nucleares informativamente independientes.
- Informaciones secundarias (causa, condición, consecuencia, finalidad, tiempo, lugar, topicalización, etc.); las contienen los SSS y SSSTop, y se materializan en relaciones de predicación equiparables a las sintácticas⁴.
- Informaciones extraproposicionales (es el caso de los marcadores del discurso, por ejemplo); son las propias de los SA, ya que estos tienen menos sustancia semántica

El siguiente ejemplo⁵ muestra la segmentación en subactos de la intervención de un hablante denominado "B":

(0) B: {SSD nos pasamos el día con la pelota en los pies SSD} {SSS pero ((no lo)) sé por qué ↑ // {SAM oye SAM} nada SSS} /// (1,44) {SAT pero SAT}

⁴ En Briz y Grupo Val.Es.Co. (2014: 60, 2003: 51) se expresa que la distinción funcional entre subactos directores y subactos subordinados puede compararse con la oposición entre el tema y el rema, o, sintácticamente, con una oración principal y una oración subordinada, o una primera oración coordinada copulativa y una segunda coordinada copulativa, respectivamente.

⁵ Para la transcripción de los ejemplos se ha empleado el sistema de transcripción semiestrecha propuesto en Briz (2004): Las barras se corresponden con pausas de duración corta "/" (inferior a 0,5 segundos), media "//" (entre 0,5 y 1 segundo) y larga "//" (superior a 1 segundo); los corchetes "[]" señalan la presencia de habla simultánea; el doble paréntesis "(())" indica una transcripción dudosa; la letra en cursiva se emplea para señalar estilo directo de habla; y las llaves "{}" marcan las fronteras entre los subactos, cuya abreviatura aparece anotada en su interior.

{SSD malos malos SSD} {SAM ;eh? SAM} / {SSD se puede ser mejor o peor SSD} {SAI ¿lo entiendes? SAI} / {SSD Julio ((Cano)) no es un crack al fútbol SSD} {SSS peroo // juega SSS} {SSD y yo no soy el a- el artista del mundo SSD} {SSS pero juegas SSS} {SAI ¿sabes? SAI}

3 Metodología

Para realizar este estudio se seleccionó como corpus de trabajo la conversación n.º 44 del Corpus Val.Es.Co. 2.0. (Cabedo y Pons, 2013). Este corpus recoge conversaciones coloquiales representativas del área geográfica de Valencia que han sido grabadas, en su mayoría, de forma secreta. La conversación 44 fue grabada en 20116, tiene un total de 10 minutos de duración y está protagonizada por dos que tienen entre 18 y 25 años de edad. La transcripción de la conversación se alineó con el audio de la grabación mediante el programa Praat (Boersma & Weenik, 2014).

Se llevaron a cabo tres segmentaciones sucesivas de la conversación 44: la segmentación prosódica en grupos fónicos, la segmentación prosódica en

grupos entonativos y la segmentación discursiva en subactos.

La segmentación prosódica en grupos fónicos contempla la pausa como el único elemento demarcador de unidades (Navarro Tomás, 1918: 29-30); dado el riesgo de confluencia de otros elementos articulatorios en los vacíos de F0 (pronunciación de sibilantes, fricativas, etc.), se estableció una duración mínima de 0,2 milésimas de segundo para considerar un silencio como frontera prosódica de grupo fónico, siguiendo el criterio fijado por Cabedo (2009, 131:137). Además, se consideró que las pausas oralizadas, que no conllevan una interrupción de la fonación, sí delimitan unidades prosódicas. La segmentación en grupos fónicos se realizó con anterioridad a la segmentación en grupos entonativos, ya que esta primera unidad engloba a segunda: los grupos fónicos utilizan el criterio pausal como única marca de frontera prosódica, mientras que los grupos entonativos utilizan las pausas y las inflexiones tonales.

Para la segmentación en grupos entonativos se estableció, además de las pausas e inflexiones tonales (Quilis et al.,1993), una serie de marcas prosódicas fronterizas para llevar a cabo una delimitación más rigurosa: los alargamientos los tonemas suspendidos, la pronunciación marcada, truncamientos fónicos, las modalidades interrogativa y exclamativa, la presencia de paralenguaje y los fragmentos en estilo directo de habla. Además, el análisis fue validado por un total de cuatro informantes7, con el fin de que la segmentación adquiriese mayor validez empírica. La segmentación en grupos entonativos se realizó con anterioridad a la discursiva debido a que la unidad grupo entonativo constituye la base sobre la que se asienta la segmentación en subactos.

La segmentación discursiva en subactos realizó conforme a los criterios teóricos propuestos por el modelo Val.Es.Co. (vid. §2.2): a partir de los grupos entonativos delimitados en la segmentación previa, se identificaron los subactos de acuerdo con los parámetros informativos fijados en el sistema.

Por último, se elaboró una base de datos mediante el programa Microsoft con el fin de recoger y cuantificar todos los resultados de las segmentaciones: se incluyeron las variables cuantitativas "grupo fónico",

que la coincidencia fuera total o mayoritaria».

7 Cabedo (2011) expresa que «[o]tro modo de segmentar el discurso oral, aunque más costoso, es el de pedir a diferentes hablantes que delimiten grupos entonativos. Posteriormente, se tendrían en cuenta únicamente aquellos casos en los

⁶ Los participantes de la conversación han dado su consentimiento para el uso de los datos.

"grupo entonativo" y "tipo de subacto", y se anotaron algunas características prosódicas a partir de las variables "tiempo de inicio del segmento", "tiempo de finalización del segmento" y "número de sílabas", a partir de las que ha sido posible extraer la duración total de los subactos y su velocidad de habla.

4 Análisis de los datos

La siguiente tabla muestra la distribución del número total de unidades obtenidas tras las segmentaciones prosódicas y la segmentación discursiva del corpus:

	Unidad	
Prosódica	GRUPO FÓNICO	335
Pr	GRUPO ENTONATIVO	495
Discusiva	SUBACTOS	505

Tabla 2 Comparativa de los resultados totales de las segmentaciones

Como se puede observar, el número total de subactos no coincide con el número total de ninguna de las dos unidades prosódicas. Si bien los resultados de las unidades subacto y grupo entonativo se encuentran bastante próximos, los resultados de la unidad grupo fónico la menoscaban como unidad prosódica de apoyo para la segmentación discursiva. Obsérvese más detenidamente la distribución del porcentaje de coincidencias entre subactos y grupos fónicos⁸:

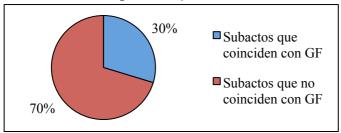


Figura 2. Distribución porcentual de las coincidencias entre subactos y grupos fónicos

Únicamente un 30 % de los subactos analizados se corresponde de forma exacta con grupos fónicos, como es el caso del siguiente ejemplo:

(1) A: {SSD ¿¡cómo [se pue(de) teneer] veintiún años y estar sacándose el bachilletaro→ SSD} // {SSS si se lo SACA SSS} / {SSS con toda su tranquilidá→ SSS} // {SSS y no hace nada más en su vida!? SSS}

En cambio, el 70 % de subactos restantes no presentan límites prosódicos que coincidan con pausas demarcativas de grupos fónicos. Son, por tanto, subactos que comprenden más de un grupo fónico, tal y como se muestra en el ejemplo

⁸ La abreviatura GF que aparece en la figura hace referencia a "grupos fónicos".

- (2), o grupos fónicos que engloban dos o más subactos, según puede apreciarse en el ejemplo (3):
- (2) A: {SSS aparte de que si te vas un fin de semana y sales el viernes de noche↑ SSS} / {SSD vas a estar allíi / puramente el sábado↑ SSD}
- (3) B: {SAT entonces eso SAT} / eeh (PAUSA ORALIZADA) / {SSD yo voy a decirle SSD} {SAI tío SAI} {SSS no vayas a la última clasee SSS} {SAI o oye SAI} {SSS ves hasta la una SSS} {SAI ; sabes? SAI} ///(2,97)

Estos resultados llevan a la conclusión de que la unidad grupo fónico no es efectiva a la hora de delimitar subactos. Parece, pues, que la unidad base que se propone desde el sistema Val.Es.Co., el grupo de entonación, es más operativa que el grupo fónico a la hora de delimitar subactos. Obsérvese a continuación los porcentajes de correspondencias entre subactos y grupos entonativos9:

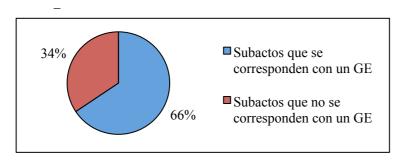


Figura 3. Distribución porcentual de las coincidencias entre subactos y grupos entonativos

El 66% de los subactos segmentados se corresponden con grupos entonativos. Esto significa que los límites prosódicos son aplicables, en más de la mitad de los casos, a la unidad mínima del modelo Val.Es.Co., tal y como se ilustra en el siguiente ejemplo:

(4) A: {SSS te oye↑ SSS} {SSD y y no vuelve a hablarte en tu vida SSD} B: {SSD alguna vez se lo hemos intentado explicar y noo SSD} // {SSD él dice que es un crack y SSD} /

El 34% de los subactos del corpus, sin embargo, no coincide con las fronteras prosódicas de los grupos de entonación. Estos subactos son, del mismo modo que en el caso de los grupos fónicos, superiores a los grupos entonativos, esto es, subactos que aglutinan dos o más grupos entonativos (5); o unidades inferiores a los grupos entonativos (6):

- (5) A: {SSS aparte de que si te vas un fin de semana y sales el viernes de noche↑ SSS} / {SSD vas a estar allíi / puramente el sábado↑ SSD} {SSS porque el domingo tampoco podéis volver→ / [tarde que tenéis que [(())] SSS
 - (6) A: {SAM (os)tras SAM} {SSD es MUY perro SSD} {SAM ¿eh? SAM} / / /

⁹ La abreviatura GE que aparece en esta figura y en las sucesivas hace referencia a "grupos entonativos".

La siguiente figura muestra las relaciones de coincidencias entre el número total de los diferentes tipos de subactos y los grupos entonativos:

200 180 160 140 no coincidencia 120 100 34 80 134 coincidencia 60 subacto+ge 40 78 60 20 28 0 **SSD** SSS SSSTop SAI SAM **SAT** Subactos sustantivos Subactos adyacentes

Figura 4. Distribución de las coincidencias entre subactos y grupos entonativos

En lo que sigue, el foco de atención se centrará en los subactos que se corresponden con grupos entonativos. Se analizará la correspondencia entre los diferentes tipos de subacto y los grupos de entonación (4.1.), y también se observarán algunas características supragsegmentales de esos subactos: la duración media de los segmentos (4.1.1.) y la velocidad media de habla (4.1.2.).

4.1 Subactos que coinciden con grupos entonativos

El grado de correspondencia entre subactos y grupos de entonación, si bien en general es alto (66 %), varía según el tipo de subacto. Las siguientes figuras muestran la distribución de las correspondencias entre los diferentes tipos de subactos y los grupos de entonación:

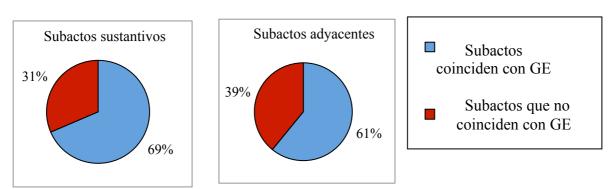


Figura 5. Distribución porcentual de coincidencias entre tipos de subacto y grupos entonativos

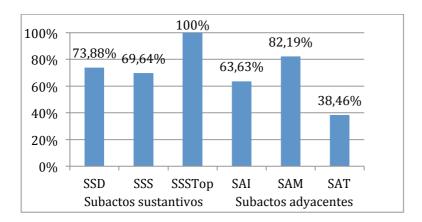


Figura 6. Distribución porcentual de los tipos de subactos que coinciden con grupos entonativos

Como se puede observar en la figura 5, la distribución general de los subactos sustantivos y adyacentes que coinciden con grupos entonativos es muy similar (69 % y 61 %, respectivamente). Sin embargo, si se atiende a la distribución de los tipos concretos de subactos (figura 6) resaltan algunos aspectos interesantes:

Por un lado, destaca, especialmente, el porcentaje absoluto de coincidencia de los subactos sustantivos topicalizados (SSSTop) con los grupos de entonación. No siempre es posible contar con las marcas prosódicas de reconocimiento de los SSSTop¹o, principalmente, debido a factores relacionados con la calidad de la grabación. En el ejemplo (7), cuya representación fónica se muestra seguidamente en la figura 7, la presencia de habla simultánea y de ruido no permiten definir con exactitud si en la intervención de A hay un tonema descendente que marca un límite prosódico y, por tanto, un SSSTop a la derecha:



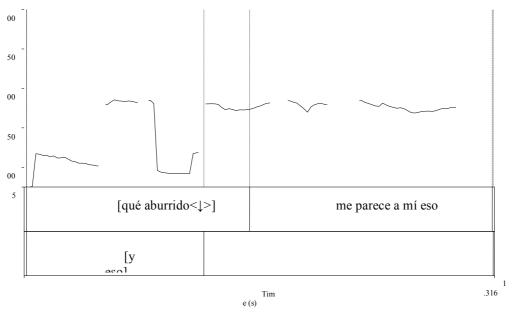


Figura 7. Representación de la curva melódica del ejemplo (7)

¹⁰ Hidalgo y Padilla (2006) realizan un análisis detallado de las características prosódicas de los SSSTop que facilita su identificación.

El orden sintáctico no marcado de los elementos de la intervención de A es lo que ha permitido, en este caso, identificar el SSSTop. Pese a la presencia de este tipo de ejemplos, los resultados del análisis revelan que, en lo que respecta a la conversación analizada, la prosodia es un factor fundamental a la hora de caracterizar este tipo de subactos y, consiguientemente, las topicalizaciónes de la información a la izquierda o a la derecha. También puede observarse en la figura 6 que el porcentaje de correspondencias entre subactos adyacentes modalizadores (SAM) y grupos entonativos es bastante elevado.

Por otro lado, también destaca el bajo porcentaje de correspondencias entre los subactos adyacentes textuales (SAT) y los subactos (38,46 %). Este factor es llamativo, ya que desde el propio modelo Val.Es.Co. se afirma que la unidad prosódica del grupo entonativo es especialmente operativa en el caso de los subactos adyacentes, dado que estos no poseen contenido proposicional. Los otros dos tipos de subactos adyacentes (interpersonales -SAI- y modales -SAM-) sí que presentan un grado de equivalencia con las fronteras prosódicas mucho más elevado, sobre todo por lo que respecta a los SAI.

4.1.1 Duración de los subactos que se corresponden con grupos entonativos

La figura 8 muestra la duración media en milésimas de segundo de los distintos segmentos discursivos que coinciden con grupos de entonación:

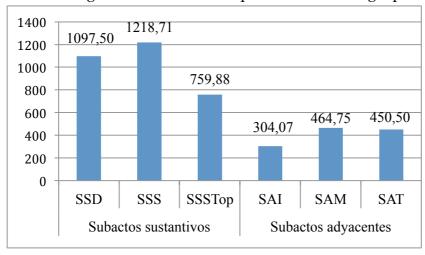


Figura 8. Duración media de los subactos que coinciden con grupos entonativos

En general, la duración media de los subactos adyacentes es inferior a la de los subactos sustantivos. Los subactos adyacentes interpersonales son los que cuentan con una duración menor (304,07 milésimas de segundo), y los subactos sustantivos subordinados son los que presentan una duración media mayor (1218,71 milésimas de segundo). Si se calcula la duración media del conjunto de subactos sustantivos (1205,18 milésimas por segundo) y del conjunto de adyacentes (406,44 milésimas de segundo), puede apreciarse que la diferencia de la duración entre ambas categorías otra es de 618,74 milésimas de segundo.

Estos resultados son coherentes si se tiene en cuenta el tipo de información que transmiten los diferentes tipos de subactos: los elementos con contenido proposicional (subactos sustantivos) tienen una mayor duración que los que no tienen contenido proposicional (subactos adyacentes). Los subactos adyacentes interpersonales (SAI) son elementos que manifiestan la interacción entre

hablante y oyente: vocativos, marcadores fáticos, elementos de regulación del contacto, etc,; los subactos adyacentes modalizadores (SAM) son expresiones que manifiestan la modalidad del hablante; los subactos adyacentes textuales (SAT) son elementos que organizan y distribuyen el flujo del habla tales como los marcadores discursivos.

4.1.2 Velocidad de habla de los subactos que coinciden con grupos entonativos

En la siguiente figura puede observarse la velocidad de habla media de los subactos que conforman grupos entonativos.

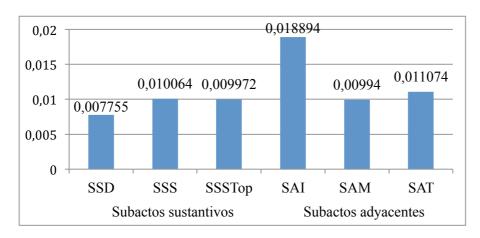


Figura 9. Velocidad de habla media de los subactos que coinciden con grupos entonativos

La medición de la velocidad de habla de los segmentos analizados se ha establecido a partir de las sílabas pronunciadas por segundo (Laver, 2002: 539-400). Según puede observarse, el tipo de subactos cuya velocidad de habla es mayor son los adyacentes interpersonales, seguidos de los adyacentes textuales y de los sustantivos subordinados. Los subactos que presentan una menor velocidad de habla son los subactos sustantivos directores.

5 Síntesis y apuntes finales

En relación a los objetivos enunciados con anterioridad (§ 1) pueden apuntarse algunas conclusiones sobre la segmentación discursiva en subactos a partir de los datos analizados:

El empleo de la unidad grupo fónico se descarta como criterio prosódico eficaz para la segmentación discursiva en subactos, ya que solo el 30 % de los subactos analizados presenta los mismos límites prosódicos que esta unidad fónica.

La unidad grupo entonativo permite, en cambio, identificar más de la mitad de los subactos de la conversación 44 (un 66 % del total de subactos). También permite identificar el 69 % de los subactos sustantivos y el 61 % de los subactos adyacentes, e incluso reconocer con bastante precisión determinados tipos de subactos, como los topicalizados (100 % de los casos) y los adyacentes modalizadores (82,19 %). Sin embargo, el criterio prosódico no es suficiente para llevar a cabo la segmentación completa del discurso en subactos, y es especialmente problemático en lo que respecta a la delimitación de subactos adyacentes textuales (38, 46 %).

Por último, pueden destacarse algunos aspectos prosódicos suprasegmentales de los grupos entonativos que coinciden con subactos: en

relación a la duración media de los subactos, los adyacentes presentan una duración menor que los sustantivos; concretamente, la diferencia entre ambas categorías de subactos radica en un total de 618,74 milésimas de segundo. La velocidad de habla media más elevada es la que presentan los subactos adyacentes interpersonales, textuales y los subactos sustantivos subordinados; los subactos sustantivos directores son, en cambio, los que tienen la menor velocidad de habla.

Los resultados de estos análisis hacen referencia a una sola conversación segmentada y responden al objetivo de forjarse como una base empírica preliminar sobre la que asentar futuras investigaciones. No son, por tanto, resultados representativos, y podrían validarse por medio de una ampliación del corpus.

Algunas de las líneas de trabajo que podrían seguirse para profundizar en la investigación de las cuestiones planteadas en este estudio son las siguientes:

- Realizar el análisis de los subactos que no se corresponden con grupos entonativos para observar sus características y profundizar en la cuestión de la aplicabilidad del criterio prosódico.
- Ampliar del corpus de trabajo para validar los resultados del análisis.
- Utilizar un sistema más sofisticado de detección de límites prosódicos, como el MESTEL (Cabedo, 2009), para asegurar una mayor objetividad en los resultados de la segmentación prosódica.
- Analizar la operatividad del criterio semántico-informativo en la delimitación de subactos.
- Contrastar los criterios propuestos por el sistema Val.Es.Co. para identificar unidades informativas menores en el discurso con los criterios que conjugan otros modelos de segmentación del discurso.

6 Referencias bibliográficas

- Boersma, P., & Weenink, D. (2014). *Praat:*Doing phonetics by computer.

 Recuperado el 20 de 3 de 2014, de

 /http://www.praat.org/
- Briz, A. (2004). "La transcripción de la conversación. El sistema Val.Es.Co.". En Almela, R. et al. (eds.), Homenaje al profesor Estanislao Ramón Trives, vol. 1, Murcia: Universidad de Murcia, 141-159.
- Briz, A. & Grupo Val.Es.Co. (2014). "Las unidades del discurso oral. La propuesta Val.Es.Co. de segmentación de la conversación (coloquial)". Estudios de Lingüística del Español, 35 (1), 11-7.
- Briz, A. (2004). "La transcripción de la conversación. El sistema Val.Es.Co.". En Almela Pérez, R. et al. (coords.), Homenaje al profesor Estanislao Ramón Trives, vol. I. Murcia: Universidad de Murcia, 141-159.
- Briz, A. & Grupo Val.Es.Co. (2003). "Un sistema de unidades para el estudio del lenguaje coloquial". *Oralia*, 6, 7-61.

- Briz y Grupo Val.Es.Co. (2003). "Un sistema de unidades para el estudio del lenguaje coloquial", *Oralia*, vol. 6, pp. 7-61.
- Brown, g. & Yule, G. (1983). *Discourse* analysis. Cambridge: Cambridge University Press.
- Cabedo, A. (2014). "On the delimitation of discursive units in colloquial Spanish: Val.Es.Co application model". En S. Pons (ed.), Discourse Segmentation in Romance Languages. Ámsterdam/Filadelfia: John Benjamins, 157-183.
- Cabedo, A. (2009). Segmentación prosódica en la conversación coloquial: sobre el grupo entonativo como mecanismo demarcativo de unidades mínimas. Valencia: Universitat de Valencia.
- Cabedo, A. & Pons, S. (eds.) (2013). *Corpus Val.Es.Co* 2.0. Recuperado el 20 de 3 de 2014 en /http://www.valesco.es/
- Cantero, F. J. (2002): *Teoría y análisis de la entonación*. Barcelona: Edicions Universitat de Barcelona.
- Chafe, W. (1994). Discourse, Consciousness, and Time: The Flow and Displacement of

- Conscious Experience in Speaking and Writin. Chicago/Londres: University of Chicago Press.
- Degand, L. & Simon, A. C. (2009). "On identifying basic discourse units in speech: theoretical and empirical issues". *Discours*, 4. Recuperado el 20 de 3 de 2014 de /http://discours.revues.org/index5 852.html/
- Hidalgo Navarro, A. (2006). Aspectos de la entonación española: viejos y nuevos enfoques. Madrid: Arco Libros.
- Hidalgo, A. & Padilla, X. (2006). "Bases para el análisis de las unidades menores del discurso oral: los subactos". *Oralia*, 9, 109-143.
- Laver, J. (2002). *Principles of phonetics*. Cambridge: Cambridge University Press.
- Lombardi, (2014). The topologic hypothesis of prominence as a cue to information structure in Italian. En Pons, S. (ed.), *Discourse Segmentation*

- in Romance Languages. Ámsterdam/Filadelfia: John Benjamins, 219-241.
- Moneglia, M. & Cresti, E. (2006). "C-ORAL-ROM: Prosodic boundaries for spontaneous speech analysis", En Kawaguchi, Y., Zaima, S. & Takagaki, T.(eds.), Spoken Language Corpus and Linguistics Informatics, Ámsterdam: John Benjamins, 89-114.
- Navarro Tomás, T. (1918). Manual de pronunciación española. Madrid: Consejo superior de investigaciones científicas.
- Pons, S. (2014). *Discourse Segmentation in Romance Languages*. Ámsterdam/Filadelfia: John Benjamins.
- Quilis, A., Cantarero, M. & Esgueva, M. (1993). "El grupo fónico y el grupo de entonación en el español hablado". *Revista de Filología Española*, 73 (1/2), 55-64.

Aspectos prosódicos en una zona de transición lingüística: la Franja de Aragón

ROMERA BARRIOS, LOURDES; FERNÁNDEZ PLANAS, ANA MA.; ROSEANO, PAOLO; ELVIRA-GARCÍA, WENDY; CARRERA SABATÉ, JOSEFINA; MARTÍNEZ CELDRÁN, EUGENIO

1 Resumen

La Franja de Aragón, comprendida entre las comunidades de Cataluña y Aragón, se reconoce como una zona de transición de diferentes variedades románicas (Giralt, 2012). En este estudio se aborda el análisis de las variedades prosódicas de catalán y castellano de dos lugares de la Franja, Fraga y Mequinenza, y la comparación con el castellano y el catalán de otras zonas de la misma área lingüística en las que se habla alguna de estas dos lenguas o bien ambas.

¹ Este trabajo se ha realizado dentro del Proyecto de investigación financiado por el MEC FFI2012-35998.

Este trabajo se inscribe en el proyecto AMPER (Contini et al., 2002) y AMPERCAT (Fernández Planas, 2005). Los datos analizados corresponden a dos hablantes, de cada uno de los puntos de encuesta, y comprenden las frases del corpus fijo establecido en este proyecto: frases declarativas absolutas, con estructura SVO; interrogativas absolutas con estructura SVO e interrogativas, también absolutas, encabezadas por la partícula *que* átona que son propias de algunas variedades del catalán.

Las frases presentan tres o cuatro acentos léxicos que recogen todas las posibilidades acentuales de estas lenguas: aguda, llana y esdrújula. El análisis acústico realizado con el programa Amper 2006 (López Bobo *et al.*, 2007), nos proporciona los datos de F0, duración e intensidad de cada vocal de las frases. Los contornos entonativos de las emisiones se han anotado según el modelo métrico y autosegmental (Pierrehumbert, 1980, Hualde 2003) y las propuestas de Cat_ToBI y Sp_ToBI (Prieto *et al.*, 2009; Estebas & Prieto, 2009).

En la primera parte del trabajo, se comparan los patrones entonativos de Fraga y Mequinenza con los obtenidos en otras zonas bilingües como Barcelona (Romera *et al.*, 2008), o Lleida y Tortosa (Fernández Planas *et al.*, 2005). Se continúa la comparación con los contornos del castellano de Zaragoza (Castañer *et al.*, 2005), como zona próxima a la Franja, y con los de otras zonas monolingües como Madrid. Este análisis comparativo nos lleva a establecer los rasgos entonativos que comparten las variedades de la Franja con las otras variedades examinadas. En la segunda parte del trabajo se realiza la comparación de la duración de las vocales en los diferentes acentos tonales de las frases, especialmente las correspondientes al tonema.

Tanto los resultados de la descripción melódica, como los de la duración, permiten justificar las convergencias mutuas entre el catalán de la Franja y el de otros puntos próximos a Fraga y Mequinenza, así como establecer las diferencias con otras variedades del castellano como las de Madrid, Zaragoza, Barcelona, Lleida o Tortosa.

2 Introdución: el catalán de la Franja

La Franja oriental de Aragón o Franja de Ponent, situada entre las comunidades de Cataluña y Valencia, se reconoce como una zona de transición de diferentes variedades románicas (Giralt, 2012). Según constatan los informes de 1996 y de 2007, EUROMOSAIC y ASOLOC respectivamente, es en esta zona donde se da el uso más elevado del catalán dentro de la comunidad lingüística catalana.

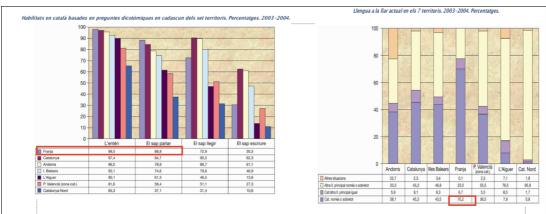


Figura 1. Uso del catalán (habilidades y uso en casa) según el informe ASOLOC

Entre los estudios sobre el catalán de la Franja cabe citar el trabajo de Galán & Moret (1995) que presenta una descripción de la lengua de Fraga y el de Giralt Latorre (2012) que aporta una visión histórico-lingüística del catalán de la Franja. Un análisis actual de la prosodia de esta zona la encontramos en los trabajos de Borràs-Comes (2013) y Prieto & Cabré (2007-2012 y 2013).

3 Objetivo, metodología y análisis

En este estudio nos planteamos averiguar si hay diferencias entonativas en la prosodia del catalán de la Franja respecto de las variedades de catalán más próximas (Lleida, Barcelona y Tortosa) y si también se pueden observar diferencias entonativas con el español de otras zonas bilingües (Mequinenza, Lleida, Barcelona) y monolingües (Zaragoza y Madrid). Y si además de las diferencias entonativas existe alguna otra característica prosódica que establezca diferencias en estas variedades dialectales.

De acuerdo con el protocolo de análisis del proyecto AMPER (Contini *et al.*, 2002; Fernández Planas, 2005), el perfil del informante corresponde a hombres y mujeres entre 25 y 55 años sin estudios superiores; se ha utilizado el corpus fijo de las denominadas frases sin expansión, un total de 27 oraciones para cada modalidad; el análisis acústico se ha realizado en el entorno Matlab con el progama AMPER 06 (López Bobo *et al.*, 2007) que proporciona datos de F0, duración e intensidad de todas las vocales de las frases analizadas. Finalmente la síntesis permite obtener las melodía de la frase desprovista de su contenido léxico.

También se ha procedido al análisis de los patrones tonales según el modelo AM (Pierrehumbert, 1980), en el que la curva de F0 se considera la representación fonética, superficial, de la estructura tonal subyacente o fonológica; esta representación subyacente consiste en secuencias estructuradas de tonos altos (H) y bajos (L) que se sitúan alrededor de las sílabas métricamente fuertes o acentuadas (*). Los patrones entonativos están formados por secuencias de uno o más acentos tonales más tonos de frontera (%) inicial y final respectivamente.

4 Descripción y comparación de los contornos tonales y de la duración

En los siguientes apartados se muestran los resultados de la variedad entonativa a través de los gráficos de los contornos melódicos de las oraciones analizadas (3.1-3.5) y de los resultados de la duración vocálica en esas oraciones (3.6-3.9). Todos los ejemplos que se aportan corresponden a voces masculinas.

4.1 Contornos entonativos del catalán de Fraga y Mequinenza

Se muestra un ejemplo del contorno melódico para cada tipo de oraciones: declarativas, interrogativas sin la partícula *que* e interrogativas con la partícula átona *que*. En la parte superior derecha de los gráficos se incluye el patrón tonal de cada frase.

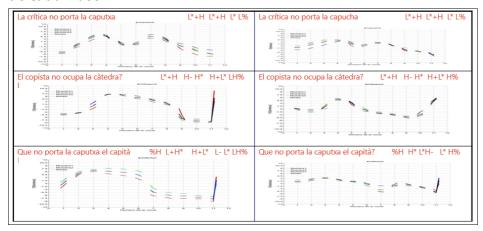


Figura 2. Contornos entonativos de las tres modalidades en Fraga (izquierda) y Mequinenza (derecha)

En las frases declarativas de ambas localidades de la Franja el contorno entonativo tiene un inicio bajo y presenta dos picos; el primero de ellos corresponde al primer acento (a la postónica cuando el sujeto es esdrújulo o llano, o a la tónica cuando es agudo), y el segundo a la postónica del verbo. El descenso final de F0, que corresponde al tonema de la frase, siempre es bajo para este tipo de modalidad.

Las interrogativas sin la partícula *que* presentan en ambos casos una subida inicial hasta la postónica del sujeto, con desacentuación en Fraga hasta la subida final muy pronunciada, mientras que en Mequinenza el final, también ascendente, no es tan brusco.

En los dos contornos de las interrogativas sin la partícula *que* puede hablarse de diferencias más claras: un inicio más alto en Mequinenza que empieza a descender tras la primera tónica, descenso que llega hasta la tónica del objeto y finaliza con una pretónica alta pero descendente, con final ascendente. En la melodía de Fraga el primer pico corresponde a la tónica del verbo, el descenso a la tónica del sujeto, mientras que la tónica final presenta una subida considerable.

4.2 Comparación con los contornos entonativos del catalán de Lleida, Barcelona y Tortosa

En los gráficos de la figura 3 se observa cómo los contornos de las modalidades declarativa e interrogativa sin la partícula que muestran una clara semejanza en los tres puntos de encuesta de Lleida, Barcelona y Tortosa, mientras que en las interrogativas con la partícula *que* los contornos difieren entre ellos: final bajo en Lleida y Barcelona y alto en Tortosa.

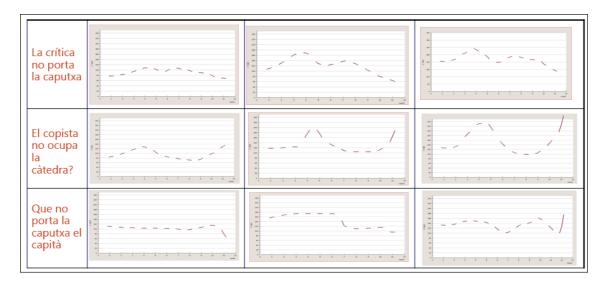


Figura 3. Contornos melódicos de Lleida (izquierda), Barcelona (centro) y Tortosa (derecha)

4.3 Los patrones tonales en catalán

La transcripción de los patrones tonales que se muestran a continuación se ha realizado según el modelo Metrico-Autosegmental (Pierrehumbert, 1980), Cat_ToBI (Prieto *et al.*, 2009) y Sp_ToBI (Prieto & Roseano, 2010). Se muestran las tres modalidades analizadas (figuras 3, 4 y 5).

4.3.1 Declarativas

SN1	sv	SN2	F. FINAL	
L*+H	L*+H	L*	L%	FRAGA
L*+H	L*+H	L*	L%	MEQUINENZA
L*+H	L*+H	L*	L%	LÉRIDA
L+H*	L+H*	H+L*	L%	TORTOSA
L+H*	desac	L*	L%	BARCELONA

Figura 4. Patrones tonales de frases declarativas

Al igual que se ha visto en los gráficos de las melodías, en los dos puntos de la Franja las declarativas presentan el mismo patrón tonal y con los otras localidades comparten el inicio bajo y el tonema bajo. Hay coincidencia con Lleida, y las diferencias con Tortosa se dan especialmente en los dos primeros acentos, y con Barcelona en la desacentuación en el sintagma verbal (Carrera *et al.*, 2010; Fernández Planas *et al.*, 2013, Romera *et al.*, en prensa).

4.3.2 Interrogativas sin la partícula que

SN1	SV	SN2	F. FINAL	
L*+H	H- H*	H+L*	LH%	FRAGA
L*+H	desac	L*	Н%	MEQUINENZA
L*+H	desac	L*	Н%	LÉRIDA
L+H*	desac	L+H*	Н%	TORTOSA
L*+H	desac	H+L*	Н%	BARCELONA

Figura 5. Patrones tonales de frases interrogativas sin la partícula que

Esta modalidad interrogativa se caracteriza por un inicio bajo y un final alto que es complejo en Fraga. La diferencia con Mequinenza la marca sobre todo la desacentuación del sintagma verbal, desacentuación que se da también en el resto de zonas.

4.3.3 Interrogativas iniciadas con la partícula átona que

F.INICIA	AL SV	SN2	SN1	F. FINAL	
%H	L+H*	H+L*	L- L*	LH%	FRAGA
%H	H*	L* H-	L*	Н%	MEQUINENZA
%H	H*	H*	H+L*	L%	LÉRIDA
%H	H+L*	Desac	L+H*	H%	TORTOSA
%H	L+H*	Desac	L*	L%	BARCELONA

Figura 6. Patrones tonales de frases interrogativas con la partícula átona que

Esta modalidad de interrogativa del catalán, con la partícula átona *que* inicial conlleva además la inversión del sujeto, por lo que su estructura difiere de las otras dos vistas hasta ahora. Este tipo de interrogativas es el que marca más la diferencia entre los distintos dialectos. Se caracteriza, como reflejan claramente los patrones, por el inicio alto; a partir de ahí las diferencias son muchas. Los dos puntos examinados de la Franja difieren en su realización, como ya vimos en los contornos entonativos (figura 2). El primer acento tonal de Fraga es bajo mientras que los patrones de los otros lugares lo presentan alto, excepto el de Barcelona. El tonema es bajo menos en Tortosa aunque el tono de frontera final presenta diferentes realizaciones: complejo en Fraga (bitonal), alto en Tortosa y Mequinenza y bajo en Lleida y Barcelona.

De acuerdo con los patrones tonales mostrados, la modalidad declarativa presenta semejanzas con las variedades de Lleida, Tortosa y Barcelona; la modalidad interrogativa sin la partícula *que* se caracteriza por las diferencias en la entonación de Fraga, mientras que la de Mequinenza es más semejante a la de Lleida. La modalidad interrogativa iniciada con la partícula átona *que* es la que presenta mayores diferencias entre todas las variedades examinadas.

4.4 Comparación con el castellano de zonas bilingües y zonas monolingües

El hecho de que la Franja sea una zona bilingüe nos permite comparar su prosodia con la del castellano; para ellos hemos tomada otras zonas también bilingües, analizadas en AMPER-CAT, Lleida y Barcelona; y otras monolingües, Madrid y Zaragoza.

En las frases declarativas (primera fila de la figura 7) los dos picos, que corresponden al sintagma nominal y al sintagma verbal son más pronunciados en Zaragoza. La modalidad de las interrogativas con la partícula átona *que* se ha incluido en Barcelona y Lleida ya que los hablantes de castellano en zonas bilingües también la tienen asumida en su prosodia.

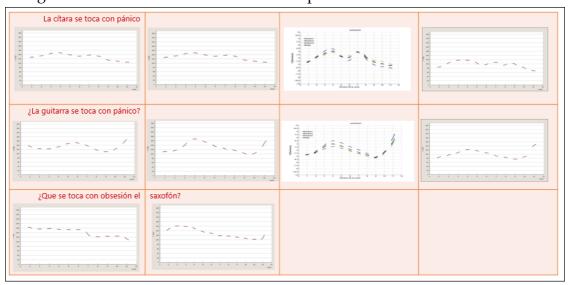


Figura 7. Contornos entonativos (de derecha a izquierda) de Barcelona, Lleida, Zaragoza y Madrid

4.5 Los patrones tonales en castellano

4.5.1 Declarativas

El patrón tonal de la zona de la Franja, Mequinenza, coincide con el de las variedades de Lleida, y uno de los dos patrones encontrados en Zaragoza. Las semejanzas entonativas de esta modalidad son su inicio bajo y su final también descendente, con un tonema bajo en la mayoría de los lugares.

SN1	SV	SN2	F. FINAL	
L*+H	L*+H	H+L*	L%	BARCELONA
L*+H	L*+H	L*	L%	LÉRIDA
L*+H	L*+H	L*	L%	MEQUINENZA
a) L*+H	L*+H	L*	L%	74046074
b) H*	L*+H	L*	L%	ZARAGOZA
L*+H	H*+L	L*	L%	MADRID

Figura 8. Patrones tonales de frases declarativas

4.5.2 Interrogativas

En las interrogativas absolutas (es la única modalidad que comparamos aquí, hemos omitido la de las interrogativas iniciada con la partícula *que*) también encontramos semejanzas en las zonas más próximas a la Franja. Mequinenza, Zaragoza y uno de los patrones de Lleida presentan el mismo patrón tonal. Barcelona y Madrid difieren de estas variedades en el tonema.

SN1	SV	SN2	F. FINAL	
L*+H	H*+L	L+H*	H%	BARCELONA
	Desac Desac			LÉRIDA
L*+H	Desac	L*	H%	MEQUINENZA
L*+H	Desacc	L*	H%	ZARAGOZA
L*+H	H*+L	H*	H%	MADRID

Figura 9. Patrones tonales de frases interrogativas

4.6 La duración de las vocales

La duración de las vocales es una de las características prosódicas presente en la diferenciación de las variedades geoprosódicas, de ahí que nos interese conocer si funciona como otro parámetro diferenciador entre las variedades examinadas. Para que una vocal se considere perceptivamente larga en relación con otras vocales adyacentes, su duranción ha de ser un 36 % superior a ellas. En esta cifra, el 36 %, sitúan Pamies y Fernández Planas (2006) el umbral de la duración.

4.6.1 La duración en Fraga y Mequinenza en catalán

En los gráficos que siguen (figura 10), correspondientes a las frases declarativas de las dos zonas de la Franja, Fraga arriba y Mequinenza abajo, se recoge la duración de las vocales de esas frases en las que se presentan las tres posibilidades acentuales finales., aguda, llana y esdrújula. Cada barra corresponde a cada una de las vocales de la frase. La estrellita encima de la barra indica la vocal con mayor duración que en estos ejemplos de la Franja coincide con la vocal tónica final de la frase; además cuando la vocal final es aguda la diferencia de duración de la vocal tónica es bastante mayor que en los otros dos casos.



Figura 10. Gráficos de duración en Fraga (arriba) y Mequinenza (abajo)

4.6.2 La duración en Lleida, Barcelona y Tortosa (catalán)

Aquí presentamos la duración con ejemplos de las tres modalidades entonativas. De una forma mayoritaria la mayor duración correponde a la vocal final de la frase, si bien esto no ocurre en todos los casos como se aprecia en las realizaciones de Lleida (que son las de la primera fila; en la fila intermedia las de Barcelona y en la última las de Tortosa).



Figura 11. duración en Lleida (L), Barcelona (B) y Tortosa (T) (catalán)

4.6.3 La duración en castellano

La comparación de la duración en castellano se ha realizado tomando como muestra los resultados de Mequinenza, lugar representativo de la Franja, y se ha contrastado con los de dos zonas bilingües, Lleida y Barcelona, y con los de otras dos zonas monolingües, Zaragoza y Madrid.

a) Zonas bilingües: Mequinenza, Lleida y Barcelona

Como también ocurría en catalán, en las realizaciones de castellano de Mequinenza las vocales tónicas finales son las que presentan mayor duración, mientras que en Lleida y Barcelona es la vocal final la más larga de toda la frase.



Figura 12. Gráficos de duración en Mequinenza (M), Lleida (L) y Barcelona (B)

b) Zonas monolingües: Zaragoza y Madrid

La duración de las vocales en Zaragoza (segunda fila) es mayor en casi todas las tónicas finales (aunque no se cumple en el ejemplo de interrogativa que se muestra en la figura 13), con lo que se asemeja claramente a la duración de Mequinenza (primera fila), mientras que en los gráficos de Madrid (tercera fila) observamos una mayor duración no solo en la tónica final sino también en la tónica del primer sintagma nominal.



Figura 13. Gráficos de duración de Mequinenza (primera fila), Zaragoza (segunda fila) y Madrid (tercera fila)

Las vocales tónicas finales de las frases de los informantes de la Franja, tanto en catalán como en castellano son las que presentan una mayor duración. Esto las asemeja con las realizaciones de Zaragoza y en cierta medida con Madrid; sin embargo en Lleida, Tortosa y Barcelona no hay coincidencia entre la vocal tónica final y una mayor duración.

5 Conclusiones

El análisis comparativo de los contornos que caracterizan el catalán de la Franja permite destacar que esta variedad comparte rasgos entonativos tanto con el catalán y el castellano de Lleida como con la zona aragonesa de Zaragoza. Se ha constatado que el final de las declarativas de Fraga presenta cierta semejanza con las de Zaragoza y que el contorno de las interrogativas de la Franja es más parecido al de las de Lleida

En relación con la duración de las vocales las realizaciones de la Franja presentan más semejanza con las de Zaragoza que con las realizaciones del resto de lugares.

La conclusión a la que es posible llegar es que las diferencias entonativas y de duración observadas, tanto en el castellano como en el catalán de la Franja pueden configurar a esta zona como una variedad de transición prosódica entre la zona catalana mediterránea y la zona continental aragonesa.

6 Bibliografía

- AMPERCAT:
 - http://www.ub.edu/labfon/amper/cast.index_ampercat.html
- Borràs-Comes, J. (2013). Català nordoccidental. En P. Prieto, & T. Cabré (coords.) (2013). L'entonació dels dialectes catalans. Barcelona: Publicacions de l'Abadia de Montserrat.
- ASOLC (2007) La situació sociolingüistica als territoris de parla catalana al inici del segle XXI. En http://www.ub.edu/slc/socio/situacioactualcatala.pdf
- Carrera Sabaté, J., Fernández Planas, A. M.ª, & Martínez Celdrán E. (2010). Declaratives i interrogatives absolutes del català en l'Atles Multimèdia de Prosòdia del l'Espai Romànic. *Caplleta*, 49, 133-167
- Castañer Martín, R. M., González Olivera, M. P., & Simón Casas, J. (2005). Aproximación al estudio de la entonación aragonesa. Estudios de Fonética Experimental, XIV, 273-293.
- Contini, M., Jean-Pierre, L., Romano, A., Roullet, S., & De Castro Moutinho, L. (2002). Un projet d'Atlas Multimédia Prosodique de l'Espace Roman. En B. Bel & I. Marlien (eds.), *Speech Prosody*. Aix-en-Provence (11-13 avril 2002), 227-230.
- Estebas Vilaplana, E., & Prieto, P. (2009): La notación prosódica en español. Una revisión del Sp_ToBI, Estudios de Fonética Experimental, XVIII, 263-283.
- **EUROMOSAIC:**
 - http://www.uoc.edu/euromosaic/web/document/catala/ca/i4/i4.htm
- Fernández Planas, A. M.ª, Carrera Sabaté, J., Román Montes de Oca, D., & Martínez Celdrán, E. (2006). Declarativas e interrogativas en Tortosa y Lleida. Comparación de su entonación. Estudios de Fonética Experimental, XV, 165-209.
- Fernández Planas, A. M.ª (2005). Aspectos generales acerca del proyecto internacional AMPER en España, en Estudios de Fonética Experimental, XIV, pp. 13-27.
- Fernández Planas, A. M., Roseano, P., Dorta, J., & Martínez Celdrán, E. (2013). ¿Hay continuidad prosódica en diferentes puntos de la Romania? El caso de algunas interrogativas. Actes del 26è Congrés Internacional de

- Lingüística i Filologia Romàniques (València, 2010). Berlin: Walter de Gruyter, vol. I, 588-599.
- Galan J., & Moret H. (1995). Estudi Descriptiu de la Llengua de Fraga. Fraga: Institut d'Estudis del Baix Cinca I.E.A.
- Giral Latorre, J. (2012). El catalán de la Franja de Aragón. Una aproximación historico- lingüística. *Archivo de Filología Aragonesa*, 68, 39-74.
- Hualde, J. I. (2003). El modelo métrico y autosegmental. En P. Prieto (coord.), *Teorías de la entonación*. Barcelona: Ariel, 155-184.
- López Bobo, M.ª J., Muñiz Cachón, C., Díaz Gómez, L., Corral Blanco, N., Brezmes Alonso, D., & Alvarellos Pedrero, M. (2007). Análisis y representación de la entonación. Replanteamiento metodológico en el marco del proyecto AMPER. En J. Dorta, & B. Fernández (eds.), La prosodia en el ámbito lingüístico románico. Madrid: La Página Ediciones, S.L. Universidad, 17-34
- Martínez Celdrán, E., & Fernández Planas, A. M.ª, (coord.) (2003-2013). AMPER-CAT Atlas Multimedia de la Prosodia del Espacio Románico. En http://stel.ub.edu/labfon/amper/in dex_ampercat_cat.html
- Martínez Celdrán, E., Fernández Planas, A. M.a, Carrera Sabaté, J. (2005a). Diferèrencies dialectals del català a partir de les oracions interrogatives absolutes amb "que". Estudios de Fonética Experimental, XIV, 327-353.
- Pamies, A., Fernández Planas, A. M.ª, Martínez Celdrán, E., Ortega Escandell, A., & Amorós Céspedes, M. C. (2002). Umbrales tonales en español peninsular. *Actas del II CFE*, 272-278.
- Pierrehumbert, J. (1980). The Phonology and Phonetics of English Intonation. Tesis Doctoral, MIT.
- Prieto, P., Lourdes Aguilar, Mascaró, I., Torres-Tamarit, F., & Vanrell M. del M. (2009). L'etiquetatge prosòdic Cat_ToBI. Estudios de Fonética Experimental, XVIII, 287-309.
- Prieto, P. & Roseano, P. (eds.) (2010).

 <u>Transcription of Intonation of the Spanish Language</u>. München: Lincom Europa.
- Prieto, P., & Cabré, T. (2007-2012) Atlas interactiu de l'entonació del català. En

- http://prosodia.upf.edu/atlesentona
- Prieto, P. & Cabré, T. (coords.) (2013). L'entonació dels dialectes catalans. Barcelona: Publicacions de l'Abadia de Montserrat.
- Ramírez Verdugo, M.ª D. (2005). Aproximación a la prosodia del habla de Madrid. *Estudios de Fonética Experimental*, XIV, 309-325.
- Romera Barrios, L., Salcioli, V., Fernández Planas, A. M.ª, Carrera, J., & Román Montes de Oca, D. (2008). The prosody of simple sentences in the Spanish of Barcelona, a Spanish-Catalan bilingual context. En L. Colantoni & J. Steele (eds.), Selected Proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology. Somerville: Cascadilla Press, 167-181.
- Romera, L., Fernández Planas, A. M.ª, García-Elvira, W., Roseano, P., Carrera, J., Ventayol A., & Martínez Celdrán, E. (en prensa). Los esquemas entonativos del fragatino: descripción y comparación con variedades románicas próximas. *Actas del XI CILG*. Pamplona, mayo de 2014.

Diferencias perceptivas entre los acentos tonales prenucleares en catalán

ROSEANO, PAOLO; FERNÁNDEZ PLANAS, ANA MA.; ELVIRA-GARCÍA, WENDY; CERDÀ MASSÓ, RAMÓN; MARTÍNEZ CELDRÁN, EUGENIO

1 Introducción

El inventario entonativo del catalán comprende, entre otros elementos, seis acentos bitonales (Prieto *et al.*, 2009), la mayoría de los cuales ya han sido caracterizados acústicamente y sometidos a pruebas de percepción. Por ejemplo, Vanrell (2011) ha concentrado sus estudios en la descripción acústica y la validación perceptiva de los acentos descendentes H+L* y ¡H+L*, mientras que Borràs *et al.* (2014) han caracterizado acústicamente y comprobado perceptivamente los acentos ascendentes con pico en la tónica (L+H* y L+¡H*). Para los acentos con pico en la postónica (L+>H* y L*+H), mientras que su caracterización fonética ya ha sido efectuada, su validación perceptiva seguía pendiente. Estudios anteriores (Roseano *et al.*, en prensa) han destacado la existencia de diferencias acústicas significativas entre los acentos de tipo L+>H*, que caracterizan el prenúcleo de las frases declarativas neutras, y los de tipo L*+H, que se encuentran en el prenúcleo de las interrogativas absolutas informativas. Más concretamente, los acentos de tipo L+>H* y L*+H difieren en cuanto a cuatro parámetros, que se resumen en la tabla 1.

	L+>H*	L*+H
Representación esquemática		
Rango del movimiento tonal	alrededor de 6 st entre valle y	alrededor de 10 st entre valle y
	pico	pico
Alineación diana tonal baja	alineada con el principio de la	alineada con el final de la
	sílaba tónica	sílaba tónica
Alineación diana tonal alta ¹	en el centro de la segunda sílaba	al final de la segunda sílaba
	postónica	postónica
Movimiento de F0 en la tónica	subida significativa de F0 en la	ninguna subida significativa
	sílaba tónica	de F0 en la sílaba tónica

Tabla 1. Representaciones esquemáticas y características acústicas y de L+>H* y L*+H

Este artículo, de acuerdo con la línea de trabajo abierta por Martínez Celdrán & Fernández Planas (2010), pretende llevar a cabo un conjunto de test de percepción con el objetivo de contestar cuatro preguntas: ¿las diferencias acústicas entre L*+H y L+>H* se perciben?; en caso afirmativo, ¿se perciben todas?; si no se perciben todas, ¿cuáles sí y cuáles no?; si se perciben todas, ¿hay unas más importantes que otras?

2 Metodología

Mediante Praat, se han llevado a cabo 8 pruebas de percepción con 50 jueces, todos hablantes nativos de catalán central que participaron de manera voluntaria y no retribuida. Las dos frases utilizadas para crear los estímulos son la declarativa neutra *La Marina mesurava pantalons* 'Marina medía pantalones' y la interrogativa absoluta informativa '*La Marina mesurava pantalons*? '¿Marina medía pantalones?'. Se ha optado por utilizar una frase con sujeto llano porque, tal y como se ha recordado en el apartado anterior, en las palabras paroxítonas como *Marina* se neutralizan las pequeñas diferencias de alineación del pico entre los acentos L+>H* y L*+H. Esto permite concentrar la atención en un conjunto más reducido de variables, que son: 1) la posición del primer valle tonal, 2) la amplitud del movimiento tonal en la sílaba tónica y 3) la amplitud del movimiento tonal entre valle y pico.

Mientras para el test 1 se han utilizado estímulos de varios hablantes con características distintas (v. apartado 3.1), para los experimentos a partir del segundo se han usado las frases del informante cuyos enunciados presentaban valores acústicos que se ajustaban de manera casi perfecta a la media de los valores medios de los datos de producción que se han presentado en la tabla 1.

3 Resultados

En esta sección se van a presentar los resultados de los ocho test, agrupados en tres bloques según las características y las finalidades de las pruebas.

3.1 Primer bloque de test

El primer bloque comprende tres pruebas en las que se han utilizado estímulos en los cuales no se había manipulado el F0 y que tenían el objetivo de

¹ Tal y como se describe en Roseano et al. (en prensa), en el caso de las palabras llanas esta diferencia de alineación queda prácticamente neutralizada.

comprobar si se percibe la diferencia entre el primer acento tonal de las frases declarativas y las interrogativas SVO, es decir entre L+>H* y L*+H.

3.1.1 Primer test

Los estímulos utilizados en el primer test fueron seis repeticiones del fragmento de frase "La Marina me-" pronunciado por 5 sujetos diferentes (4 mujeres y 1 hombre), en 2 modalidades (declarativa e interrogativa). Los estímulos en cuestión se presentaron en orden aleatorio y alternados con un número igual de distractores. La presencia de distractores y voces diferentes hacía que este experimento no fuera fácil para los jueces. Las respuestas obtenidas han sido 3 000 (50 jueces x 5 sujetos x 6 frases x 2 modalidades).

Para cada estímulo, los jueces tenían que decidir si el fragmento de frase que escuchaban correspondía al comienzo de la frase declarativa *La Marina mesurava pantalons* o de la frase interrogativa *La Marina mesurava pantalons?* Los resultados son los que se presentan en la figura 1.

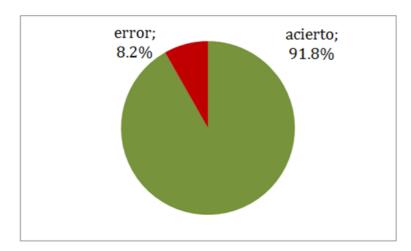


Figura 1. Resultados del primer test de percepción

Tal como se puede observar en el gráfico, hay un 91,8 % de identificaciones correctas de la modalidad oracional, lo que significa que, en general, las diferencias acústicas entre el primer acento de las interrogativas (L*+H) y las declarativas (L+>H*) se perciben. Además, cabe destacar que todos los jueces han alcanzado un nivel de acierto muy por encima del azar; por tanto, en los análisis siguientes no se ha excluido a ninguno de ellos.

3.1.2 Segundo test

En el test 2, que sigue el método del *gating*, los jueces escuchaban fragmentos de diferente longitud del comienzo de la frase *La Marina mesurava pantalons.*/? y tenían que decidir si era el principio de la frase declarativa o de la interrogativa. En total, había 12 estímulos, es decir 6 *gates* de cada modalidad (v. figura 2). Los estímulos se presentaron dos veces, alternados con un número igual de distractores. Las respuestas obtenidas fueron 1 200 (50 jueces x 6 frases x 2 modalidades x 2 repeticiones).

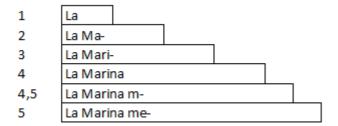


Figura 2. Gates del segundo test de percepción

Los resultados de la prueba (figura 3) permiten destacar que para los *gates* hasta la tónica (incluida) las respuestas son al azar². Para los *gates* a partir de la primera postónica, el porcentaje de aciertos es significativamente más alto que el nivel del azar; los resultados mejoran ulteriormente en las postónicas siguientes.

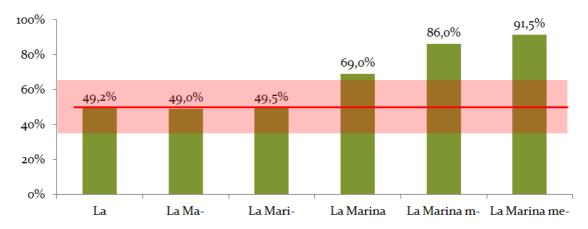


Figura 3. Resultados del segundo test de percepción (porcentaje de aciertos)

3.1.3 Tercer test

El test 3 también sigue el método del *gating*, pero con una diferencia: los jueces escuchaban fragmentos de longitud diferente del comienzo de la frase interrogativa y de la frase declarativa uno tras otro y tenían que decidir cuál de los dos era el comienzo de la frase interrogativa. Había 6 estímulos, es decir 6 pares de fragmentos (figura 4) en los dos órdenes de presentación (decl+int, int+decl). Los estímulos no estaban alternados con distractores. Todo ello hacía que este test resultara previsiblemente más fácil en comparación con el primero. Cada estímulo se presentó dos veces y las respuestas obtenidas fueron 600 (50 jueces x 6 pares de fragmentos x 2 órdenes de presentación).

² En este test y los siguientes el nivel del azar no se concibe exclusivamente como el que corresponde al azar puro (es decir a la distribucion 50-50 %, que corresponde a la línea roja horizontal en la figura), sino que se define como aquella franja (que en las figuras se marca con un rectángulo rojo) que no difiere estadísticamente del nivel del azar puro. El cálculo de la significatividad de las diferencias se ha efectuado mediante el Khi-quadrado y se ha fijado el umbral en p = 0.05.

1	La La
2	La Ma La Ma-
3	La Mari La Mari-
4	La Marina La Marina
4,5	La Marina m La Marina m-
5	La Marina me La Marina me-

Figura 4. Gates y pares de estímulos del tercer test de percepción

Los resultados de la prueba (figura 5), en general, confirman los resultados del test 2: para los *gates* hasta la tónica (incluida), la prueba del χ^2 revela que las respuestas no son significativamente mejores que el nivel de azar, a pesar de que las condiciones experimentales fueran más favorables en este último test. Por otra parte, se confirma (e incluso se mejora) el nivel de acierto en las postónicas.

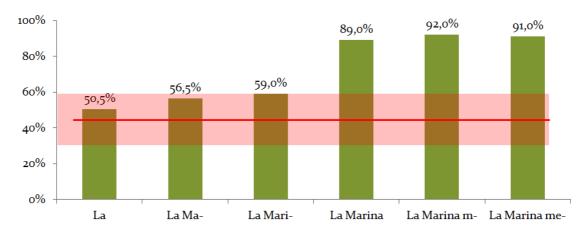


Figura 5. Resultados del segundo tercer de percepción (porcentaje de aciertos)

3.1.4 Discusión de los resultados del primer bloque de test

Los resultados de los primeros tres test, en conjunto, permiten concluir, en primer lugar, que las diferencias entre L*+H y L+>H* se perciben. En segundo lugar, cabe destacar que el nivel de aciertos es significativamente más alto que el nivel del azar cuando los jueces escuchan los estímulos hasta las sílabas postónicas. Por lo tanto, el contenido entonativo de las sílabas postónicas parece tener un papel clave a la hora de diferenciar entre los dos acentos en cuestión.

3.2 Segundo bloque de test

El segundo bloque comprende tres test realizados con estímulos en los que se ha manipulado el F0 y que tienen el objetivo de comprobar si las características tonales de la sílaba tónica y las postónicas vehiculan por separado informaciones sobre la modalidad oracional.

3.2.1 Cuarto test

Para valorar si la información entonativa contenida solo en las sílabas postónicas es suficiente para identificar la modalidad oracional, se presentaron estímulos donde se había sustituido por ruido la sílaba tónica y el ataque de la

postónica (figura 6). Había 2 estímulos (*La Ma_a m-*) y cada uno de ellos se presentó dos veces. Los estímulos se alternaban con el mismo número de distractores. Las respuestas obtenidas fueron 200 (50 jueces x 2 estímulos x 2 repeticiones).

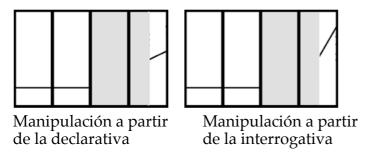


Figura 6. Representación esquemática de los estímulos del cuarto test de percepción

Los resultados revelan que el nivel de acierto (67,3 %) es significativamente más alto que el nivel del azar (figura 7). Por lo tanto, se puede concluir que la información tonal contenida en las sílabas postónicas es suficiente para percibir correctamente la modalidad oracional. No se han observado diferencias significativas entre los estímulos generados a partir de la frase interrogativa y a partir de la declarativa.

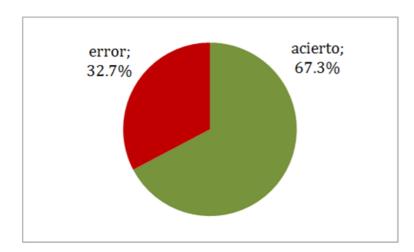
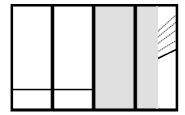


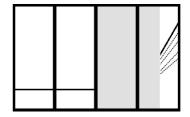
Figura 7. Resultados del cuarto test de percepción

3.2.2 Quinto test

Si nos concentramos en la información tonal contenida en las sílabas postónicas (que en el apartado 3.1 hemos dicho que puede ser suficiente para identificar la modalidad oracional), podemos intentar averiguar si la percepción es categorial y si hay un umbral en el cual se produce el cambio de percepción. Con este fin, se presentaron estímulos donde se había sustituido la sílaba tónica y el ataque de la postónica por ruido blanco y se había manipulado el rango del movimiento en dos direcciones y en 5 pasos de igual amplitud en st (figura 8).

En total había 10 estímulos (*La Ma__a m-*) y cada uno de ellos se presentó dos veces. Los estímulos se alternaban con el mismo número de distractores y las respuestas obtenidas fueron 1 000 (50 jueces x 5 estímulos x 2 modalidades x 2 repeticiones).





Manipulación a partir de la declarativa

Manipulación a partir de la interrogativa

Figura 8. Representación esquemática de los estímulos del quinto test de percepción

Los resultados del test (figura 9) permiten observar que hay un cambio de percepción de modalidad que tiene lugar entre el 2.º y el 3.º paso, es decir alrededor, aproximadamente, de los 7,8 st. También se observa que, cuando el valor del pico está entre 7,3 st y 8,3 st, las respuestas se sitúan en el nivel del azar (es decir, dentro del rectángulo rojo que aparece en la figura 9). No se han observado diferencias significativas entre los estímulos generados a partir de la frase interrogativa o a partir de la declarativa.

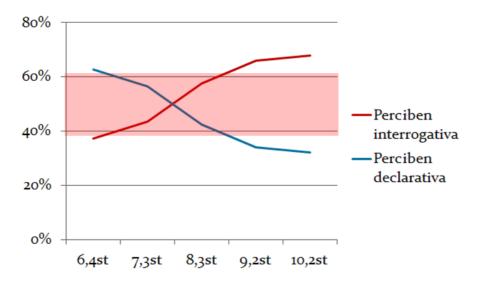
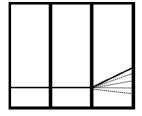
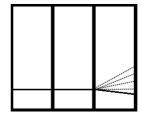


Figura 9. Resultados del cuarto test de percepción

3.2.3 Sexto test

Si nos concentramos en la información tonal contenida en la sílaba tónica (que en principio, de acuerdo con los resultados del test 2, no parecería tener importancia), podemos intentar confirmar si esta se percibe o no, averiguar si la percepción es categorial y descubrir si hay un umbral donde se produce el cambio de percepción. Con este fin, se presentaron estímulos en los que se había manipulado el rango del movimiento tonal en la sílaba tónica en dos direcciones y en 5 pasos calculados en st (figura 10). En total había 10 estímulos (*La Mari-*) y cada uno de ellos se presentó dos veces. Los estímulos se alternaban con el mismo número de distractores y las respuestas obtenidas fueron 1 000 (50 jueces x 5 estímulos x 2 modalidades x 2 repeticiones).





Manipulación a partir de la declarativa

Manipulación a partir de la interrogativa

Figura 10. Representación esquemática de los estímulos del sexto test de percepción

Los resultados (figura 11) revelan que, si no hay ninguna subida de F0 en la tónica (es decir, hasta el paso de +0,85 st excluido), la percepción no supera el nivel de azar, es decir que la respuesta cae en el rectángulo rojo de la figura 11. Por lo contrario, si en la tónica hay una subida de F0 >0,85 st, la frase se percibe como declarativa. La diferencia respecto a los resultados de los test 2 y 3, donde la información tonal de la tónica no se percibía, se puede explicar a partir de las diferentes condiciones experimentales, más favorables en este test que en los anteriores. Tal como era de esperar, no se han observado diferencias significativas entre los estímulos generados a partir de la frase interrogativa o a partir de la declarativa.

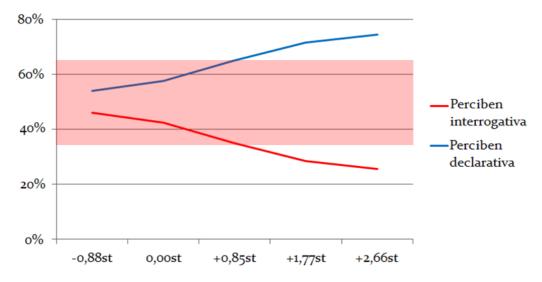


Figura 11. Resultados del sexto test de percepción

3.2.4 Discusión de los resultados del segundo bloque de test

Los resultados de los test 5 y 6 confirman que los movimientos de F0 en las sílabas postónicas por sí solos son suficientes para percibir correctamente la modalidad oracional (el cambio de percepción se produce alrededor de los 7,8 st). En cuanto a los movimientos de F0 en la sílaba tónica, una subida de F0 implica una interpretación como declarativa. Una bajada o un mantenimiento, por el contrario, no son interpretables, es decir, que dan lugar a respuestas al azar.

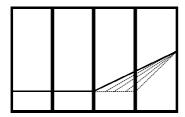
3.3 Tercer bloque de test

Las dos pruebas del tercer bloque utilizan estímulos manipulados y contradictorios, ya que combinan características de las declarativas y de las

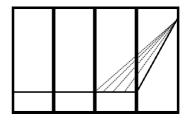
interrogativas con el fin de comprobar si, cuando la sílaba tónica y las postónicas vehiculan informaciones tonales no congruentes, una de las dos prevalece.

3.3.1 Séptimo test

Para comprobar qué consecuencias tiene la presencia de informaciones tonales contradictorias vehiculadas por la altura del pico y la alineación del valle, se manipuló la alineación del valle en 5 pasos equidistantes, numerados de 1 a 5, y en dos direcciones (figura 12). Había 10 estímulos (*La Marina m-*) y cada uno de ellos se presentó dos veces. Los estímulos estaban alternados con un igual número de distractores y las respuestas obtenidas fueron 1 000.



Manipulación a partir de la declarativa



Manipulación a partir de la interrogativa

Figura 12. Representación esquemática de los estímulos del séptimo test de percepción

Para la interpretación de los resultados del test, es oportuno comentar separadamente las respuestas obtenidas a partir de los estímulos declarativos (figura 13, imagen de la izquierda) e interrogativos (figura 13, imagen de la derecha).

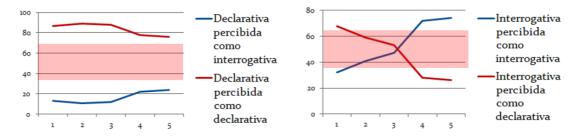
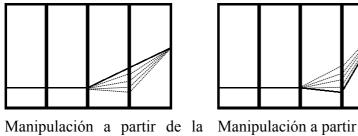


Figura 13. Resultados del séptimo test de percepción

Los datos relativos a los estímulos declarativos revelan que no hay cambio de percepción: las frases siempre se perciben como declarativas. Por lo tanto, independientemente de la alineación del valle, un pico tonal bajo solamente es compatible con la interpretación como declarativa (en este sentido, un tono H sería *marcado*). Si se analizan las respuestas relativas a los estímulos interrogativos, se observa que hay cambio de percepción alrededor del tercer paso, es decir, el centro de la sílaba. En otras palabras, un pico alto es compatible tanto con la interpretación como declarativa como con la interpretación como declarativa como declarativa (en este sentido, un tono ¡H sería *no marcado*). En resumidas cuentas, la presencia de un pico bajo prevalece sobre la alineación del valle, que a su vez prevalece sobre la presencia de un pico alto.

3.3.2 Octavo test

Para comprobar qué consecuencias tiene la presencia de informaciones tonales contradictorias vehiculadas por la altura del pico y el movimiento tonal en la sílaba tónica, se manipuló el movimiento tonal en la tónica en 5 pasos calculados en st y en dos direcciones (figura 14). Había 10 estímulos (*La Marina m-*) y cada uno de ellos se presentó dos veces. Los estímulos se alternaban con el mismo número de distractores y las respuestas obtenidas fueron 1 000.



Manipulación a partir de la Mani declarativa de la

Manipulación a partir de la interrogativa

Figura 14. Representación esquemática de los estímulos del octavo test de percepción

Para la interpretación de los resultados del octavo test, es oportuno comentar separadamente las respuestas obtenidas a partir de los estímulos declarativos (figura 15, imagen de la izquierda) e interrogativos (figura 15, imagen de la derecha).

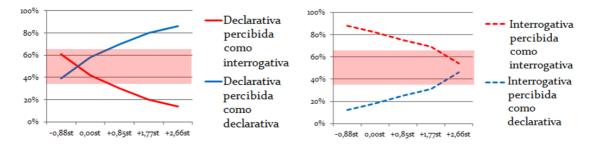


Figura 15. Resultados del octavo test de percepción

Los datos relativos a los estímulos declarativos revelan que no hay un cambio real de percepción, ya que las líneas se cruzan dentro de la franja del azar. Lo que se observa es que si el pico es bajo y la F0 en la tónica no sube (es decir, si el estímulo es congruente), se percibe como interrogativa. Por lo contrario, si el pico es bajo y F0 en la tónica sube (es decir, si el estímulo no es congruente), los jueces responden al azar. Si se analizan las respuestas relativas a los estímulos interrogativos, se observa que no hay ningún cambio de percepción. Se puede notar como, si el pico es alto y la subida en la tónica es reducida o inexistente (es decir, si el estímulo es congruente), el estímulo se percibe como interrogativo. Sin embargo, si el pico es alto y la subida en la tónica es marcada (es decir, si el estímulo no es congruente), las respuestas se dan al azar. En síntesis, si los estímulos son congruentes, se perciben correctamente; si no son congruentes, las respuestas son al azar.

3.3.3 Discusión de los resultados del tercer bloque de test

Las respuestas de los test 7 y 8, en su conjunto, revelan, en primer lugar, que si en un fragmento hay información entonativa contradictoria, la amplitud del pico y del movimiento tonal en la tónica se neutralizan. En segundo lugar, permiten concluir que la presencia de un pico bajo prevalece sobre la alineación valle, que a su vez prevalece sobre la presencia de un pico alto.

4 Conclusiones

Para presentar las conclusiones generales del conjunto de test de percepción, es oportuno volver a las preguntas que se han formulado en la introducción de este trabajo. En cuanto a la primera de ellas, se ha visto que las diferencias acústicas entre los acentos de tipo L*+H y L+>H* no solo son acústica y estadísticamente significativas como habían postulado Roseano *et al.* (en prensa), sino que también lo son desde el punto de vista perceptivo. En relación con la segunda y la tercera de las preguntas iniciales, se confirma que todas las diferencias que se han considerado (alineación del valle tonal, rango del movimiento tonal en la tónica, amplitud del movimiento tonal entre el valle y el pico) se perciben. Finalmente, en cuanto a la cuarta pregunta, es decir, si algunos de los parámetros acústicos considerados son perceptivamente más relevantes que otros, los resultados permiten afirmar que el rango del pico parece más importante que la posición del valle.

Los test que se han presentado conllevan unas reflexiones sobre la actual caracterización del sistema de transcripción Cat_ToBI (Prieto *et al.*, 2009), en la cual la diferencia entre los dos tipos de acento en cuestión hace referencia exclusivamente a la alineación de la diana tonal baja. Por una parte, en la próxima revisión de sistema de Cat_ToBI sería oportuno reformular las descripciones de los acentos en cuestión de manera que quede adecuadamente reflejada la importancia del rango de los movimientos tonales. Por otra parte, para que la transcripción entonativa sea más transparente (es decir, más fiel a la realidad acústica) y más congruente con los elementos perceptivamente distintivos (es decir, con la fonología), habría que plantearse si no sería más oportuno transcribir los dos acentos como L+>H* y L*+¡H en vez de L+>H* y L*+H.

5 Bibliografía

Borràs-Comes, J., Vanrell, M. M., & Prieto, P. (2014). The role of pitch range in establishing intonational contrasts. Journal of the International Phonetic Association, 44(1), 1-20.

Martínez Celdrán, E. y Fernández Planas, A. M. (2010). ¿Son diferentes las estructuras L*+H y L+>H* perceptivamente? Presentación en el II Congrés sobre entonació del català i Cat_ToBI, Barcelona, Universitat Pompeu Fabra, 9 de 7 de 2010.

Prieto, P., Aguilar, L., Mascaró, I., Torres Tamarit, F. J. & Vanrell, M. M. (2009). L'etiquetatge prosòdic Cat_ToBI. Estudios de Fonética Experimental, 18, 287-309.

Roseano, P., Fernández Planas, A. M.ª, Elvira-García, W., Cerdà Massó, R., & Martínez Celdrán, E. (en prensa). Caracterització acústica dels accents prenuclears de les interrogatives absolutes i les declaratives neutres en català central. Estudios de Fonética Experimental, 23.

Vanrell, M. M. (2011). The phonological relevance of *tonal scaling in the intonational grammar of Catalan*. Tesis doctoral. Barcelona: Universitat Autònoma de Barcelona.

Agradecimientos: Este trabajo se ha realizado con una ayuda del MCI (FFI2012-35998).

La relación de las características del lenguaje corporal y la voz en el discurso mendaz

SÁNCHEZ CONDE, CRISTINA

1 Introducción

No son muchos los estudios sobre gesto y prosodia que tratan sobre la diferencia entre la verdad y la mentira. De ellos, casi ninguno trata sobre la relación entre ambos campos en estos dos tipos de discurso. Algunas investigaciones parecen demostrar que existe alguna característica que los diferencia. Lo que ha supuesto un quebradero de cabeza es descifrar cuál o cuáles son esas características que aún hoy en día resultan poco claras. Autores como DePaulo et al. (2003) han estudiado la relación que parece existir entre gesto y distintos temas como la sintaxis. En su estudio, se buscaron en distintas categorías indicadores de mentira. La investigación concluyó, entre otras cosas, que los enunciados falsos son más cortos, tienen menos detalles y presentan signos de hacer un esfuerzo mental, ya que las mentiras exigen ser pensadas por no ser hechos vividos. Por esa razón, no son tan fluidas. Se ha intentado durante tanto tiempo encontrar la precisa característica inequívoca y distintiva entre la verdad de la mentira, que casi no se ha reparado en definir esa "particularidad" como un conjunto de propiedades. La hipótesis que ha llevado a realizar este trabajo es que no existe una única peculiaridad en el discurso que determine la mentira, sino un conjunto que puede no funcionar por separado, ya que dependen de múltiples factores, por ejemplo, emocionales. En este trabajo se pretenden estudiar las variaciones observadas en nuestro corpus entre los discursos mendaz y veraz en el ámbito prosódico y gestual.

Ya temprano, en los años 70, las investigaciones concluían que el tono de voz cambia durante un enunciado con intención de engaño. En un experimento llevado a cabo por Streeter, Krauss, Geller, Olson & Apple (1977), se concluyó que en aquellos enunciados donde había una mentira, la frecuencia fundamental era más alta que en los enunciados ciertos. Ya en este estudio diferenciaron enunciados falsos en los que hubiese excitación o estrés de las que no. El resultado mostraba que los enunciados dichos en una situación de más estrés provocaba el aumento de la frecuencia fundamental.

Sobre las características de la voz en la mentira también se ha teorizado en diferentes temas. Van Edwards (2013) afirma que es cuestión de sexo, y la frecuencia fundamental se eleva en la mujer, pero baja en el hombre durante este tipo de enunciado. Sin embargo, la mayoría de los estudios afirman que el tono de voz es más alto cuando hay mentira sin especificar sexo. Según Ekman, esto puede deberse a que las personas que mienten tienen miedo de que les pillen. Mediante un experimento, el autor quiso demostrar que el aumento del tono de voz no es señal de mentira, sino de miedo, enfado e incluso excitación. Precisamente el sentimiento de miedo hace que el hablante cometa más errores y produzca más pausas. Por lo tanto es importante tener en cuenta a la hora de analizar el discurso el número de errores y el número y tiempo de las pausas. Sobre las pausas puntualiza que o bien pueden ser muy largas, o más cortas y frecuentes.

Van Edwards también añade dos puntos que difieren en la verdad y en la mentira y que también se tendrán en cuenta en el presente trabajo: Según la autora, en la mentira se utilizan 1/3 más de palabras que en la verdad, y que la velocidad tampoco es igual entre ambos discursos. Apunta que en la mentira la velocidad suele ser mayor o menor y esto ocurre porque la gente que miente trata de procesar lo que están diciendo.

La teoría de que el tono de voz es más alto durante la mentira está bastante arraigada. Esa creencia puede influir en el comportamiento de las personas a la hora de mentir (Villar, Arciuli & Paterson, 2012) En un estudio realizado por Villar, Arciuli & Paterson (2012) se tuvo en cuenta este factor y se experimentó sobre la relación del pensamiento sobre el tono de voz en la mentira y el comportamiento del tono de voz de la persona. Su estudio resulta llamativo a la hora de mostrar los resultados: Los participantes en el experimento que creían que el tono de voz aumentaba en la mentira mostraban una diferencia significantemente más alta en la frecuencia fundamental durante los enunciados mendaces que en los verdaderos. Los informantes cuyo pensamiento era que el tono de voz no aumentaba en la mentira, presentaban, aunque no de manera significante, una frecuencia fundamental un poco más baja en los enunciados falsos. Por último, los participantes que no sabían si el tono de voz aumentaba en la mentira, también mostraron un ligero descenso en la frecuencia fundamental en los enunciados mendaces, aunque tampoco de forma significante.

La mayoría de investigaciones actuales sobre la F0 en la mentira no tienen en cuenta diversos factores que pueden ser determinantes a la hora de hacer un análisis. Además del estrés o ansiedad, existen otras causas que pueden hacer variar los resultados en las investigaciones, como las creencias de cada persona. También el idioma parece ser un factor importante que puede alterar las conclusiones de una investigación: La mayoría de estudios publicados que demuestran que la frecuencia fundamental sí es significativamente más alta durante el discurso mendaz son sobre el inglés. Sin embargo, Spence, Arciuli &

Villar (2012), demuestran que en italiano no hay una diferencia significativa entre el tono de voz durante falsos enunciados y verdaderos.

Estos estudios no vienen a demostrar más que la existencia de contradicciones entre investigaciones. Por lo tanto, no hay aún hoy en día un marcador de mentira que sea completamente fiable.

Para el estudio del lenguaje corporal en general y la mentira en particular, es preciso remarcar que no todo el mundo tiene un comportamiento igual a la hora de expresarse (Ekman, 2009) Algunas personas no comenten "errores" a la hora de mentir que puedan delatarlos. Si bien es cierto que el hecho de que no incurran en ningún signo característico que podamos etiquetar como típico de la mentira, no es prueba en ningún momento de que el discurso que se emita sea veraz. Al mismo tiempo, tampoco un solo gesto que sea frecuente durante la mentira es signo de que un enunciado sea falso. "No hay una única expresión para cada emoción, sino docenas y, para algunas emociones hay cientos de expresiones". (Ekman, 2009: 127) Por esta misma razón es importante destacar el hecho de que no todas las personas se expresan de igual manera. Sin olvidarse de acentuar este hecho, Ekman enumera algunas pistas que son típicas durante la mentira: Numerosas pausas largas y cortas entre palabras, numerosos errores, pocos ilustradores, numerosos manipuladores, signos de miedo, angustia o enfado en las expresiones faciales pese a los sentimientos reales, etc. (Ekman, 2009: 166).

Dentro de los gestos faciales, Ekman reconoce la acción de levantar o bajar las cejas como una de las más frecuentes y que se usan de forma habitual como acompañamiento de un enunciado enfático o acentuado. En otros casos, levantar las cejas también es señal de miedo. Por el otro lado, una persona que miente también puede mostrar este gesto ya estará intranquilo por si lo descubren. Esto viene a demostrar una vez más que no todos los gestos tienen una correlación directa con una única emoción. También demuestra para poder ser lo más categóricos posibles a la hora de interpretar si un enunciado es falso o verdadero, es necesario sumar otras características que avalen dicha hipótesis.

Se dice que los ojos nunca mienten, y Ekman dedica un apartado especial a la actividad de los mismos. Para nuestro estudio no nos centraremos en ellos (aunque sí en la mirada), pero sí es importante subrayar algunos aspectos que nos facilitan el análisis del corpus. Este autor relaciona la dirección de la mirada con ciertas emociones y así una mirada hacia abajo se relaciona con la emoción de la tristeza, vergüenza o culpa. Mirar a lo lejos se relaciona con las dos últimas emociones señaladas y con el asco. Ocultar estas emociones con la mirada es fácil, pues controlar la dirección de la mirada no es ningún problema. Por lo tanto, aunque sea útil, tampoco podemos asegurar que un enunciado es veraz o mendaz solamente por esta característica.

Sobre el tiempo, Ekman explica que es importante determinar la duración del gesto y cuánto tiempo tarda en aparecer. Un gesto muy duradero señala una posible mentira, a menos que el hablante esté en el punto máximo de una experiencia. Aún en este último caso, según el autor, las expresiones raramente son muy largas, en vez de eso, aparecen muchas cortas.

La sonrisa es otro punto relevante a tener en cuenta para nuestro estudio, ya que en el corpus se da con frecuencia en algunos informantes. Ekman define a la sonrisa como un gesto difícil de tratar, pues hay múltiples y con distintos significados. "Las sonrisas reales se relacionan con el músculo cigomático mayor, mientras que las producidas por el miedo se relacionan al risorio que tira de las comisuras de los labios en forma horizontal y hacia las orejas de manera que los labios quedan estirados formando una figura rectangular". (Ekman, 2009: 151) Ekman destaca que el movimiento del músculo risorio

también puede darse en una risa verdadera. Para poder distinguir una sonrisa falsa es necesario fijarse en otras características: La más común es que se vea acompañada por unas cejas levantadas y estiradas.

En relación al resto del cuerpo, Ekman establece dos tipos de gestos característicos: los ilustradores y los manipuladores. Los ilustradores son los que David McNeill clasifica como icónicos y metafóricos (ver apartado "Metodología") que recrean imágenes con el cuerpo. Según Ekman, este tipo de gestos tienen a disminuir considerablemente en los enunciados mendaces, pero son muy difíciles de tratar, pues hay enunciados más dados que otros a realizar este tipo de gestos. Los manipuladores son gestos que conllevan un movimiento en el que una parte del cuerpo "manipula" de alguna forma a otra, como rascarse, tocarse una mano, etc. En estos gestos, existen dos partes: el receptor, que casi siempre es una mano pero puede ser otra parte del cuerpo, o incluso objetos. El manipulador es normalmente una mano aunque puede ser cualquier cosa, como por ejemplo los dientes mordiendo el labio (Ekman, 2009). El autor establece que este tipo de gestos aumenta con la incomodidad, como el de una persona al mentir. También destaca que puede ocurrir lo contrario: Cuando una persona miente intenta no hacer ningún gesto manipulador, pues socialmente se cree que expresar esta clase de gestos es prueba de mentira. Por lo tanto, tampoco es una característica que pueda definirse categóricamente como un signo de la mentira.

Las características que Paul Ekman nos explica como posibles pistas para detectar mentiras nos sirven de gran ayuda en conjunto, pero no por separado. Ya el propio autor reconoce que para un único gesto hay múltiples interpretaciones, que no podemos tomar solamente una característica como prueba determinante de un tipo de enunciado. Es necesario, pues, reunir una serie de características típicas del discurso mendaz para poder concluir de manera más precisa que un enunciado particular es o no mentira. "Una persona que detecte mentiras no debe nunca confiar en una sola prueba de que hay mentira, pues hay debe haber muchas. Las pistas de la cara deben confirmarse con pistas de la voz, palabras o resto del cuerpo, o incluso con otros gestos de la propia cara. Una sola prueba no debe ser interpretada a menos que sea muy repetida, o mejor aún, confirmada por algún otro gesto facial". (Ekman, 2009: 147) Sobre ese principio gira este trabajo: Conocer las características gestuales y de la voz que se relacionan entre sí para poder detectar que un enunciado sea falso.

En la presente investigación se pretende analizar la verdad y la mentira en el campo de la voz y el gesto en combinación para descubrir si existe alguna relación entre ellos y para observar los posibles patrones de coordinación que existan. Se contó para ello con una metodología que permite comparar las características gestuales y prosódicas en la verdad y la mentira y que además permite también buscar la frecuencia de aparición de distintos patrones.

2 Metodología

2.1 Informantes

Para el estudio han participado 4 personas (tres hombres y una mujer) de entre 22 y 25 años. Se les informó vagamente del propósito del experimento sin entrar en detalles para evitar la influencia de su conocimiento sobre el mismo a la hora de crear el corpus.

2.2 Procedimiento

En los estudios más recientes sobre el efecto de la mentira en el tono de voz se demuestra que el tono de voz puede verse afectado en ámbitos de ansiedad o estrés. Por ese motivo se intentó excluir ese factor a la hora de realizar el corpus (Streeter, Krauss, Geller, Olson & Apple, 1977; Pollina, Vakoch & Wurm, 1998; DePaulo, Lindsay, Malone, Muhlenbruck, Charlton & Cooper, 2003; Spence, Arciuli & Villar, 2012; Villar, Arciuli & Paterson, 2012) Para llevar a cabo la exclusión, se propusieron distintas metodologías para la colección de datos. El presente experimento se presentó en dos partes a modo de juego para los informantes y se grabó a cada uno de ellos delante de los demás. Se procedió de esta guisa para evitar incurrir en el aumento de ansiedad, o nervios a la hora de producir falsos enunciados como bien sugieren estudios anteriores. En la primera parte se les mostró diferentes obras de arte pictóricas (una a cada uno) para que procediesen a su descripción incorporando datos ficticios en ellos. Para aumentar la espontaneidad, cada informante tenía que intentar evitar que los otros tres no identificasen qué datos habían falseado. En la segunda parte, se siguió el mismo procedimiento, pero el objeto a describir ya no era una imagen previamente dada, sino una persona conocida por cada informante y desconocida para el resto.

2.3 Análisis

David McNeill (1992) planteó una organización de los gestos muy útil para distinguir clases de gestos, cómo clasificarlos y etiquetarlos. Este estudio se apoya en esta clasificación para transcribir el corpus gestual, y a la vez, sigue las directrices de Roseano, González, Borràs-Comes & Prieto (2013) sobre las aportaciones de McNeill. La metodología que sigue Roseano, González, Borràs-Comes & Prieto (2013) también trabaja con la clasificación de otros autores como Allwood *et al.* o Nonhebel *et al.*, que complementan de forma más detallada la transcripción de los gestos para el análisis de los mismos. Entendemos que cuanto más detallado sea el análisis, más detallados serán los resultados para este y futuros estudios sobre el tema. A continuación, presentamos dicha clasificación y pautas:

Para el etiquetado de gestos manuales se utiliza la codificación descrita en las siguientes tablas:

LH	Mano Izquierda	Α	Mano cerrada
RH	Mano Derecha	В	Mano abierta
2SH	Ambas manos haciendo lo mismo	G	Índice extendido
2DH	Cada mano hace un gesto diferente	X1	Forma no determinada

Tabla 1. Uso específico y forma de la mano (Roseano, González, Borràs-Comes & Prieto, 2013 y McNeill, 1992)

¹ Esta etiqueta no se incluye en el inventario de McNeill. Roseano, González, Borràs-Comes & Prieto (2013) la incorporan a propósito de su investigación. Se decidió integrarlo también para el análisis del corpus de este estudio por su gran utilidad a la hora de etiquetar un gesto manual que resulta dudoso en cuanto a forma.

PTU/FTU	Palma/Dedo hacia arriba
PTD/FTD	Palma/Dedo hacia abajo
PTC/FTC	Palma/Dedo hacia el centro
PAB/FAB	Palma/Dedo de adentro hacia fuera
PTB/FTB	Palma/Dedo de afuera hacia dentro
PAC/FAC	Palma/Dedo del centro hacia la derecha/izquierda

Tabla 2. Orientación de la palma de la mano (P) o del dedo (F) (Roseano, González, Borràs-Comes & Prieto, 2013 y McNeill, 1992)

Roseano, González, Borràs-Comes & Prieto (2013) también incluyen un nuevo apartado en el etiquetado de gestos manuales: movimientos gestuales, que además de localizar el gesto en el espacio, también señala la dirección del movimiento y que también se utiliza en este estudio ya que puede resultar relevante en el análisis.

En cuanto a la función del gesto, se utilizó también la clasificación de McNeill (1992) utilizada en Roseano, González, Borràs-Comes & Prieto (2013)

Golpe	"Golpecitos de la(s) mano(s) hacia arriba y abajo o hacia adelante y atrás calibrándose
	rítmicamente con los picos tonales del discurso" (McNeill, 2010: 4)
Señalamiento	"Dedo índice extendido, aunque casi todas las partes del cuerpo que puedan extenderse o
	agarrar un objeto pueden usarse" (McNeill, 2010: 4)
Icónico	"Representación de aspectos del tópico del discurso que lo acompaña" (De Ruiter, 2000: 285)
	"Estos gestos presentan imágenes de entidades o acciones concretas" (McNeill, 2010: 4)
Metafórico	"Los gestos no se limitan a representaciones de algo concreto. También pueden representar
	contenido abstracto, imaginando lo inimaginable. En un gesto metafórico, el significado
	abstracto se presenta como si tuviese forma y/u ocupase espacio" (McNeill, 2010: 4)

Tabla 3. Función del gesto (Roseano, González, Borràs-Comes & Prieto, 2013; McNeill, 1992 y 2010 y De Ruiter, 2000)

En este trabajo incorporamos una función extra que necesitamos a la hora de analizar el corpus. Lo determinamos "acompañamiento" pues la única función que podría asignársele es la de acompañar al discurso o enfatizarlo.

Para clasificar los gestos, tanto de la cabeza y hombros, como de la boca, seguimos también las pautas que aparecen en Roseano, González, Borràs-Comes & Prieto (2013) y que se basan en Allwood, Cerrato, Dybkjaer, Jokinen, Navaretta & Paggio (2005) para cabeza y hombros y en Nonhebel, Crasborn & Van Der Kooij (2004) para la boca.

Cabeza Cabeza	
Nod (Asentir)	Bajar y levantar la cabeza ligeramente y de forma breve
Headshake (Sacudir la cabeza)	Girar la cabeza de izquierda a derecha o viceversa de forma repetida.
Turn (Girar)	Girar la cabeza a la derecha o a la izquierda y mantenerla en esa posición.
Tilt (Ladear)	Mover la cabeza lateralmente a una posición inclinada
Hombros	
Shrug (Encoger los hombros)	Levantar los hombros
Swing (Balancear)	Mover de un lado a otro de forma
Boca	
Mouth stretched down (Boca estirada hacia abajo)	Los labios están estirados y las comisuras dispuestas hacia abajo
Mouth stretched (Boca estirada)	Los labios están estirados y las comisuras están dispuestas a los lados de la cara
Lower lip forward (Labio inferior hacia adelante)	El labio inferior está adelantado
Corners up (Comisuras alzadas)	Las comisuras están hacia arriba

Tabla 4. Cabeza, hombros y boca (Roseano, González, Borràs-Comes & Prieto, 2013 según Allwood, Cerrato, Dybkjaer, Jokinen, Navaretta & Paggio, 2005)

Para el estudio y análisis de la mirada creamos una codificación propia y adaptada al corpus:

Look Down (LD)	Mirada hacia abajo
Look Away to the Left (LAL)	Desviación de la mirada hacia la izquierda
Look Away to the Right (LAR)	Desviación de la mirada hacia la derecha
Roof (R)	Mirada hacia arriba o al techo
Gesture (G)	Mirada hacia el gesto que se está haciendo
Image (IM)	Mirada hacia la imagen que describen
Camera (CAM)	Mirada hacia la cámara
Audience (Aud)	Mirada hacia la gente presente en la sala

Tabla 5. Mirada

3 Resultados

3.1 Frecuencia fundamental

Calculamos a través de diferentes test T-Student para muestras independientes si había diferencia en la F0 entre los enunciados veraces y mendaces de cada informante. Llamaba la atención que en la mentira, la mayoría de los casos elevaban o bajaban considerablemente la F0 teniendo en cuenta la media de F0 de los enunciados veraces de cada hablante. Por esa razón, se calculó si esos enunciados mendaces que elevaban o bajaban a partir de la media de F0 de los enunciados donde hay verdad de cada hablante. Tres de los cuatro informantes mostraron que sí producían un tono de voz estadísticamente significativo más alto en los enunciados mendaces que superaban la media de F0 de los veraces. (t(2) = 2,423, p <0,136; t(5) = 2,879, p <0,035; t(2) = 5,583, p <0,031; t(5) = 2,834, p <0,037) Por otra parte, también tres informantes mostraron que sí se produce bajada significativa del tono de voz mendaz teniendo en cuenta la media de F0 de los enunciados veraces. (t(4) = 5,648, p <0,005; t(2) = -1,721, p <0,227; t(8) = -3,482, p <0,008; t(2) = -6,895, p <0,020)

3.2 Pausas entre enunciados

Se realizaron dos pruebas T-Student para cada hablante teniendo en cuenta el contexto de los enunciados. En la primera prueba, se halló si existía diferencia de tiempo después de los enunciados falsos cuando el siguiente es uno verdadero, y lo comparamos con el tiempo que hay después de un enunciado verdadero seguido de otro verdadero. Solamente uno de los resultados marcó una diferencia significante. (t(1,138) = 2,494, p <0,218; t(5) = 1,880, p <0,119, t(6) = -5,081, p <0,002; t(12) = 3,413, p <0,005149) En la segunda prueba, se halló lo opuesto: si la diferencia de tiempo antes de un enunciado falso cuando el que le precede es verdadero es significativo comparándolo con el tiempo entre dos enunciados veraces. En este caso, también solamente uno de los informantes mostró que sí había diferencia, (t(5) = 0,726, p <0,497; t(5) = -1,037, p <0,347; t(6) = -6,915, p <0,000452; t(15) = 0,609, p <0,552) y resultó ser el mismo que mostró diferencia de tiempo después de los enunciados falsos.

3.3 Pausas dentro de los enunciados

Sobre las diferencias entre las pausas dentro de los enunciados se estudiaron dos puntos diferentes. El primero trató con la diferencia temporal en las pausas de los enunciados veraces y mendaces de tres de los cuatro informantes (uno de ellos no presentó ninguna pausa dentro de ningún enunciado). Se dedujo que no había ninguna diferencia significante entre el tiempo de los dos tipos de pausa. (t(3) = 0,058, p <0,958; t(13) = 0,028, p <0,978; t(6) = 1,602, p <0,160) En el segundo punto se calculó qué porcentaje de pausas hay en ambos tipos de enunciados. De las pausas que se producen dentro de los enunciados, el 80 %, 78,6 % y 25 % en cada informante están dentro de un enunciado mendaz.

3.4 Número de sílabas de los enunciados y duración

Hallamos la diferencia entre el número de sílabas de la verdad y la mentira mediante test T-Student, y la estadística nos sugiere que no hay ninguna diferencia significativa entre ambos tipos de discurso. (t(18) = 0,455, p <0,654; t(17) = -0,889, p <0,386; t(20) = 1,013, p <0,323; t(27) = 1,047, p <0,304).

Tampoco se halló diferencia significativa alguna al comparar también mediante T-Student la duración de los dos tipos de enunciados. (t(18) = 0,079, p <0,938; t(17) = -1,176, p <0,256, t(20) = -0,246, p <0,808; t(27) = 0,781, p <0,442).

3.5 Velocidad

La velocidad se calculó en número de sílabas por segundo. Se compararon las cifras obtenidas entre los discursos veraz y mendaz, pero no se hallaron diferencias significativas. (t(16) = -0,105, p <0,918; t(10,619) = 1,026, p <0,328; t(29) = -0,871, p <0,394; t(27) = 1,107, p <0,278).

3.6 Tasa de errores

Para hallar el porcentaje de errores (repeticiones, palabras del estilo "eeeh", etc.) calculamos el número de errores totales en cada uno de los dos discursos de cada informante. Tres de los hablantes mostraron una tasa de error mayor durante la mentira (60,3 %, 68,7 % y 52,4 %). Sin embargo uno de los informantes muestra más errores durante su discurso veraz (63,3 %) Cabe destacar que esta persona es la misma que utiliza más pausas dentro de los enunciados veraces que en los mendaces.

3.7 Lenguaje corporal

Esta parte del estudio se centra más en una valoración cualitativa que cuantitativa, pues los factores son muy variables.

Es importante volver a destacar las palabras de Paul Ekman que nos advierten que no solamente existe una expresión para cada emoción. Es decir, un mismo gesto podremos encontrarlo en el contexto de la verdad o la mentira. Es el conjunto de expresiones el que nos lleva a poder detectar que un enunciado es falso. Teniendo esto en cuenta, y que no todas las personas se expresan de igual manera, podemos concluir de forma general que hay miles de combinaciones de gestos que nos indican que algo puede ser verdadero o que puede ser falso. Ninguna expresión que sea típicamente señal de mentira es

exclusiva de los enunciados de la misma categoría. Así, encontramos estos gestos también en enunciados ciertos. Lo que nos lleva a determinar que un discurso es veraz o mendaz es el conjunto, y en ese conjunto, una serie de combinaciones que se repiten y que pueden depender de cada persona. En el corpus utilizado para este estudio, pudimos ver tales características. Cada uno de los informantes tenía una forma de expresar la mentira de diferente forma, pero todos tenían algo en común: los gestos que utilizaron más o menos frecuentes en cada uno de los hablantes, eran los mismos. Cada informante combina los gestos de diferentes formas o con diferente regularidad. En los siguientes puntos se resume lo más destacable:

- Se utilizan más ilustradores en la verdad y más manipuladores en la mentira.
- En la verdad se utilizan más o menos las mismas combinaciones gestuales de la mano en cada informante, mientras que en la mentira casi no existe ningún patrón común.
- En la mentira se desvía la mirada y/o se intenta no mirar al público y en la verdad se desvía solamente en situaciones muy particulares.
- Si alguna combinación pierde alguno de sus gestos (combinación gestual de manos, movimientos de cabeza y hombros, mirada, etc.), se incorpora algún otro elemento que siga reforzando la hipótesis de que un enunciado es falso en la mayoría de los casos.
- El número de señalamientos llega a doblarse en la mentira en algunos casos.
- En la mentira se frunce el ceño mucho más.
- La F0 disminuye o aumenta en la mentira con respecto a la media en la verdad y en algunos casos coincide con el tipo de patrón gestual utilizado.

3.8 Prosodia y lenguaje corporal

Llama la atención que todos los informantes presentan una curiosa relación entre gesto y voz durante el discurso mendaz. Se observó que cuanto más se distancia la F0 en la mentira de la media en la verdad, disminuye el número de expresiones corporales típicas de la mentira. Es decir, que la cantidad de expresiones corporales es inversamente proporcional al aumento o bajada de F0 desde su media en la verdad. Cuantos menos gestos utiliza, más se aleja la F0 de la media.

4 Conclusión

Convencidos de que no hay una única característica que pueda discriminar un enunciado veraz de otro mendaz de forma categórica, nos sumergimos en la búsqueda del conjunto de particularidades tanto gestuales como de la voz que se combinan para poder hacer una distinción entre los dos tipos de discurso. Concluimos que sí existe una relación entre estos dos factores, aunque siempre teniendo en cuenta que cada persona expresa la mentira de forma diferente, e incluso la misma persona varía a la hora de expresarla. También tuvimos en cuenta que las mismas características se pueden dar en cualquier tipo de enunciado, y que por eso debíamos prestar más importancia al conjunto de las mismas que al hecho de la existencia de una sola.

En cuanto a la voz, se determinó que en la mentira la frecuencia fundamental aumenta, pero también puede disminuir. Se estudió también el tiempo entre dos enunciados cuando el primero era mentira y el segundo verdad, y viceversa con el tiempo entre dos enunciados veraces. Solo uno de los informantes mostró una diferencia significativa en el tiempo antes de una mentira, y dos después de ella. Detectamos que dos de tres realizaban más pausas en la mentira que en la

verdad y uno no las realizaba. En cuanto a la velocidad, duración de los enunciados y longitud en sílabas, no se encontró diferencia alguna significativa entre los enunciados falsos y ciertos. Sí percibimos que tres de los cuatro informantes cometían más errores durante la mentira.

Sobre el lenguaje corporal, comprobamos que los informantes tienen la tendencia a repetir ciertas combinaciones de características a la hora de mentir, y que si una falta, otra ocupa su lugar. Pudimos comprobar una serie de peculiaridades que se repetían de forma frecuente. Comprobamos que cada gesto no se podía analizar de forma independiente como bien sugería Ekman y observamos las múltiples combinaciones que se dan durante la mentira.

Existe una relación inversa entre el tono de voz y los gestos. En la mentira, cuando más disminuye o se eleva la frecuencia fundamental de la media en la verdad, encontramos menos gestos típicos del discurso mendaz encontramos.

5 Bibliografía

- Allwood, J., Cerrato, L., Dybkjaer, L., Jokinen, K., Navarretta, C., & Paggio, P. (2005). The MUMIN Multimodal Coding Scheme. NorFA yearbook 2005, 129-157. http://sskkii.gu.se/jens/publication s/bfiles/B70.pdf
- Borsma, P., Weenink, D. (2013). Praat:
 Doing Phonetics by Computer
 (Version 5.3.39). En
 http://www.fon.hum.uva.nl/praat/
- DePaulo, B. M., Lindsay, J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to Deception. Psychological Bulletin, 129(1), 74-118. En http://www.angelfire.com/lieclass/ C2D.pdf
- De Ruiter, J. P. (2000). The production of gesture and speech. Language and Gesture (1.ª ed.). Cambridge: Cambridge University Press, 284-311. En http://www.unibielefeld.de/lili/personen/jruiter/downloads/DeRuiter_GestureBook_2000.pdf
- Ekman, P. (2009). Telling Lies: Clues to deceit in the marketplace, politics, and marriage. New York/London: W.W. Norton & Company.
- Ladefoged, P., & Maddieson, I. (1996). The Sounds of the World's Languages. Oxford: Blackwell.
- Lausberg, H., & Sloetjes, H. (2009). Coding gestural behavior with NEUROGES-ELAN system. Behavior Research Methods, Instruments & Computers. En http://tla.mpi.nl/tools/tlatools/elan/
- McNeill, D. (1992). Hand and Mind. Chicago/London: University of Chicago Press.

- McNeill, D. (2010). Gesture: A Psycholinguistic Approach. The Cambridge Encyclopedia of the Language Sciences (ed. Patrick Colm Hogan).
- http://mcneilllab.uchicago.edu/pdfs/gest ure.a_psycholinguistic_approach.ca mbridge.encyclop.pdf
- Nonhebel, A., Crasborn, O., & Van Der Kooij, E. (2004). Sign language transcription conventions for the ECHO project. BSL and NGT mouth annotation. Radboud University Nijmegen. http://signlang.ruhosting.nl/echo/docs/ECHO _transcr_mouth.pdf
- Pollina, D. A., Vakoch, D. A., & Wurm, L. H. (1998). Formant Structure of Vowels Spoken Durong Attempted Deception. Polygraph, 27(2). En http://www.clas.wayne.edu/multimedia/usercontent/File/Wurm/Pollina_Vakoch_Wurm_1998.pdf
- Roseano, P., González, M., Borràs-Comes, & Prieto, P. (no publicado, 2013). Communicating epistemic stance: How speech and gesture patterns reflect epistemicity and evidentiality. Universitat Pompeu Fabra.
- Spence, K., Arciuli, J., & Villar, G.(2012).

 The role of pitch and speech rate as markers of deception in Italian speech, Australasian Speech Science and Technology Association (ASSTA)
- http://assta.org/sst/SST-12/SST2012/PDF/AUTHOR/ST1200 35.PDF
- SPSS Inc. (2012). IBM SPSS Statistics version 21. Chicago.
- Streeter, L. A., Krauss, R. M., Geller, V., Olson, C., & Apple, W. (1997). Pitch

- Changes during Attempted Deception. Journal of Personality and Social Psychology, 35(5), 345-350. En http://www.cs.columbia.edu/~julia/papers/Streeter_Krauss_Geller_et_al_1977.pdf
- Van Edwards, V. (2013). Human Lie Detection & Body Language 101.USA. Science of People. En http://www.scienceofpeople.com/
- Villar, G., Arciuli, J., & Paterson H. (2012).

 Vocal Pitch Production during Lying:
 Beliefs about Deception Matter.
 Psychiatry, Psychology and Law,
 20(1), 123-132. En
 http://www.tandfonline.com/doi/a
 bs/10.1080/13218719.2011.633320#.U
 gaF2JLgRv8

Variación fonética de los tonos estándar del chino mandarín. Un estudio de corpus

YAO, JUNMING; GARRIDO ALMIÑANA, JUAN MARÍA

1 Introducción

El chino mandarín es una lengua tonal. Tiene cuatro tonos estándar y el tono neutro. Dichos tonos tienen función léxica y se aplican sobre la sílaba. Las marcas tonales son "—, /, \/, \" y no marcado respectivamente en *pinyin*. Por ejemplo, 妈 mā (mamá), 麻 má (lino), 马 mǎ (caballo), 骂 mà (reñir) y 吗 ma (partícula).

Según las descripciones tradicionales, los cuatro tonos estándar tienen los siguientes contornos: el primero, alto y llano, el segundo, ascendente, el tercero, descendente y ascendente y el cuarto, descendente, como reflejan las marcas tonales. Se suelen describir estos contornos con el sistema de cinco niveles de Chao (1930) como ilustramos en la figura 1.

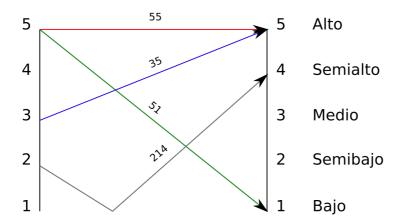


Figura 1. Contornos tonales de los cuatro tonos estándar del chino mandarín, según Chao (1930)

El tono neutro se produce con menor duración que otros tonos estándar (Lin, 1990), pero no tiene un contorno fijo sino que cambia según sus tonos precedentes (Cao, 1992; Chao, 1968; Comité, 2004; Wang, 2003). Dicho tono no entra en nuestro estudio actual.

Muchas veces en el habla continua los tonos no coinciden con los contornos estándar e incluso son totalmente contrarios a ellos. En la figura 2 tenemos un enunciado realizado por un hablante nativo. El tono de cada sílaba está anotado con los números "1, 2, 3, 4" y "0" para referirse a los tonos primero, segundo, tercero, cuarto y neutro. Se puede observar que los tonos sombreados tienen contornos muy distintos a los estándar definidos.

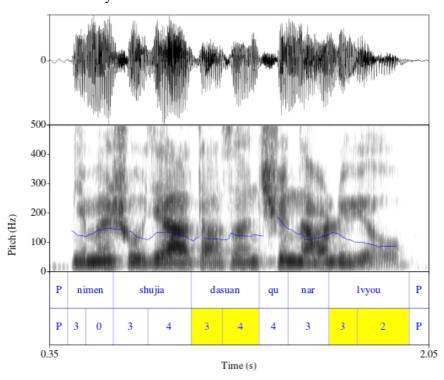


Figura 2. Ejemplo de la realización de los tonos en el enunciado "nimen shujià dasuan qu nar luyóu?", pronunciado por un hablante nativo de chino mandarín.

Aunque existen estudios sobre la variación fonética de los tonos del chino mandarín (Lin, 2006; Shi, 2009; Shih y Kochanski, 2000; Xu, 1993; Zuo, 2002), este tema es aún una cuestión pendiente, entre otros motivos porque requiere el manejo de una gran cantidad de datos, procedentes de diferentes locutores, que hacen muy difícil su análisis mediante métodos manuales.

El presente trabajo es un estudio acústico de la variación fonética de los cuatro tonos estándar del chino mandarín, que se enmarca en un proyecto más amplio de descripción de las curvas melódicas de dicho idioma, a partir del análisis de un corpus amplio de enunciados leídos por diferentes locutores. En concreto, tiene tres objetivos: describir la realización fonética de los cuatro tonos, estudiar sus variaciones y analizar los factores relacionados.

Siguiendo estos objetivos, planteamos dos hipótesis:

- La realización fonética de los cuatro tonos del chino mandarín difiere de sus contornos estándar en el habla continua.
- Los contornos de los cuatro tonos del chino mandarín varían en función de factores como el sandhi, el contexto tonal y la posición de la sílaba en el grupo entonativo.

Para verificar nuestras hipótesis, realizamos un experimento acústico sobre un corpus amplio de oraciones, y uno perceptivo, para validar los resultados del estudio acústico.

2 Análisis acústico

2.1 Corpus

El corpus que se empleó para este estudio incluye 651 oraciones, en total 15 873 sílabas, extraídas tanto de enunciados aislados como de párrafos. Este corpus fue seleccionado de un corpus más amplio de 278 enunciados aislados y 140 párrafos neutros con todo tipo de combinaciones tonales y cinco modalidades (declarativa, interrogativa absoluta, interrogativa parcial, exclamativa e imperativa). Fue grabado con una frecuencia de muestreo de 44 o 48 kHz y 16 o 24 bits de resolución (fue grabado en dos sesiones diferentes) por nueve hablantes nativos de chino mandarín, tres hombres y seis mujeres de entre 23 y 34 años.

Para cada fichero *wav* del corpus creamos un fichero *TextGrid* en *Praat* que contiene una serie de estratos (*tiers*) con segmentaciones de palabras, sílabas, grupos entonativos y oraciones, y con etiquetados de *pinyin*, tonos, pausas y factores que vamos a estudiar para la variación tonal (prominencia silábica, sílaba *sandhi*, contexto tonal y modalidad oracional). En la figura 3 se presenta un ejemplo.

Sobre la prominencia de las palabras bisílabas del chino mandarín, existen afirmaciones totalmente contrarias, por ejemplo fuerte-débil (Duanmu, 2000; Kochanski et al., 2003) o débil-fuerte (Chao, 1968; Lin, 1990). Nuestro etiquetado de las sílabas prominentes se basa en la duración, ya que es una característica acústica muy importante y un parámetro objetivo de las sílabas acentuadas en dicha lengua (Chao, 1932; Shen, 1990; Wang, 1993; Yao, 2009).

El etiquetado de las sílabas *sandhi* se basa en estudios tradicionales sobre diferentes combinaciones de dos tonos consecutivos (Comité, 2004; Lin, 1990; Wang, 2003; Wu, 1993). En el caso del tercer tono, existen *no sandhi, sandhi fonético* (delante del primero, segundo y cuarto tono) y *sandhi fonémico* (delante de otro tercer tono).

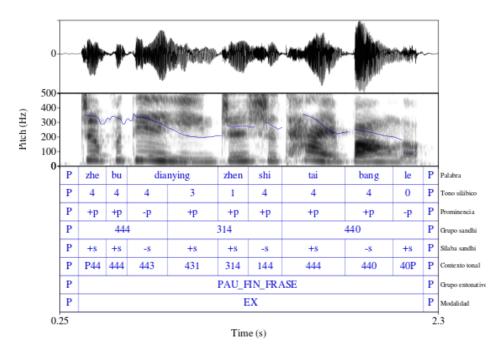


Figura 3. Segmentación y TextGrid con factores que se analizan para los contornos tonales. Enunciado: "zhè bù diànyĭng zhēn shì tài bàng le!", pronunciado por una hablante nativa de chino mandarín

2.2 Procesado automático

El corpus anotado fue procesado de forma automática, dado el gran volumen de datos, para obtener representaciones acústicas de la evolución de todos los tonos incluidos en el mismo. Todo el corpus fue estilizado y anotado con MelAn, el sistema automático de estilización, anotación y modelización presentado en Garrido (2010) que permite la obtención de listados de los patrones de F0 presentes en el corpus analizado, con indicación del número de apariciones de cada patrón en el mismo. De esta manera se obtuvieron *TextGrids* etiquetados para todos los grupos entonativos como refleja la figura 4. Desarrollada inicialmente para el procesado de las curvas melódicas del español, la herramienta fue modificada para extraer los patrones en el ámbito de la sílaba, unidad natural de los tonos en chino mandarín. Aparte de otras características, MelAn utiliza dos niveles tonales (P y V), y uno adicional (P+ y V-) para cada nivel, para la anotación de las curvas melódicas. Los patrones extraídos utilizan la anotación de #1#2_..._#1#2, donde #1 puede ser P, P+, V o V–. Indica el nivel tonal del punto de inflexión. #2 es I, M o F. Indica la posición del punto de inflexión dentro del núcleo de la sílaba (Inicial, Media o Final). Si un punto de inflexión se encuentra antes del núcleo, se considera como inicial. Si está después del núcleo, se considera como final. Por ejemplo, VI_PF refleja un patrón ascendente que tiene el punto de inflexión bajo al inicio del núcleo de la sílaba y el punto alto al final del núcleo. El experimento de percepción presentado en Yao y Garrido (2010) permitió comprobar que los contornos generados empleando este sistema de codificación resultaban muy semejantes a los originales desde el punto de vista perceptivo, lo que parece indicar que no son necesarios los cinco niveles empleados tradicionalmente para la descripción de los cuatro tonos.

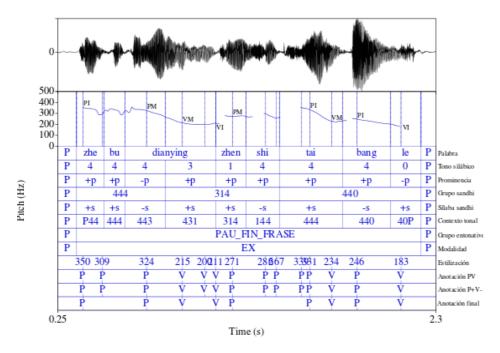


Figura 4. Estilización y anotación del enunciado "zhè bù diànyĭng zhēn shì tài bàng le!", pronunciado por una hablante nativa de chino mandarín

MelAn también permitió extrajer automáticamente todos los patrones de los cuatro tonos del corpus en función de los diferentes factores. Se generaron tablas de patrones con estos datos para analizar los patrones más frecuentes bajo distintas condiciones y para estudiar la influencia de cada factor.

2.3 Resultados

Patrones más frecuentes según tonos

En la tabla 1 se muestran los cinco patrones más frecuentes de cada uno de los cuatro tonos obtenidos de nuestro corpus sin tener en cuenta otros factores.

Tono	Patrón	Frecuencia	Tono	Patrón	Frecuencia
1 (2799)	PI	28,76 (805)	3 (2602)	VM	15,37 (400)
	0	24,90 (697)		0	12,07 (314)
	VI_PM	8,75 (245)		PI_VM	11,80 (307)
	PF	7,15 (200)		VI	10,11 (263)
	PI_PF	6,07 (170)		VF	5,38 (140)
2 (3347)	VM	21,84 (731)	4 (5805)	PI	27,46 (1594)
	0	18,29 (612)		0	12,16 (706)
	VI	11,41 (382)		PI_VF	10,30 (598)
	VF	6,66 (223)		PM	7,44 (432)
	PI_VM	6,54 (219)		PI_VM	6,74 (391)

Tabla 1. Patrones más frecuentes de los cuatro tonos del corpus estudiado sin tener en cuenta otros factores. Entre paréntesis figura el número de casos

Se observa que los patrones más frecuentes de los cuatro tonos los dividen en dos grupos generalmente: dos tonos altos y dos bajos, teniendo cada uno sólo un punto de inflexión en la mayoría de los casos. En otras palabras, no reflejan los contornos estándar sino que parecen tonos de registro, los cuales se distinguen por la altura tonal en vez del contorno. Una reflexión sobre ello es

que en el habla continua probablemente no sea tan importante la trayectoria del tono en cada sílaba como se suele creer.

Escogiendo el patrón más frecuente de cada tono, podríamos considerar que su realización fonética es PI, VM, VM y PI respectivamente. La figura 5 presenta un ejemplo de realización de cada tono.

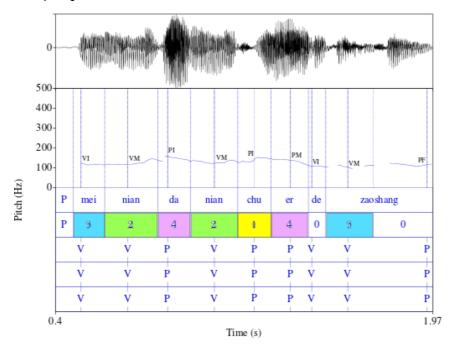


Figura 5. Ejemplo de la realización fonética de los tonos. Enunciado: "měi nián dà nián chū èr de zăoshang...", pronunciado por un hablante nativo de chino mandarín

Factores relevantes

Del mismo modo, se generaron tablas con patrones más frecuentes de cada tono en función de la modalidad oracional, la prominencia silábica, el sandhi, el contexto tonal y la posición de la sílaba dentro del grupo entonativo. Los datos obtenidos mostraron que el patrón más frecuente de cada tono no varía sustancialmente según la modalidad del enunciado ni tampoco según la prominencia de la sílaba.

En cambio, la posición de la sílaba sí afecta a ello. La tabla 2 muestra el patrón más frecuente de cada tono en distintas posiciones dentro del grupo entonativo. Se percibe que en la posición final y final de enunciado¹, los tonos mantienen sus contornos estándar, lo cual coincide con las afirmaciones de Lin (2006) sobre los llamados tonos de límite.

Sobre la influencia del efecto sandhi, sólo se observa en el tercer tono. Cuando no es sandhi, tiene el patrón más frecuente PI VM, cuando es sandhi fonético, corresponde al patrón VM y cuando es sandhi fonémico, el patrón más frecuente tiende a ser VM_PF. La figura 6 muestra un ejemplo.

La influencia del contexto tonal también solo se encuentra en el tercer tono. Se observan variaciones de los patrones más frecuentes según cambia el tono precedente y el tono próximo del tercer tono, pero no se han encontrado regularidades en el presente trabajo.

¹ Usamos el término final de enunciado cuando el final del grupo entonativo coincide con el final de una oración.

	Posición	Posición				
Tono	INI	INT	FIN	FEN		
1	PI	PI	PI_PF	PI_PF		
2	VI	VM	VM_PF	VM_PF		
3	VI	VM	PI_VM	PI_VM		
4	PI	PI	PI VF	PI VF		

Tabla 2. Patrones más frecuentes de los cuatro tonos según la posición en el grupo entonativo. INI: inicial. INT: interior. FIN: final. FEN: final de enunciado

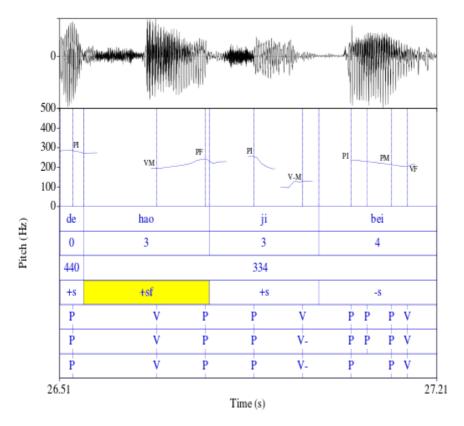


Figura 6. Ejemplo del patrón del tercer tono sandhi fonémico

Prueba de percepción

Para comprobar si los patrones más frecuentes obtenidos automáticamente por MelAn sirven para construir frases con tonos adecuados, realizamos una prueba de percepción con estímulos sintetizados.

Se evaluó el grado de adecuación de la realización fonética de los tonos en una escala del 1 al 5, siendo el 1 *poco adecuado* y el 5 *perfecto*.

3 Estímulos

Se eligieron 13 frases y párrafos de un locutor y los mismos 13 de una locutora, representativos en el corpus de todos los factores relevantes para la definición de los tonos, y se resintetizaron con un nuevo contorno de F0 contruido con los patrones más frecuentes extraídos según cada factor. Para la síntesis, también se utilizaron los valores medios de la declinación de los grupos entonativos de cada locutor individual. Se obtuvieron así 26 estímulos en total. La generación de estos estímulos se realizó con una versión de ModProso (Garrido, 2013), aplicando el método *Overlap-Add*.

Los estímulos se presentaron en orden aleatorio con acceso web para que los sujetos pudieran realizar la prueba remotamente.

En la figura 7 se observan la forma de onda, espectrograma y contornos de F0 original y sintetizado de uno de los enunciados empleados en la prueba.

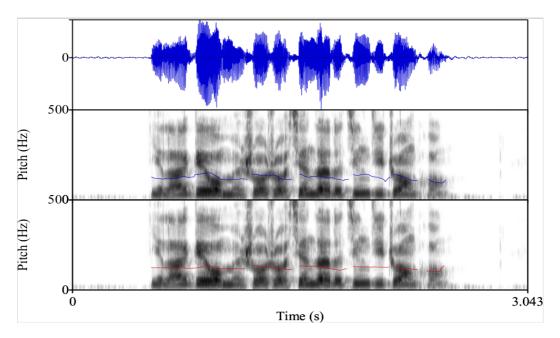


Figura 7. Forma de onda, espectrograma y contornos de F0 original y sintetizado del enunciado "nĭ lái gěi women shuoshuo xiànzai de jīngjì zhuangkuang"

3.1 Sujetos

Participaron en la prueba 21 oyentes chinos, de los cuales un 80 % son profesores o locutores de radio, profesiones que requieren una certificación oficial de nivel alto en chino mandarín.

Los sujetos desconocían el origen de los estímulos. Podían escucharlos tantas veces como fuera necesario. Se les pidió que utilizaran auriculares durante el proceso para mejorar y homogeneizar las condiciones de la prueba.

3.2 Resultados

La tabla 3 muestra la puntuación media de los 21 sujetos obtenida para cada estímulo. Los resultados están separados por locutores.

	Puntuaciones medias					
Estímulo	Locutor	Locutora	Los dos			
1	4,47	4,56	4,51			
2	3,78	3,79	3,79			
3	3,93	3,23	3,58			
4	3,66	3,44	3,55			
5	3,57	3.12	3,34			
6	4,80	4,78	4,79			
7	4,01	3,60	3,80			
8	4,15	3,60	3,88			
9	3,87	3,49	3,68			
10	3,84	3,42	3,63			
11	4,02	4,10	4,06			
12	4,21	3,92	4,06			
13	3,22	3,27	3,25			
Todos	3,96	3,72	3,84			

Tabla 3. Puntuación media obtenida para cada estímulo según el locutor en la prueba de percepción

La media global del conjunto de los estímulos (los dos locutores conjuntamente) fue 3,84, la correspondiente a los estímulos del locutor 3,96 y 3,72 los de la locutora. Dado que la escala de la puntuación es de 1 a 5, puede considerarse que una nota mayor que 3 es aceptable. Es decir, en general los tonos de los estímulos sintetizados son bastante adecuados, además, los resultados del informante masculino son mejores que los del femenino.

Todas las puntuaciones de los estímulos están por encima de 3. Además, un 46 % son mayores que 4 en el caso del locutor y un 23 % en el de la locutora.

Aunque algunos enunciados recibieron notas relativamente más bajas que otros, por ejemplo el 5 de la locutora, en general los resultados son satisfactorios.

4 Conclusiones

Después de todo el análisis puede afirmarse que se han cumplido los objetivos y verificado las hipótesis planteadas al principio de este trabajo. Destacan algunos aspectos:

- No coincidencia entre los patrones más frecuentes de la realización fonética de los tonos con los contornos estándar.
- Un punto de inflexión P o V en la mayoría de los patrones más frecuentes salvo al final del grupo entonativo.
- Tendencia a mantener los contornos estándar al final del grupo entonativo (tonos de límite).
- Influencia de sandhi y el contexto tonal en el tercer tono.

Por otro lado, los resultados del presente estudio, sobre todo la buena calificación en la prueba de percepción, refuerzan la idea de considerar solo dos niveles tonales en vez de cinco para la descripción de los tonos del chino mandarín.

Finalmente, vale la pena destacar las ventajas del uso de una herramienta automática como MelAn en los estudios de fonética experimental cuando se emplean corpus grandes, ya que permiten estudiar múltiples factores con mayor fiabilidad de los resultados estadísticos, al poderse manejar una mayor cantidad de datos.

Para futuros trabajos, se podría analizar a fondo las influencias de los diferentes factores en la variación de los tonos, por ejemplo el caso del cambio del tercer tono por el contexto tonal.

5 Bibliografía

- Cao, J. (1992). On neutral-tone syllables in Mandarin Chinese. *Canadian Acoustics*, 3-1992.
- Chao, Y. (1930). A system of "tone-letters". Le Maître Phonétique, troisième série 30, 24-27, Avrill-Juin. Republicado en Wu Z., & Zhao X. (eds.), Linguistic Essays by Yuenren Chao, 2006, 98-102. Beijing: The Commercial Press.
- Chao, Y. (1932). A preliminary study of English intonation (with American variants) and its Chinese equivalents. BIHP, *The Ts'ai Yüan P'ei Anniversary Volume*. Republicado en Wu Z., & Zhao X. (eds.), *Linguistic Essays by*

- *Yuenren Chao*, 2006, 110-197. Beijing: The Commercial Press.
- Chao, Y. (1968). *A Grammar of Spoken Chinese*. Berkeley: University of California Press.
- Comité Nacional de Lengua y Escritura (2004).《普通话水平测试实施纲要》 [Normativa para el examen de Putonghua]. Beijing: The Commercial Press.
- Duanmu, S. (2000). The Phonology of Standard Chinese. Oxford: Oxford University Press.
- Garrido, J. M. (2010). A tool for automatic F0 stylisation, annotation and

- modelling of large corpora. Speech Prosody 2010, Chicago.
- Garrido, J. M. (2013). ModProso: A Praat-Based tool for F0 Prediction and Modification. *Proceedings of TRASP* 2013, 38-41.
- Kochanski, G. P., Shih, C. y Jing, H. (2003). Quantitative measurement prosodic strength in Mandarin. Speech Communication, 41, 625-645.
- Lin, M. (2006). Interrogative vs. declarative and the boundary tone in Standard Chinese. 《中国语文》 Language] 4, 364-376.
- Lin, T. (1990). 《语音探索集稿》[Essays on *Phonetic Research*]. Beijing: BLCUP.
- Shi, F. (2009). Exploration of Experimental Phonology. Beijing: Peking University Press.
- Shih, C. y Kochanski, G. P. (2000). Chinese tone modeling with Stem-ML. the International Proceedings of Conference on Spoken Language Processing 2000, Beijing.
- Shen, X. (1990). The Prosody of Mandarin Chinese. California: University of California Press.
- Wang, L. (1993). 四十年来汉语语音研究综述 [A summary on Chinese phonetic studies in the last forty years]. En J. Liu y J. Hou (Eds.),《中国语文研究四十年纪 念文集》, 174-180. Beijing: BLCUP.
- Wang, L. (2003). 《播音员主持人训练手册》 [Manual formación de presentadores y locutores]. Beijing: CUCP.
- Wu, Z. (1993). 普通话声调实验研究四十年 [Forty years of experimentalstudies on Putonghua's tone system]. En J. Liu y J. Hou (Eds.),《中国语文研究四十年纪 念文集》, 169-173. Beijing: BLCUP.
- Xu, Y. (1993). Contextual Tonal Variation in Mandarin Chinese. Ph.D. dissertation, University of Connecticut.
- Yao, J. (2009). Prominencia silábica de las palabras bisílabas en chino mandarín en un contexto de foco ancho. Trabajo de Máster, UIMP.
- Yao, J. & Garrido, J. M. (2010). Validación perceptiva de un sistema deanotación automática de contornos de F0 aplicado al chino mandarín. IX Congreso Internacional de lingüística General, Valladolid.
- Zuo, P. (2002). Tonal coarticulation: F0 Contextual realization

Mandarin Chinese tones. Ph.D., IMS Phonetik.

Parte II. Enseñanza de segundas lenguas

Propostes didàctiques d'entonació de cortesia atenuant per al nivell b1 de català com a llengua estrangera

DEVÍS HERRAIZ, EMPAR; BARTOLÍ RIGOL, MARTA

1 Introducció

La cortesia en el diàleg quotidià és un fenomen expressat a través de recursos morfològics, sintàctics, lèxics i semàntics, però la participació del component fònic també compleix un paper primordial en el moment d'interpretar un enunciat com a cortès o descortès. A nivell suprasegmental, l'entonació de cortesia és de suma importància pel seu marcat caràcter cultural i perquè ofereix formes que permeten evitar possibles conflictes i establir una relació basada en la cordialitat i la cooperació entre els interlocutors, cosa que, indubtablement, beneficia les relacions socials i interculturals. El desconeixement dels usos més freqüents en una llengua estrangera pot provocar, en canvi, malentesos que dificultin les relacions socials.

L'estudi formalitzat de la cortesia verbal comença els anys 70 i 80 amb l'aparició de les tres teories fundadores: la de R. Lakoff (1973), la de G. N. Leech (1983) i la de P. Brown i S. Levinson (1987). Des de llavors, l'assumpte ha suscitat gran interès en diferents llengües, en canvi, en àmbit català comptem només amb les referències de Payrató (1988), Nadeu (2008) i Nadeu & Prieto (2011).

L'objectiu d'aquesta recerca ha estat, doncs, cobrir aquest buit analitzant els trets melòdics de cortesia atenuant que utilitzen els parlants de català en els seus intercanvis, amb la finalitat d'oferir recursos didàctics als docents de català com a llengua estrangera de nivell B1, orientats a evitar els conflictes provocats

per malentesos de tipus melòdic. Es presentaran els resultats obtinguts, a partir de l'anàlisi acústica de les dades i la seva validació perceptiva, i una sèrie de propostes didàctiques que considerem de gran utilitat docent.

Controlar el funcionament dels trets melòdics cortesos atenuants augmentarà l'efectivitat en la comunicació, la satisfacció de l'interlocutor i, a nivell professional, la qualitat del servei i la formació cultural del treballador.

2 Metodologia

Per a l'anàlisi de les dades s'ha aplicat el mètode d'Anàlisisi Melòdica de la Parla que presenta F. J. Cantero (2002) en el seu llibre Teoría y Análisis de la entonación. El punt de partida és la definició de l'entonació a partir de les variacions de freqüència fonamental (F0) que compleixen una funció lingüística durant l'emissió de veu (Cantero, 2002: 18). L'autor considera que l'entonació funciona en tres nivells:

1) Pre-lingüístic: Comprèn els fenòmens de l'accent, el ritme i la melodia, que funcionen conjuntament com a elements fònics coestructuradors del discurs, al marge de qualsevol altra dimensió significativa.

En aquest nivell, es considera l'entonació com un "contenidor" lingüístic, integrant i delimitant les unitats del discurs. És el que Quilis (1993) anomenava "funció delimitadora" i "funció integradora" de l'entonació. Els trets melòdics característics d'aquest nivell són trets fonètics, ben concrets.

2) Lingüístic: Comprèn les característiques melòdiques, el rendiment fonològic de les quals, permet caracteritzar i distingir les unitats funcionals de l'entonació, els seus "signes lingüístics".

En català van ser identificats vuit tonemes resultants de la combinació dels trets fonològics¹ (/±interrogatiu, ±suspès, ±emfàtic/) que conformen els quatre tipus d'entonació descrits per Cantero (neutra, interrogativa, suspesa i emfàtica), els patrones melòdics dels quals constitueixen les melodies típiques del català (Font-Rotchès, 2007). Fins ara, s'han trobat vuit patrons melòdics diferents per al català, amb les seves variants i els seus marges de dispersió.

3) Paralingüístic: Comprèn les variants melòdiques de tonemes, dins dels amplis marges de dispersió de cada un d'ells.

Aquestes melodies permeten expressar emocions particulars, trets discursius idiolectals o de la personalitat del parlant, i fins i tot poden aparèixer variablement codificades (com trobem, per exemple, en l'expressió de la "cortesia": Devís, 2011).

El mètode per a l'anàlisi dels tres nivells entonatius permet:

- segmentar la corba entonativa en segments tonals (normalment, un segment tonal per vocal, excepte en les inflexions);
 - analitzar l'entonació independent d'altres nivells d'anàlisi lingüística;
- analitzar tots els fenòmens tonals de la parla inclosos en la melodia (accent, ritme i entonació).

I comprèn dues fases:

Fase acústica (descriptiva):

- extracció de la F0 de l'enunciat;
- determinació dels valors de F0 vocàlics; s'identifiquen les vocals i es transcriu el seu valor mitjà;

¹ Els trets fonològics descrits en aquest nivell són trets abstractes, merament opositius.

- la successió de valors vocàlics genera una corba melòdica essencial que elimina els valors irrellevants;
- i, en últim lloc, cada patró melòdic obtingut en Hz s'estandarditza en percentatges per construir una melodia independent de les característiques del parlant.

Fase perceptiva (experimental):

- se sintetitza la melodia resultant mitjançant el programa d'anàlisi acústica PRAAT: doing phonetics by computer i se substitueixen els valors tonals pels valors estandarditzats;
- es modifiquen únicament els valors, la rellevància dels quals volem comprovar;
- es realitza un experiment perceptiu que permet validar la melodia sintetitzada i/o falsar la hipòtesi melòdica.

3 Corpus

El corpus utilitzat és el "Corpus oral de parla espontània" elaborat per la professora Font Rotchés entre el 1999 i el 2000. Està constituït per 580 enunciats, produïts en un context de diàleg per 160 informants catalanoparlants, d'ambdós sexes, d'edats compreses entre els 18 i els 70 anys i amb dedicacions professionals molt variades. Compta amb 47 hores de material audiovisual procedents de 20 programes diferents de televisió catalana (TV3 i Canal 33) i de la franja en català de televisió espanyola (TVE2). Els informants procedeixen de les diferents poblacions de Catalunya i, en menor nombre, del País Valencià i de Mallorca.

4 Descripició dels tres melòdics atenuants en català

4.1 Codi de cortesia²

Les dades analitzades en la fase acústica van permetre elaborar la següent hipòtesi sobre els trets melòdics responsables de mitigar enunciats que podrien contenir algun tipus "d'agressivitat" lèxicogramatical implícita.

1) Des del punt de vista de l'entonació lingüística3.

Inflexions finals
+ suspensa:
(amb ascens final entre el 15% i el 70%)
+ neutra (amb un descens no superior al 40%)
circumflexa ascendent-descendent

2) Des del punt de vista de l'entonació paralingüística.

² Els trets melòdics de la cortesia no formen part del codi lingüístic de la llengua (que en l'entonació serien exclusivament els trets fonològics) compartit per tots els parlants de l'idioma. És a dir, no tots els parlants marquen la cortesia de la mateixa manera, ni obeeixen al mateix codi. No obstant això, podem considerar que la cortesia s'inscriu en un nivell d'anàlisi de l'entonació que està molt a prop de constituir un codi estable o semiestable: socialment compartit, molt a prop del codi lingüístic (Cantero & Mateo, 2011).

³ L'entonació lingüística comprèn les característiques melòdiques, el rendiment fonològic de les quals permet caracteritzar i distingir les unitats funcionals de l'entonació, els seus "signes lingüístics" (Cantero, 2002).

⁴ L'entonació paralingüística permet expressar emocions particulars, trets discursius idiolectals o de la personalitat del parlant, i fins i tot pot estar variablement codificada, com en el cas de la cortesia (Cantero, 2002).

Trets d'èmfasi (focus ample)				
Prominència en àtones (amb ascens no superiors al 30 %)				
Declinacions molt planes				
Registre tonal baix				
Primer pic desplaçat o absent				
Trets d'èmfasi de paraula				
(focus estret)				
Èmfasi de paraula amb inflexió circumflexa				

4.1.1 Examples

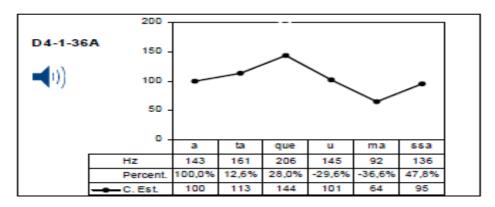


Figura 1. Melodia estandarditzada de l'enunciat: "Ataqueu massa"

La figura 1 presenta una inflexió final suspensa (+ 47,8%), des del punt de vista de l'entonació lingüística⁵, que, en aquest context, a nivell pragmàtic funciona com a tret atenuant per mitigar l'agressivitat implícita d'aquest enunciat en una situació de confrontació. Aquesta transposició funcional del contorn entonatiu (+ suspens) que a nivell pragmàtic actua com a mecanisme atenuant ja havia estat identificat precedentment per Hidalgo (2009). Les prominències en àtones, en canvi, des del punt de vista de l'entonació paralingüística⁶, es converteixen en claus acústiques rellevants per poder distingir un enunciat cortès d'un de neutre, són èmfasis de focus ample molt recurrents en tots els exemples analitzats, que generen un ritme melòdic que distingeix i caracteritza les emissions corteses atenuants.

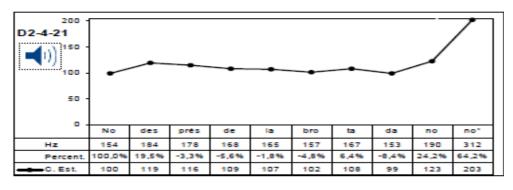


Figura 2. Melodia estandarditzada de l'enunciat: "No, després de la brotada no"

⁵ Per nivell lingüístic ens referim al segon nivell d'entonació descrit per Cantero (2002). 6 El nivell paralingüístic és el tercer nivell d'entonació descrit per Cantero (2002).

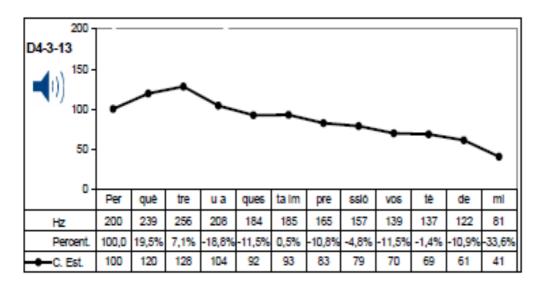


Figura 3. Melodia estandarditzada de l'enunciat: "Per què treu aquesta impressió vostè de mi"

Aquests trets identificats a la fase acústica del corpus en català coincideixen, essencialment, amb els identificats per a l'espanyol (Devís, 2011); en canvi, aquí no van aparèixer altres trets que en espanyol són molt rellevants, com les inflexions internes i les inflexions finals interrogatives. Per aquest motiu, vam decidir incorporar a la fase perceptiva també aquests últims, amb l'objectiu d'observar si també eren rendibles per al català. La introducció i la seva validació com a trets atenuants també per al català ens va permetre concloure que l'entonació de cortesia atenuant del català i de l'espanyol peninsular no són dos codis idiomàtics clarament diferenciats, sinó més aviat un únic codi cultural compartit.

Els models melòdics més rendibles en català capaços de convertir, per exemple, ordres en precs i confrontacions en cooperacions són els següents:

- 1. inflexió final interrogativa
- 2. èmfasi de paraula amb inflexió circumflexa
- 3. combinació d'inflexions internes amb inflexió final suspensa
- 4. combinació d'inflexions internes amb inflexió final circumflexa
- 5. prominències en vocals àtones (simples o combinades amb inflexió final circumflexa)
 - 6. inflexió final suspensa.

Aquestes conclusions són les que van servir de guia per elaborar les següents propostes didàctiques, útils per a l'ensenyament de l'entonació catalana de cortesia atenuant a parlants estrangers de nivell B1.

5 Proposta d'activitats

En les pàgines següents proporcionem una mostra d'activitats que es podrien fer amb un grup d'alumnes de nivell B1. Abans de posar-les en pràctica faríem una activitat prèvia general amb els estudiants per tal que poguessin escoltar algunes intervencions en què es fes ús de l'entonació de cortesia i descortesia. A part de l'àudio, faríem alguna pràctica oral amb algunes frases que proporcionaria el professor i amb tot el grup classe.

Activitat 1: No em toquis les coses

Objectiu: pràctica de l'entonació de cortesia

Funcions lingüístiques: demanar un favor, respondre a una petició, expressar que

s'està tip d'una situació i demanar disculpes.

Nivell: B1

Temps: 30-40 minuts⁷ **Destreses:** interacció

Tipus d'agrupament a classe: en parelles

Material: targetes

En aquesta activitat plantegem als estudiants una situació en què primer hauran de demanar un favor de manera atenuada (cortesa) i després de forma menys cortesa. Aprofitem per repassar o introduir diverses estructures útils per demanar favors i respondre a una petició, així com d'altres per expressar que estem tips o farts d'alguna cosa i demanar disculpes. N'exposem algunes tot seguit:

Demanar un favor:

- Pots/podries fer-me un favor?
- Podria demanar-te un favor?
- Em fas un favor?
- Necessito demanar-te un favor.
- Pots/podries...? És que...

Per respondre a la petició:

- És clar que sí!
- Clar!
- Per descomptat!
- Compta amb mi!
- I tant!

Per expressar que s'està tip d'alguna cosa:

- Estic fart de...
- Estic fins al capdamunt de..
- N'estic molt fart.../N'estic molt tip...

Demanar disculpes:

- Em sap greu...
- No t'ho prenguis així...
- Jo només volia...
- És que a mi m'agrada tant l'ordre...
- No sabia que...
- Em pensava que t'agradaria...

La situació que presentem als estudiants té lloc en una oficina. Es tracta d'un company o companya de feina que té la mania d'arreglar l'escriptori als altres. Els alumnes s'agrupen en parelles i donem a cada un d'ells una targeta amb la informació del paper que han de representar.

⁷ El temps és orientatiu, depèn una mica del repàs de les estructures lingüístiques i del que calgui comentar.

Personatge A:

El teu company d'oficina té la mania de posar ordre a les teves coses i canviar-te de lloc els objectes i papers de la taula. Alguna vegada li ho has comentat, però no et fa cas i continua igual. N'estàs una mica tip/a ja.

Personatge B

T'agrada l'ordre i t'encanta arreglar i decorar taules, la teva i la dels teus companys. Per descomptat que ho fas perquè sigui més còmode trobar el que busquen i perquè a l'oficina se sentin a gust. Fa poc vas comprar-los alguns cactus.

Les indicacions que donem als estudiants són les següents:

- 1. El personatge A ha de demanar a B un favor de manera cortesa (és a dir, que no li remeni l'escriptori).
 - 2. El personatge B l'accepta amablement.
- 3. Es repeteix la situació, però aquesta vegada A n'està fart i s'expressa llavors de forma descortesa.
- 4. El personatge B demana disculpes i es mostra una mica dolgut perquè en la seva acció només hi havia bones intencions.

Al final de l'activitat seria convenient que dues parelles representessin el diàleg davant la resta de la classe i entre tots, podria valorar-se'n l'actuació.

Activitat 2: El veí pesat

Objectiu: pràctica de l'entonació de cortesia

Funció lingüística: demanar un favor, preguntar per informació, respondre a una ordre o petició amablement i aviar algú.

Nivell: B1

Temps: 30-40 minuts **Destreses:** interacció

Tipus d'agrupament a classe: en gran grup

Material: targetes

En aquesta activitat els alumnes hauran de fer ús de l'entonació de cortesia per tal d'aviar algú, en aquest cas, un veí o veïna molt pesats. Hi haurà dos papers: el veí pesat, que demanarà informació o favors, i el veí cansat, que l'aviarà.

Pel que fa a les funcions lingüístiques que posaríem em pràctica podrien ser les que exposem a continuació:

Demanar informació8:

- *Saps si...?*
- Què...?
- Quin/quina/quins/quines...?
- On...?
- *Quan...?*

Demanar un favor:

- Necessitaria...
- Podries...

⁸ Ens referim a demanar informació sobre el veïnat en general.

- Si us plau...
- Resulta que... i és que...

Aviar algú:

- Disculpa'm, però és que me n'he d'anar.
- Perdona, però és que tinc pressa.
- Ostres, quina hora és!! Mare meva! Si jo havia de...
- Si de cas, parlem en un altre moment, perquè ara...
- Em sap greu, però t'he de deixar...

Desenvolupament

El primer que farem a l'aula és presentar la situació als alumnes, i per això, els farem algunes preguntes sobre els seus veïns; per exemple, els podem preguntar si han tingut bones o males experiències amb veïns i si en tenen algun de molt tafaner o pesat, etc. Els demanem que en parelles hi pensin i ho comentin o bé, en parlem entre tots.

A continuació podem fer una petita llista sobre tipus de veïns o situacions possibles i els preguntem com reaccionen o com reaccionarien ells a aquestes situacions. Cal ser cortesos perquè els continuarem tenint com a veïns.

Llista possible:

- veí/veïna que t'espia,
- que et fa un excés de preguntes i intenta coincidir amb tu,
- que et va al darrere,
- que li falta sempre alguna cosa a casa,
- que et demana el trepant, no te'l torna, i li has de demanar cada vegada.

Els preguntem com aviarien educadament un veí que els entreté a l'escala en un moment en què tenen pressa. Comentem les estructures que podrien utilitzar i el tipus d'entonació i fem també algunes proves.

Després de comentar-ho entre tots, els proposem un joc que consisteix a fer el paper de veí pesat. Es reparteixen situacions escrites en targetes a un o dos alumnes, que són els encarregats de molestar els altres.

L'activitat es pot desenvolupar en gran grup. Els alumnes que tenen el paper de veïns (dos o tres) es col·loquen al centre de l'aula i es dirigeixen a un alumne en concret per fer-li algun comentari o demanar-li alguna cosa. L'altre se l'ha de treure de sobre educadament.

Com que l'activitat es desenvolupa en gran grup, el professor o la professora pot valorar les interaccions després de cada intervenció.

Activitat 3: Ara no puc parlar

Objectiu: pràctica de l'entonació de cortesia

Funció lingüística: aviar algú

Nivell: B1

Temps: 30-40 minuts **Destreses:** interacció

Tipus d'agrupament a classe: en parelles

Material: targetes

En aquesta activitat treballarem també la manera d'aviar algú de manera cortesa i efectiva, però ho farem per telèfon. Caldrà utilitzar una entonació molt cortesa, perquè es tracta d'un parent o molt bon amic que ens està explicant mil

i una històries en un moment en què nosaltres hem de marxar i no tenim gaire temps per dedicar-li.

Els exponents de la funció "aviar algú", que són més o menys els mateixos que hem exposat en l'activitat anterior, són els següents:

Aviar algú:

- Ai, perdona, t'he de deixar, és que me n'he d'anar...
- Escolta, una cosa, és que...
- Hem de veure'ns un dia i parlar-ne tranquil·lament, però ara... em sap greu, però...
 - Et trucaré més tard, és que ara...

Desenvolupament

Com a introducció a l'activitat, preguntarem als alumnes si coneixen el verb "enrotllar-se" i "empatollar-se". Si no els coneixen, els donarem algun exemple perquè en dedueixin el significat.

Si disposem d'un ordinador a l'aula, podem mostrar-los algun fragment de pel·lícula o de serial, en què algú s'enrotlli, per exemple, podríem triar el personatge de David Güell de Plats bruts.

Un cop ja saben de què es tracta, podem preguntar-los si en coneixen sinònims o equivalents, per exemple, tenir xerrera, tenir xerrameca, etc.

A continuació els podem preguntar a ells directament si es troben sovint en la situació d'haver d'aguantar un "rotllo" a algú. Els podem demanar que ens expliquin algun cas personal. També els preguntem si tenen paciència o si saben com tallar la conversa delicadament... i com se senten quan ho fan i tenen la impressió que l'altre s'ha molestat per la "seva falta d'atenció."

Ens tornarem a mirar les estructures i l'entonació per tal de tenir models a l'hora de practicar-les. Després, formaran parelles i els repartirem dues situacions. Hi haurà dos papers tal com mostrem a continuació: A i B.

Personatge A:

Un dels teus millors amics/de les teves millors amigues ha tornat a renyir amb el seu xicot/la seva xicota de sempre. És una situació cíclica. No li dónes cap importància, perquè cada tres mesos estan igual. Et truca perquè està entestat/da a explicar-te tot el que els ha passat, amb tots els ets i uts. Ja li has prestat atenció unes quantes i no tens ganes de sentir la història, tenint en compte que es reconciliaran en un setmana... (POTS TRIAR ENTRE EL MOTIU 1 O 2 PER AVIAR-LO O AVIAR-LA).

- 1. Estàs fet/a pols, fa dies que programes ficar-te a la banyera i dissoldre aquella pastilleta de móres salvatges que fa tan bona olor... i que només fas un cop cada... ni te'n recordes... L'has d'aviar com sigui.
- 2. Estàs fet/a pols, l'únic que vols és seure al sofà i prendre't una cerveseta ben fresqueta i potser fer una repassada a uns quants canals de la tele.

Personatge B:

Estàs abatut/da, trist/a, desolat/da... i al mateix temps totalment capficat/da en el que ha passat. Has tornat a discutir amb el teu xicot/la teva xicota i necessites explicar-li a la teva millor amiga/al teu millor amic. Ho necessites, t'has de desfogar explicant-ho i potser també necessites que et doni alguna solució.

Activitat 4: Sóc profe

Objectiu: pràctica de l'entonació de cortesia **Funció lingüística:** rebutjar una petició

Nivell: B1

Temps: 30-40 minuts **Destreses:** interacció

Tipus d'agrupament a classe: en parelles

Material: pissarra

En aquesta activitat aprendrem a respondre a una petició que considerem inadequada i a la qual no volem accedir.

Desenvolupament

Per començar faríem una activitat prèvia en la qual demanaríem als alumnes que pensessin en situacions a la feina, al veïnat, entre coneguts, en les quals els han demanat alguna cosa que han considerat inadequada, perquè no era la seva feina o la tasca que els corresponia. Si als alumnes no se'ls acut res, podem preguntar-los si algun cop a la feina els ha demanat algú que els porti el cafè, no com a favor, si no com si fos la seva obligació.

Els comentem que situarem el context en una escola d'idiomes per a adults i el conflicte serà amb el professor. Plantegem el cas d'un professor a qui els alumnes demanen el següent:

- 1. que els faci fotocòpies del llibre en exclusiva perquè encara no el tenen a la llibreria.
 - 2. que els telefoni per dir-los què han fet a classe i quins són els deures.
 - 3. que els corregeixi exercicis extres.
 - 4. que els expliqui abans de començar la classe què han fet el dia abans.

Podem treballar aquestes situacions i també aquelles que els alumnes hagin comentat en l'activitat prèvia. Es pot fer una llista de peticions i escriure-la a la pissarra. A continuació ens mirem unes quantes estructures que es podrien posar en pràctica. Les que proposem són les següents:

Expressar una necessitat o un desig:

(de manera directa)

- Necessito que...
- Vull que...
- Oi que em podries...

Rebutjar una petició:

- No puc fer això que em demanes... perquè...
- Hauries de... perquè jo no t'ho podré fer...
- Et suggereixo que parlis amb un company del curs i ...
- Jo de tu... perquè jo...

Amb els alumnes es poden afegir algunes estructures més que serveixin per rebutjar una petició. Els mostrem també quin tipus d'entonació de cortesia hauríem d'utilitzar en aquests casos. Després els demanem que en parelles triïn qualsevol de les situacions de la llista i la interpretin. Per acabar, un parell de parelles poden representar el diàleg que han establert davant de tota la classe.

6 Conclusions

L'estudi de les funcions pragmàtiques del component melòdic posseeix escassa tradició en l'àmbit català. En el cas de les funcions corteses, el component fònic ha estat el gran oblidat. Els estudis realitzats eviten mencionar-lo, o ho fan de forma superficial, o bé tracten de buscar solucions descriptives de compromís. La investigació que hem realitzat tracta de sanar aquest buit classificant els trets de cortesia atenuant del català, a partir de l'anàlisi acústica i perceptiva d'un corpus de parla espontània i, al mateix temps, oferir unes propostes didàctiques pensades per a alumnes estrangers de català d'un nivell B1. Les propostes han estat dividides en dues parts. En la primera es presenta el contingut mitjançant una activitat de percepció que permet la introducció de diverses mostres d'entonació de (des)cortesia. A continuació es realitza una pràctica d'identificació per proporcionar un major control als alumnes sobre el contingut que volem treballar. Es tracta d'una pràctica en què el professor pot intervenir i estar pendent de les produccions dels estudiants per identificar possibles dificultats o problemes, tant en la producció com en la percepció de l'entonació de (des)cortesia. A la segona part, oferim la possibilitat de realitzar diferents activitats interactives centrades en l'entonació, que permeten també el repàs o la introducció de funcions lingüístiques que encaixen en les diferents situacions o contextos creats en les activitats.

Aquest tipus d'activitats centrades en l'entonació de cortesia atenuant no apareixen en els manuals d'aprenentatge de català per a estrangers i considerem que és un contingut que resultaria convenient introduir, especialment en els manuals de català amb finalitats professionals, ja que la cortesia, en moltes ocasions, és la clau per a la rendibilitat d'un negoci o, simplement, perquè la comunicació sigui eficaç. Un client, per exemple, que ha estat cortesament tractat, tindrà molta més confiança cap al dependent i l'empresa, el producte, la informació o el servei que hagi rebut. Per tant, la cortesia es converteix en una eina de comunicació que pot desencadenar un torrent d'informació importantíssima per al bon funcionament d'un negoci.

7 Bibliografia

- Brown, P., & Levinson S. (1987 [1978]). Politeness. Some universals in language usage. Cambridge University Press.
- Cantero Serena, F. J. (2002). *Teoría y análisis* de la entonación. Barcelona: Edicions de la Universitat de Barcelona.
- Cantero Serena, F. J., & Mateo, M. (2011). Análisis melódico del habla: complejidad y entonación en el discurso, *Oralia* 14, 105-127.
- Devís Herraiz, E. (2011). La entonación de (des)cortesía en el español coloquial, *Phonica* 7, 36-79.
- Font-Rotchés, D. (2007). *L'entonació del català*. Barcelona: Publicacions de l'Abadia de Montserrat.
- Hidalgo, A. (2009). Modalización (des)cortés y prosodia: estado de la cuestión en el ámbito hispánico. Boletín de Filología de la Universidad de Chile XLIV/1, 161-195.

- Lakoff, R. T. (1973). The logic of politeness; or minding your p's and q's, dins Papers from the seventh regional meeting of the Chicago Linguistic Society. Chicago: University Press, 292-305.
- Leech, G. (1983). *Principles of Pragmatics*. Londres: Longman.
- Nadeu, M. (2008). The contribution of tonal scaling to the perception of politeness in Catalan. Barcelona, Universitat Pompeu Fabra, dins
- http://prosodia.upf.edu/home/arxiu/tesi s/master/tesina_nadeu.pdf (26 gener 2015).
- Nadeu, M., & Prieto P. (2011). Pitch range, gestural information, and perceived politeness in Catalan. *Journal of Pragmatics* 43(3), 841-854.
- Payrató, L. (1988). *Català col·loquial*. València, Universitat de València.

Percepción auditiva y visual como estrategias de adquisición de sonidos nuevos del español por estudiantes rusos

DMÍTRIEVA, ANNA

1 Introducción

Cada idioma dispone de un inventario fonológico único. Es por ello que a la hora de aprender una lengua extranjera nos tengamos que enfrentar a la adquisición de sonidos nuevos. En este proceso es muy probable que estos sonidos "ajenos" se confundan con otros que ya existen en nuestra lengua materna, si son fonéticamente cercanos y por lo tanto perceptivamente parecidos. También hay que tener en cuenta el factor de la percepción visual ya que este puede influir en el estudiante a la hora de tener que elegir un sonido.

2 Premisas y objetivos

Varios fonetistas han demostrado que la transferencia negativa entre la lengua materna y la lengua extranjera puede estar condicionada por dos factores: la percepción "filtrada" y la base articulatoria (Trubetzkoy, 1973; Flege, 1987; Gil Fernández, 2007; Bondarko, 1987). El primero indica que el estudiante de un idioma extranjero puede percibir los sonidos de forma incorrecta debido a la "criba" fonológica, asociando inconscientemente los sonidos nuevos a los que ya existen en su lengua materna. En este caso los sonidos de la lengua extranjera pueden ser sustituidos en el habla del estudiante por los sonidos propios de su lengua materna. Si, por el contrario, la percepción es correcta, el

segundo factor indica que debido a la base articulatoria de su propio idioma le resulta difícil adquirir y producir los sonidos nuevos. En este sentido, el estudiante puede conformarse con una ligera modificación del sonido para que "suene" diferente pero no llegaría a la pronunciación correcta.

Según Flege (1987), en el proceso de adquisición de un idioma extranjero hay una serie de sonidos que provoca más problemas que otros y que con mayor frecuencia lleva a la transferencia negativa tanto a nivel perceptivo como a nivel articulatorio (los dos factores indicados anteriormente). Se trata de los denominados "sonidos similares": los sonidos que articulatoria y acústicamente se encuentran bastante cerca a los sonidos de la lengua materna y que por lo tanto son más fáciles de confundir o ignorar por los estudiantes.

A ello cabe añadir un factor más: la representación gráfica del sonido en los dos idiomas. Varias investigaciones (Escudero, 2010) han demostrado que la similitud gráfica de las letras que corresponden a los sonidos puede tanto facilitar como dificultar la adquisición/realización de estos sonidos. Cuando una palabra gráficamente parecida en los dos idiomas tiene el mismo significado pero se pronuncia de distinta manera, ocurre con mucha frecuencia que esta palabra induce al estudiante a escoger inconscientemente la pronunciación de su lengua materna, en vez de utilizar la pronunciación correcta, incluso de un nivel bastante alto de conocimiento del idioma extranjero. Al contrario, cuando las palabras carecen de similitud gráfica, el estudiante puede modificar su pronunciación correctamente según las normas del idioma extranjero.

En la investigación llevada a cabo se trataba de estudiar la transferencia negativa entre el ruso y el español basada en las diferencias de los dos sistemas fonológicos y en otros posibles factores. Uno de los objetivos fue el de analizar cómo un grupo homogéneo de estudiantes rusos de español realizaba en su habla sonidos que pueden considerarse "similares".

Se suponía que los estudiantes se apoyaban en distintas estrategias según qué grupo de sonidos, mostrando distintas realizaciones de estos, bien intercambiándolos por los sonidos de su lengua materna, bien modificándolos. Se han considerado dos condiciones: a) similitud perceptiva/articulatoria de los sonidos y b) similitud gráfica de las letras que representan estos sonidos en los dos idiomas.

3 Metodología

3.1 Corpus

El sistema fonológico del ruso tiene 36 consonantes frente a 19 consonantes en español. La característica más peculiar del ruso es la presencia de fonemas palatalizados. Esa palatalización cambia mucho la calidad del sonido: esto es debido a que en la producción de esos sonidos participa casi toda la superficie de la lengua. Por otra parte, a diferencia del español, el sistema ruso tiene numerosas parejas con el rasgo distintivo de sonoridad y no dispone de ningún sonido interdental.

	Bilabial	Labio- dental	Inter- dental	Dental	Alveolar	Palato- alveolar	Palatal	Velar
Nasal	m m ^j			(2)	n n ^j		n	1.1
Oclusivo	pp ^j bb ^j			tt ^j dd ^j			J	$\mathbf{k} \mathbf{k}^{\mathbf{j}} \mathbf{g} \mathbf{g}^{\mathbf{j}}$
Fricativo		ff ^j vv ^j	θ		$\frac{\mathbf{s} \mathbf{s}^{\mathbf{j}}}{\widehat{\mathbf{t}}\widehat{\mathbf{s}}} \mathbf{z} \mathbf{z}^{\mathbf{j}}$	$\frac{\int_{0}^{1} \int_{0}^{1} f(x)}{\int_{0}^{1} f(x)} dx$		x x ^j
Aproxi- mante							j	
Vibrante multiple					r r ^j			
Vibrante simple					r			
Lateral aproxi- mante					11 ^j		(Å)	

Tabla 1. Sistemas consonánticos del ruso y el español (Fuentes: Bondarko, 1998; Gil Fernández, 2007; Quilis, 1997; Hualde, 2005)

En la tabla 1 se presentan los dos sistemas consonánticos en comparación: en negro los sonidos que comparten los dos idiomas, en rojo los sonidos propios del español y en azul los sonidos propios del ruso.

Los seis sonidos "nuevos" del español son bastante cercanos a algún sonido ruso acústica o articulatoriamente:

- fricativa interdental θ vs. fricativa alveolar s
- nasal palatal n vs. nasal alveolar palatalizada n^j
- aproximante lateral l vs. aproximante lateral velarizada ł
- oclusiva palatal j · vs. aproximante palatal j
- vibrante múltiple r y simple r vs. alófonos libres del mismo fonema
- lateral aproximante palatal Λ ² vs. aprox. lat. velarizada palatalizada ¹

De estos seis grupos fueron elegidos los dos más peculiares (el de la interdental θ y el de la nasal palatal p) para analizar el fenómeno de la transferencia negativa para los sonidos similares. El corpus contenía 8 palabras con la nasal palatal y 70 con la interdental (incluyendo 17 palabras con la grafía "z" y 53 con la grafía "c"), dado que la segunda tiene considerablemente mayor frecuencia que la primera.

3.2 Participantes

En el experimento ha participado un grupo de estudiantes de nivel B1 de español, que pertenecían al mismo perfil socio-lingüístico y que durante 3 años estuvieron asistiendo al mismo curso de ELE para adultos en un centro de la lengua española en San-Petersburgo (Rusia), impartido por profesores hispanohablantes.

En total en el grupo había 7 estudiantes, de los cuales fueron preseleccionadas 4 mujeres que presentaban un grupo muy homogéneo (30-45 años, el mismo lugar de origen y residencia, el mismo nivel de estudios y de conocimiento de lenguas extranjeras, el mismo tiempo y lugar de estudios del español). La tabla 2 presenta el perfil de los informantes. Una persona de las no seleccionadas tenía el perfil semejante al grupo anterior, pero con un

¹ Según la clasificación de Gil Fernández (2007); se realiza como aproximante palatal j en varias posiciones. 2 En las zonas donde se mantiene.

conocimiento del español algo más amplio debido a su asistencia a cursos de corto plazo en España en varias ocasiones. Esta estudiante sirvió como un informante de control para el grupo de análisis. Todos los participantes han advertido que a lo largo de 2 años de aprendizaje del español nunca les habían corregido la pronunciación adquirida o les habían ofrecido ejercicios con el objetivo de trabajar la pronunciación. Por otra parte, una de las informantes también fue observada fuera del contexto del experimento y demostró los mismos resultados que los obtenidos durante esta investigación. Aparte ha sido grabada una informante española, de 25 años, lingüista, procedente de Sevilla, como un informante de control.

	Informante 1	Informante 2	Informante 3	Informante 5
Edad	33	45	44	40
Lugar de nacimiento y residencia	San Petersburgo	San Petersburgo	San Petersburgo	San Petersburgo
Formación	Superior	Superior	Superior	Superior
Actividad profesional	Economía	Marketing	IT	Administración empresarial
Estudios del español	"Adelante", 2.5 años	1 año clases <u>partic</u> ; "Adelante", 3 años	"Adelante", 2.5 años	"Adelante", 3 años
Conocimiento de otro idioma extranjera	Inglés	Inglés	Inglés, Francés (solo en colegio)	Inglés

Tabla 2. Informantes

3.3 El procedimiento de recogida de datos

El procedimiento de grabación del material incluía tres fases: lectura de un texto fonéticamente representativo, lectura de un listado de palabras de mayor interés fonológico, y habla espontanea en forma de respuestas a las preguntas basadas en el texto. La lectura fue elegida como la fuente principal por dos objetivos: 1) obtener los mismos datos para todas las informantes; y 2) considerar la influencia de la percepción visual de los elementos gráficos sobre la pronunciación.

Equipamiento: la grabación se realizó en las condiciones normales, con silencio de fondo, utilizando una grabadora digital Ritmix RR-600 (bitrate 128 kb/sec) y un micrófono externo situado a una distancia de 20 cm de la boca. Las grabaciones de todos los informantes se guardaron en formato *.wav (frecuencia de muestreo: 44100 Hz, cuantización de 16-bit) y se analizaron con Praat (Boersma, Weenink, 2014).

4 Resultados

4.1 Realización de la nasal palatal [n]:

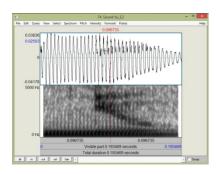


Figura 1. La palabra "ñu", pronunciada por la informante española: nasal palatal

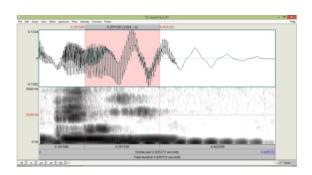


Figura 2. La palabra "ñu", pronunciada por la inform. 1: nasal palatalizada

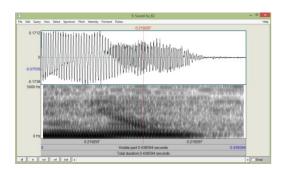


Figura 3. La palabra "ñu", pronunciada por la inform. 2: nasal palatalizada

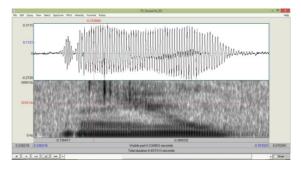


Figura 4. La palabra "ñu", pronunciada por la inform. 5: nasal palatalizada acompañada por una palatal aproximante

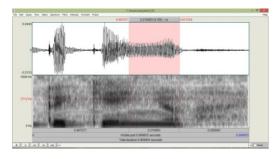


Figura 5. La palabra "pequeño", pronunciada por la inform. 2: nasal palatalizada

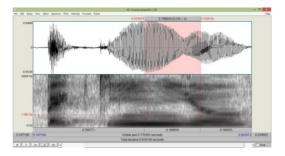


Figura 6. La palabra "pequeño", pronunciada por la inform. 5: nasal palatalizada acompañada por una aproximante palatal

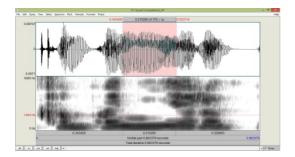


Figura 7. La palabra "compañeros", pronunciada

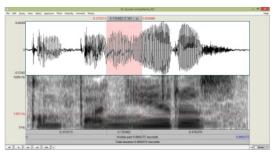


Figura 8. La palabra "compañeros", pronunciada

por la inform. 1: nasal palatalizada, acompañada por una aproximante palatal

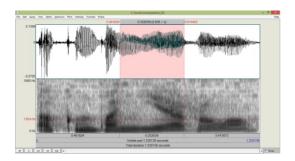


Figura 9. La palabra "acompañarlo", pronunciada por la inform. 2: nasal palatalizada, acompañada por una aproximante palatal

por la inform. 5: nasal palatalizada, acompañada por una aproximante palatal

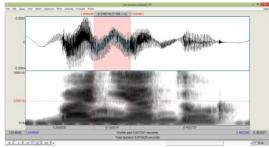


Figura 10. La palabra "bañador", pronunciada por la inform. 1: nasal palatalizada, acompañada por una aproximante palatal

4.2 Realización de la fricativa interdental $[\theta]$:

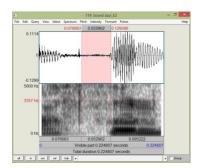


Figura 11. La palabra "azul", pronunciada por la inform. española: fricativa interdental sorda

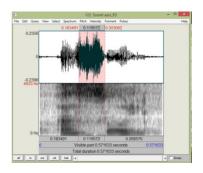


Figura 12. La palabra "azul", pronunciada por la inform. 3: fricativa alveolar sorda

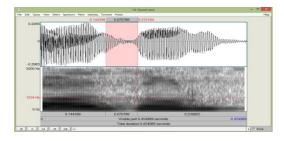


Figura 13. La palabra "azul", pronunciada por la inform. 2: fricativa alveolar sonora

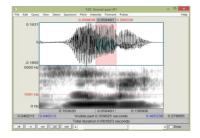


Figura 14. La palabra "azul", pronunciada por la inform. 1: fricativa alveolar sonora

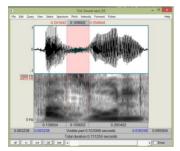


Figura 15. La palabra "azul", pronunciada por la inform. 5: fricativa alveolar sonora

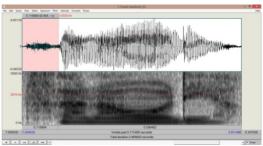


Figura 16. La palabra "zanahoria", pronunciada por la inf. española: fricativa interdental sorda

Figura 17. La palabra "zanahoria", pronunciada por la inform. 1: fricativa alveolar sonora

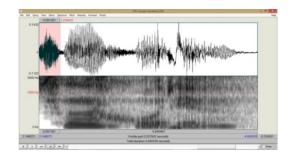


Figura 18. La palabra "zanahoria", pronunciada por la inform. 3: fricativa alveolar sorda

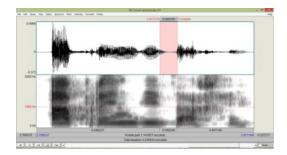


Figura 19. La palabra "aprendizaje", pronunciada por la inform. 1: fricativa alveolar sonora

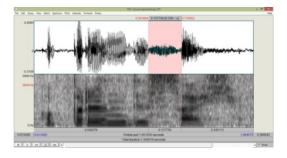


Figura 20. La palabra "aprendizaje", pronunciada por la inform. 5: fricativa alveolar sorda

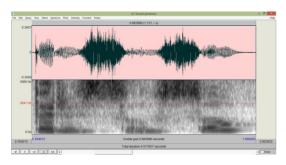


Figura 21. La palabra "perezoso", pronunciada por la inform. 2: fricativa alveolar sorda

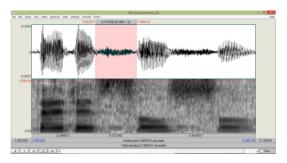


Figura 22. La palabra "perezoso", pronunciada por la inform. 5: fricativa alveolar sorda

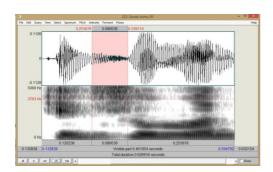


Figura 23. La palabra "zumo", pronunciada por la inform. 1: fricativa alveolar sonora

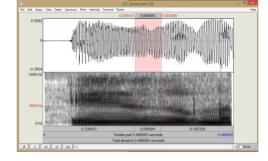


Figura 24. La palabra "zumo", pronunciada por la inform. 2: fricativa alveolar sonora

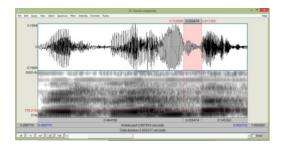


Figura 25. La palabra "vergüenza", pronunciada por la inform. 2: fricativa alveolar sonora

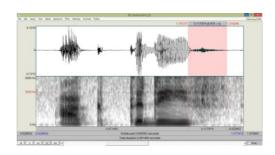


Figura 26. La palabra "actriz", pronunciada por la inform. española: fricativa interdental sorda

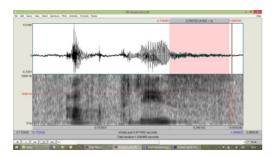


Figura 26. La palabra "actriz", pronunciada por la inform. 5: fricativa alveolar sorda

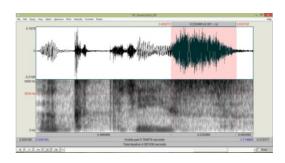


Figura 27. La palabra "actriz", pronunciada por la inform. 3: fricativa alveolar sorda

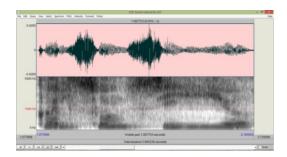


Figura 28. La palabra "calzoncillo", pronunciada por la inform. 3: dos fricativas alveolares sordas

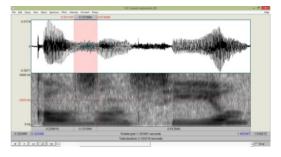


Figura 29. La palabra "calzoncillo", pronunciada por la inform. 5: fricativa alveolar sonora y fricativa alveolar sorda

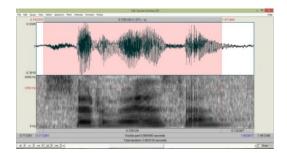


Figura 30. La palabra "cerveza", pronunciada por la inform. 5: fricativa alveolar sorda y fricativa alveolar sonora

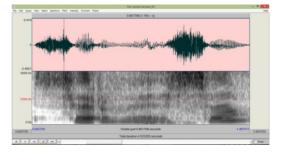
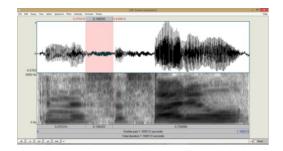


Figura 31. La palabra "cerveza", pronunciada por la inform. 3: dos fricativas alveolares sordas



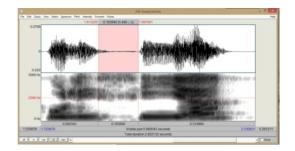


Figura 32. La palabra "innecesario", pronunciada por la inform. 2: fricativa alveolar sorda

Figura 33. La palabra "policía", pronunciada por la inform. 1: fricativa alveolar sorda

5 Discusión de resultados

5.1 Realización de [n]

El ruso dispone solamente de un fonema palatal, que es aproximante /j/. Por otra parte, en ruso disponemos de un fonema alveolar nasal palatalizado /n^j/ que al oído poco entrenado le puede parecer una nasal palatal [n] española. Algunos estudiantes han demostrado no poder distinguir la diferencia ya que no se entrenaron para oírlo y producirlo correctamente. Por lo tanto han seguido utilizando el sonido de su lengua materna, la [n^j] palatalizada. Otros han demostrado percibir que hay alguna diferencia entre el sonido propio de su idioma y el sonido español, han modificado el sonido al producirlo para que sea distinto a la nasal palatalizada rusa, y han empleado la aproximante palatal [j] para acompañar a la [n^j]. Las dos palabras donde se han producido más casos de aproximación a una pronunciación correcta han sido: la forma monosílaba " $\tilde{n}u$ " [nu] donde la única consonante ha concentrado en sí todo el esfuerzo de pronunciación; y la forma altamente frecuente "pequeño" [pekeno].

Estrategias: los resultados han demostrado que los estudiantes utilizan la estrategia de percepción auditiva: 1) en el caso de no percibir la diferencia, utilizan el sonido alveolar palatalizado (sustitución); 2) en el caso de percibir la diferencia, utilizan un sonido alveolar palatalizado acompañado por un sonido palatal aproximante (modificación).

5.2 Realización de $[\theta]$

Nos enfrentamos con dos problemas: 1) el ruso no dispone de ningún fonema interdental; 2) la ortografía rusa dispone de una letra "3" (en la transliteración "z") que corresponde a la fricativa alveolar sonora [z].

También hay que tener en cuenta que en ruso los dos fonemas alveolares fricativos (sordo /s/ y sonoro/z/) aparecen como alófonos combinatorios uno de otro: ante una sorda o en el final absoluto todas se ensordecen, ante una sonora, todas se sonorizan. Por lo tanto, los dos sonidos tienen un enlace muy estrecho en la base articulatoria del hablante ruso.

Los informantes se han dividido en dos grupos. El primer grupo ha sustituido en la mayoría de los casos los sonidos interdentales por un sonido

alveolar sordo [s], para ambas grafías. El segundo grupo ha sustituido la interdental representada por una grafía "z" por una alveolar fricativa sonora (o en algunos casos una [s] considerablemente sonorizada), salvo en la posición del final absoluto ("pez", "actriz") lo que corresponde a la pauta de realización de este sonido en ruso en esta posición de ensordecimiento obligatorio. En la posición final y en todos los casos cuando el sonido interdental esté presentado con una grafía "c", este grupo de estudiantes ha sustituido la interdental por una alveolar sorda [s], igual que el primer grupo. Ambos grupos han utilizado un sonido palatalizado ante las vocales anteriores [i], [e].

Estrategias: El primer grupo ha utilizado mayoritariamente la estrategia de percepción auditiva. El segundo grupo ha utilizado la estrategia de percepción visual. Ambos grupos han demostrado que no perciben el sonido interdental correctamente y no han conseguido modificar su pronunciación para acercarse a la pronunciación correcta española (sustitución).

Por otra parte, es importante tener en cuenta que el fenómeno de "seseo" está bastante extendido en algunas partes del mundo hispanohablante y que en algunas circunstancias los profesores, procedentes de estas zonas dialectales, pueden carecer del sonido interdental en su habla e ignorar su carencia en el habla de sus estudiantes. En cualquier caso, lo que nos interesa resaltar es que la aparición de un sonido alveolar sonoro en lugar de uno sordo se puede considerar como una falta de destreza articulatoria por parte del estudiante ruso que aprende español ya que interfiere la pronunciación correcta de las palabras y añade al acento un rasgo muy destacado.

6 Conclusión

Los resultados del experimento han apoyado la hipótesis y han demostrado que los estudiantes eligen estrategias distintas al enfrentarse a estos dos sonidos nuevos, similares a los de su lengua materna (figura 34). En la realización de la nasal palatal, los participantes de manera unánime han utilizado el sonido más parecido de su sistema materno, apoyándose en su oído (sustitución y modificación), mientras que en la interdental los participantes se han basado tanto en su oído como en la forma gráfica que veían (sustitución).

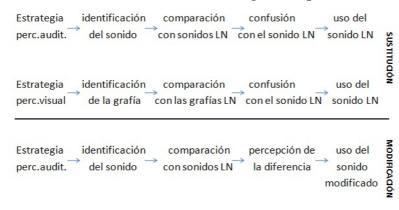


Figura 34. Estrategias de percepción

Estas estrategias demuestran que los sonidos nuevos del idioma extranjero requieren una atención muy especial a la hora del aprendizaje, ya que los estudiantes que llegan a un nivel intermedio de conocimiento del idioma extranjero siguen apoyándose en su lengua materna cuando pronuncian los

sonidos, y se basan tanto en su oído "filtrado" como en la grafía familiar. Sin trabajar cuidadosamente la pronunciación y la percepción de los sonidos nuevos resulta muy difícil adquirirlos correctamente.

7 Bibliografía

- Boersma, P., & Weenink, D. (2014). *Praat:*doing phonetics by computer
 [Computer program]. Versión 5.4,
 recuperada el 4 de 10 de 2014, de
 http://www.praat.org/
- Bondarko, L. V. (1987). *Interferentsiya* zvukovyx sistem. Leningrado: Izdatelstvo Leningradskogo universiteta. (esp.: *Interferencia de los sistemas fónicos*).
- Escudero, P., & Wanrooij, K. (2010). The effect of L1 orthography on nonnative vowel perception. *Language and Speech*. 53, 343-65.
- Flege, J. E. (1987). The production of 'new' and 'similar' phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*. 15, 47-65.
- Gil Fernández, J. (2007). Fonética para profesores de español: de la teoría a la práctica. Madrid: Arco/Libros.
- Trubetzkoy, N. S. (1973). *Principios de fonología*. Madrid: Cincel.

La presencia de la lengua escrita en los procesos de aprendizaje y adquisición de la pronunciación del español como lengua extranjera: ¿podríamos hablar de interferencia?

GIRALT LORENZ, MARTA; CANTERO SERENA, FRANCISCO JOSÉ

1 Introducción

Los procesos que acontecen en el aprendizaje y adquisición de una lengua extranjera en un contexto formal instruccional de enseñanza, suelen estar circunscritos a enfoques pedagógicos en los que se trabaja con la lengua oral y la lengua escrita como medios lingüísticos para conseguir que los aprendices desarrollen y alcancen una competencia comunicativa de la lengua que están aprendiendo.

A pesar de que el aprendizaje de la lengua oral y la adquisición de las competencias fónicas se trabajan principalmente mediante la oralidad, la lengua escrita siempre está presente en dichos procesos. Así, nuestro estudio presenta datos de una investigación que se realizó con aprendices de español lengua extranjera de nivel inicial que siguieron un enfoque oral en un contexto formal

de instrucción desde el inicio de su proceso de aprendizaje y adquisición de la lengua.

Los datos con los que contamos (grabaciones de los alumnos en distintos momentos de su proceso de aprendizaje y adquisición de la pronunciación) muestran evidencias sobre el desarrollo de las competencias fónicas de los alumnos, así como de factores relacionados con el aprendizaje de la pronunciación. A partir del análisis de las grabaciones y de su comparación, puede testimoniarse si existe o no una interferencia de la lengua escrita durante los procesos de aprendizaje y adquisición de la pronunciación, al mismo tiempo que se establecen evidencias sobre los procesos de adquisición de la pronunciación en E/LE.

Así, las dos preguntas de investigación que nos planteamos al iniciar el estudio son las siguientes:

- ¿Qué procesos de aprendizaje y adquisición de la pronunciación en E/LE se dan cuando se sigue un enfoque pedagógico basado en la oralidad?
- ¿Existe alguna diferencia en las producciones orales de los estudiantes cuando se ha seguido un enfoque oral y cuando se ha seguido un enfoque en el que se usa la lengua escrita?

1.1 Factores de aprendizaje y adquisición de la pronunciación

Al hablar de los mecanismos que entran en juego en la adquisición de una lengua debemos mencionar la reestructuración, la precisión y la fluidez (ritmo y entonación). En la enseñanza de la pronunciación mediante un enfoque oral, el énfasis recaería en la fluidez y por tanto, nos encontraríamos ante actividades y tareas de aprendizaje que orientarían al alumnado hacia el desarrollo de la fluidez, aunque no por eso abandonaríamos los mecanismos de precisión, o lo que es lo mismo, las tareas orientadas hacia el desarrollo de la corrección. La corrección y la fluidez no son contrarios sino complementarios (Brumfit, 1984; Skehan, 1993).

Un enfoque basado en la oralidad va a perseguir el desarrollo de la fluidez. Los aprendices necesitan la oportunidad de comunicarse para poder desarrollar su fluidez oral (Ellis, 1990). Siendo la fluidez uno de los mecanismos al que debe darse más énfasis en el proceso comunicativo para la adquisición y aprendizaje de la lengua oral, la destreza de la expresión oral será la que cobre prioridad sobre las otras cuando trabajemos mediante un enfoque oral.

El desarrollo de la fluidez puede considerarse uno de los primeros logros en los estadios iniciales del aprendizaje de la pronunciación y la lengua oral, mientras que la precisión y la corrección pueden alcanzarse una vez el aprendiz tenga la posibilidad de prestar atención a la forma de los elementos gramaticales, fonológicos y discursivos en sus propias producciones orales (Douglas Brown, 2001).

Si nos remitimos a las teorías de adquisición de segundas lenguas, ya Krashen y Terrell, en 1983, intuían que para la adquisición de la pronunciación, ésta no debía enseñarse a partir de un sistema de reglas sino que precisamente la práctica oral es lo que permite la adquisición en pronunciación, teniendo en cuenta que, primeramente, los estudiantes reciben un *input* comprensible y que, después tienen la oportunidad de practicar oralmente con la lengua:

Toda la práctica oral que se realice al trabajar mediante un enfoque oral va a estar siempre exclusivamente basada en la oralidad, es decir, se va a prescindir de la lengua escrita. Algunos autores (Celce-Murcia, 1996) ya apuntan que para

coger fluidez debe practicarse con actividades sin lengua escrita. A partir de la oralidad se potencia no solo la producción oral en los aprendices sino que también se desarrollan las estrategias de percepción, es decir, las estrategias de comprensión oral. Sabemos que en el proceso de comprensión oral el hablante / oyente (en su L1 y por consiguiente en L2) realiza un proceso de restitución para entender los enunciados del discurso oral. Esta restitución por un lado se apoya en el contexto (Poch, 2000) pero, por otro, se apoya en la representación mental de la forma escrita de las palabras que los hablantes/oyentes poseen. A consecuencia de todo esto, la percepción auditiva está fuertemente mediatizada por la escritura (Cantero, 1997). Trabajando sin la lengua escrita y mediante la oralidad, no solo evitaremos la interferencia sino que será posible desarrollar estrategias de comprensión en los hablantes/oyentes que aprenden español.

Asimismo, también es necesario apuntar que la interacción es un elemento crucial en todo el proceso de aprendizaje y adquisición de la lengua oral y la pronunciación. La interacción es la actividad comunicativa más importante puesto que la adquisición de la lengua es producto, esencialmente, de la interacción comunicativa entre dos interlocutores (Cantero, 2008). Así, la competencia interactiva que los alumnos vayan desarrollando va a ejercer de eje y motor de las competencias comunicativas específicas (competencia lingüística, competencia discursiva, competencia cultural y competencia estratégica) constituyendo el formato básico de la comunicación oral. A las dimensiones fónicas de cada una de estas competencias se les ha llamado competencias fónicas (Cantero, 2014), siendo estas, unas de las competencias que tienen un papel central al hablar del aprendizaje y adquisición de la pronunciación mediante un enfoque oral.

1.2 La mediación de la lengua escrita como factor crítico

Entendemos la mediación de la lengua escrita como un elemento que interfiere de manera negativa en la adquisición y el aprendizaje de la pronunciación de una lengua extranjera, en nuestro caso el español. Nótense aquí una cuestión de cierta importancia. Consideramos que la mediación de la lengua escrita como medio lingüístico es un factor crítico para la enseñanza y el aprendizaje/ adquisición de la pronunciación de una lengua extranjera pero no lo es para el de otras destrezas lingüísticas.

Convendría hacer mención también de la llamada competencia ortoépica, competencia que se basa en el aprendizaje de las reglas de las correspondencias entre grafías y sonidos. Las distintas perspectivas que se han preocupado de atender el tema de la adquisición fónica y de los procesos que intervienen en ésta nunca han tenido en cuenta la forma en que se realiza el aprendizaje que va a conducir a la adquisición fónica y, por tanto, ninguna de ellas ha contemplado la posibilidad de que el factor que influye negativamente en la adquisición fónica sea algo externo al aprendiz (Cantero, 1999).

En el aprendizaje de una lengua extranjera se ha usado siempre como mediación la lengua escrita, es decir, los alumnos no suelen estar expuestos únicamente a la lengua oral. Como afirma Torremocho (2004), en la mayoría de libros de texto de e/le, la lengua oral suele presentarse mediante un texto escrito y la lectura de éste, al mismo tiempo que se escucha.

Los hablantes nativos de una lengua acostumbran a relacionar los sonidos de su propia lengua con el código escrito y sistema de reglas ortográficas de ésta (Port, 2007), así como también lo hacen en el momento de enfrentarse al aprendizaje de una lengua extranjera (Zampini, 1994; Stokes, 2001). En este

contexto de aprendizaje y de adquisición es donde hablaríamos de competencia ortoépica.

Además de estas cuestiones que ponen de relieve la relación grafía-sonido, la aproximación a la lengua oral y pronunciación mediante el código escrito inhibe al aprendiz y no le deja espacio para el desarrollo de sus producciones orales y el desarrollo de su competencia discursiva. En los contextos formales de aprendizaje de una lengua extranjera la mediación de la lengua escrita puede actuar como refuerzo a la corrección, preponderando la forma sobre el significado. Los aprendices, especialmente cuando priorizan la corrección frente a la fluidez, suelen apoyarse en la lengua escrita en el momento en el que se disponen a realizar una intervención oral, hecho que deja entrever cómo la mediación de la escritura puede ser un pequeño obstáculo para la aparición de secuencias potencialmente adquisitivas. La lengua escrita ofrece al alumno de forma directa la forma lingüística de lo que quiere producir oralmente, sin posibilidad de llevar a cabo ningún tipo de estrategia cognitiva o de impidiendo, aparición comunicación, además, la de intercambios comunicativos espontáneos.

Consideramos por tanto que una de las grandes interferencias que sufre la interlengua fónica del aprendiz en el aprendizaje y la adquisición de la lengua oral, se produce a consecuencia del uso que se hace de la lengua escrita como medio y ruta de acceso a la oralidad en un contexto de aprendizaje formal.

2 Metodología

Para este estudio se han manejado datos de tipo cuantitativo, que enmarcamos dentro del paradigma de investigación *quasi*-experimental, los cuales se recogieron en las grabaciones que se realizaron de las pruebas orales. Para el análisis de los datos de las grabaciones y su valoración se han establecido 5 categorías de análisis.

Las cinco categorías que van a tenerse en cuenta para el análisis del desarrollo y adquisición de las tres competencias: productiva, perceptiva e interactiva, (MECR: 2001), son: inteligibilidad; fluidez; interacción; creatividad y pronunciación. Para cada uno de los criterios se han utilizado tablas de niveles de análisis. Esos niveles de análisis se han planteado en base a una escala graduada. De esta forma, la actuación de los alumnos se ubica en distintos puntos de un *continuum* en función de los 5 criterios mencionados.

2.1 Contextualización del estudio

La investigación-acción de donde proceden nuestros datos, se basa en una acción didáctica que se llevó a cabo durante un semestre con alumnos adultos de español del curso de Inicial 1 (principiantes absolutos). La acción se implementó en 4 grupos distintos, con un total de 50 alumnos. El número de horas de clase fue de 30 horas por grupo y la propuesta didáctica implementada siguió un enfoque oral durante las 10 primeras horas de clase de forma íntegra, mientras que en las 20 siguientes horas de clase se alternó entre sesiones en la que se trabajaba mediante un enfoque oral y sesiones en la que se hacía uso de la lengua escrita.

En cuanto a la recogida de datos, a lo largo del desarrollo de la acción didáctica se realizaron diferentes grabaciones de las producciones de los alumnos.

2.2 Corpus de datos

Las pruebas orales se efectuaron en los cuatro grupos de Inicial 1 en los que se llevó a cabo la acción didáctica. Al finalizar las seis primeras sesiones en las que no se usó el soporte escrito, se realizó una prueba oral para empezar a evaluar el nivel de lengua oral aprendido / adquirido durante las seis primeras sesiones. Esta prueba evaluadora consistió en una tarea final en la que los alumnos dramatizaron un diálogo en parejas, intercambiando la información personal de un personaje ficticio o famoso que ellos mismos habían creado. Se decidió tomar como muestra de registros orales la representación de un diálogo real puesto que se considera que el diálogo es la forma clásica y más sencilla de la comunicación discursiva (Calsamilgia & Tusón, 1999).

Se consideró que cada una de las intervenciones de los alumnos formaba una unidad dialogal (Calsamiglia & Tusón, 1999) con una interacción, una secuencia o episodio y un intercambio. Precisamente en este intercambio es donde se halla una de las formas más típica de los intercambios orales dentro del discurso oral: los pares adyacentes (Sacks, Schegloff & Jefferson, 1974), formato que se tuvo en cuenta en el momento de analizar las grabaciones y las intervenciones de los alumnos.

Para registrar los diálogos, los alumnos se sentaron de dos en dos y la profesora colocó una grabadora audio-digital frente a ellos, en la mesa, con lo que a medida que iban realizando el diálogo este quedaba registrado. La grabadora que se utilizó era de tipo digital con lo que los datos obtenidos fueron de gran calidad, sin ruido ambiental ni interferencias de ningún tipo.

La segunda prueba oral, también se registró en grabadora siguiendo el mismo procedimiento. Los alumnos de nuevo tenían que escenificar un diálogo, no solo intercambiando su información personal sino que podían intercambiar información sobre otros contenidos trabajados durante el curso. Estas segundas grabaciones se realizaron en la última sesión de curso y los alumnos escenificaron el mismo diálogo en los cuatro grupos.

3 Resultados

3.1 Análisis y comparación de las grabaciones

El análisis de las grabaciones constituye una comparación global entre las dos grabaciones tras haber aplicado las 5 categorías de análisis a cada una de las producciones de los alumnos en ambas grabaciones. Para llevar a cabo la comparación se ha tenido en cuenta la valoración global de las dos grabaciones desde el punto de vista de los valores que se ha otorgado a los distintos criterios de análisis. Por este motivo, iremos presentando la valoración general de cada uno de los cinco criterios. Esta comparación nos va a obligar a realizar comparaciones específicas: comparación de los distintos rasgos de interacción que los informantes llevan a cabo en la primera y en la segunda grabación; comparación de los distintos rasgos de creatividad que aparecen en sus discursos durante ambos registros; comparación de los fenómenos fonéticos, de transferencia de la L1 y L2, y de interferencia de la lengua escrita.

La inteligibilidad del discurso oral de los informantes es buena en la primera grabación, tras 5 sesiones de clase, y es total en la segunda grabación tras 15 sesiones de clase. La fluidez ha sido bien valorada tanto en la primera como en la segunda grabación. El único rasgo en detrimento de la fluidez, en sendas grabaciones, han sido las autocorrecciones y las autorreparaciones que los

aprendices han llevado a cabo. Este hecho, a pesar de repercutir en la fluidez, es considerado como un indicio de desarrollo de la competencia comunicativa y del proceso de desarrollo de adquisición de la lengua.

La interacción ha estado presente en las dos grabaciones, los aprendices han sido capaces de interactuar en los dos intercambios registrados. Han demostrado poseer estrategias de interacción y haber iniciado el proceso de desarrollo de la competencia interactiva. Presentamos a continuación la tabla 1 en la que puede compararse el uso de elementos y recursos para interactuar que los alumnos han utilizado en la primera y la segunda grabación. Puede apreciarse como en la segunda grabación, si tenemos en cuenta los 4 grupos, el uso de los recursos para interactuar y el de los marcadores discursivos es mayor. Los aprendices han ampliado el tipo de recursos utilizados, aunque la diferencia no sea demasiado grande.

PRIMERA GRABACIÓN		SEGUNDA GRABACIÓN	
ELEMENTOS UTILIZADOS (Marcadores discursivos)	RASGOS DE INTERACCIÓN	MARCADORES DISURSIVOS UTILIZADOS	RASGOS DE INTERACCIÓN
	Toma del turno de palabra e iniciativa en los intercambios e intervenciones		Toma del turno de palabra y de la iniciativa en el intercambio
muy bien, muy bien	Muestra de interés ante lo que le dice su interlocutor	repeticiones, vocalizaciones	Muestra de interés hacia lo que comunica el interlocutor
Repetición	Aclaraciones si el interlocutor no entiende	muy bien, muy, bien, Ah! Sí!, ¿si?	Preguntas de aclaración
Repetición para marcar que se ha entendido, vale, vale; Ok; Oh ¡Si!; vocalizaciones.	Confirmación de que han entendido la información de su interlocutor		Confirmación de que se ha entendido
Bien, bien; Ah! ¿Si?; Ahh!!; ¡qué bien!; entonaciones	Muestra de interés e implicación en la conversación		Muestra de interés ante lo que le dice su interlocutor
También; vocalizaciones.	Marca de énfasis en las informaciones coincidentes		Marca de énfasis en las informaciones coincidentes
	Inclusión de preguntas para hacer que el intercambio sea más significativo		
(¡Oh!Periodista!).	Muestras de sorpresa		Muestras de sorpresa
			Colaboración ante un lapsus o bloqueo del interlocutor
Encantado;	Respeto de los guiones culturales en relación a los ofrecimientos y aceptación de éstos y en las presentaciones, sobre todo en el uso de elementos de cortesía.		Respeto de las normas sociales y de la cortesía.
(por ejemplo, no preguntar la edad)	Respeto a ciertos guiones culturales		

Tabla 1. Comparación del uso de los distintos recursos para interactuar

En ambas grabaciones los informantes presentan un perfil de interlocutores cooperativos y colaborativos, aunque es en la segunda cuando esto puede considerarse como un rasgo totalmente generalizado. En cuanto a la creatividad, puede afirmarse que se ha producido un avance desde la primera grabación a la segunda grabación, especialmente si se tiene en cuenta el número de alumnos que presentan rasgos de creatividad en su discurso, y los valores que han alcanzado en el análisis de este criterio.

A continuación vamos a mostrar en la tabla 2 elementos que reflejan la creatividad en los intercambios comunicativos de los aprendices. Cada una de

las listas corresponde a la primera y a la segunda grabación. De esta forma, podremos observar si el repertorio es mayor, menor o igual entre sendas grabaciones.

Rasgos de creatividad Rasgos de creatividad SEGUNDA GRABACIÓN PRIMERA GRABACIÓN - Sentido del humor. - Muestras de humor y de complicidad entre los - Realización de entonaciones creativas. - Improvisaciones. interlocutores. - Entonaciones muy creativas - Comentarios e intervenciones que están fuera del para expresar sentimiento, emoción. -Espontaneidad y naturalidad en las intervenciones. - Intervenciones que están fuera del guión del intercambio que los alumnos han preparado. En algunos casos, los contenidos lingüísticos que usan los informantes se han trabajado en el curso pero existen otros casos en los que los aprendices están usando contenidos que no se han cubierto en clase. - La forma como llevan a cabo la interacción los aprendices se ha considerado también como un rasgo de creatividad puesto que es un reflejo de cómo se están poniendo en marcha muchos procesos de comunicación.

Tabla 2. Elementos que reflejan creatividad

Tras la comparación, encontramos que los rasgos de creatividad que aparecen en los 4 grupos en la primera grabación y en la segunda grabación, son muy similares. La diferencia fundamental es que en la segunda grabación hay muchos más informantes que presentan rasgos de creatividad en su discurso y, por tanto, muchos más aprendices que demuestran sentirse muy cómodos realizando el intercambio comunicativo.

El último criterio que vamos a comentar es el de la pronunciación. Al comparar las dos grabaciones es posible hallar diferencias fundamentales. Por un lado, en la primera grabación no se ha detectado ninguna interferencia de la lengua escrita. En cambio, en la segunda grabación, algunos de los informantes presentaban en sus producciones orales interferencia de la lengua escrita.

INTERFERENCIA DE LA LE
Aspiración de la letra "h"
Realización de la [z] como [tʃ]: [z] 〈[tʃ]
Realización de la fricativa velar [j] como sonora: [c] 〈 [ʒ].
Realización de la "e" como [i]: [e] ([i]
Realización de la le como [1]. [e] ([1]
Realización de la [l] como [l].

Tabla 3: Interferencia de la lengua escrita

Junto a esta interferencia, presente únicamente en la segunda grabación, se ha detectado también en la segunda grabación que algunos de los informantes poseían ciertas transferencias en la pronunciación de algunos sonidos que provenían de su L2, que en todos los casos era el francés. Concretamente la transferencia provenía de la pronunciación de las vocales. Sin embargo, en la primera grabación no se ha encontrado ningún ejemplo de esta transferencia. En cuanto a los fenómenos que se han dado en sendas grabaciones, puede hablarse de la transferencia de determinados rasgos de la L1 de los estudiantes, en el caso del inglés (aunque también existe algún informante con interferencia del francés como L1, el alemán o el polaco).

Estos rasgos en la pronunciación de los informantes son los predecibles desde la perspectiva de la fonética contrastiva y han aparecido tanto en la primera como en la segunda grabación. Podríamos establecer como fenómenos más repetidos en las muestras analizadas los que aparecen en la tabla 4.

Realización de la [r] retrofleja.	
Diptongaciones.	
Fricativización de los sonidos [b] y [t].	
Realización de la "ll" como "l": [l] ([l].	
Realización de la "ñ" como "n": [η]{[n].	

Tabla 4: Rasgos de pronunciación de los informantes

4 Conclusión

Como conclusión a los resultados del estudio presentado, podemos constatar que al seguir un enfoque oral, los alumnos aprenden a hablar y desarrollan los distintos elementos relacionadas con la pronunciación: fluidez; interacción; creatividad, al mismo tiempo que van adquiriendo una pronunciación inteligible y acurada sin interferencias de la lengua escrita. Se ha evidenciado como la interlengua de los aprendices no posee interferencias de la lengua escrita cuando se sigue un enfoque oral mientras que esas interferencias sí aparecen cuando se usa la mediación de la lengua escrita. Los procesos de aprendizaje que se activan en el aula al seguir un enfoque pedagógico basado totalmente en la oralidad, favorecen y promueven la iniciación de la adquisición de la pronunciación, sin que haya ninguna interferencia de la lengua escrita. En definitiva, un enfoque oral fomenta la adquisición de la lengua oral y la pronunciación, o lo que es lo mismo, pone en funcionamiento el desarrollo de la competencia fónica, como parte del engranaje de las competencias comunicativas específicas en una lengua extranjera.

5 Bibliografía

- Brumfit, C. J. (1984). *Communicative* methodology in language teaching. Cambridge: Cambridge University
- Cantero, F. J. (1997). De la fonética instrumental a la enseñanza de la pronunciación, en F. J Cantero, A. Mendoza & C. Romea (eds.), Didáctica de la lengua y la literatura para una sociedad plurilingüe del siglo XXI. Barcelona: Publicacions de la Universitat de Barcelona.
- Cantero, F. J. (1999). La mediació lectoescriptora en l'ensenyament de llengües, en *Jornades sobre* l'ensenyament de la llengua oral. Barcelona: Universitat de Barcelona.
- Cantero, F. J. (2008). Complejidad y competencia comunicativa. *Revista Horizontes de Lingüística Aplicada, v. 7, n.º 1, 71-87.*
- Cantero, F. J. (2014). Adquisición de competencias fónicas, en Y. Congosto, A. Salvador, M. L. Montero Curiel (eds.), Fonética

- *experimental. Educación Superior e Investigación.* Colección Bibliotheca Philologica. Madrid: Arco/Libros.
- Casalmiglia, H., & Tusón, A. (1999). *Las cosas del decir*. Barcelona: Ariel.
- Celce-Murcia, M., Brinton, D. M, & Goodwin, J. M. (1996). Teaching Pronunciation. A Reference of English to Speakers of Other Languages. Cambridge University Press.
- Council for Cultural Cooperation (2001).

 Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Strasbourg: Council of Europe. (2002): Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación. Instituto Cervantes. En http://cvc.cervantes.es/obref/marco
- Douglas Brown, H. (2001). Teaching by Principles. An Interactive Approach to Language Pedagogy. Adison Wesley Longman: NY.
- Ellis, R. (1990). *Instructed Second Language Acquisition*. Oxford: Oxford University Press.
- Krashen, S. D., & Terell, T. D. (1983). *The* natural approach: Language acquisition in the classroom. London: Prentice Hall Europe.
- Poch, D. (2000). Sensación física y realidad mental: de la onda sonora al significado de los enunciados. *CARABELA*, 49, 5-15, Madrid: S.G.E.L.
- Port, R. E. (2007). The graphical basis of phones and phonemes en Ocke-Schwen Bohn & Murria J. Munro (Eds.) Language Experience in Second Language speech Learning. Aarhus University/Simon Fraser University.
- Sacks, H., Schegloff, E. M, & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, *50*, 696-735.
- Skehan, P. (1993). A Framework for the Implementation of Task-Based Learning. Plenary Lecture: IATEFL, Swansea. Conferencia Mecanografiado.
- Torremocho, I. (2004). Los textos orales en clase de ELE. *Glosas Didácticas*, 12, 121-126.
- Zampini, Mary L. (1994). The Rol of Native Language Transfer and Task Formality in the Acquisition of Spanish Spirantization. *Hispania*, 77: 470-81.

Análisis descriptivo de las estrategias de aprendizaje y creencias en los procesos de adquisición y aprendizaje de la pronunciación en E/LE

GIRALT LORENZ, MARTA

1 Introducción

En este artículo se presentan datos que provienen de un estudio que se realizó con estudiantes de E/LE de nivel inicial con el objetivo de analizar y describir qué tipo de estrategias de aprendizaje y de creencias emergen durante los procesos de adquisición y aprendizaje de la pronunciación en E/LE.

Partimos de la premisa de que enseñar pronunciación se corresponde con enseñar lengua oral. Enseñar a pronunciar no puede desligarse del acto de habla como tal, en cuanto a elemento del proceso de comunicación y, por tanto, solamente va a poder adquirirse una competencia fónica cuando se use la lengua oral en procesos de comunicación significativos mediados completamente por la lengua oral. Así, favoreciendo y promoviendo la oralidad como único medio lingüístico de comunicación en el aula, los estudiantes centran su atención en esta, hecho que propicia un clima más favorable para un mayor aprendizaje/adquisición de la pronunciación. Se ha demostrado que,

cuando los aprendices monitorizan en alguno de los aspectos del aprendizaje, se potencia la reflexión sobre esos aspectos y, esto permite un mayor aprendizaje.

El hecho de que los estudiantes se inicien al aprendizaje de E/LE a partir de un enfoque pedagógico basado únicamente en la oralidad, por un lado, puede propiciar que se activen procesos y estrategias relacionados con la adquisición y aprendizaje de la pronunciación pero, por otro lado, puede poner en tela de juicio algunas de las creencias que los alumnos posean sobre lo qué es el aprendizaje de una lengua extranjera y cómo se lleva a cabo.

Partiendo de este planteamiento, nuestras preguntas de investigación serán dos:

- 1. ¿Qué tipo de estrategias de aprendizaje llevan a cabo los aprendices cuando se sigue un enfoque pedagógico centrado únicamente en la oralidad?
- 2. ¿Subyace algún tipo de creencia en los estudiantes relacionada con el aprendizaje de la pronunciación?

1.1 Estrategias de aprendizaje

Diversos son los autores que han tratado este tema, aunque nosotros hemos optado por tomar el modelo que Rebecca Oxford estableció en 1990. Entendemos por estrategia de aprendizaje las técnicas o mecanismos a los que puede recurrir el aprendiz para adquirir conocimientos (Rubin, 1975)¹.

Partiendo de la clasificación de Oxford, podemos hablar de estrategias directas e indirectas. Las estrategias directas serían las estrategias de memoria (asociaciones, repaso, actuar); estrategias cognitivas (practicar, recibir, enviar) y estrategias compensatorias (adivinar el sentido, superar carencias). Las estrategias indirectas serían las estrategias metacognitivas (implicación en el proceso de aprendizaje), las estrategias afectivas (reducir ansiedad, relajación) y las estrategias sociales (interacción y cooperación para el aprendizaje).

1.2 Creencias de los alumnos

Para analizar y describir las creencias que poseen los aprendices que forman parte de nuestro estudio, es imprescindible delimitar el concepto de creencia. Las creencias pueden entenderse desde el ámbito de la filosofía. Según Pajares (1992: 309), el término creencia suele aparecer bajo distintas denominaciones:

...attitudes, values, judgments, axioms, opinions, ideology, perceptions, conceptions, conceptual systems, preconceptions, dispositions, implicit theories, explicit theories, personal theories, internal mental processes, action strategies, rules of practice, practical principles, perspectives, repertories of understanding, and social strategy [...].

Como seres humanos e individuos de una sociedad tenemos creencias relacionadas con los distintos ámbitos de nuestra vida. La creencias las podemos entender como diversas ideas que una persona va formándose

¹ Vamos a seguir el modelo de Rebeca Oxford para la taxonomía y clasificación de las estrategias, pero hemos citado la definición de estrategias de aprendizaje que Rubin acuñó en 1975, por ser este uno de los pioneros en el estudio de este tema

respecto a un tema. A partir de nuestras creencias, nos enfrentaremos al mundo de una forma determinada.

En el ámbito de la educación y la enseñanza de lenguas, los estudiantes de una lengua extranjera poseen ideas preconcebidas sobre lo qué significa aprender una lengua y cómo hacerlo. A continuación, presentamos algunas creencias, ampliamente extendidas, acerca de lo qué es el aprendizaje de una lengua extranjera:

Creencias de lo que es el aprendizaje de una lengua extranjera:

- Los niños aprenden lenguas con mayor facilidad que los adultos
- Como mejor se aprende una lengua es en el país donde ésta se habla
- Las personas que hablan más de una lengua son muy inteligentes
- Para aprender una lengua hay que practicar mucho

Cuestionario BALLI (Beliefs About Language Learning Inventory), diseñado por Horwitz (1987)

A esta lista podríamos añadir creencias sobre cómo los estudiantes creen qué es una clase de lengua o qué creencias tienen respecto a la lengua oral y a la pronunciación.

Como definición de creencia en el ámbito de la enseñanza y aprendizaje de lenguas vamos a seguir la definición de Richardson según la cual las creencias son "Formas de comprender, premisas o proposiciones sobre el mundo, sostenidas psicológicamente, que son tenidas por verdaderas" (Richardson, 1996: 103).

2 Metodología

Este estudio se basa en el modelo de la investigación-acción, como una de las alternativas posibles de las que el profesor dispone, para llevar a cabo investigación en el aula (Brown & Rodgers, 2002).

La metodología específica que se va a seguir, dado que la investigaciónacción no posee una metodología propia (Cohen & Manion, 1990), va a ser de carácter cualitativo. Contamos con datos provenientes de un cuestionario que completaron los alumnos y un diario de clase elaborado por la profesora que se ha analizado siguiendo técnicas de análisis del discurso.

La investigación-acción de donde proceden los datos, se basa en una acción didáctica que se llevó a cabo en el Instituto Cervantes de Dublín (Irlanda) durante un semestre. Los participantes en la acción didáctica fueron alumnos adultos de español del curso de Inicial 1 (principiantes absolutos). La acción se implementó con un total de 50 alumnos. El número de horas de clase fue de 30 horas (cada sesión era de dos horas) y la propuesta didáctica implementada siguió un enfoque oral durante las 10 primeras horas de clase de forma íntegra, mientras que en las 20 siguientes horas de clase se alteró entre sesiones en la que se trabajaba mediante un enfoque oral y sesiones en la que se hacía uso de la lengua escrita. En la sesión número 6 se introdujo el alfabeto junto con las reglas ortoépicas.

El cuestionario con el que contamos en nuestro corpus de datos nos sirve para conocer el perfil de los alumnos. Los estudiantes lo cumplimentaron, el primer día de clase, a partir de datos de su información personal (nombre, nacionalidad, edad, profesión) y de datos más relacionados con el hecho de ser aprendices de una lengua, en este caso el español (motivación para estudiar español, anteriores experiencias de aprendizaje, competencia comunicativa en otra lengua, contacto con la lengua meta [LM]).

A partir de estos datos fue posible saber cuál era el contexto extraacadémico de los aprendices, qué grado de familiaridad poseían con la LM y si habían tenido experiencia como aprendices en otra lengua. Esta última información nos interesaba para poder establecer si los aprendices habían tenido la oportunidad de desarrollar diferentes estrategias de aprendizaje y además, llevaban consigo un bagaje que les permitiera tener unas creencias muy consolidadas.

La elaboración del diario de clase se realizó durante la duración de la acción didáctica. La profesora, a lo largo de cada sesión y al final de estas fue elaborando un diario en el que se incluyeron observaciones, sentimientos, reacciones, interpretaciones, reflexiones, presentimientos, hipótesis y explicaciones (Kemmis & McTaggart, 1988). En el diario se creó lo que denominamos "anecdotario"², y que según su propio nombre indica, es el lugar donde quedaron registradas las anécdotas que ocurrieron en cada una de las sesiones: reflexiones de los alumnos, incidencias, interacción entre los aprendices, etc., datos que nos han permitido conseguir información acerca del proceso dentro del aula.

3 Resultados

3.1 Cuestionario

La edad de los alumnos oscilaba entre los 20 y los 60, a pesar de que había un gran porcentaje de alumnos que se hallaban entre los 20 y los 40 años, a excepción de un grupo que poseía un número elevando de alumnos en la franja de los 60 años.

La mayoría de los alumnos desempeñaban profesiones en áreas como la educación, las tecnologías, la contabilidad, las finanzas, los negocios, las ciencias, la arquitectura, el marketing, la ingeniería, el secretariado o el periodismo. A pesar de haber una variedad de profesiones muy heterogénea, el perfil de todos ellos respondía a un tipo de alumno con cierta formación académica y una considerable experiencia de aprendizaje a través de la lengua escrita.

La mayoría de estudiantes eran de origen irlandés pero también encontramos otras nacionalidades: polacos, alemanes, una estudiante francesa y una estudiante turca.

Al ser las vacaciones y los viajes una de las mayores motivaciones para estudiar español, muchos de los alumnos habían visitado España al iniciar el curso, y estaban familiarizados con la lengua.

En cuanto a sus experiencias de aprendizaje de una L2, contamos con alumnos que mayoritariamente han estudiado antes una lengua extranjera.

3.2 Diario de la profesora

Para el análisis del registro de anécdotas del diario, vamos a analizar los datos a partir de las reflexiones y las preguntas de los alumnos.

² Nótese aquí que el término "anécdota" y "anecdotario" debe entenderse desde la perspectiva de la disciplina del análisis del discurso y la investigación etnográfica de aula. No por ser anécdotas son hechos poco importantes que acontecen en el aula si no que los anecdotarios son las descripciones escritas de las observaciones de los hechos significativos que ocurren en el aula.

3.2.1 Reflexiones de los alumnos

A partir de las reflexiones que algunos alumnos llevaron a cabo durante el desarrollo de la acción didáctica, puede apreciarse como en todo momento fueron muy conscientes del desarrollo y proceso de su aprendizaje.

Por un lado, reflexionan sobre la efectividad de algunas de las actividades, reflejo de que están siendo conscientes de lo que están aprendiendo. Junto a éstas, apreciamos también referencias a sus estilos de aprendizaje, es decir, un enfoque basado en la oralidad promueve en los estudiantes un alto grado de reflexión en cuanto a la manera como están aprendiendo la lengua oral y la pronunciación (véase ejemplo 1, en tabla 1). Este ejemplo pone de manifiesto la metarreflexión de la aprendiz, la cual es consciente del estilo visual de su aprendizaje. En algunos casos, hay alumnos que prefieren un enfoque donde se priorice la oralidad y la pronunciación, mientras que en otros casos, las preferencias se decantan por seguir un enfoque que no carezca de la mediación escrita.

Posiblemente, para muchos estudiantes el hecho de iniciarse en la enseñanza y aprendizaje de una lengua solamente a partir de la oralidad es algo innovador y muy diferente a las experiencias de aprendizaje que han tenido anteriormente, y a las que están acostumbrados. De hecho, muchas de las intervenciones de los alumnos, muestran las creencias que subyacen en cuanto a lo qué consideran que es una clase de lengua extranjera (véase ejemplo 2, en tabla 1). En el ejemplo 2 se evidencia la creencia de que una clase de lengua es una clase en la que no se practica demasiado la lengua oral, a no ser que sea una clase específica de conversación.

En otros casos, hay comentarios que muestran que un enfoque oral puede causar incomodidad en determinado tipo de aprendices. Los estudiantes que no se encuentran demasiado cómodos trabajando desde la propia oralidad son los que le dan más peso a la escritura. Ya sea por creencias, por estilos de aprendizaje o por ambos.

Mencionaremos ahora el caso de un estudiante que anotó en su cuaderno la correspondencia entre algunas de las vocales españolas y las vocales inglesas. Dicha correspondencia se basaba sobre todo en las grafías, puesto que cuando la profesora le explica que la [e] en español se pronuncia como la de la palabra "elefante", el alumno le dice que prefiere hacer la correspondencia entre los sonidos de las vocales de su lengua y los sonidos de las vocales del español. Esta correspondencia se basa en la asociación grafía en inglés-sonido que representa. Este hecho, por un lado, muestra cómo el alumno está llevando a cabo sus propias estrategias de aprendizaje, pero por otro lado, muestran cómo incluso para detectar la correspondencia de sonidos entre las dos lenguas necesita apoyarse en la escritura.

Otra reflexión que debe subrayarse proviene de dos alumnos, los cuales comentaron a la profesora que el trabajar mediante un enfoque oral les suponía un proceso de ajuste entre oralidad y escritura. Cuando los alumnos habían trabajado unos contenidos en clase de forma oral, después tenían que reajustar la imagen de la forma sonora de la palabra, que habían percibido en clase, a la forma ortográfica. Por tanto, dichos alumnos empezaban a ser conscientes de que el español no es una *lengua fonética*, como muchas veces se ha creído (Jackson & Rubio, 1969: 2; Clarke, 1982: 117), y que la asociación entre grafía y sonido no es una correspondencia tan obvia para el aprendiz de español, sobre todo cuando el punto de partida es la oralidad y la percepción de la lengua que se está estudiando.

En cuanto a los aspectos relacionados con la pronunciación, puede observarse como algunos alumnos van más allá y piensan el porqué han producido de una manera y no de otra (véase ejemplo 3, en tabla 1). Se cuestionan las causas de sus dificultades en la pronunciación, encontrando una explicación que va más allá de explicaciones basadas en la imposibilidad e impotencia que un alumno puede sufrir cuando se inicia en una lengua extranjera. El ejemplo 3, además, muestra cómo las reflexiones de los alumnos sobre sus procesos de aprendizaje de la pronunciación van desde la articulatoria hasta la acentuación y la entonación.

Finalmente, el último punto que incluimos en este análisis está relacionado con reflexiones de los alumnos que demuestran cómo durante las sesiones en las que se sigue un enfoque oral, los alumnos desarrollan diversas estrategias de aprendizaje, sobre todo en lo concerniente a la pronunciación:

- 1.- Estrategias de asociación entre nuevos sonidos de la lengua meta y los conocimientos previos del aprendiz (véase ejemplo 4, en tabla 1).
- 2.- Estrategias de aprendizaje que realizan al asociar el sonido que han percibido en español con la correspondencia grafía-sonido de su lengua (véase ejemplo 5, en tabla 1).

El comentario de una alumna: "Si no lo veo escrito, no se me graba y no lo aprendo".	Ejemplo 1
Dos estudiantes al presentarles el enfoque metodológico que se iba a seguir en el curso comentaron que les interesaba mucho porque en realidad querían hacer una clase de conversación. Es decir, su concepto de clase de español convencional no parece que incluye la práctica oral.	Ejemplo 2
Dos alumnas al realizar una actividad sobre el acento de las palabras en español, continuamente producen todas las palabras agudas. Al terminar le comentan a la profesora que se debe a que son del norte de Irlanda.	Ejemplo 3
Un alumno le pregunta a la profesora si el sonido [η] es como el sonido del nombre del huracán "El Niño".	Ejemplo 4
[0] la escriben como "th"; [$^{\Lambda}$] la escriben como "y", finalmente, asocian nuevos sonidos o sonidos que les causan dificultad a palabras que ya han aprendido de forma oral en sesiones anteriores, y que contienen esos sonidos.	Ejemplo 5

Tabla 1. Reflexiones de los alumnos

3.2.2 Preguntas de los alumnos

A partir del registro de anécdotas que recoge preguntas que los alumnos fueron haciendo a medida que transcurría la marcha de la acción, hemos podido establecer distintas clases de fenómenos que pueden apreciarse y que versan sobre distintos temas y aspectos. Para presentarlos de forma organizada y sistemática, añadimos una parrilla (véase tabla 2) en la que aparecerá el tema al que se refiere el fenómeno y los ejemplos que se recogen, acompañados de un sucinto comentario explicativo.

TEMA	FENÓMENO					
Estilos de aprendizaje	- Insistencia en preguntar "¿Cómo se escribe?" durante las sesiones en las que no se hacía uso del código escrito, especialmente en el grupo 4. Nótese que era el grupo donde había gente más mayor.					
Creencias de los alumnos ante la ortografía	de escribir las palabras tal y como las perciben. La palabra escrita tiene mucho peso en cuanto modelo de corrección que queda fijado "por escrito".					
Estrategias y procesos de aprendizaje relacionadas con el aprendizaje y adquisición de la pronunciación						

Tabla 2. Preguntas de los alumnos

4 Discusión

Tras analizar los resultados que nos ha aportado el diario y los temas que aparecen a partir de las preguntas y las reflexiones de los alumnos, puede afirmarse que cuestiones relacionadas con los estilos y estrategias de aprendizaje de los alumnos emergen continuamente. Los ejemplos sobre estrategias de asociación de sonidos muestran qué tipo de estrategias de aprendizaje se activan. La existencia de procesos de concienciación sobre el proceso de aprendizaje de la lengua oral acontecen continuamente en el aula. Encontramos varias referencias en el diario de la profesora que evidencian la concienciación que existe entre los aprendices de la diferenciación entre grafía/sonido.

Podríamos hablar de estrategias de aprendizaje de memoria (asociación de sonidos LM y grafías L1, asociación sonidos LM y L1, asociación de sonidos en la LM), estrategias de aprendizaje cognitivas (discriminación de grupos fónicos, identificación de las dificultades a nivel de articulatoria y prosódico) y estrategias de aprendizaje metacognitivas (reflexiones metacognitivas sobre sus estilos de aprendizaje).

Los ejemplos presentados muestran como para la elaboración e integración del conocimiento sobre los contenidos de pronunciación —lo referidos al sistema fonético, a la entonación, al ritmo— los aprendices llevan a cabo estrategias de aprendizaje tales como el razonamiento inductivo o el análisis contrastivo. Es decir, a partir de lo que oyen, de la monitorización en ciertos fenómenos y de la reflexión sobre éstos, los aprendices establecen sus propias hipótesis acerca del funcionamiento de la lengua: de cómo es un sonido, de cómo se entona en español, etc. Lo mismo ocurre con la comparación del sistema de la lengua meta con la L1 o incluso con una posible L2 del aprendiz.

La estrategias que llevan al razonamiento deductivo vienen siempre después de las que acompañan al razonamiento inductivo, puesto que, al seguir un modelo didáctico basado únicamente en la oralidad, no se presentan reglas de funcionamiento de la lengua relacionadas con la pronunciación. Éste es un factor importante en la interpretación de los resultados. La ausencia de instrucción explícita puede dar cuenta de los casos registrados de transferencia de reglas desde la L1 de los participantes, de tal modo que podríamos hablar de

una interferencia de la lengua escrita así como también del desconocimiento de las reglas de las correspondencias grafía-sonido en español.

En cuanto a los procesos de aprendizaje de la pronunciación y la lengua oral es también importante señalar que desde el primer momento en el diario de la profesora aparecen referencias sobre cómo al trabajar mediante un enfoque oral, los aprendices están constantemente focalizando en la pronunciación (estrategias cognitivas). Tras el análisis de los datos, también puede afirmarse que las creencias de los alumnos respecto al aprendizaje de la pronunciación coinciden en la mayoría de los casos, sin embargo, hay una disparidad de creencias relacionadas con el uso o no de la lengua escrita. Existen muchos casos de alumnos que no se sienten cómodos sin tener acceso (perceptivamente o productivamente) a la forma escrita de lo que está practicando oralmente. Los estilos de aprendizaje –visual, auditivo, etc.– en este sentido se corresponden bastante con el tipo de creencia respecto a la lengua escrita y a la lengua oral.

5 Conclusiones

Las dinámicas que se establecen en el aula cuando se sigue un enfoque oral enriquecen los procesos de aprendizaje y adquisición de la lengua oral y la pronunciación, motivando a los alumnos y despertando su interés por cuestiones relacionadas con la lengua oral y la pronunciación.

Distintas estrategias de aprendizaje para la pronunciación –de memoria, cognitivas y metacognitivas– se activan: relación de sonidos de las dos lenguas; relación de grafías de la L1 con los sonidos de la LM; reflexiones metacognitivas sobre cómo están aprendiendo y sus estilos de aprendizaje; razonamiento inductivo sobre cómo se articula un sonido o sobre el ritmo y la entonación.

Las creencias que poseen la mayoría de los alumnos en cuanto al aprendizaje de la pronunciación dan prioridad a la práctica de la lengua oral pero sin prescindir de la lengua escrita. Esta última es considerada como un elemento imprescindible en los procesos de aprendizaje de una lengua para algunos de los aprendices, a pesar de que el objetivo sea el aprendizaje de la pronunciación.

6 Bibliografía

- Brown, J. D, & Rodgers. (2002). *Doing Second Language Research*. Oxford:
 Oxford University Press.
- Clarke, E. (1982). *Mastering Spanish*, London, Mcmillan press.
- Cohen, L., & Manion, L. (1990). *Métodos de la investigación educativa*. Madrid: La Muralla.
- Horwitz, E. K. (1987). Surveying student beliefs about language learning. En A. Wenden & R. Rubin (eds.), *Learner* strategies in language learning. London: Prentice Hall International.
- Jackson, E., & Rubio, A. (1969). *Spanish made simple*, London, Heineman.
- Kemmis, S., & McTaggart, R. (eds.) (1988). The Action Research Planner (3.ª edición). Geelong, Australia: Deakin University Press.
- Oxford, R. (1990). Language learning strategies What every teacher

- should know, Boston, Heinle & Heinle Publishers.
- Pajares, M. F. (1992). Teachers' beliefs and educational research: cleaning up a messy construct, *Review of Educational Research*, 62, 307-332.
- Richardson, V. (1996). The role of attitudes and beliefs in learning to teach, en J.Sikula, T. J. Buttery, y Ed. Guyton. (eds.), *Handbook of Research on Teacher Education* (2.ª edición). New York: Macmillan, 102-119.
- Rubin, J. (1975): What the "Good Language Learner" can teach us. *TESOL Quarterly*, 9, 41-51.

L1-L2map: una base de datos fónica para la enseñanza de la pronunciación de segundas lenguas

MARTÍNEZ-PARICIO, VIOLETA; KOREMAN, JACQUES; HUSBY, OLAF

1 Introducción

El objetivo de este trabajo es presentar las principales propiedades y utilidades de L1-L2*map*, una base de datos multilingüe que permite establecer comparaciones entre los inventarios fónicos de dos lenguas cualesquiera (de entre las documentadas en el sistema). Aunque esta base de datos puede resultar útil a lingüistas interesados en aspectos teóricos de tipología fonética y fonológica, la finalidad de este trabajo es describir y explicar la aportación de L1-L2*map* al mundo de la lingüística aplicada; concretamente, al ámbito de la enseñanza de la pronunciación de segundas lenguas. En este sentido, destacaremos cómo la información segmental presente en L1-L2*map* puede resultar de gran valor a la hora de diseñar ejercicios de pronunciación de segundas lenguas más personalizados (i.e. ajustados a las necesidades específicas de los hablantes en función de su lengua nativa).

Numerosas investigaciones sobre la adquisición de segundas lenguas y, específicamente, sobre la adquisición de su pronunciación, coinciden en destacar la gran complejidad que rodea el proceso de adquisición del sistema fónico de una lengua distinta a la nativa (véase Broselow y Kang, 2013 y referencias allí citadas). Entre los muchos y variados factores que interactúan en

este proceso, los expertos (p. ej. Lado, 1957; Major, 2008) resaltan el importante papel que desempeña la lengua nativa (L1) del hablante. En la actualidad sabemos que la L1 no es el único factor que determina la adquisición de la pronunciación de una segunda lengua (L2), sin embargo, es uno de los factores que lo condiciona fuertemente: aquellos sonidos de la L2 que se desea aprender ausentes en la L1 del hablante, suelen ser por lo general (aunque no siempre) más difíciles de adquirir. Por tanto, desde el punto de vista didáctico sería muy útil saber cuáles son esas áreas fónicas que presentan mayor grado de dificultad para cada aprendiz y así poder desarrollar ejercicios de pronunciación más ajustados a las necesidades de cada hablante. Desafortunadamente, esta atención personalizada no puede realizarse en el aula, donde muchas veces conviven hablantes de numerosas y variadas lenguas nativas y el profesor no siempre conoce las peculiaridades fónicas de cada lengua. Con la intención de superar este obstáculo, en la Universidad de Ciencia y Tecnología de Noruega surgió la idea de desarrollar una base de datos que permitiera establecer las principales diferencias y similitudes fónicas entre dos lenguas, una posible L1 y una posible L2.

En los apartados que siguen nos ocuparemos, en primer lugar, de presentar las principales propiedades actuales de esta base de datos. Expondremos las fuentes de las que se nutre e ilustraremos la manera en que se visualizan los datos comparativos en esta plataforma digital (sección 2). A continuación, en la sección 3 se presenta un programa informático para la práctica de la pronunciación de segundas lenguas, CALST, que extrae automáticamente información de L1-L2map. A partir de esta información, el programa CALST selecciona y ofrece al usuario unos ejercicios de pronunciación de L2 determinados, en función de las necesidades particulares de cada hablante. Aunque en la actualidad este programa permite practicar únicamente la pronunciación del noruego, en esta sección veremos cómo se pueden integrar ejercicios en CALST que permitirán también practicar la pronunciación de otras L2s. Asimismo, dado que las dos herramientas digitales (L1-L2map y CALST) están limitadas en la actualidad a aspectos segmentales del lenguaje, en la sección 4 se discuten futuras expansiones previstas para ambas plataformas. Finalmente, en la sección 5 se concluye.

2 L1-L2*map*

L1-L2map ha sido desarrollada en la Universidad de Ciencia y Tecnología de Noruega, en Trondheim, gracias a la colaboración de fonetistas e informáticos del mismo centro. Esta base de datos multilingüe permite comparar el número y los tipos de segmentos (vocales y consonantes) que conforman el inventario fónico de dos lenguas cualesquiera, de entre un total de 500 lenguas (aproximadamente) para las que existe información en nuestro sistema (Koreman *et al.*, 2011). La base de datos es de libre acceso y puede consultarse en la red de manera gratuita en el siguiente enlace electrónico: http://calst.no/L1-L2map.

2.1 Fuentes de L1-L2map

L1-L2map se nutre en su mayor parte de la información segmental presente en otras dos grandes bases de datos: UPSID (Maddieson 1984) y LAPSyD (Maddieson *et al.*, 2011). Estas bases de datos, también disponibles en internet, presentan información sobre los inventarios fónicos de varios centenares de lenguas y ofrecen, además, información adicional sobre algunos aspectos básicos de su fonología y fonética (p. ej. distribución alofónica de algunos

sonidos, reglas fonológicas, posición del acento, inventario de tonos, etc.). La principal diferencia entre estas bases de datos y L1-L2map es que esta última permite seleccionar dos lenguas y comparar sus segmentos. En el siguiente apartado veremos cómo se presenta esta información en la base de datos. Además, L1-L2map ha incorporado información segmental de otras lenguas, para las que no existía información en UPSID y LAPSyD. Por último, L1-L2map está implementada como una wiki lo que permite a otros expertos (previa autorización) añadir información sobre más lenguas, así como corregir los datos existentes sobre las lenguas ya presentes en la base de datos. En numerosas ocasiones, los autores de las gramáticas y artículos de referencia sobre los inventarios fónicos de las lenguas no están de acuerdo en torno al conjunto exacto de sonidos que son pertinentes en dicha lengua. Por ejemplo, a veces no está del todo claro el estatus fonémico de un sonido o las fuentes de una misma lengua varían en la interpretación de algún rasgo caracterizador de un sonido (p. ej. su punto de articulación, su modo de articulación, etc.). Por ello, y aun siendo conscientes de las limitaciones que supone trabajar con tantos, y tan variados, datos lingüísticos, pensamos que resulta positivo que estos datos puedan ser eventualmente corregidos, bien por hablantes nativos de la propia lengua o especialistas en su fonética o fonología.

2.2 Visualización de los datos en L1-L2map

En la pantalla inicial de la base de datos, el usuario puede seleccionar las dos lenguas que desee comparar y, de este modo, consultar automáticamente sus similitudes y diferencias segmentales. Por ejemplo, en el contexto de enseñanza de una lengua extranjera, el profesor de dicha lengua puede seleccionar alguna de las lenguas nativas de sus estudiantes (L1) y compararla con la lengua meta que desea enseñarles (L2). Una vez seleccionadas las dos lenguas –en este caso una potencial L1 y L2–, la información segmental se presenta en unos cuadros de transcripción fonética, inspirados en las tablas del alfabeto internacional de fonética (AFI). Para señalar las similitudes y diferencias entre las dos lenguas, empleamos un código de colores sencillo, que aparece resumido en la siguiente leyenda.

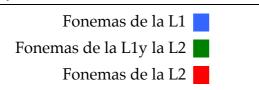


Figura 1. Leyenda con el código de colores para indicar las diferencias y similitudes del inventario fonémico de las lenguas.

Como se indica en la figura 1, los fonemas que son pertinentes y están presentes en las dos lenguas (L1 y L2) aparecen en verde; los que solo ocurren en la L1, se indican en azul. Por último, aquellos segmentos presentes únicamente en la L2, se marcan en rojo. Estos últimos son frecuentemente los más interesantes desde el punto de vista de la enseñanza de una L2, ya que son los menos familiares para el hablante y, por tanto, aquellos que probablemente tenga que practicar en mayor medida en su pronunciación.

Para ilustrar mejor cómo se visualizan los datos en nuestra base de datos, obsérvese la siguiente tabla generada por L1-L2*map*. En ella se compara el

inventario fónico del chino mandarín (L1) con el del español (L2).¹ En la parte superior de la tabla vemos que "Mandarin" aparece sombreado en azul, pues en este caso funciona como L1, y "Spanish" aparece sombreado en rojo, pues en este contexto particular es la L2. Los fonemas consonánticos y los vocálicos aparecen en distintas tablas. Abajo solo indicamos la comparación de los fonemas consonánticos de estas dos lenguas.

Languages - Sounds -		Mandarin				Spanish						Username		
Consonants		Consonants (other)			Vowels	s Language in						formation		
	Labial			Coronal										
View positions	Bilabia	al	Labio- dental	Dental Alveolar		Palato- alveolar	Retroflex		Palatal		Velar		Uvular	
Plosive	p p ^h	b			t t ^h	d						k kh	g	
Nasa	1 1	m				n					л		ŋ	
Tril	ı					r								
Tap, Flag	0					ſ								
Fricative	e		f	θ	s			ş	Z,	ç	į	х		χ
Latera fricative														
Approximan	t										j ų		w	
Latera approximan						1								
Lateral flag														

Tabla 1. Comparación de los fonemas consonánticos del chino mandarín (L1) y el español (L2)

Mediante esta tabla comparativa, el profesor de español (L2) puede saber inmediatamente qué segmentos presentarán, probablemente, mayor dificultad para los hablantes nativos de chino mandarín. Por ejemplo, en esta tabla, las oclusivas sonoras aparecen en rojo (/b, d, g/) lo que indica su ausencia en el chino mandarín. Esta ausencia suele ir ligada a la falta de familiaridad con la realización de esta serie de sonidos. Efectivamente, dicha dificultad ha sido constatada en diversas investigaciones sobre la adquisición del español por parte de hablantes de chino mandarín. Incluso en niveles avanzados de español, los hablantes chinos suelen ensordecer en numerosas ocasiones estas consonantes sonoras, lo que puede interferir en la inteligibilidad del hablante (p. ej. estu[t]io en lugar de estu[ð]io, Blecua y Esteve, 2014). Asimismo, otra fuente de dificultad para los hablantes de chino mandarín que desean aprender el español es la realización de las líquidas. Como puede observarse en la tabla, los hablantes chinos no poseen ninguna variante rótica en su lengua nativa, lo que lleva a confundir estos segmentos con la otra líquida del sistema del español, la lateral. Aunque los hablantes de chino mandarín sí poseen un fonema lateral, la falta del contraste lateral vs. rótica en su sistema dificulta la percepción y producción de estos segmentos. En definitiva, estas tablas básicas proporcionan un buen punto de partida para aproximarse a la enseñanza de la pronunciación de una L2.

¹ La variedad del español aquí representada es la correspondiente a la norma septentrional estándar yeísta, en la que el fonema $/\lambda$ / ha desaparecido y pasa a realizarse con cualquiera de las realizaciones documentas para la realización de la /i/.

Concretamente, una vez establecidos los segmentos y contrastes que *a priori* pueden presentar un mayor grado de dificultad para cada hablante (por estar ausentes en su L1), el profesor de la L2 podrá focalizar sus esfuerzos en practicar y corregir la producción y percepción de dichos sonidos, así como la de sonidos similares con los que contrastan mínimamente. Por ejemplo, un hablante de español (L1) que desee practicar la pronunciación del inglés (L2), tendrá que adquirir correctamente, entre otros, el fonema labiodental fricativo sonoro /v/ del inglés, ausente en su lengua nativa (y por tanto, señalado en rojo en las tablas de L1-L2map). Con esta finalidad, el profesor podrá realizar ejercicios sencillos de repetición y escucha en clase en los que el hablante se vea expuesto a palabras y expresiones que contengan este sonido (p. ej. [v]an 'furgoneta'). Asimismo, podrá compararse la realización de este y otros fonemas similares a los que se opone mínimamente, como por ejemplo la labiodental fricativa sorda /f/ (p. ej. [f]an 'abanico') o la bilabial oclusiva sonora /b/ (p. ej. [b]an 'prohibir'). De esta manera, el aprendiz se puede ir familiarizando poco a poco con el nuevo sonido (p. ej. [v]) e interiorizar las diferencias de significado que acarrean la presencia/ausencia de uno u otro sonido.

Sin embargo, esta atención personalizada en la que el profesor concentra sus esfuerzos en practicar la producción y percepción de aquellos sonidos y contrastes más dificultosos no puede realizarse en el aula, donde muchas veces conviven hablantes de numerosas y variadas lenguas nativas (con distintas necesidades) y, además, se deben practicar también otras destrezas lingüísticas. En ese sentido, y para facilitar la práctica de la pronunciación, en la Universidad de Ciencia y Tecnología de Noruega desarrollamos un segundo programa informático: CALST (Computer Assisted Listening and Speaking Tutor). Como veremos en los siguientes apartados, este programa es capaz de extraer automáticamente la información relativa a las diferencias de la L1 del usuario y la L2 que este desea aprender, para luego ofrecerle ejercicios de pronunciación más ajustados a sus necesidades particulares

3 CALST

CALST es una plataforma digital que permite practicar la pronunciación, la escucha y la ortografía de los dialectos más importantes del noruego. Está inspirado en VILLE (Wik, 2011), un programa informático para la práctica de la pronunciación del sueco. Frente a otros programas similares de enseñanza de la pronunciación de una L2, lo realmente novedoso de CALST es que se nutre del análisis contrastivo llevado a cabo en L1-L2*map*. De este modo, cada hablante se centra en practicar aquellas áreas de la pronunciación del noruego qué más necesite trabajar. Así, un hablante cuya lengua nativa es el español, el alemán o el vietnamita no serán expuestos a los mismos ejercicios, ya que cada uno se enfrenta a distintas dificultades, dadas las particularidades fónicas de su lengua nativa.

CALST está disponible de manera gratuita en la red (http://calst.no). Cualquier persona que desee aprender noruego puede utilizar esta plataforma para practicar su pronunciación, no solo en las horas de clase, sino en cualquier momento que este lo desee. La primera vez que un usuario se conecta a CALST, debe indicar cuál es su lengua nativa y qué dialecto del noruego desea aprender. El noruego, a diferencia de otras lenguas, no posee una pronunciación estándar. Los distintos dialectos y variedades lingüísticas características de una zona se usan a diario y, por lo tanto, es crucial para los estudiantes que viven en Noruega familiarizarse con el dialecto utilizado en su lugar de residencia. Una vez registradas la L1 del usuario y el dialecto meta que

desea aprender (L2), el programa liga automáticamente los sonidos en rojo (aquellos que solo ocurren en la L2) de la tabla generada por L1-L2*map* a ejercicios de pronunciación y escucha que le ayuden a perfeccionar la producción y percepción de estos sonidos.

En la actualidad, los ejercicios de escucha y pronunciación solo permiten practicar la pronunciación de los dialectos del noruego. Sin embargo, CALST es modular y, por ello, una vez se desarrollen y graben ejercicios para la práctica de la pronunciación de otras lenguas (p. ej. alemán, español, francés) estos podrán integrarse automáticamente en el sistema. Así, estudiantes de otras lenguas también podrán beneficiarse del análisis contrastivo llevado a cabo en L1-L2map y practicar las áreas de pronunciación de la L2 que encuentren más problemáticas. De hecho, ya hay planes concretos para elaborar los ejercicios pertinentes para el español y el inglés.

3.1 Ejercicios en CALST

Como se indicó al inicio del artículo, el principal objetivo de este artículo era presentar las propiedades de L1-L2map y ejemplificar su utilidad en el ámbito de la enseñanza de la pronunciación de L2s. No daremos por tanto cuenta detallada de los distintos tipos de ejercicios que conforman CALST. Tan solo ilustraremos cómo funciona el programa y su relación con L1-L2map mediante la muestra de un ejercicio básico de escucha. Cualquier lector interesado en conocer con más detalle las características de CALST puede registrarse en la plataforma y examinar por sí mismo los distintos tipos de ejercicios de pronunciación, escucha y ortografía que conforman el programa. Cabe matizar, sin embargo, que existen dos grandes bloques de ejercicios: (i) ejercicios de vocabulario básico, en los que se practica la pronunciación, escucha y ortografía de unas mil palabras o expresiones básicas y (ii) ejercicios destinados a familiarizar al hablante con aquellos sonidos y contrastes que le son menos familiares. Los primeros son iguales para todos los hablantes y los segundos varían en función de la L1 del hablante. A continuación, presentamos un ejercicio tipo de escucha de este segundo bloque de ejercicios, pensado para un hablante de español.

Una vez se han introducido los datos de la L1 y la L2 del usuario, L1-L2map realiza la comparación de los segmentos y establece qué segmentos no están presentes en la L1 del hablante. Estos segmentos, indicados en rojo en la base de datos, son ligados automáticamente a ejercicios de pronunciación que inciden en su práctica. Por ejemplo, el sonido que corresponde a la realización del fonema /v/ no existe en español (L1), pero sí en noruego (L2) y por tanto el aprendiz de noruego tendrá que practicar su correcta realización y distinción en la cadena hablada. Para ello, un tipo de ejercicio de escucha típico en CALST ofrece al usuario dos palabras que constituyen un par mínimo como por ejemplo [v]itt 'blanco' y [b]itt 'mordisco'. En un tipo de ejercicio, referido como AXB e ilustrado abajo en la figura 2, el aprendiz escucha primero la palabra de la izquierda, a continuación la de la derecha (que difiera mínimamente de la anterior) y en tercer lugar vuelve a escuchar una de las dos palabras. Debe decidir entonces cuál de las dos palabras ha oído en tercer lugar. Ambas palabras se dan en su forma ortográfica y se puede seleccionar una opción en al que la traducción de estas al inglés se presenta debajo de ellas (como en el ejemplo). El contraste entre [v] y [b] no solo se ejemplificará y practicará con más series de pares mínimos en los que los sonidos aparecen en posición de ataque silábico, sino que habrá ejercicios complementarios en los que estos

sonidos se encuentren en posición de coda silábica. Esto es muy importante dado que en muchas lenguas, la realización de un fonema y su contraste a otros fonemas puede percibirse fácilmente en posición de ataque silábico, pero en posición de final de sílaba, algunos hablantes de lenguas que no admiten todas sus consonantes en la posición de coda, suelen presentar más dificultades a la hora de distinguir ciertas oposiciones. Por ello, es necesario practicar la realización de los segmentos en los distintos tipos de posiciones silábicas.

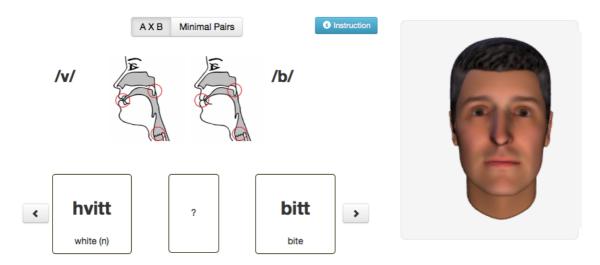


Figura 2. Ejemplo de un ejercicio de percepción en CALST con un par mínimo en el que contrastan /v/ y /b/

4 Futuras expansiones de L1-L2map y CALST

Las propiedades segmentales del lenguaje constituyen solo una parte de los rasgos característicos de la pronunciación de una lengua. Para evitar tener acento extranjero en una L2 y, más importante, para asegurar su comprensión e inteligibilidad, el hablante debe practicar también la producción y percepción de los rasgos suprasegmentales característicos de la L2 (p. ej. acento, tonos léxicos y patrones entonativos). También deberá adquirir las nuevas restricciones combinatorias de los segmentos en la L2. Por ejemplo, aunque un hablante español posee los dos segmentos /s/ y /k/, cuando estos aparecen en posición de ataque complejo en lenguas como el inglés (p. ej. skype), suele tener dificultades en su producción (ya que esta combinación no es propia de su lengua nativa). Para resolver esa dificultad, el hablante español suele insertar una vocal precediendo a la secuencia sk-, de tal modo que el ataque complejo se rompe, p. ej. [e]s.kype, [ə]skype (los puntos indican los límites silábicos). En este caso y similares, las restricciones silábicas propias de la L1 suelen a veces transferirse e influir en la pronunciación de la L2. Hablantes de distintas lenguas emplearán unas u otras estrategias (elisión, inserción de vocales, metátesis) al enfrentarse a combinaciones segmentales poco familiares y, por ello, tendrá que trabajar en su correcta pronunciación y distinción. Lo mismo sucede con la realización del acento, los tonos léxicos y los patrones entonativos: en muchas ocasiones el aprendiz debe familiarizarse con las nuevas propiedades prosódicas de la L2.

Por todos estos motivos, estamos trabajando en la incorporación de estos datos en L1-L2*map*, para que permita comparar también las propiedades prosódicas de las lenguas (Martínez-Paricio *et al.*, 2014). De este modo, CALST podrá extraer la información relevante y proporcionar a cada aprendiz

ejercicios más elaborados donde no solo se practiquen las propiedades segmentales del lenguaje, sino también las suprasegmentales.

5 Conclusiones

En este trabajo se han descrito las propiedades fundamentales de la base de datos multilingüe L1-L2*map*. Hemos visto cómo esta plataforma puede ser útil en el ámbito de la enseñanza de la pronunciación de segundas lenguas. Concretamente, hemos visto como la información más relevante desde el punto de vista didáctico en L1-L2*map* puede ser usada directamente por los informáticos y profesores de una L2 en el desarrollo de un programa informático que permite practicar la pronunciación, CALST. Aunque ambas plataformas contienen en la actualidad únicamente información segmental y están centradas en el noruego, hemos anticipado cómo trataremos de expandir ambos sistemas para que incorporen también información suprasegmental y permitan practicar también la pronunciación de otras L2.

6 Bibliografía

- Blecua Falgueras, B. & Esteve Martín, M. T. (2014). Realizaciones de las oclusivas y aproximantes del español por parte de hablantes chinos. Presentación oral en el VI CIFE (Valencia, 5 de noviembre).
- Broselow, E., & Kang, Y. (2013). Phonology and Speech. En J. Herschensohn & M. Young- Scholten (eds.), *The Cambridge Handbook of Second Language Acquisition*, 529-544. Cambridge: Cambridge University Press.
- Harris, J. (1969). *Spanish Phonology*. Cambridge: MIT Press.
- Koreman, J., Bech, Ø., Husby, O., & Wik, P. (2011). L1-L2map: a tool for multilingual contrastive analysis. Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS2011), Hong Kong.
- Koreman, J., Wik, P., Husby, O., & Albertsen, E. (2013). Universal contrastive analysis as a learning principle in CAPT. En *Proceedings of the workshop on Speech and Language Technology in Education (SLaTE 2013)*, 172-177.

- Maddieson, I. (1984). *Patterns of Sounds*. Cambridge: Cambridge University Press.
- Maddieson I., Flavier S., Marsico E., & Pellegrino F. (2011). LAPSyD: Lyon-Albuquerque Phonological Systems Databases, Version 1.0. http://www.lapsyd.ddl.ish-lyon.cnrs.fr/
- Major, R. C. (2008). Transfer in second language phonology. En J. G. H Edwards y M. L. Zampin (eds.). *Phonology and second language acquisition*, 36, 63-94. John Benjamins Publishing.
- Martínez-Paricio Martínez-Paricio, V., J. Koreman, O. Husby, J. Abrahamsen, & Ø. Bech. Expanding CALST: multilingual analysis of L1-L2 phonotactics for language teaching, Pronunciation in Second Language Learning and Teaching (PSLLT 2014), Santa Barbara, EE. UU., 5-6 septiembre 2014.
- Navarro Tomás, T. (1918). Manual de pronunciación española. Madrid: CSIC.
- Wik, P. (2011). *The Virtual Language Teacher*. Tesis doctoral, KTH School of Computer Science and Communication, 2011.

Acquisition of L2 Portuguese stops by Chinese-speaking learners - Considerations

NUNES, ANA MARGARIDA BELÉM

1 Introduction

It is well known that Chinese's are lately becoming more interested about Portuguese Language and Culture due to the vast possibilities of business with not only Portugal but especially with Brazil and Angola. Being a Professor of Portuguese at the University of Macau, China, I am always confronted with the differences of the languages and the problems that learners (Mandarin and/or Cantonese speakers) have to face. This research has an experimental nature and its focus is the acquisition of Portuguese consonants by Chinese learners of Portuguese as a Second Language. In this particular case it will be analyzed the European Portuguese (from now on EP) and not the Brazilian variety that the students have also access.

Second language acquisition is a complex process that involves, among other things, the simultaneous acquisition of a new sound system and lexicon and the relation between speech production and speech perception is important (Fowler, Carol A., pp. 633), so it is clear that these students are most probably writing the same way that they perceive the phones and by that, the words. Nevertheless and given the scarcity of studies about European Portuguese as a Second Language for Chinese speakers very few is yet known about the major difficulties of these students. Chinese speakers learning Portuguese do have all the same and common linguistic conflicts that all adults have when experiencing a new language: Word order; animacy; grammar rules and exceptions; number and gender of the words (something that is a great dilemma for Chinese learners, once that they seem not to have this distinction so marked as in EP); pronunciation and the center of this research phonetic

difficulties. The phenomenon described as foreign accent may be a consequence of differences between the production of one or the other language phonetic level. It is quite common that speakers of two different languages that share a certain similarity in a particular phoneme set for that segment the same phonetic realization in both.

Studies report high levels of anxiety on learning a foreign language (MacIntyre and Gardner, 1994) "the feeling of tension and apprehension specifically associated with Second Language contexts, including speaking, listening and learning". If this is true for all the students can be even more evident for Chinese students. According to personal experience and as result of some literature review, it is known that Chinese students do not like to be exposed so they will mainly interact with the professor if they fill very confident about their answers, meaning that anxiety on speaking or writing in a foreign language and the fear of failing or the sense of self-esteem it is noticeable on the students that the texts are here analyzed.

2 Second Language acquisition

According to a study carried out by McQueen and Cutler (2006) 'listeners come equipped to deal with variability in speech sounds – whether that variability arises from differences among speakers or within a given speaker over time' meaning that listeners can modify their own perception of utterances produced by a speaker if they feel some unusual characteristics on their speech. In the present analysis the speakers are all adults (all with an average age around the 20 years old) so the main point it is also correlated with the fact that some adults trying to learn a foreign language, in this concrete set of circumstances, in a country where the language is not spoken very often or in a daily basis (in Macau it is used English as "língua franca"), they are "exposed to much less language input than an average child." Meaning that all the material that an adult as access is not so concentrated and not so dynamic or real than for a child so it is obvious that " a general deficiency of input may well explain many cases of adult foreign language errors", according to Robert Bley-Vroman (The logical problem of foreign language learning, pp. 41-68, 1989).

In adults the complete success on learning a foreign language it is very difficult perhaps it does not even exist and in its "Interlanguage" article, Serlinker (1972) stated that the "rare cases of apparent complete success could be regarded as peripheral to the enterprise of second language acquisition theory". Serlinker (1972) believed that only 5% of adults would have a complete success in second language acquisition. However what can be said about the Chinese students of Portuguese as a second language is that some of them, and considering that both languages are typologically different, develop a competence and performance near native-speakers. The degree of achievement on adults has a substantial variation; they tend to fail in different degrees even if they have a more constant exposure to the language, in the specific analyzes that is described in this paper students study for four Scholl years to have their Bachelor in Portuguese Studies. The variation on their accomplishment is strictly related with their learning strategies according to Robert Bley-Vroman (1989) but also with their motivation, future expectations and communicative needs, such variation is universally accepted is a type of "general problem solving".

Many studies are focused on the requirements needed to work on L2 pronunciation lessons: Focusing on tasks of production and perception of the sounds of foreign language. Escudero (2000) shows that the improvement of the

perception of L1 and L2 is influence by the input nature to which students are exposed. For this reason there are two dominant theories on the perception of L2: the Model of Learning (Speech Learning Model - SLM) of Flege (1995) and the Perceptual Assimilation Model - PAM), Best (1995). These models attempt to explain how adult speakers of L2 may learn to perceive non-native contrasts and how their particular perception is different from the perception of native L1. The two theories point to the linguistic experience as the predominant factor. It is supported by the literature that activities such talking and listening are close. In the public domain it is important that the speaker knows if the listener can perceive clearly what is being said. The relation between speech production and speech perception is also important in the private domain in which 'articulatory mechanisms support talking, and perceptual mechanisms support listening to speech' (Fowler, pp. 633). As stated by Mitterer and Cutler (2006) to understand someone's speech and to learn a language it is necessary to recognize words. This means being able to distinguish between words that are very close such as for European Portuguese: [patu] (duck) and [batu] (beat) or [fake] (knife) and [vake] (cow).

It is known that the phonological elements have their primary differences in the vocal tract and not in the mind (production versus perception) in linguistic terms there are significant actions of the vocal tract that are called gestures (Fowler, Carol A., pp. 634). Those different actions or gestures are dynamic and show the difference between pairs like [t] and [d]; [p] and [m]; [p] and [b] and also between [f] and [v]. The differences between minimal pairs are simple to understand if listeners perceive gestures. Nevertheless the pairs that we present just differ on the voiced/unvoiced parameter and not that much on the articulation for European Portuguese. According to a study carried out by McQueen and Cutler (2006) 'listeners come equipped to deal with variability in speech sounds – whether that variability arises from differences among speakers or within a given speaker over time' meaning that listeners can modify their own perception of utterances produced by a speaker if they feel some unusual characteristics on their speech.

One common characteristic of adolescents and adults who learn a second language in a late stage is the different way that they accent some phonetic segments and sequences of the new language. The phonetic realization of phonological structures in the L2 is different from native language patterns. When producing sentences in L2, speakers often produce phonetic segments and sequences that appear to be a product of complex interaction between L1 and L2 phonetic realization rules (Spanish accented English Speakers; American-accented Japanese speakers). Studies have shown that adult listeners' perception of non-native consonant and vowel contrasts have demonstrated a poor performance when compared with native language listeners for several of the phonetic contrasts. On a study carried out by Flege and Wang (1989) it was reported that native Chinese learners of English could perceive voicing contrasts in final stop consonants in contexts where they had a preceding vowel, closure voicing and release cues. However, when the final consonants were unreleased their performance wasn't so great, and the native speakers could maintain their perceptual differentiation.

All phonetic features that can serve to distinguish phonological segments (contrasts in certain languages) can be differentiated acoustically. For example voicing contrasts between oral stops consonants in English) pet-bet; bet-bed, are differentiated by temporal parameters, like VOT – Voice Onset Time for initial stops, duration of consonant closure for medial stops, preceding vowel

duration for medial and final stops and also by spectral characteristics. Speech as an acoustical signal is a code in which the phonetic segments and sequences are specified by context-dependent and each language specific acoustic parameters.

Cross-language studies with infants that were exposed to different languages were examined to determine if acoustic parameters associated with non-native phonetic contrasts were discriminated. For example voiced and unvoiced, aspirated syllable – p and b – for English and the stop phones /p/ and /b/ for Spanish). The monolingual Spanish speakers showed good discrimination of the aspirated consonants but the same did no happen with the other contrasts. Several studies, for example comparing infants with adults with contrasts in Hindi and English languages, showed that infants perform better and can discriminate better the place of articulation contrast. In adults it seems that native-language phonetic perception is strong and mechanic. The ability to extract the phonetic message from the acoustical signal even in non-optimal situations (noise, unfamiliar talkers, and distracting tasks) requires cognitive resources on the part of the native listener, on a perception task. Languagespecific patterns of performance are not related to differences in basic auditory capabilities of adult speakers of different languages, they reflect the languagespecific patterns and the acoustic-phonetic relation that they know. All the sequences and its categorization have become automatic even if the listening conditions are not optimal, Strange (2006).

3 Some aspects of Portuguese and Chinese phonetics

Everybody that want to learn (or even teach) a foreign language should know about the importance of general phonetics in what concerns pronunciation as well as grammar, Ball (2009). Phonation is the term used to describe the different modifications to airflow as it passes trough the glottis within the larynx, Ball, 2009. To produce different sounds the vocal folds have distinct behaviors in the case of the consonants that are studied in this paper it should been explained that to produce the plosives [p, t, k] that are considered voiceless consonants, the vocal folds are usually apart and they will not vibrate. On the other hand to produce [b, d, g] the vocal folds are close together and vibrating when the air passes through them.

In the specific case of European Portuguese and the example of the 6 stops or plosives [p, t, k] and [b, d, g] that just differ on the degree of closure between articulators, giving therefore origin to voiceless and voiced consonants respectively, it is curious to underline that those sounds are usually and often misperceived by children at the stage that they start learning to read and write (around 6 to 8 years old). This confusion, even in native speakers, may give an idea about the irregular characteristics of these stops acoustics'. It is even observe, as noticed with Chinese learners of Portuguese, that the pair $(t \rightarrow d)$ is the one that occurs more frequently. As such this can be also a possible explanation to the big difficulties that foreigners have with the Portuguese system of stops, once that it seems very hard to pronounce or hear the difference between voiced and voiceless stops. Accordingly and to summarize the European Portuguese plosive's articulation: [b] voiced consonant, the airstream is impeded (or "stopped") by the lips (voiced bilabial stop); [d] voiced alveolar stop; [g] voiced glottal stop; on the other hand we have the all the voiceless pairs [p, t, k]: bilabial stop; alveolar stop and glottal stop respectively.

The case of Chinese it is necessary to fix a standard to unify the pronunciation across China. The Chinese alphabet is simply a phonetic system for representing the National Pronunciation of Chinese, according to Chin-Shi, Li (reprint of 1922). Once that Peking is the capital, the Pekingese pronunciation is seen as standard, yet there are two varieties Mandarin (Southern and Northern Mandarin). According to Chin-Shi, Li (reprint of 1922) Cantonese is predominant abroad. Cantonese has been the one Chinese dialect most influenced by European language as a result of the contact with English in Hong Kong and Portuguese in Macau. Spoken Chinese has nineteen consonants plus three palatals and some syllabic consonants and even when a 'syllable begins with a vowel, there is sometimes a consonant-like articulatory gesture before the vowel', (Duanmu, 2007). Some differences that can be noticed are that Chinese has aspired sounds of the consonants [p, t, k, ts] what does not happen in European Portuguese. It is also interesting to notice that in Chinese according to the findings of D. Zhou and Wu (1963:22), the tongue tip when pronouncing the dentals [ts, tsh, s, t, th, n, l] can be on "either the upper or the lower teeth", which may perhaps help understand why Chinese students do not grasp the correct articulation of some sounds. In a study by Fu (1956:4) notes that some Beijing speakers use interdentals instead of the dentals [ts, th,

s], (Duanmu, 2007), those articulation differences do not cause any distinction between words, since those distinctions are tonal.

There exist several representation systems of the Chinese sounds and characters: the "Cantonese Pinyin" created by Yu Bingzhao in 1971, adjusted later, 1990, by the "Institute of Language in Education". In 1993 was created the "Jyutping" system developed by the Hong Kong Government and described as "multifunctional, systematic, user-friendly, compatible with all possible modern Cantonese sounds, and solely based on alphanumeric characters without any diacritics and strange symbols. There is also the "S.L. Wong Romanization" created by Wong Shik Ling based on the International Phonetic Alphabet and published in 1941 and, finally the "Yale Romanization", system created by the University of Yale to represent Cantonese, mandarin, Korean and Japanese. Systems are represented as follows:

IPA	Jyutping	S.L.Wong Romanization	Cantonese Pinyin	Yale Romanization
[p]	[b]	[b]	[b]	[b]
$[p^h]$	[p]	[p]	[p]	[p]
[m]	[m]	[m]	[m]	[m]
[f]	[f]	[f]	[f]	[f]
[t]	[d]	[d]	[d]	[d]
[t ^h]	[t]	[t]	[t]	[t]
[n]	[n]	[n]	[n]	[n]
[1]	[1]	[1]	[1]	[1]
[k]	[g]	[g]	[g]	[g]
[k ^h]	[k]	[k]	[k]	[k]
[ŋ]	[ng]	[ng]	[ng]	[ng]

Table 1: Representing Chinese Romanization systems

The phonetic symbols used to describe the consonants are almost all the same once that the phonemes /m/, /f/, /n/ e /1/ are pronounced the same

way in Portuguese and Cantonese. However as it is possible to observe in the Cantonese phonetic system we cannot find the voiced consonants /b, d, g/.

Therefore based on some phonetic transcriptions of Chinese words, it is noticeable the similarity between sounds in Chinese that in European Portuguese are minimal pairs representing different words. For example in Chinese (written in Pinyin) we may have:

Word	Phonetic transcription	meaning
bin1	[pʲin]	guest
pin1	[p ^{hj} in]	piece together
dai4	[tai]	bag
tai1	[tʰai]	embryo
guai4	[k ^w ai]	strange
kuai4	[k ^{hw} ai]	fast

Table 2: Examples from Chinese language

Those examples come to show that the difference between consonants that in European Portuguese are considered completely different and are minimal pairs in Chinese is not so clear, Nunes & Martins, 2013. The difference between dai4 and tai1 is only the tone, or if we see the phonetic transcription is given by the aspiration of the consonant. This phenomenon does not occur in European Portuguese, where we do not have tones or aspiration of consonants. According to Duanmu, 2007, does not exist an agreement whether we can consider [tai] and [thai] a minimal pair, once that this depends if [th] is considered one sound only. Of course that the idea of minimal pair is that two words just differ in one sound only that can distinguish them. Even though we consider that this examples can probably help us to explain the confusion that Chinese learners of Portuguese make with pairs like [tre] (aunt) and [dre] (day) or [cole] (collar) and [kole] (glue) where they present difficulties on perceiving the initial sounds. In European Portuguese the contrast is on the voiced – unvoiced characteristic; [p, t, k], (unvoiced) [b, d, g] (voiced).

German, English or French learners do not have this problem however, Spanish learners of Portuguese, due to the proximity between both languages, show difficulties in terms of perception of these minimal pairs. According to Nunes et al. (2006) the proximity of languages can represent an obstacle to the acquisition of some phonetic segments, namely of European Portuguese. Therefore phonological similarities favor transfer.

4 Methodology

4.1 Corpus

The presented *corpus* was collected from 3 different written works of 25 informants all undergraduate students at the University of Macao attending, at

the time, the third year of the Portuguese Studies Bachelor during the second semester of the Academic year of 2013-2014. During the third year students are already expected to have a great performance in Portuguese with a deeper knowledge of lexicon, grammar and several linguistics aspects (syntax, semantics, prosody, phonetics, etc.) All the works were, as usual corrected and graded, and only the specific cases (words) were highlighted to be described here. All the students will remain anonymous. It is important to emphasize that the *corpus* is very limited; being only a pilot study that hopefully will conduct to the creation of several well supported and different *corpora* to sustain and explain some of the first ideas included here and explain much more and better other problems, constrains and difficulties that are usually felt by Chinese students who want to learn Portuguese.

The most part of the *corpus* presents the more common difficulties that they exhibit in what concerns to European Portuguese minimal pairs [p, t, k] / [b, d, g], but it was also of particular interest to just show some other specific occurrences' on this students such as the confusion between [l] and [r]. One of the concerns to do this work was to provide the students with an ample time to write and even the possibility of using the dictionary, at the same time the subjects of the written works were always of their concern or even open topics where they could write about their own interests. For this study it was not take into consideration the presence of stressed or unstressed vowels, nasals or final position.

In the next 4 tables are exposed 8 different wrong replacements of consonants and selected words where students present discrimination problems:

When [d] becomes a	Correct word	When [t] becomes a	Correct word
[t]		[d]	
Funtamentais	Fundamentais	Pergundar	Perguntar
	(Fundamentals)		(ask)
Matrasta	Madrasta	Enquando	Enquanto
	(Stepmother)		(while)
Secuntária	Secundária	Lendamente	Lentamente
	(Secondary)		(slowly)
Muto	Mudo	Indensificar	Intensificar
	(Mute)		(intesify)
Ententimento	Entendimento	Gado	Gato
	(Understanding)	(Cattle)	(Cat)
Cortão	Cordão	Caudela	Cautela
	(Cord)		(Caution)
Responteu	Respondeu	Imudável	Imutável
	(Answered)		(Immutable)
Ententer-se	Entender-se	Afasdatas	Afastadas
	(Understand each		(apart)
	other)		-
Patrão	Padrão	Vandagens	Vantagens
(Boss)	(Standard)	Varidageris	(Advantage)
Rápita	Rápida		
	(Quick)		
Rapitamente	Rapidamente		
	(Quickly)		
Estúpitas	Estúpidas		
	(Stupid)		
Arrepentem	Arrependem		
	(Regret)		_

Table 3: [t] instead of [d] and [d] instead of [t]

When [k] becomes a [g]	Correct word	When [g] becomes a [k]	Correct word
Compligado	Complicado	Códico	Código
	(Complicated)		(Code)
Signifigados	Significados	Cama	Gama
	(Meanings)	(Bed)	(Range)
Agredito	Acredito	Bricassem	Brigassem
	(Believe)		(Quarreled)
		Conjucação	Conjugação
			(Conjugation)

Table 4: [g] instead of [k] and [k] instead of [t]

When [b] becomes a [p]	Correct word	When [p] becomes a [b]	Correct word
Pompo	Bombo	Exemblo	Exemplo
_	(Drum)		(Example)
Bote	Pote		
(Smal boat)	(Pot)		
Pico	Bico		
(peak)	(Nozzle)		
Recepe	Recebe		
_	(Receives)		

Table 5: [p] in place of [b] and [b] in place of [p]

When [l] becomes a [r]	Correct word	When [r] becomes a [l]	Correct word
Simpres	Simples	Memolizar	Memorizar
	(Simple)		(Remember)
Aperativa	Apelativa (Appealing)	Pobleza	Pobreza (Poverty)
Faram	Falam (Speak)	Valiação	Variação (Variation)
		Dola	Dora (Dora)

Table 6: [r] in place of [l] and [l] in place of [r]

5 Observation and Results

It is necessary to refer that this is a pilot study which has to be complemented with a wider corpus for written problems, oral misunderstood and perception discrimination difficulties. Meaning that a complete study where it can be analyzed the 3 major skills of the students and understand if they are exactly the same or if it varies according to the task they are performing. It is known that German, English or French learners do not have this kind of problems in what concerns to those consonants. Yet, Spanish learners of Portuguese, for instance, due to the proximity between both languages, show difficulties in terms of perception of minimal pairs, namely in what concerns to vowels. According to Nunes et al. (2006) the proximity of languages can represent an obstacle to the acquisition of a foreign language. Phonological similarities favor transfer. The results show a clear difference on the number of words, therefore more problems with those pairs, on distinguish the pair [d], [t] followed by the pair [k], [g] and [l], [r]. It is also necessary to analyze the context of occurrence, which was not in study here. Nevertheless it is possible to observe that the biggest problems are on differentiating [d] and [t] and it can be seen that all this changes and misunderstood can cause communication problems and difficulties, once that some words only changing that consonant have a complete different meaning, and others do not even exist in European Portuguese. The next step it will be to study in an accurate study if this problem co-exist on perception and orally and in all this performances in which contexts are they prevalent (CVC, CCV or initial context).

6 Conclusion

The results, though not very representative, may show already that there is interference from the Chinese phonetic system on the production of certain Portuguese sounds. Studies on the acquisition of phonetic aspects of a Second language are important once that they bring to our knowledge some patterns of production and perception that researchers and Professors can work with and try to implement and or modify in order to improve students acquisition. It has to be stated that studies contemplating the acquisition of Portuguese language by Chinese and Cantonese speakers is still very scarce, so what we may have here it is probably a beginning of research on this field. Different strategies are all welcome and should be use frequently in the classroom context, when our main goal is to give the necessary confidence to the learners and make them become more fluent on a foreign language.

7 Future Work

As it was already referred, it is intended to create a larger *corpus* based on the information that was found in this preliminary study. This new *corpus* will be more vast with several and different examples of all the minimal pairs and contexts (words and text), taking also in consideration the presence or absent of stressed or unstressed vowels and even its articulation.

It will be important to test not only written production but also perception and oral performances and compare the results of all the tasks. Comparing results of each test between students from the very first year and the third and fourth year on Portuguese studies, it will also be interesting to compare if Mandarin and Cantonese speakers present the same difficulties and how do they improve over the years.

8 References

- Ball, M. J., & Rahilly, J. (2009). *Phonetics The Science of Speech*, Hodder Education.
- Bauer, R. S., & Benedict, P. K. (1997). Modern Cantonese Phonology, Trends in Linguistics. Studies and monographs [TILSM]. Berlin: Mouton de Gruyter.
- Bickford, A. C., Floyd, R., (2006) Articulatory Phonetics-tools for analyzing the world's languages, fourth Edition, SIL International, Dallas, Texas.
- Doughty, C. J. (2003). Instructed SLA: Constraints, Compensations, and Enhancement, pp.256-310, The Handbook of Second language acquisition, edited by, Catherine J. Doughty and Michael H. Long, Blackwell publishing.

- Duanmu, S. (2007). The Phonology of Standard Chinese, Oxford University Press, Second edition.
- Fowler, C. A. (2008) Galanticci, Bruno, The Relation of Speech Perception and Speech Production, pp.633-652, The Handbook of Speech Perception, Blackwell publishing.
- Gass, S. M., Schachter, J. (1989) Linguistic Perspectives on Second Language acquisition, Cambridge Applied Linguistics.
- Gui, M.Ch. (2005). The phonology of Gangzhou Cantonese.

 Munchen:Lincom Europa.
- Hayes-Harb, R. (2007) Lexical and statistical evidence in the acquisition of second language acquisition of second language phonemes, Second

- Language research, 23, pp.65-94, SAGE publications,.
- Labov, W. (1991) Sociolinguistics Patterns, University of Pennsylvania Press, Inc..
- Laver, J. (1994). *Principles of phonetics*. Cambridge: Cambridge University Press.
- Martins, M. R. (2002). Fonética do Português, Trinta anos de investigação, Caminho.
- Nunes, A. M. B., & Martins, C. (2013). Perception of European Portuguese minimal pairs by Chinese Speakers first approaches. Linguistic Studies of Human Language Book, Chapter III, 157-148, ATINER.
- O'Brien, I., Segalowitz, N., Collentine, J., Freed, B., & Collentine, J. (2006). Phonological memory and lexical narrative, and grammatical skills in second language oral production by adult learners, Applied Psycholinguistics 27, pp. 377-402.
- Pisoni, D., Remez, B., & Robert E. (2005) The Handbook of Speech Perception, Blackwell Publishing.
- Scott, S. K. (2005). *The Neurobiology of Speech Perception*. Twenty first century psycholinguistics, tour cornerstones, edited by Anne Cutler.
- Shi, L. Ch. (1922). *Chinese phonetic system* and language (English translation). Reprint, Second Edition. Commercial Press Limited: Shangai.
- Zhu, X., & Nong, A. (2010). *Fonética*. Beijing: The Commercial Press.

Enseñanza de E/LE a alumnos de edades tempranas en una escuela de Irlanda del Norte

PIQUER VIVES, INMA

1 Introducción

El objetivo de la propuesta de introducir la lengua española en un centro de educación infantil y primaria de Irlanda del Norte es el de demostrar que el éxito de una propuesta de enseñanza precoz de la lengua (Early Language Learning) radica en la eficacia e idoneidad de la metodología utilizada por encima de otros parámetros tales como la edad de los alumnos o la duración de la misma.

En nuestro país, pese a que la enseñanza de una lengua extranjera en la etapa de educación infantil no sea contemplada en los currículos oficiales, el número de centros que ofrece a sus alumnos la posibilidad de aprender inglés en párvulos de tres, cuatro y cinco años de edad crece año tras año. Algunos estudios afirman que dicha iniciativa carece de impacto ya que el alumnado de mayor edad tiene un ritmo de aprendizaje superior al joven (Muñoz *et al.*, 2003). Dichos estudios afirman que, en el caso del aprendizaje del inglés como lengua extranjera, tan solo tendría sentido el avance en la introducción de la lengua si las condiciones de instrucción fuesen similares a las condiciones naturales de adquisición del lenguaje. Es decir, mediante programas de inmersión lingüística. Realidad poco probable en los centros de educación infantil y primaria de nuestro país.

Por otro lado, algunos de los que, pese a las voces que defienden una introducción de la lengua extranjera más tardía, optaron por diseñar e implementar un programa de introducción del inglés en infantil, han acabado abandonando bajo las presiones de los recortes en educación, la falta de formación del profesorado en el este campo, la oferta limitada de propuestas funcionales y rigurosas o la negativa de los miembros de la plantilla a asumir tal responsabilidad.

Sin embargo, muchos centros que, de forma tímida, se aventuraron en esta iniciativa hace ya varios años, han acabado consolidándola en la escuela, divulgando los resultados obtenidos, integrando su propuesta en el proyecto lingüístico de centro, velando por garantizar su continuidad en cursos superiores y obteniendo resultados que no dejan de sorprender a profesionales, familiares y alumnos que se congratulan ante las habilidades comunicativas plurilingües desarrolladas por los alumnos que aprenden una lengua extranjera en la etapa de educación infantil. El profesorado responsable de esta labor, parte de la idea que un primer contacto de calidad con la lengua garantizará a los alumnos un acercamiento progresivo que repercutirá de forma positiva en sus aprendizajes futuros (Pierce *et al.*, 2014).

Múltiples estudios realizados por el doctor de origen valenciano Álvaro Pascual-Leone, profesor de medicina en la Harvard Medical School de Boston y su equipo constatan que, además de las múltiples ventajas que supone el aprendizaje de una o más lenguas extranjeras para evitar el envejecimiento cerebral, se reconoce la necesidad de iniciar dicho aprendizaje en la niñez (Bak, Nissan, Allerhand & Deary, 2014).

Este es el caso del colegio público Sant Cristòfor de Begues (Barcelona), donde fui artífice y testigo de las experiencias citadas anteriormente y donde llegué a la convicción de que el avance en la introducción de la lengua extranjera, sin responder a las exigencias de los programas de inmersión lingüística, permite la obtención de resultados plausibles por medicación del seguimiento de una metodología adecuada que ponga el énfasis en la actitud y la motivación del alumnado hacia el aprendizaje de la lengua extranjera.

La triangulación de los datos recogidos en el proceso de observación permitió la finalización del proceso de investigación que se presentó en la defensa pública de la tesis doctoral titulada "Iniciació a l'anglès com a llengua estrangera a l'educació infantil" (Piquer, 2005).

Tras años de trabajo e investigación en las aulas del parvulario diversificando metodologías, perfeccionando actividades y ampliando el fondo de recursos dirigidos a la enseñanza del inglés como lengua extranjera en la etapa de educación infantil, la calidad de los resultados obtenidos fue motivo suficiente para pensar en nuevas posibilidades de instrucción en contextos diferentes a partir del diseño, implementación y generalización de propuestas metodológicas.

Con este objetivo en mente, recurrí a algunos de mis contactos en Irlanda del Norte y aprovechando la financiación facilitada por la OAPEE (Organismo Autónomo de Programas Educativos Europeos) a través de la obtención de una beca del PAP (Programa de Aprendizaje Permanente) denominada Job Shadowing planifiqué y diseñé una propuesta de formación intensiva en lengua española en la escuela de verano del colegio de educación primaria Ballycarrickmaddy Primary School de Lisburn (Irlanda del Norte).

2 Contexto de la experiencia

La propuesta diseñada para introducir el español como lengua extranjera en el colegio Ballycarrickmaddy Primary School durante el verano de 2013 en el marco del programa "Summer School" que el centro ofrece a las familias y a sus de los alumnos, tuvo una duración de una semana y la frecuencia diaria de tres horas de instrucción en lengua española. El grupo de alumnos que asistió a las sesiones de español fue mixto, lo hizo de forma irregular y jamás superó la cifra de diez alumnos por sesión. Las edades de los mismos oscilaban entre seis y siete años, respondían a niveles cognitivos distintos y el conocimiento previo que tenían de la lengua española era inexistente o residual (algunos alumnos habían asistido a clases de español en horario extraescolar con anterioridad donde habían aprendido algunas palabras aisladas). Sin embargo, el nivel de motivación de los alumnos por el aprendizaje del español era alto en todos los casos y las expectativas por el curso elevadas.

Ballycarrickmaddy Primary School es un centro pequeño, de una sola línea, que en los últimos años ha crecido sensiblemente. Es un centro de educación primaria ubicado en un entorno rural perteneciente a la población de Upper Ballynderry, cercana a la localidad de Lisburn. El equipo de profesionales del centro, liderado por su director Mr. Colin Ford, comparten el lema de la escuela, "Learning for Life" y consideran que es responsabilidad del centro ofrecer a los alumnos y a sus familias recursos tales como: proyectos y actividades que promuevan la implicación de toda la comunidad educativa, aprendizaje por proyectos cooperativos, una infraestructura que permita el uso de programas telemáticos, espacios de reunión y trabajo que faciliten la participación de las familias y agentes de la comunidad, apostar por la enseñanza del español como lengua extranjera en horario escolar y extraescolar, una amplia oferta de actividades fuera del horario escolar que permita a la familias conciliar la vida laboral con la familiar, etc.

Conocedor de la inminente introducción de una lengua extranjera en el currículum británcio de educación primaria, Mr. Colin Ford se mostró muy interesado en la posibilidad de ofrecer un curso intensivo de español en el verano de 2013 a los alumnos de KS2 de su centro que pudieran estar interesados. Se partió de la premisa que un primer contacto de calidad con la lengua extranjera tendría ventajas de cara a la futura implementación del nuevo Currículum Nacional Británico en 2014, curso a partir del cual la enseñanza de una lengua extranjera es obligatoria desde el segundo curso de educación primaria.

3 Propuesta didáctica

La propuesta diseñada e implementada en el centro Ballycarrickmaddy Primary School de Irlanda del Norte tuvo lugar gracias al acuerdo existente entre las personas encargadas de llevarla a cabo, el apoyo e implicación de las familias de los alumnos en la iniciativa y a la experiencia, en el campo de la enseñanza del inglés como lengua extranjera en la etapa de educación infantil en escuelas de nuestro país, de la profesora que diseñó e implementó la propuesta en el centro.

Las condiciones de implementación de la propuesta piloto no permitieron atender aspectos de especial relevancia tales como: principios de temporización adecuados, tratamiento integrado de las lenguas que se enseñan en el centro, evaluación y continuidad de la propuesta, formación del profesorado en la materia, etc.

Pese a las carencias y dificultades que suponía implementar una propuesta de estas características en el centro, se valoraron los aspectos a favor del pilotaje de una propuesta de enseñanza del español como lengua extranjera a alumnos de segundo curso de educación primaria y se optó por aprovechar una oportunidad única de experimentación y aprendizaje en el aula.

Una vez descritas las carencias de la propuesta, se presentarán sus puntos fuertes así como todos los aspectos tenidos en cuenta para su diseño y posterior implementación en el aula.

3.1 Marco teórico

Los alumnos de entre cinco y siete años de edad tienen una fuerte necesidad de comunicarse y, por lo tanto, un gran interés en aprender estructuras y formas de comunicación que les permita relacionarse con los demás. A esta edad, el aprendizaje de una nueva lengua goza de una fuerte carga motivacional para el alumno que hace que la considere tanto o más atractiva que la propia o propias.

Las investigaciones realizadas en el las áreas de aprendizaje precoz de lenguas y multiculturalidad (Cummins, 1976; Swain, 1981; Genesee, 1994) muestran seis niveles de progreso:

- 1. Aumento de la autoestima del individuo consciente de su aprendizaje.
- 2. Se favorece al desarrollo cognitivo.
- 3. Desarrollo de la competencia lingüística del alumno.
- 4. Facilidad para el aprendizaje de otras lenguas y desarrollo de la competencia comunicativa.
 - 5. Desarrollo de una sensibilidad social superior.
 - 6. Se favorece la creatividad.

Con alumnos de estas edades, la lengua extranjera juega un papel educativo, sobre todo en lo que respecta a la vertiente afectiva hacia la cultura de la misma. A pequeñas dosis y en relación con: las posibilidades cognitivas, estado madurativo y comportamientos culturales determinados correspondientes a la edad del niño, es posible y aconsejable formarlo en la comprensión y apreciación de culturas diversas favoreciendo la comprensión internacional y la fraternidad universal (Titone, 1992).

Existe además, evidencia empírica suficiente para reconocer la superioridad auditiva y oral de los niños respecto a los alumnos de más edad (Taylor, 1974; Schumann, 1975). Los alumnos que se encuentran en la etapa educativa de infantil y primeros ciclos de primaria imitan y reproducen los sonidos de una nueva lengua con mayor precisión, el aumento de la edad muestra una disminución en la calidad de la misma (Volterra and Taeschner, 1978).

Una ventaja adicional para los estudiantes jóvenes es su tendencia natural a entusiasmarse con nuevos retos y propuestas (Robinson, 2009), que sumada al uso de recursos de gestión de aula adecuados por parte del profesorado (Cowley, 2003; Groves, 2009; Tomlinson & Imbeau, 2011; Kaiser & Rasminsky, 2011) y al seguimiento de una metodología de enseñanza eficaz inspirada en los formatos de Bruner (Bruner, 1984): hacer cosas en lengua extranjera en contextos compartidos a partir de sesiones de breve duración y de elevada frecuencia, hacen posible que la enseñanza y el aprendizaje de una lengua extranjera a edades tempranas permita la obtención de resultados exitosos pese al no seguimiento de un programa de inmersión lingüística.

Sobre la figura del profesional recae la responsabilidad de crear los formatos de acción e interacción conjunta que permitan a los alumnos aprender a actuar lingüísticamente en espacios habilitados para esta finalidad. Es primordial la recreación en el aula de situaciones extraídas de la vida cotidiana. Dichas simulaciones deben gozar del máximo realismo posible y responder a los parámetros del discurso: finalidad, papel social de los interlocutores, tiempo, espacio social, tema y canal (Ruiz, 2000).

La tarea principal del docente consistirá en generar comunicación contextualizada en el aula. Esto permitirá que los alumnos puedan comunicarse en la nueva lengua sin dejar de hacer las actividades habituales y cotidianas del día a día en la escuela. Se velará porque los alumnos se acerquen a la lengua y cultura extranjera, así como a nuevos modelos de comunicación humana, no muy alejados de la propia (Piquer, 2006).

Én el diseño de la propuesta, además de atender factores tales como la temporización, el contexto de enseñanza y aprendizaje de la lengua extranjera, los recursos materiales y humanos necesarios, las actividades didácticas y de evaluación, fue prioritario pensar en aspectos tales como: la detección y selección de una cantidad suficiente y variada de input lingüístico, el uso referentes visuales en el aula y la realización de actividades que fomentasen la comunicación en contextos significativos de aprendizaje de la lengua extranjera (Krashen, 1994).

3.2 Temporización y recursos

La propuesta didáctica fue diseñada partiendo de la premisa de ofrecer a los alumnos una oferta amplia de actividades diversificadas y significativas que favorecieran el uso de la lengua extranjera en el aula desde el primer momento (Siqués & Vila, 2014). El pilotaje de la propuesta didáctica se realizó durante cinco sesiones de frecuencia diaria y tres horas de duración cada una. Los espacios utilizados para su realización fueron dos aulas de educación primaria y una sala de usos múltiples. El hecho de disponer de varios espacios de trabajo permitía a los alumnos establecer una asociación espacio-recurso-tarea, hecho que agilizaba su implicación y concentración en cada una de las actividades propuestas. En cuanto a los recursos materiales utilizados, la coordinación previa entre la maestra y los profesionales del centro permitió disponer de todo el material necesario para la correcta realización de las actividades planificadas. En el caso de actividades improvisadas en función de las situaciones acontecidas, se dispuso del material y los recursos al alcance.

3.3 Bloques de actividades

Se reprodujo el mismo formato de instrucción en todas las sesiones. Dicho formato consistía en una secuenciación temporal dividida en cuatro grandes bloques de actividades. La distribución temporal de los bloques se hizo de forma equitativa, variando sensiblemente la duración del último dependiendo del tiempo disponible. Las actividades que exigían más atención y concentración a los alumnos y que ofrecían a la maestra un rol más activo eran las que se ubicaban en la franja horaria de primera hora de la mañana. A medida que avanzaba la sesión, las actividades que implicaban al alumnado más activamente y que exigían su participación y protagonismo iban tomando relevancia.

3.3.1 Presentación de vocabulario

Las actividades del bloque de la primera franja horaria son aquellas que permiten presentar y repasar el vocabulario básico mediante imágenes, canciones y juegos manipulativos. Esta aproximación al vocabulario objeto de aprendizaje desde múltiples enfoques, facilita a los alumnos un aprendizaje gradual respetando ritmos de aprendizaje e inteligencias de los alumnos (Gardner, 2003). Algunos de los recursos utilizados en este bloque son las fotografías o bits (Doman, 1984), los videos musicales, los referentes visuales distribuidos en el aula, los juegos manipulativos, etc. La selección de materiales del interés de los alumnos y el uso de imágenes que evocan el vocabulario trabajado permite a los alumnos recordar y aprender vocabulario y estructuras que podrán utilizar en el segundo bloque de actividades de la sesión de forma contextualizada y significativa (Krashen, 1994).

3.3.2 Uso funcional del vocabulario aprendido

El segundo bloque de actividades de la sesión se enmarca en un contexto ficticio recreado a imagen y semejanza de uno real cercano y conocido por los alumnos. Se trata de reproducir una tienda en el aula en la que los alumnos pueden adquirir el material que necesitarán para confeccionar un mural. En este contexto, el uso del vocabulario y estructuras aprendidas en el bloque anterior es continuo. Es mediante la repetición del mismo que las estructuras aprendidas se afianzan y el vocabulario adquirido cobra sentido y significatividad para los alumnos. La participación de los alumnos en esta actividad de juego simbólico permite la adquisición del vocabulario de forma subconsciente, así pues, se confirma la hipótesis de Adquisición-Aprendizaje presentada en la teoría del Natural Approach de Krashen y Terrell (Krashen & Terrell, 1983).

3.3.3 Conocimiento de la cultura asociada a la lengua

El tercer bloque de actividades difiere de los anteriores en objetivo y formato. En la realización de las actividades de este bloque, la maestra es permisiva con el uso de la L1 y permite que los alumnos la utilicen para expresarse abiertamente. El objetivo de este bloque de actividades es el de acercar la cultura asociada a la lengua objeto de aprendizaje a los alumnos mediante la proyección de videos de promoción turística, canciones con una fuerte carga cultural y la recreación en el aula de conversaciones y debates sobre aspectos culturales relevantes y la comparación de los mismos con los de la cultura propia. El acercamiento cultural en la clase de lengua extranjera vincula afectiva y emocionalmente a los alumnos al folclore, las costumbres y las tradiciones asociadas a la lengua extranjera (Megías, 2008).

3.3.4 Dramatización en lengua extranjera

El cuarto y último bloque de actividades es de duración variable dependiendo el tiempo restante y tiene la función de permitir a los alumnos expresarse libremente a la vez que dramatizan, por imitación, canciones propias del folclore de la lengua objeto de aprendizaje. La función catártica de este bloque temático permite a los alumnos seguir implicados en el proceso de aprendizaje de la lengua extranjera sin acusar, de forma significativa, el cansancio acumulado tras toda una mañana de exposición a la lengua extranjera.

El uso del vídeo como recurso que permite la dramatización de canciones infantiles propias de la lengua española por repetición e imitación, no solo facilita la tarea de la maestra, sino que ofrece a los alumnos modelos distintos de habla de la lengua española. Se trata, además, de un recurso idóneo para la catarsis y la libre expresión en lengua extranjera para los alumnos. Se pudo comprobar que una selección de recursos audiovisuales alineada con los intereses de los alumnos y trabajada en el aula, permite a los niños y a sus familias practicar en casa y compartir los aprendizajes adquiridos. Para hacerlo posible, los enlaces a las actividades realizadas en el aula mediante el uso del video, se pueden encontrar en la página web del centro

(http://www.ballycarrickmaddyprimaryschool.org/summer-spanishschool.html).

4 Divulgación y diseminación de la propuesta

El diseño e implementación de la propuesta llevada a cabo en Ballycarrickmaddy Primary School de Irlanda del Norte se hizo teniendo en cuenta las opciones de difusión de la misma.

Se pensó en las familias de los alumnos participantes, y en las del centro en general, como primer colectivo interesado en conocer los resultados obtenidos con la implementación de la propuesta, así como sus posibilidades de continuidad en el ámbito familiar.

4.1 El contexto familiar

Las familias de los alumnos participantes en el Spanish Summer Course llevado a cabo en Ballycarrickmaddy Primary School en verano de 2013 mostraron interés en que sus hijos participaran en dicho programa. En algunos casos, la asistencia de los alumnos al curso no tuvo continuidad. En otros, sin embargo, fue una prioridad que las familias se afanaron en mostrar con su agradecimiento y predisposición a colaborar en todo cuanto fuese necesario.

4.1.1 Información previa sobre la propuesta

Las familias de los alumnos con los que se planificó realizar la propuesta de enseñanza del inglés como lengua extranjera fueron informadas hacia final de curso para que pudieran planificar su calendario estival de actividades teniendo en cuenta la posibilidad de participar en esta actividad piloto de forma gratuita. Entre sus compromisos, se pidieron los permisos necesarios para que la práctica pudiera ser grabada en video con finalidades académicas y de divulgación de la experiencia.

4.1.2 Página web del centro

La información sobre la propuesta fue también anunciada en la web del centro, abriéndose un espacio en la misma donde se facilitaron los enlaces a los videos de las canciones que se trabajarían en clase. Fueron muchas las familias que aprovecharon este recurso para compartir con sus hijos algunas de las actividades más motivadoras y destacadas de la propuesta.

4.1.3 Invitación a las familias

Con la finalidad de mostrar a las familias los resultados obtenidos y aprendizajes adquiridos, el final de la última sesión de la propuesta fue una clase abierta en la que las familias de los alumnos fueros invitados a presenciar una pequeña muestra de las distintas actividades realizadas durante el curso. Se proyectaron fotografías de distintos momentos del curso, se mostraron los trabajos realizados, los alumnos hicieron una pequeña representación del juego simbólico realizado en el aula e hicieron una dramatización de algunas de las canciones aprendidas. Por último, la profesora hizo la entrega de los diplomas de participación en el curso que los mismos alumnos habían decorado en clase.

Tras la edición de las horas de metraje de las que se dispuso una vez finalizado el curso, las familias pudieron ver en la celebración de la denominada Open Night for Parents del mes de noviembre, un breve documental de menos de quince minutos de duración en el que se presentó el trabajo realizado en el aula con los alumnos durante el Spanish Summer Course.

4.2 El contexto local

La prensa local fue invitada el último día del curso para hacer un breve reportaje. Por motivos ajenos a la organización, no fue posible anunciar el curso en la prensa local.

Sin embargo, la buena práctica en la enseñanza del español como lengua extranjera llevada a cabo en Ballycarrickmaddy Primary School ha sido presentada en varias ocasiones en encuentros mantenidos por directores de diferentes centros de la zona y en muestras de buenas prácticas escolares.

Tras la aceptación del programa piloto, las familias del centro y el equipo directivo han buscado alternativas que permitan dar continuidad a dicha propuesta. El centro colabora con el instituto de educación secundaria local para que, una vez por semana, alumnos de este centro visiten la escuela de educación primaria y ayuden al profesorado con la enseñanza del español. Algunas familias se han ofrecido voluntarias para facilitar el desplazamiento de los alumnos. Dicho programa de Aprendizaje Servicio (APS) tiene continuidad, consolidándose en el centro y sentando un precedente que otros centros de Irlanda del Norte quieren imitar.

4.3 El contexto internacional

Todavía no se han programado iniciativas de difusión en el contexto educativo británico pero, teniendo en cuenta la reciente implementación de los nuevos currículums en las escuelas de Gran Bretaña y la obligatoriedad de la enseñanza de una lengua extranjera desde KS2, existe la voluntad de buscar vías de difusión de la experiencia llevada a cabo en Ballycarrickmaddy Primary School en verano de 2013 en vistas a ofrecer al profesorado y los equipos directivos un ejemplo de buena práctica de aula que permita su extrapolación a contextos de características similares.

La publicación de artículos en revistas reconocidas a nivel nacional e internacional y la presentación de la experiencia en cursos de formación del profesorado, congresos y conferencias son algunas de la medidas de difusión planificadas y en vías de realización.

5 Conclusiones

El diseño e implementación de la propuesta de enseñanza de español lengua extranjera aplicada a alumnos de edades tempranas durante un periodo limitado de tiempo nos permite afirmar que factores tales como la duración del

programa, la edad de los aprendices, o el contexto de aprendizaje de la lengua, pese a su relevancia en el proceso de aprendizaje, pasan a un segundo plano ante aspectos verdaderamente relevantes tales como la adecuación de la metodología utilizada, la destreza del docente en este campo y la selección y uso de recursos alineados con las exigencias de la metodología de enseñanza utilizada.

6 Bibliografía

- Bak, T., Nissan, J., Allerhand, M., & Deary, I. (2014). Does bilingualism influence cognitive aging? *Annals of Neurology*. 75(6): 803-973. http://onlinelibrary.wiley.com/doi/10.1002/ana.v75.6/issuetoc
- Cowley, S. (2003). Sue cowley's teaching clinic. London: Continuum.
- Cummins, J. (1976). The influence of bilingualism on cognitive growth: a sythesis of
- research findings and explanatory hypotheses. *Working Papers on Bilingualism*, 9: 1-43.
- Doman, G. (1984). Cómo multiplicar la inteligencia de su bebé. Madrid: Edaf.
- Gardner, H. (2003). Multiple Intelligences after twenty years. *Paper presented at the American Educational Research Association*. Chicago, Illinois.
- Genesee, F. (1989): Early bilingual development: one language or two? *Journal of*
- Child Language. 16(1-2): 161-179.
- Groves, Eric (2009). The everything classroom management book: a teacher's guide to an organized, productive, and calm classroom. Avon, MA: Adams Media.
- Kaiser, B., & Rasminsky, J. S. (2011). Challenging behavior in young children: understanding, preventing, and responding effectively. Boston, MA: Allyn and Bacon.
- Krashen, S. & Terrell, T. (1983). The Natural Approach: Language Acquisition in the Classroom. Hayward, CA: Alemany Press.
- Krashen, S. (1994). The input hypothesis and its rivals. In N.Ellis ed., *Implicit and Explicit Learning of Languages*. London: Academic Press, 45-77
- Megías, M. (2008). Hola!: A holistic approach to foreign/second language learning for kids. En S. Sánchez & R. Duran (eds.), Didactic approaches for teachers of English in an international context. Salamanca.

- Ediciones Universidad de Salamanca, 11-37
- Muñoz, C; Victori, M., et al. (2003). En torno a los efectos de la edad en el aprendizaje escolar de una lengua extranjera. Eduling, 1. En http://diposit.ub.edu/dspace/bitstream/2445/11541/1/Munozetal2002.pdf
- Pierce, L., et al. (2014): Mapping the unconscious maintenance of a lost first language. PNAS. En http://www.pnas.org/content/early/2014/11/12/1409411111
- Piquer, I. (2006) Aprender inglés en la escuela desde los tres años. Revista Porta Linguarum. 6: 115-128. En http://www.ugr.es/~portalin/articulos/PL_numero6/piquer.pdf
- Piquer, I. (2005). Foreign Language Teaching to Very Young Learners. Keys for Drawing up a Research and Teaching Proposal. *Revista APAC*, 56.
- Robinson, K. (2009). The Element. How finding your passion changes everything. Nueva York: Penguin Group.
- Ruíz, U. (2000): Didáctica de la segunda lengua en educación infantil y primaria. Madrid: Síntesis Educación.
- Schumann, J. (1975). Affective factors and the problem of age in Second Language Acquisition. *Language Learning*, 25: 209-225.
- Siqués, C., & Vila, I. (2014). L'ensenyament de l'anglès a l'educació infantil: mites i reptes. English, what else? Perspectiva Escolar, 373: 11-16.
- Swain, M. (1981). Time and Timing in Bilingual Education. *Language Learning*, 31(1): 1-16.
- Taylor, B. (1974). Toward a theory of Second Language Acquisition. *Language Learning*, 24: 23-35.
- Tomlinson, C., & Imbeau, M. (2011). *Leading* and managing a differentiated classroom. Alexandria, VA: ASCD.
- Volterra, V. & Taeschner, T. (1978). The acquisition and development of

language by bilingual children. *Journal of Child Language*, 5: 311-326.

La percepción de oposiciones fonético / fonológicas en L2

ROMITO, LUCIANO; FRONTERA, MANUELA

1 Introducción

La experiencia de la lengua materna modela la capacidad de percibir y adquirir sonidos no familiares, puesto que todas las características fonético/fonológicas de un idioma intervienen e interfieren en el proceso de percepción de los sonidos de un idioma extranjero. Eso depende en primer lugar de los sistemas fonológicos que subyacen a la ejecución lingüística, los cuales son específicos de cada lengua y actúan como un tamiz, filtrando las características fonéticas relevantes para la identificación de los distintos fonemas, seguidas por aquellas necesarias para garantizar la comprensión comunicativa y finalmente las relacionadas con la personalidad del hablante. Por eso, ante un idioma extranjero L2, el hablante tendrá que adaptar su tamiz fonológico a las características de la lengua de destino, para poder extrapolar las informaciones indispensables para la identificación de los sonidos: "since this sieve is not suited for the foreign language, numerous mistakes and misinterpretation are the result" (Trubetzkoy, 1969: 52). Eso significa que cada aprendiente tiende a interpretar sonidos nuevos, tanto compartidos como no por su sistema nativo, a través de la clave de lectura representada por su propio conjunto de reglas y estructuras fonológicas: inevitablemente, esto llevará a cometer errores de identificación, causa de los mismos errores durante la producción concreta. Es indispensable por lo tanto conocer los obstáculos que

pueden proporcionar la fonología y la fonética nativas a lo largo de las fases de percepción e identificación de nuevos sonidos.

En el aprendizaje de un idioma, la percepción desempeña un papel fundamental: en los procesos que regulan la adquisición del idioma materno, la producción tiene su eje en la competencia fonológica, pero cuando se aprende una L2, la percepción es la primera interfaz entre lengua de origen y lengua de llegada, así que sólo después de la actuación de los procesos perceptivos se pueden distinguir y discriminar los sonidos a nivel fonético y finalmente reconducirlos a categorías fonológicas. Resulta fundamental entender las razones por las que algunos sonidos se interpretan e interiorizan de manera equivocada, haciendo que el hablante siga cometiendo errores en el acto de habla y no consiga corregir o llenar varios vacíos consolidados en el tiempo; es importante, después de haber entendido las causas de esas 'alteraciones lingüísticas', intervenir a priori, tanto sobre el aprendiente precoz como sobre el tardío, intentando que se superen esos obstáculos y proporcionando soluciones para la correcta discriminación de sonidos nuevos.

1.1 Perceptual Assimilation Model y PAM/L2

Catherine T. Best (1994, 1995, 2007) ha intentado sacar a la luz esa cuestión, creando un modelo perceptivo (Perceptual Assimilation Model, PAM y PAM/L2) que explica cómo y por qué los aprendientes de una L2 están predispuestos a discriminar oposiciones lingüísticas no nativas de manera previsible. Las predicciones alcanzadas por Best explican cómo los oyentes tienden a categorizar o asimilar fonos no nativos en relación con las categorías fonológicas de sus propias lenguas maternas, y cómo discriminarán oposiciones de sonidos no nativos. Con el apoyo de las teorías de la percepción directa (Gibson, 1966) y la fonología articulatoria (Goldstein and Fowler, 2003), Best ilustra un modelo basado en la idea central de percepción realista de sonidos lingüísticos considerados como gestos articulatorios, estímulos distales captados por el oyente sin recurrir a conocimientos implícitos o innatos: el oyente percibe informaciones relacionadas a los movimientos articulatorios producidos por el hablante e intenta reconducirlos a sonidos ya asimilados del sistema lingüístico de L1. Best demuestra que cada oyente, sometido a la escucha de sonidos no nativos desconocidos, debido al efecto de la influencia fonológica actuada por parte del sistema lingüístico de la lengua materna, asimilará sonidos no nativos a los más cercanos desde un punto de vista fonético y articulatorio en la L1. La semejanza/discrepancia de los fonos L2 con respecto a L1 será evaluada sobre características espaciales vinculadas a los articuladores activados durante su realización y a los niveles de constricción determinados por los mismos movimientos. Las imágenes que siguen resumen los tipos de asimilaciones y de categorizaciones previstos por el PAM:

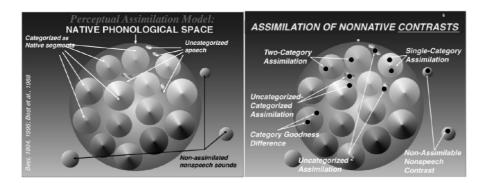


Figura 3. Espacio fonológico nativo y asimilación de las oposiciones no nativas

Como este trabajo se centra en la percepción de sonidos no tanto de un idioma extranjero cualquiera, sino de un idioma de aprendizaje, o sea de una lengua segunda (L2), hay que tener en cuenta la ampliación del modelo perceptivo PAM, es decir el PAM-PAM/L2 (Vayra, Avesani, Best, Bohn, 2008, 2012). Si además de la percepción de oventes totalmente inexpertos de sonidos lingüísticos no nativos, se analiza el comportamiento de oyentes acostumbrados a percibir sonidos de un idioma extranjero de interés, hay que considerar otros parámetros que intervienen en los mismos procesos, aunque las categorías PAM permanecen flexibles y adaptables también a ese tipo de contexto. En general, y eso vale tanto para los hablantes más o menos expertos de L2 como para los hablantes inexpertos, la calidad perceptiva vinculada a los hablantes siempre se verá afectada por la distancia fonológico/fonética entre la L1 de partida y la lengua en examen: paradójicamente, cuanto mayor sea la distancia entre los dos idiomas, menos serán las dificultades de los oyentes en identificar y reconocer sonidos claramente diferentes de los nativos, y que difícilmente se puedan confundir o atribuir a categorías ya interiorizadas. En el caso de sonidos compartidos por los sistemas de los hablantes, la proximidad fonética y fonológica con la L2 podría representar una ventaja, puesto que los sonidos se realizan de manera muy parecida y no representan un obstáculo durante el proceso de adquisición. Además, los problemas de discriminación perceptiva entre sonidos bastante parecidos nacen de idiomas fonotácticamente cercanos a la L2, que pueden no tanto influir sino interferir en la percepción y categorización de los sonidos no nativos. Por lo general, se estima que la percepción puede mejorar de manera directamente proporcional al nivel de competencia lingüística del hablante: esa competencia concierne tanto los hablantes aprendientes de un idioma extranjero (con la experiencia de más años de estudio y de contacto directo y prolongado con el idioma, en contextos de lengua nativa, también mejoran los éxitos en la percepción y la producción), como con la adquisición espontánea del idioma a lo largo de la vida, vinculada al avance de edad (Best identifica etapas fundamentales en la relación entre edad y adquisición del idioma). La ventaja de un oyente aprendiente respecto a un inexperto puede hallarse, a nivel perceptivo, en el hecho de que el primero analizará el estímulo disfrutando de informaciones no solamente articulatorias y fonéticas sino también fonológicas, mientras que el segundo se limitará a la percepción de sonidos producidos a través de gestos articulatorios de la señal vocal, sin un soporte fonológico consolidado. En la teoría se confirma además un punto importante que tiene que ver con la influencia de la L1 sobre la L2. Si es verdad que la distancia entre los sistemas fonéticos y fonológicos de los dos idiomas puede afectar mucho al comportamiento perceptivo del aprendiente, es también cierto que las categorías fonémicas de L1 siempre impactarán sobre la creación de categorías L2, pasando por delante del nivel lingüístico y de la exposición más o menos temprana al nuevo idioma.

2 Objetivos e hipótesis

Este estudio se propone averiguar la capacidad de discriminación de oposiciones fonológicas y fonéticas del español, por parte de hablantes de italiano L1 aprendientes de español L2. Se fijará de manera específica en los resultados obtenidos a partir de hablantes con distintos niveles de conocimiento y competencia lingüística. Según las hipótesis iniciales, se esperará (i) mayor éxito en la discriminación de oposiciones que engloban sonidos nativos, (ii) respecto a aquellas que plantean por lo menos un sonido no nativo, además de (iii) una evidente mejora del rendimiento en la discriminación, vinculada al progreso de los años de estudio del idioma.

Dichas hipótesis se constatarán a través de un experimento de discriminación perceptiva que consta de tres oposiciones distintas: /b/-/d/; $/b/-/\beta/$; $/x/-/\gamma/$ (oposiciones tanto fonológicas como fonéticas). En ellas aparecen sonidos compartidos por los dos sistemas fonológicos y sonidos no nativos para los hablantes de italiano L1.

En el caso de la primera oposición se averiguará la percepción de sonidos idénticos compartidos entre L1 y L2, o sea la oclusiva bilabial sonora /b/ y la oclusiva alveolar sonora /d/. En este caso, sería posible confirmar la hipótesis según la cual oposiciones de sonidos divergentes de la misma manera que en L1 no van a causar ningún problema a los oyentes, y estos asimilarán de manera excelente cada fonema al que le corresponda en su propio inventario fonológico. Se debería encontrar el mismo comportamiento por cada nivel de competencia lingüística en L2.

La segunda oposición contiene un sonido conocido por los hablantes L1 que es la oclusiva bilabial sonora /b/ y otro no nativo, la correspondiente fricativa (o aproximante) bilabial sonora /β/: según las previsiones PAM-PAM/L2 los hablantes italianos, no teniendo en su sistema lingüístico el alófono fricativo del español, deberían asimilarlo al sonido articulatoriamente más próximo, o sea la oclusiva realizada en el punto correspondiente. Esto sería un caso en que ambos sonidos pueden ser asimilados a la misma categoría nativa, pero como realizaciones más o menos buenas de la categoría ideal con la cual se identifica el sonido conocido. Los sonidos podrían así resultar muy parecidos, sobre todo para el oído de los aprendientes menos expertos; el sonido nativo se identificará con éxito, el no nativo, percibido como diferente, podrá causar incertidumbre e indecisión, aunque se espere un reconocimiento correcto por parte de la mayoría de los oyentes. Al tratarse de aprendientes del idioma español, nuestros oyentes no deberían tener muchas dificultades en la discriminación de los dos sonidos, o a lo sumo proporcionar unas vacilaciones en los casos de competencia lingüística menor. Por parte de los oyentes aprendientes expertos se espera una discriminación excelente, puesto que deberían haber interiorizado el empleo de la variante alofónica en determinados contextos y saber distinguirla de la realización oclusiva.

La tercera oposición tiene que ver con dos sonidos ambos ausentes en el inventario fonológico del italiano, o sea las fricativas velares sorda y sonora /x/ y / γ /, donde la primera es un fonema del español, la segunda una realización alofónica de la oclusiva velar sonora que corresponde. Tratándose de sonidos

ambos ausentes en el sistema fonológico y fonético del italiano, ninguno de los dos no será categorizado, pero si por un lado los oyentes deberían reconocer la fricativa sorda fácilmente, como parte de una específica categoría del español, por otro podrían dudar en la identificación del sonido sonoro, que representa una variante más reconocible por los hablantes más expertos; los otros podrían reconducirla al sonido más próximo en italiano, o sea la oclusiva velar sonora /g/, con punto de articulación común. Como este trabajo se detiene sobre la discriminación perceptiva de las oposiciones, es posible predecir que los sonidos, aun causando más incertidumbre respecto a las precedentes oposiciones, van a ser bien discriminadas por parte de los oyentes. Se puede atribuir un margen de error: (i) a la poca familiaridad con sonidos ambos no nativos, que podría tener que ver otra vez con los aprendientes menos expertos; (ii) a la proximidad entre los sonidos, puesto que en los casos de contrastes no asimilables a categorías nativas, la fonología materna no representa un obstáculo para la discriminación, sino que resulta determinante la diferencia entre los dos sonidos.

Con la subdivisión de los oyentes en función de su conocimiento del idioma, se plantea otra variable, que previsiblemente interactúa con la de la dificultad discriminativa, proporcionada por cada oposición lingüística; así las hipótesis que se pretenden contrastar incluyen en primer lugar una diferencia notable en el rendimiento de estudiantes con competencia lingüística distinta, con la probabilidad de descubrir mejoras considerables al progresar de los años de estudio del idioma; segundo, se prevé que las oposiciones se discriminen con un nivel de éxito de tipo $\frac{b}{-d} > \frac{b}{-d}$

3 Metodología

La realización del test de discriminación perceptiva consta de dos etapas: la primera etapa prepara las grabaciones y la organización de los estímulos que formarán parte del test, la segunda es la de la distribución del test a los oyentes, que tendrán que rellenar un cuestionario estructurado de manera adecuada para este experimento.

Se ha grabado a una mujer adulta de origen aragonés, hablante nativa de español L1. La selección de esta hablante no implica la presencia de variables lingüísticas estrictamente regionales, sino que representa un hablante 'ideal' de español/castellano L1 estándar. El experimento se ha sido realizado con estudiantes de español de la Universidad de la Calabria (Cosenza). Se trata de individuos de entre 19 y 41 años de edad, procedentes de la carrera en idiomas extranjeros, entre los cuales 14 son de primer año, 12 de segundo, 7 de cuarto y 19 de quinto (52 oyentes en total)¹. Las grabaciones efectuadas sobre la hablante nativa de español L1 son oposiciones de sílabas CV, donde las vocales empleadas son todas aquellas compartidas entre los sistemas del italiano y del español ([a], [e], [i], [o], [u]). Cada vocal se utiliza para 10 repeticiones, por un total de 150 (50 repeticiones por cada oposición consonántica). Las oposiciones se presentan en grupos de 15 que engloban cada par una sola vez, de manera desordenada y no secuencial, con posición diferente en cada bloque². Entre todas las grabaciones se ha seleccionado las mejores realizaciones por cada

De los 52 estudiantes solo 6 declaran tener relaciones estrechas con hispano-hablantes o haber estudiado o trabajado en el extranjero (países hispanófonos) por lo menos un mes (2 estudiantes) y por un máximo de 15 meses (1 estudiante de quinto año); solamente 3 personas habían estudiado español ya en el instituto. Ninguno de ellos tiene problemas de audición o de percepción auditiva.

² Las grabaciones se efectúan en sala insonorizada con grabador EDIROL R-09 24bit wave/MP3 y micrófono Sennheiser E835. Las grabaciones obtenidas han sido analizadas y etiquetadas a través del software Sound Forge 4.5.

sílaba, utilizadas para la construcción del experimento perceptivo. Se ha intentado utilizar las producciones más controladas tanto en duración como desde el punto de vista de la entonación.

Los estímulos ofrecidos a los oyentes son de tipo AXB, donde A y B son las sílabas que contienen las consonantes en oposición examinadas y X es otra sílaba con la misma consonante que A o B, pero asociada a una vocal diferente. Se proporcionan 8 bloques de 12 tripletes (por un total de 96 estímulos) de tipo AAB-ABB-BBA-BAA, en posición casual pero distribuidos en cada bloque de la misma manera. En cada bloque la misma oposición aparece cuatro veces, pero una sola vez en sílabas con la misma vocal, dos veces con X asociada a A y dos veces a B. Entre las sílabas de un triplete hay un intervalo de un segundo, entre cada triplete y otro un intervalo de 3,5 segundos; cuando todos los oyentes han completado las respuestas de un bloque, se activa la reproducción del bloque siguiente.

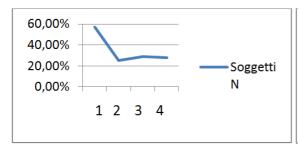
Los 52 oyentes ejecutan el test en grupos de 3 o 4 en sala insonorizada. Cada sujeto dispone de un esquema en papel³ donde se presentan los tripletes divididos en bloques y donde tendrá que marcar las sílabas que contienen el mismo sonido consonántico. La tarea de los estudiantes es la de asociar el sonido consonántico del estímulo X al de la sílaba A o B.

Los análisis de los datos son de tipo descriptivo e inferencial. Los primeros quieren averiguar el grado de competencia lingüística, en términos de 'certidumbre' del oyente, de cada grupo, así como la aceptabilidad de las previsiones de PAM-PAM/L2 sobre el éxito en las discriminaciones perceptivas, estableciendo como parámetro de juicio las varias oposiciones y los bloques de escucha (para evidenciar la medida en que los grupos mejoran su rendimiento a lo largo de la prueba). Los segundos se realizan por medio del software SPSS. Para contrastar las hipótesis iniciales (test no paramétricos) se emplean dos métodos de elaboración de datos, o sea ANOVA y el test Kruskal-Wallis. El objetivo de esos test es establecer el comportamiento de los distintos grupos a través de las respuestas a las categorías de estímulos, y entender cuáles proporcionan mayores dificultades en los oyentes y cuales reciben un rendimiento mejor.

4 Resultados

Los primeros análisis se basan en los porcentajes de respuestas N (nulas) sacadas del test de discriminación. La no respuesta al estímulo se considera, por cierto, como índice de inseguridad e incertidumbre, aunque en un test como el que se ha propuesto puede mejorar a lo largo de su desarrollo, una vez que los sujetos hayan ganado familiaridad con la prueba escuchando más veces los estímulos. Por esta razón se ha calculado primero el valor medio de sujetos que, por cada grupo, no hayan completado todo el test; se ha estimado, después, la media de no respuestas, tanto para el primer bloque de estímulos como para el octavo, para comprobar la variación de respuestas nulas a lo largo de la prueba entera.

³ Véase el Apéndice.



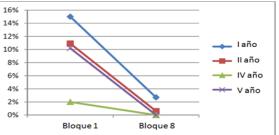


Gráfico 1. Porcentajes de sujetos que no han contestado al test de manera completa

Gráfico 2. Porcentaje de N en los Bloques 1 y 8 de cada grupo

Como se puede entender por estos datos, es evidente que la competencia lingüística en el aprendizaje de una L2 desarrolla un papel fundamental en la capacidad de reconocer oposiciones en la lengua de llegada. Si los oyentes del primer grupo (I año) demuestran menor seguridad, evitando más respuestas, ese porcentaje se reduce mucho en los grupos sucesivos (véase gráfico 1); nótese en particular la diferencia entre estudiantes de primero y segundo año (57,1-25 %), los últimos denotan una gran mejora y mayores dominio y seguridad en la competencia. Los porcentajes de los grupos de estudiantes de cuarto y quinto año se sitúan aproximadamente en el mismo nivel (28,5-27,7 %), a confirmación de una prontitud mayor, consecuencia de la competencia vinculada a aprendientes de L2 de nivel avanzado.

Es posible notar, además, que la posibilidad de respuestas no exhibidas, como se preveía, siempre es más alta en el primer bloque de estímulos y mucho más baja o, para los aprendientes avanzados, nula, en las respuestas al último bloque de estímulos (véase Gráfico 2): aún este dato confirma que en la primera parte de la prueba los estudiantes seguían titubeantes, pero han ganado seguridad y conciencia a lo largo de la prueba. En esta primera etapa de los análisis, se puede afirmar que las hipótesis iniciales sobre la seguridad en la competencia de los hablantes han sido confirmadas.

El análisis siguiente se fija en el nivel de éxito alcanzado por cada grupo en la discriminación de las distintas oposiciones. Se intentará confirmar las hipótesis de las teorías PAM e PAM-PAM/L2 sobre la percepción correcta de sonidos nativos frente a sonidos no nativos y la influencia del nivel de competencia lingüística y de proximidad entre los idiomas. Se analiza a continuación el comportamiento de cada grupo en relación con las tres categorías de oposición: la oposición que engloba sonidos ambos compartidos por los dos sistemas lingüísticos, o sea nativos, es la que, como se había previsto, se discrimina mejor; supuestamente los oyentes no encuentran dificultades para percibir estos sonidos, reconduciéndolos a sonidos ya conocidos de su propia lengua materna. Sin embargo, existe un pequeño porcentaje de errores que habrían podido ser producidos en las respuestas a los primeros bloques, los más problemáticos para cada estudiante, mientras se familiarizaba con el test. Los porcentajes relativos a las otras oposiciones sugieren en primer lugar una incertidumbre mayor para los pares de sonidos nativo-no nativo $/b/-/\beta/$, los cuales, siendo percibidos por los oyentes como sonidos muy parecidos entre ellos, se relacionan con un número más alto de no respuestas, probablemente causadas por indecisión; en las oposiciones /x/-/γ/, aumenta el porcentaje de respuestas no exactas: es verdad que los sonidos no nativos, al no resultar asimilables a ninguna categoría nativa, se discriminan gracias a la distancia entre ellos, y de hecho resultan muy próximos, en cuanto que sólo se distinguen por una oposición de sonoridad. Como preveían las teorías de base confirmadas

por estos análisis, se trata de hecho de las oposiciones más difíciles de discriminar.

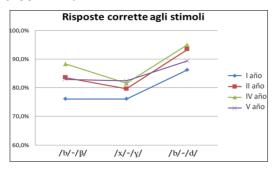


Gráfico 3. Porcentajes de respuestas correctas a los estímulos por parte de cada grupo;

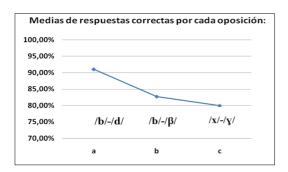


Gráfico 4. Porcentajes medios de discriminaciones correctas por cada oposición

Los análisis estadísticos ANOVA incluyen en primer lugar los resultados obtenidos por los primeros dos grupos de oyentes, de I y II año. La tabla de estadísticas descriptivas muestra los valores de desviación estándar en base a las respuestas exactas en cada categoría de oposición; a corresponde a la oposición /b/-/d/, b a la oposición /b/-/ β / y c a /x/-/ γ /. Los valores de desviación más altos se encuentran en los estímulos constituidos por oposiciones que presentan sonidos no nativos, o sea en b (/b/-/ β /) y c (/x/-/ γ /). Para el primer grupo la desviación baja en relación a los estímulos constituidos por oposiciones nativas, a (/b/-/d/): un valor más bajo es índice de mayor homogeneidad en las respuestas, de hecho se espera encontrar valores significativos en la desviación cuando los oyentes se enfrenten a una tarea más difícil para ellos. Por lo que se refiere al segundo año, se puede apreciar con respecto al grupo precedente una disminución del valor vinculado a las respuestas a los estímulos de tipo b: esto plantea una mejora, con el progreso de la competencia lingüística, en la capacidad de responder a estímulos de sonidos nativos-no nativos (/b/-/β/). La desviación mayor se puede hallar otra vez en las respuestas a los estímulos de sonidos no nativos c (/x/-/ γ /), y eso en cuanto representan motivo de incertidumbre aún para los estudiantes más expertos.

	I año	II año
$a = \frac{b}{-d}$	- desviación (3,8)	- desviación (3,4)
$b = \frac{b}{-\beta}$	+ desviación (5,8)	+ desviación (3,7)
$c = /x/-/\gamma/$	+ desviación (5,6)	+ desviación (6)

Tabla 1. Comparaciones descriptivas con variables en los estímulos (grupos I y II año)

Para averiguar el nivel de homogeneidad de respuestas a los estímulos, dentro de cada grupo, se recurre al test de homogeneidad de Levene: el test resulta significativo solo para el grupo de segundo año, que presenta divergencias, a confirmación de los datos descriptivos, en las respuestas a los estímulos c, o sea de sonidos no nativos $(/x/-/\gamma/)$. El test Anova Univariata (tabla 2) permite confirmar estos datos una vez más.

		Somma dei quadrati	df	Media dei quadrati	F	Sig.
IAnnoTr	Fra gruppi	176,333	2	88,167	3,297	,057
	Entro gruppi	561,625	21	26,744		
	Totale	737,958	23			
IIAnnoTr	Fra gruppi	186,583	2	93,292	4,460	,024
	Entro gruppi	439,250	21	20,917		
	Totale	625,833	23			

Tabla 2. ANOVA Univariata. Esta prueba evidencia la falta de homogeneidad en las respuestas

El nivel de significatividad para este test es de 0,05. Mientras para el grupo de I año no hay significatividad, esta sí se puede apreciar en el grupo II año (tabla 3): el comportamiento de los grupos con respecto a los estímulos es homogéneo, sin embargo existen unas pequeñas diferencias, analizadas en detalle a través de una prueba post hoc de comparaciones múltiples:

Variabile dipendente	(I) stimoli	(J) stimoli	Differenza fra medie (I-J)	Errore std.	Sig.
IIAnnoTr	b	С	1,87500	2,28674	,695
		а	-4,75000	2,28674	,119
	С	b	-1,87500	2,28674	,695
		a	-6,62500 [*]	2,28674	,022
	а	b	4,75000	2,28674	,119
		С	6,62500*	2,28674	,022

Tabla 3. HSD de Tuckey para las comparaciones entre las respuestas asociadas a cada estímulo

Como no hay significatividad para el grupo I año, nos fijamos más en los datos obtenidos para el grupo II año. Los controles cruzados de las respuestas respecto a los estímulos proporcionan resultados significativos exclusivamente para la comparación de c (/x/-/ γ /) a a (/b/-/d/) y viceversa: los estímulos de sonidos ambos no nativos provocan más confusión y menos homogeneidad en las respuestas; los estímulos de sonidos ambos nativos (a), conocidos por los oyentes, se caracterizan al contrario por un comportamiento más homogéneo y

con menos variación respecto a los otros estímulos. Efectivamente se trata de las categorías de estímulos más lejanos entre sí, tanto en términos fonológicos como fonéticos y, sobre todo, con comportamiento perceptivo en las antípodas por parte de no nativos. La comparación entre las medias de b y a (sonidos nativos/no nativos respecto a sonidos ambos nativos) y de b y c (sonidos nativos/no nativos respecto a sonidos ambos no nativos) y viceversa, averigua la hipótesis nula, con valores inferiores al nivel de significatividad establecido, en cuanto las respuestas a estos estímulos no denotan grandes diferencias.

Los datos relativos a los grupos de IV y V año se analizan por medio del test Kruskal-Wallis, test no paramétrico y empleado para tratar muestras de grupos no homogéneos entre ellos. El comportamiento de los estudiantes de los dos grupos, con relación a los estímulos, es muy parecido, sobre todo por lo que se refiere a las respuestas proporcionadas para los estímulos a, que, recuérdese, constan de sonidos ambos nativos, /b/-/d/, conocidos por los oyentes y compartidos por su sistema fonológico. Los rangos vinculados a esos estímulos además son los más altos, confirmación de que se trata de aquellos que obtienen la respuesta mejor. Por lo que hace a los estímulos de tipo b (/b/-/ β /) existe una pequeño margen de diferencia entre los grupos, puesto que el V año ofrece una respuesta mejor; y con todo no se trata de una diferencia significativa. La distancia más considerable entre los grupos se puede observar en los rangos de las respuestas a los estímulos de tipo c (/x/-/ γ /, sonidos no nativos). El grupo V año ofrece una respuesta mejor para esta categoría, lo cual plantea una mejora gradual asociada a competencias más avanzadas, causa de mayor seguridad en los oyentes. Si se examinan los niveles de significatividad también para estos dos grupos (tabla 4), el grupo V año da respuestas bastante coherentes sin alcanzar valores significativos, mientras que el grupo IV año evidencia un comportamiento menos homogéneo en las respuestas.

	I	V Año	V Año
Chi-quadrato	7,770		5,895
df	l,	2	2
Sig. Asint.		,021	,052

Tabla 4. Test Kruskal Wallis con variable de agrupación en los estímulos

Como para los grupos precedentes, se examinan en el detalle los valores significativos, pero en este caso, no pudiendo emplear test de comparaciones múltiples, se utiliza la prueba pairwise. Se analizan sólo los datos significativos (IV año). El V año denota de todas formas una mejora, con respecto al otro grupo, en las respuestas a *b*, aunque no estadísticamente significativo.

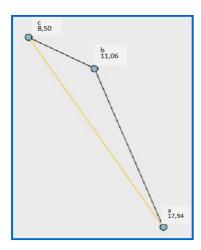


Gráfico 5. Comparaciones pairwise: cada nudo muestra el rango medio de estímulos

El test permite deducir que, como para los grupos anteriores, las diferencias más significativas, en las respuestas a los distintos estímulos, siempre tienen que ver con la comparación entre estímulos c (sonidos ambos no nativos) y a (sonidos ambos nativos); las otras comparaciones no resultan significativas y confirman comportamientos homogéneos en las respuestas hacia a (/b/-/d/) y b (/b/-/ β /), que comparten la presencia de sonidos no nativos, o de sonidos nativos y no nativos muy próximos unos a otros, que siempre pueden causar indecisión, determinando tendencias parecidas en las respuestas. La comparación entre los rangos de a (/b/-/d/) y c (/x/-/ γ /) evidencia una distancia mayor con respecto de la comparación que se acaba de comentar, pero menor con respecto de la única significativa (c – a), la cual engloba sonidos que sí resultan simples para analizar (/b/-/d/), pero también otros más complejos (/x/- γ / γ); esta es por la razón por la cual la tendencia en las respuestas proporciondas para esos estímulos diverge significativamente, confirmando las hipótesis inciales del presente análisis.

5 Conclusiones

Este estudio quiere representar un pequeño paso en la exploración de los procesos perceptivos, que caracterizan el recorrido de adquisición de una segunda lengua, con el objetivo futuro de conseguir, a través de nuevos análisis sobre los procesos de categorización y producción, herramientas útiles y concretas para el empleo de nuevos métodos de enseñanza que tengan en cuenta –más allá de las usuales técnicas didácticas– los aspectos estrictamente fonéticos, para conseguir una mejora considerable en los procesos de percepción y producción de una L2, a niveles más o menos altos de competencia lingüística.

6 Bibliografía

Best, C. T., (1994). The Emergence of Native-Language Phonological Influences in Infants: A Perceptual Assimilation Model. The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words, 167-224, recuperado de

http://www.haskins.yale.edu/publications/pub-b.html el 25/07/14.

Best, C. T., (1995). A direct realist view of cross-language speech perception. Speech Perception and Linguistic Experience, 171-206, recuperado de

- http://www.haskins.yale.edu/publications/pub-b.html el 25/07/14.
- Best, C. T., Tyler, M. D., (2007). Nonnative and second-language speech perception: Commonalities and complementarities. Second language speech learning: The role of language experience in speech perception and production, 13-34, recuperado de http://www.haskins.yale.edu/publications/pub-b.html el 25/07/14.
- Gibson, J. J., (1966). *The senses considered as perceptual systems*, Boston, MA: Houghton-Mifflin.
- Goldstein, L., & Fowler, C. (2003). Articulatory phonology: a phonology for public language use. *Phonetics and Phonology in Language Comprehension*

- and Production: Differences and Similarities, 159-207.
- Trubetzkoy, N. S., (1939). *Principles of Phonology*, translation of Baltaxe, C. A. M., Berkeley, University of California Press, 1969.
- Vayra, M., Avesani, C., Best, C. T., Bohn, O. S., (2012). Non solo dettagli fonetici, non solo categorie fonologiche: l'interazione tra fonetica e fonologia nella percezione di suoni non-nativi, *Studi e Saggi Linguistici*, 119-146.
- Vayra, M., Avesani, C., Best, C. T., Bohn, O. S., (2008). Fonologia e acquisizione. In che modo l'esperienza della lingua materna plasma la percezione dei suoni del linguaggio?. *Processi fonetici e categorie fonologiche nell'acquisizione dell'italiano*, 15-41.

7 Apéndice

Iniziali (nome e cognome)		Sesso []M []F
Età		Anno di corso
Esperienze all'estero (paesi ispanofoni)	[] SÍ	[]NO
Note		

Ascolterai 8 blocchi di sillabe. Ciascun blocco è composto da 12 triplette di tipo A - X - B. Indica per ogni tripletta se la *consonante* del suono intermedio X è uguale a quella della sillaba A o B.

Es.: se pensi che X sia uguale ad A: AXB oppure AXB se pensi che X sia uguale a B: AXB oppure AXB

Ascolterai ogni blocco una sola volta. Tra una tripletta e l'altra disponi di alcuni secondi per poter segnare la tua risposta.

Blocco 1						
1. AXB	2. A X B					
3. AXB	4. A X B					
5. AXB	6. AXB					
7. AXB	8. AXB					
9. AXB	10. AXB					
11. AXB	12. AXB					

Blocco 2	
1. AXB	2. AXB
3. AXB	4. A X B
5. AXB	6. A X B
7. AXB	8. A X B
9. AXB	10. AXB
11. AXB	12. AXB

Blocco 3	
1. AXB	2. AXB
3. AXB	4. AXB
5. A X B	6. AXB
7. A X B	8. AXB
9. AXB	10. AXB
11. AXB	12. A X B

Blocco 4	
1. AXB	2. AXB
3. AXB	4. A X B
5. AXB	6. A X B
7. AXB	8. AXB
9. AXB	10. AXB
11. AXB	12. AXB

Blocco 5	
1. AXB	2. A X B
3. AXB	4. A X B
5. AXB	6. AXB
7. AXB	8. A X B
9. AXB	10. AXB
11. AXB	12. A X B

Blocco 6	
1. AXB	2. A X B
3. AXB	4. A X B
5. AXB	6. A X B
7. AXB	8. A X B
9. AXB	10. AXB
11. AXB	12. A X B

Blocco 7	
1. A X B	2. AXB
3. A X B	4. AXB
5. A X B	6. A X B
7. A X B	8. A X B
9. A X B	10. AXB
11. AXB	12. AXB

Blocco 8	
1. AXB	2. AXB
3. AXB	4. AXB
5. AXB	6. A X B
7. AXB	8. A X B
9. AXB	10. AXB
11. AXB	12. A X B

Aspectos prosódicos de las interrogativas en aprendientes italianos de E/LE

SAVY, RENATA; LUQUE MOYA, JOSE ANTONIO

1 Introducción

1.1 El aprendizaje prosódico

Dentro del aula de E/LE, las sesiones que el docente de español como L2 tiende a dedicarle a la pronunciación y a la entonación suelen ser escasas, pues en la mayoría de los casos la prosodia es entendida como algo secundario en los procesos de adquisición, primando sobre otros aspectos como pueden ser la gramática, el léxico o la ortografía. Se trata de un terreno al que actualmente se le empieza a dar la importancia que se merece. De entre las diferentes destrezas (escuchar, hablar, leer y escribir), la que más nos interesan son las dos primeras, donde la prosodia se hace patente. La enseñanza de hablar y escuchar en el aula es fundamental desde una perspectiva comunicativa y funcional, aunque ello plantee dificultades a la hora de encontrar métodos efectivos, secuenciar contenidos o proponer ítems de evaluación. Quizá por estas dificultades que se plantean, el hablar y escuchar habían sido relegados a un segundo plano como contenidos de aprendizaje, primando la enseñanza/aprendizaje de contenidos fundamentalmente de carácter escrito (Cortés Moreno, 2001).

Quizá influenciado por los propios materiales que existen sobre L2 o por sus propias convicciones, el docente de español como L2 deja de lado las cuestiones entonativas, considerándolo algo secundario, algo que se adquiere a través de

las audiciones hechas en clase, oyendo al profesor u a hablantes nativos¹. La enseñanza de la prosodia es, por tanto, uno de los aspectos a los que menos tiempo (por no decir ninguno) se le dedica. Se trata de algo paradójico ya que es uno de los campos más difíciles y comprometedores para la transmisión del mensaje y, por ende, uno de los que más incide en el éxito o fracaso de la comunicación; además de facilitar la adquisición y el uso de funciones pragmáticas (Sosa, 1999).

Varias pueden ser las razones de la situación de la enseñanza de la prosodia en el aula. Por una parte, la llegada del enfoque comunicativo que relegó la enseñanza de la pronunciación a cuando el alumno podía comunicarse. Por otra parte, en el caso de los profesores no nativos, la falta de autoconfianza en aspectos que jamás aprendieron en su proceso de formación.

A pesar de que las obras dedicadas a la prosodia son numerosas, las obras destinadas a la didáctica de la prosodia son, por contra, escasas (véase Carbó *et al.*, 2003). El espacio destinado a la pronunciación en los manuales² suele ser reducido, siendo frecuente dejarla para el final del libro o del curso.

El resultado de esta situación es que en el proceso de aprendizaje lingüístico de una L2 no se tome en consideración la competencia "metaprosódica" de los aprendientes; los aprendientes se basan en una percepción impresionista de la entonación y tienden a transferir patrones rítmico-entonativos de su L1 (fenómeno conocido como "transfer" prosodico)³.

1.2 Presentación del proyecto

El presente proyecto forma parte de un trabajo más amplio en el que se analizan las realizaciones prosódicas de diversos tipos de "peticiones" (Alfano & Savy, 2013; en prensa; Luque Moya & Savy, en prensa); con el fin de verificar y describir los rasgos entonativos ligados a las diversas funciones pragmáticas, también en clave contrastiva (italiano-español) y con enfoque didáctico. En él se lleva a cabo un análisis y descripción de un tipo particular de petición: la pregunta 'polar' o 'absoluta' (es decir las que inducen una respuesta 'sí' o 'no'), ampliamente estudiada desde un punto de vista prosódico en muchas lenguas, incluidas el italiano y el español (Dorta & Hernández, 2005; Face, 2004; Congosto Martín, 2011; Font-Rotchés & Cantero Serena, en prensa).

La hipótesis de la que partimos es que no existe un mismo desarrollo de la competencia prosódica como de la competencia lingüística general de los aprendientes. Para corroborar o refutar esta idea y verificar fenómenos de transferencia prosódica de la L1 sobre la L2, se lleva a cabo una monitorización de la competencia prosódica en diferentes niveles de aprendizaje.

2 Metodología de investigación

2.1 Informantes

Todos los informantes son mujeres, estudiantes universitarias de E/LE de la provincia italiana de Salerno, con las mismas características anagráficas. Se han seleccionado exclusivamente a los estudiantes expuestos al aprendizaje en el

¹ Esta problemática ya ha sido puesta en evidencia recientemente por autores como Cantero (2003) Carbó, Llisterri, Machuca, de la Mota, Riera & Ríos (2003); Gil (2012); Hidalgo & Cabedo (2012).

² Véase Cortés Moreno (2001) para un análisis de ejercicios de entonación en los manuales de E/LE.

³ Véase, entre otros, Gut (2003); Goad & White (2006); Mennen (2007); Rasier & Hiligsmann (2007); Nguyên, Ingram & Pensalfini (2008); Swerts & Zerbian (2010).

aula, sin experiencia directa en contexto de inmersión, diferenciándolos en función de 3 niveles de lengua, correspondientes a los niveles del M.C.E.R: básico (A2/B1, 10 sujetos), intermedio (B2, 7s.) y superior (C1/C2, 8s.).

Para construir un modelo prosódico de la lengua *target* con la que comparar las producciones de los aprendientes, se han registrado también a las cuatro profesoras de E/LE de los sujetos: todas proceden de diferentes regiones de España y que viven en Italia desde hace 20 años aproximadamente.

2.2 Método de elicitación

El método de elicitación que se ha empleado para la extracción de los datos es la lectura contextualizada de diversos tipos de peticiones, en la que cada enunciado está enmarcado en un contexto situacional descrito y construido en forma de diálogo. El mismo método se ha empleado para las grabaciones de las dos lenguas objeto de la comparación. A los informantes se les ha pedido leer mentalmente los diversos escenarios y a continuación interpretar varias veces en voz alta solo las interrogaciones, fijándose en la situación comunicativa propuesta y acercándose lo máximo posible a una realización espontánea⁴.

2.3 Corpus de enunciados

Entre los diversos tipos de peticiones elicitadas con este método, como se ha dicho antes en este trabajo, solo consideraremos las preguntas absolutas, a partir de ahora llamadas *Query_Yes/No* (Q_Y), según el esquema de anotación pragmática Pr.A.T.I.D. utilizado (véase Savy, 2010). Los enunciados se realizan a través de diferentes estructuras sintácticas, frecuentes y comunes a ambas lenguas. Las principales son tres:

- estructura S-V-O, del tipo "¿el colibrí come semillas?"
- estructura S-V-C (predicado nominal), del tipo "¿el albergue es cómodo?"
- estructura V-S (estructura presentativa), del tipo "¿hay un águila?"

Estas estructuras sintácticas dan lugar a diversas estructuras informativas, cuya organización puede ser altamente relevante para la realización prosódica⁵. Los primeros dos tipos se corresponden con estructuras del tipo *Topic+Comment* (T/C), mientras que la última presenta solo *Comment* (C).

Para determinar si los patrones prosódicos varían en función de la estructura acentual de la palabra, los enunciados han sido construidos variando el perfil acentual de las palabras *target* (para este trabajo, los ítems lexicales en función de sujeto) considerando: 3 agudas, 3 llanas y 3 esdrújulas. El corpus resulta constituido por 27 enunciados diferentes, por un total de 675 ocurrencias producidas por los aprendientes, repetidas en las dos lenguas. A este, se suman 108 producciones de los docentes de E/LE, por un total de 1 458 enunciados.

⁴ El motivo de utilizar la lectura contextualizada, incluso con el riesgo de alejarnos de un modelo realmente espontáneo del habla, se explica por la necesidad de obtener producciones idénticas para cada informante de cada nivel para obtener una comparación prosódica puntual.

⁵ Para la relación entre prosodia y estructura informativa, véase, entre los numerosos estudios, sobre el italiano, Cresti (2000), D'Imperio, Elordieta, Frota, Prieto & Vígario (2005), Giordano & Crocco (2006); y referido al español, De la Mota (1997); García-Lecumberri (1995), Face (2001), Hualde (2002).

2.4 Análisis prosódico

Para el análisis prosódico se ha descrito la curva de f_0 a través del sistema de anotación IntSint (Hirst & Di Cristo, 1998). Después se ha procedido a la descripción de algunos fenómenos relevantes para una comparación, bien entre estructuras informativas y sintácticas diversas, bien entre lenguas diferentes.

El método de elaboración de datos considera, por una parte, el patrón global de la curva melódica (P), por otra parte, el contorno terminal (CT) de la curva a partir de la última vocal tónica, al cual se le reconoce en diferentes modelos⁶ el papel principal de vehículo de la modalidad o función pragmática. Se han distinguido tres tipos principales de CT: descendente (L), plano (S), ascendente (H), el de tipo H es el más frecuente y presenta diversas realizaciones (fig. 1) que, desde un punto de vista puramente descriptivo, se pueden distinguir en tres perfiles que llamaremos H_1 (ligeramente ascendente), H_2 (medianamente ascendente) y H_3 (fuertemente ascendente).

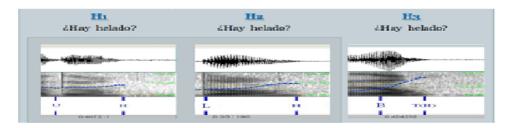


Figura 1. Realizaciones del CT de tipo H

La clasificación de los tres perfiles se basa en la secuencia de las etiquetas⁷, donde H₁ corresponde a la secuencia UH o SH; H₂ a la secuencia LH; H
3 a la secuencia BH, LT o BT. Sin embargo, debido a la escasa frecuencia de la solución H₂ y a la proximidad entre H₁ y H₂, se optó por agruparlas como H₁ que se opone a H₃, caracterizado por una brusca ascensión. Por otra parte, tenemos en cuenta dos fenómenos relevantes: la realización del acento principal y el perfil de los elementos de Topic descritos separadamente porque con frecuencia son realizados en una unidad prosódica separada (Crocco & Savy, 2007). En cuanto al acento (tanto de la unidad prosódica global como de la unidad de Topic) se ha partido de la descripción fonética de *IntSint*, dando lugar a una clasificación en 4 tipos fundamentales⁸: el de tipo H (que comprende las etiquetas T, H, U, U-H, U-T); el de tipo L (etiquetas B, L, D, D-L, D-B); el de tipo HL (etiquetas T-D, T-L, H-D, H-L, H-B, T-B) y, finalmente, el de tipo LH (etiquetas B-U, B-H, L-U, L-H, L-T, B-T). En lo que respecta a la unidad de Topic, se ha descrito el patrón global, el acento y el confín⁹ que presenta. La figura 2 resume ejemplificando los diferentes parámetros analizados:

⁶Véase el modelo de Cantero (2002) y Cantero & Font Rotchés (2009) junto con las contribuciones incluidas en este volumen, que distinguen en la curva melódica una parte inicial (anacrusis), una parte central (cuerpo) y una parte final (inflexión final).

⁷ En estado actual del trabajo, no se han efectuado mediciones acústicas ni verificaciones perceptivas de las diferencias entre los tres perfiles.

⁸ Que son reconocidos en la descripción autosegmental, pero sin ninguna hipótesis o control de fenómenos de alineamiento respecto al núcleo.

⁹ El confín de la unidad de Topic tiene semejanza formal, pero no funcional, con el CT y es descrito en los términos de "tono de frontera" del modelo autosegmental. (por ejemplo, Pierrehumbert, 1980).

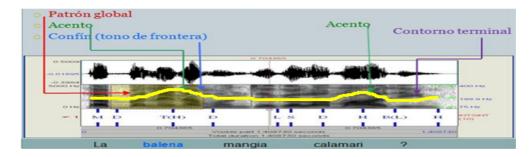


Figura 2. Muestra de los parámetros analizados

2.5 Método de comparación

Como primer paso, se han construido los modelos en L1 para las dos lenguas que nos van a servir de referencia para la comparación prosódica españolitaliano. El modelo españolio es el que presentan las profesoras de E/LE de los sujetos registrados. El modelo italiano ha sido extraído a partir de una muestra de las producciones de 10 informantes. Los modelos se han comparado entre ellos (según los parámetros descritos en §2.4) para extraer las semejanzas y las diferencias e individuar los parámetros necesarios para la comparación con las producciones de la L2. Finalmente, las producciones de cada hablante, para cada nivel y estructura sintáctico-informativa sometida a análisis, han sido comparadas tanto con el modelo de la L1 como con el de la lengua *target*.

3 Modelos L1 y L2

3.1 Comparación de las producciones de los modelos IT-ES

Vamos a mostrar esquemáticamente el resultado de la comparación de los distintos tipos de interrogativas entre el italiano y el español a través de la curva estilizada de ambos modelos acompañada de una descripción de los parámetros que señalábamos en la página anterior, haciendo especial hincapié en lo que tienen en común y lo que los opone.

Las dos estructuras SVO-SVC se corresponden con estructuras informativas de tipo T/C y presentan un perfil próximo (figs. 3 y 4); la única diferencia relevante está en la realización del acento nuclear que presenta diversas variantes, sobre todo en español, donde en la estructura SVC se realiza un acento del tipo HL que modifica el perfil de la curva en la parte preterminal; sin embargo, tal desviación no se repite sistemáticamente. Por ello, de ahora en adelante, las consideraremos como una única unidad. Por otra parte, el perfil de la unidad de Topic realizada por el sujeto no presenta diferencias sustanciales en función de la estructura acentual del *ítem* lexical (aguda, llana y esdrújula)

¹⁰ Para los modelos españoles que nos van a servir de base para llevar a cabo la comparación, hay que recalcar el hecho de que se trata de sujetos que viven en Italia desde hace 20 años y que proceden de diferentes regiones de España. Así se explica que no coincidan con los modelos generales descritos en la bibliografía (Estebas-Vilaplana & Prieto, 2010; Quilis, 1993).

¹¹ El modelo italiano que aquí presentamos no coincide con las descripciones generales que han hecho algunos autores como Crocco (2006) o D'Imperio (2001) sobre el napolitano ya que a pesar de tratarse de una misma región, el salernitano presenta un patrón diferente al napolitano. Así, el salernitano se caracteriza por un Topic muy marcado, debido a su ritmo trocaico, y un CT ascendente, frente a la cadencia napolitana.

por lo que en las páginas que siguen omitiremos una discusión detallada sobre este aspecto.

MODELO ITALIANO SVO	_	MODELO ESPAÑOL SVO
P : ascendente- descendente	Topic: M-T-D T-> tónica	P: plano -descendente Topic: M-T T-> confin
Acento: L o H	CT: H (H ₁)	Acento: L o H CT: H (H ₃)

Figura 3. Estructura SVO. Modelos estilizados IT-ES (estructura informativa T/C)

MODELO ITALIANO SVC		MODELO ESPAÑOL SVC	
P : ascendente- descendente	Topic: M-T-L T-> tónica	P: plano -descendente Topic: M-T T-> confin	
Acento: H	CT: H (H ₁)	Acento: HL CT: H (H ₃)	

Figura 4. Estructura SVC. Modelos estilizados IT-ES (estructura informativa T/C)

En cuanto a la estructura sintáctica VS (*Comment*), en el caso del italiano se han distinguido dos modelos: el primero caracterizado por un patrón ascendente-descendente y acento de tipo HL; el segundo, en cambio, presenta un patrón ascendente-descendente-ascendente, con un acento LH. En el segundo caso, por lo tanto, la salida no ocupa solo el CT sino que alcanza el patrón global, incluyendo el acento. También en este caso las diferencias de estructura acentual del *ítem* sujeto no modifican el perfil global ni el CT del enunciado.

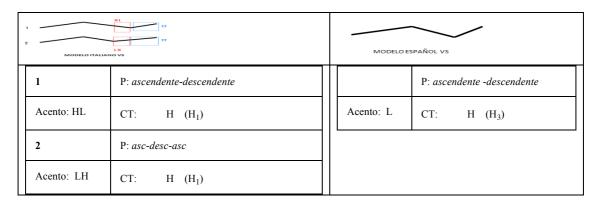


Figura 5. Estructura vs. modelos estilizados IT-ES (estructura informativa Comment)

La comparación entre ambos modelos italiano-español evidencia una serie de semejanzas y diferencias, como manifiesta la siguiente tabla (1). Lo que nos interesa es analizar lo que es diferente entre ambas lenguas.

QUERY YES/NO			
Topic+Comment SVO-SVC	Comment VS		
TOPIC: Diferente.			
P: Diferente.	P: Igual.		
Acento: Igual ¹² .	Acento: Igual.		
CT: Diferente.	CT: Diferente.		

Tabla 1. Semejanzas y diferencias entre ES-IT en las estructuras informativas

4 Realizaciones de aprendientes de E/LE

En las páginas que siguen se presentan los porcentajes comentados de las soluciones que ofrecen los aprendientes italianos de E/LE, comparándolas con el modelo español según lo descrito en §2.5, prestando especial atención a lo que difiere entre los modelos españoles e italianos, y divididas en: nivel, estructura informativa/sintáctica. En el caso de la estructura informativa T/C (estructura sintáctica SVO), la comparación se va a centrar en lo que opone ambos modelos, esto es, el T, el P y el CT. En la tabla 2 se recogen los porcentajes de las soluciones que se acercan al italiano, al español o a otras soluciones.

		I				T		CT	
Nivel	IT	ES	Otras sol.	IT	ES	Otras sol.	IT	ES	Otras sol.
A2/B1	90 %	10 %	0%	77%	9%	14%	72%	1%	27%
B2	95%	5%	0%	83%	8%	9%	83%	0%	17%
C1/C2	77%	23%	0%	65%	14%	21%	47%	13%	40%

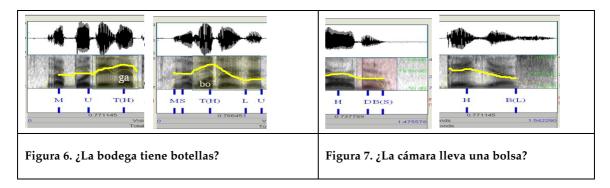
Tabla 2. Soluciones de aprendientes de E/LE para estructura T/C

Como puede observarse, el nivel inicial y el intermedio ofrecen datos muy próximos entre sí, razón por la que se van a comentar juntos. En cambio, al nivel superior (C1-C2), cuyos datos resultan más interesantes, será comentado aparte.

Los dos primeros niveles dan como solución un calco de los patrones prosódicos del modelo italiano sobre el español (*transfer*). En el caso del P, este calco afecta a más del 90 %, frente a un pequeño porcentaje de realizaciones (5-10 %), probablemente casuales, que se acerca al español. En cuanto al Topic y CT, los porcentajes de *transfer* de modelo de L1 son ligeramente más bajos (entre el 72 % y el 83 %); sin embargo, esta disminución no se hace a favor del modelo español, sino de otras soluciones: en el Topic, hacia un desplazamiento del pico de f₀ y, en el CT, hacia perfiles planos (S) o descendientes (L).

Los porcentajes de *transfer* prosódico de la L1 a la L2 de los aprendientes que se encuentran al final de su proceso de aprendizaje (C1-C2) disminuyen en cierta medida, pero con resultados sorprendentes: para P la disminución es relativa (más del 75 % de los enunciados mantiene el perfil de L1). En contra, para el Topic y CT, los porcentajes se reducen a cerca de la mitad, pero tal disminución no conlleva un aumento de las soluciones del español (13-14 %), sino que va en aras de una serie de soluciones alternativas: en el caso del Topic, un desplazamiento del pico de f₀ sobre la sílaba postónica o pretónica (fig. 6); en el caso del CT, hacia perfiles planos (S) o descendientes (L) (fig. 7).

¹² Con algunas variaciones no sistemáticas.



A la vista de los datos del nivel superior (C1-C2), podemos señalar que no existe correlación entre las soluciones que presentan un acento de Topic "español" y las que presentan un CT de tipo H₃, es decir, no se han encontrado ejemplos en los que coincidan ambas soluciones "españolas".

Por lo que respecta a la estructura presentativa VS con estructura informativa *Comment,* la comparación se va a centrar en lo que difieren ambos modelos: el CT. La tabla 3 muestra los porcentajes para cada uno de los niveles.

	CT			
Nivel	IT	ES	Otras soluciones	
A2/B1	66 %	2 %	32 %	
B2	76 %	2 %	22 %	
C1/C2	34 %	18 %	48 %	

Tabla 3. Soluciones de aprendientes de E/LE para estructura Comment.

En los niveles inicial (A2-B1) e intermedio (B2) sigue siendo elevado el porcentaje de calco de L1 sobre L2 (66 %), frente a un mínimo porcentaje de realizaciones (2 %), probablemente casuales, que se acerca al modelo español. Más llamativo es el hecho de que, comparado con la estructura anterior, el calco de L1 haya bajado en un 6 %, sin embargo, esta cifra no se ha traducido en un ascenso significativo del modelo L2, sino que se ha versado sobre el porcentaje de 'soluciones diversas' (como antes, de tipo plano [S] o descendiente [L]).

Finalmente, en los informantes del nivel superior (C1-C2), el porcentaje de las realizaciones que mantienen los patrones de L1 ha descendido a un 34 %, no obstante, este descenso no va a favor del modelo español (18 %), sino que lo ha hecho en pro de soluciones como S y L, ambas con un 24 % de las soluciones totales. Por último, cabe poner en evidencia que en el nivel inicial e intermedio las muestras de CT de tipo H₃ (modelo español) son prácticamente inexistentes (2 %). Es ya en el nivel superior donde empiezan a aparecer ejemplos (fig. 8), aunque las muestras son escasas (18 %).

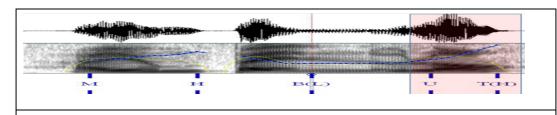


Figura 8: ¿Hay una cámara?

5 Discusión

Un resumen general de los resultados obtenidos confirma nuestra hipótesis (v. §1.2) sobre la ausencia de desarrollo lineal y homogéneo de la competencia prosódica que acompaña el proceso de aprendizaje de una lengua extranjera: si los aprendientes iniciales producen, como es previsible, una transferencia total de los patrones prosódicos de la L1, ningún incremento consciente se registra en los niveles intermedios ni en los avanzados. Los aprendientes de niveles superiores, de hecho, presentan en ocasiones soluciones diversas que, sin embargo, no permiten observar una asimilación consciente y regular del modelo de la lengua target. El único dato extraíble nos induce a hipotetizar un alejamiento de los patrones prosódicos de la lengua nativa que solo casualmente pueden encontrarse soluciones del modelo español o que están a medio camino entre el modelo español e italiano. Esta situación intermedia nos sugiere que sí estamos ante el inicio de un cambio prosódico pero que tal cambio no está dirigido hacia un modelo y se da de manera "desorganizada" y impresionista. El resultado es, por lo tanto, peor si se compara con los procesos de adquisición espontánea de L2; para este fin será útil analizar el comportamiento prosódico de los aprendientes de E/LE que sí han residido en España durante un largo periodo, como planteamos hacer en lo sucesivo de nuestra investigación. Los datos arrojados en este trabajo, de momento, ponen de manifiesto la necesidad de materiales de E/LE que se ocupen de manera directa de cuestiones prosódicas, al igual que imponen una reflexión sobre la oportunidad de desarrollar competencias metaprosódicas tanto en los aprendientes como, incluso más, en los docentes de lengua extranjera.

Bibliografía

- Alfano, I., & Savy, R. (2013). Los patrones entonativos como marcas de frontera: un análisis de algunos tipos de peticiones en italiano y en español. En M. V. Calvi, A. Cancellier & E. Liverani (eds.), Frontiere, soglie e interazioni: i linguaggi ispanici nella tradizione e nella contemporaneità, vol. II-Lingua, Collana Labirinti, 152, 7-26. Trento: Università degli Studi di Trento.
- Alfano, I., & Savy, R. (en prensa, 2014). Pragmática y prosodia de las preguntas pronominales en italiano y en español. Diálogo de la Lengua. Revista de investigación en filología y lingüística, 6.
- Cantero, F.J. (2003). Fonética y didáctica de la pronunciación. En A. Mendoza (ed.), Didáctica de la lengua y la literatura. Madrid: Prentice Hall, 545-572.
- Cantero, F. J., & Font Rotchés, D. (2009). Protocolo para el análisis melódico del habla. Estudios de Fonética Experimental, XVIII, 17-32.
- Carbó, C., Llisterri, J., Machuca M., de la Mota, C., Riera, M., & Ríos A. (2003). Estándar oral y enseñanza de la

- pronunciación del español como primera lengua y como lengua extranjera, ELUA. Estudios Lingüística, 17, 161-179.
- Cresti, E. (2000). Corpus di italiano parlato, Firenze: Accademia della Crusca.
- C. Crocco, (2006).Prosodic informational aspects of polar questions in Neapolitan Italian. En R. Hoffmann & H. Mixdorff (eds.), Proceedings of Speech Prosody, Dresda: **TUCpress**
- Crocco, C., & Savy, R. (2007). Topic in dialogue: prosodic and syntactic features. En Proceedings of Interspeech 2007. Antwerp: ISCA, 114-117.
- Congosto Martín, Y. (2011). Continuum entonativo: declarativas interrogativas absolutas en cuatro variedades del español peninsular y americano. Revista Internacional de Lingüística Iberoamericana (RILI), vol. IX, 1(17), 75-90.
- Cortés Moreno, M. (2001). El papel de la prosodia en la enseñanza de la lengua extranjera: Una revisión de los materiales didácticos. Madrid: Edinumen.

- D'Imperio, M. (2001). Focus and tonal structure in Neapolitan Italian. *Speech Communication*, 33, 339-356.
- D'Imperio, M., Elordieta, G., Frota, S., Prieto, P., & Vigário, M. (2005), Intonational phrasing in Romance: The role ofsyntactic and prosodic structure. En S. Frota, M. Vigário & M. J. Freitas (eds.), *Prosodies*. The Hague: Mouton de Gruyter, 59-98.
- De la Mota, C. (1997). Prosody of sentences with contrastive new information in Spanish. En A. Botinis (ed.), Intonation: theory, models and applications, Proceedings of an ESCA Workshop. ISCA Archive, 75-78. En http://www.isca-speech.org/archive.
- Dorta, J., & Hernández, B. (2005). Análisis prosódico de un corpus de habla experimental: interrogativas absolutas con expansión en el objeto vs. sin expansión. Estudios de Fonética Experimental, XIV, 67-123.
- Estebas-Vilaplana, E., & Prieto, P. (2010).

 Castilian Spanish Intonation. En P.
 Prieto & P. Roseano (eds.),
 Transcription of intonation of the
 Spanish language. Munich: Lincom
 Europa, 17-48.
- Face, T. L. (2001). Focus and early peak alignment in Spanish intonation. *Probus*, 13, 223-246.
- Face, T. L. (2004). The intonation of absolute interrogatives in Castilian Spanish. *Southwest Journal of Linguistics*, 23(2), 65-79.
- Font-Rotchés, D., & Cantero Serena, F. J. (en prensa). Análisis contrastivo de la entonación de la interrogativas absolutas del castellano y del catalán. En Y. Congosto (ed.), Fonética experimental, educación superior e investigación. Cáceres/Sevilla: Univ. de Extremadura/Univ. de Sevilla.
- García-Lecumberri, M. (1995). Intonational Signalling of Information Structure in English and Spanish. A comparative Study, PhD Diss., University of London.
- Gil, J. (ed.) (2012). Aproximación a la enseñanza de la pronunciación en el aula de español. Madrid: Edinumen.
- Giordano, R., & Crocco, C. (2006). Sul rapporto tra articolazione intonativa e articolazione informativa. In F. Albano Leoni & R. Giordano (eds.), *Italiano Parlato: analisi di un dialogo*. Napoli: Liguori, 159-188.
- Goad, H., & White, L. (2006). Prosodic Transfer: L1 Effects on the

- Production of L2 Determiners. En D. Bamman, T. Magnitskaiaand & C. Zaller (eds.), Proceedings of the 30th Annual Boston University Conference on Language Development. Somerville, MA: Cascadilla Press, 213-224.
- Gut, U. (2003). Prosody in second language speech production: the role of the native language. Zeitschrift für Fremdsprachen Lehren und Lernen, 32, 133-152.
- Hidalgo, N., & Cabedo, A. (2012). La enseñanza de la entonación en el aula de ELE. Madrid: Arco Libros.
- Hirst, D., & Di Cristo, A. (1998). A survey of intonation systems. *En Intonation Systems: A Survey of Twenty Languages*. Cambridge, UK: CUP.
- Hualde, J. I. (2002). Intonation in Spanish and the other Ibero-Romance languages: Overview and status quaestionis. En C. Wiltshire & J. Camps (eds.), Romance phonology and variation. Selected papers from the 30 th Linguistic Symposium on Romance Languages. Amsterdam: John Benjamin, 101-116.
- Luque Moya, J. A., & Savy, R. (en prensa).

 Un método para la enseñanza asistida de la prosodia en aprendientes italófonos de ELE. En I. Arroyo, S. Musto, & V. Ripa (eds.), La codifica e la sua interpretazione. Livelli di codifica e di interpretazione degli enunciati.

 Napoli: Edizioni Scientifiche Italiane.
- Mennen, I. (2007). Phonological and phonetic influences in non-native intonation. En J. Trouvain & U. Gut (eds.), Non-native Prosody: Phonetic Descriptions and Teaching Practice. Mouton De Gruyter, 53-76.
- Nguyen, A. T. T., Ingram, J., & Pensalfini, R. (2008). Prosodic transfer in Vietnamese acquisition of English contrastive stress patterns. *Journal of Phonetics*, 36, 158-190.
- Pierrehumbert, J. B. (1980). The Phonology and Phonetics of English Intonation. PhD dissertation, MIT.
- Sosa, J. (1999). La entonación del español, su estructura fónica, variabilidad y dialectología. Madrid: Cátedra.
- Quilis, A. (1993): Tratado de fonología y fonética españolas. Madrid: Gredos.
- Rasier, L., & Hiligsmann, Ph. (2007). Prosodic transfer. Theoretical and Methodological Issues. *Nouveaux cahiers de linguistique Française*, 28, 41-66.

- Savy, R. (2010). Pr.A.T.I.D: a coding scheme for pragmatic annotation of dialogues. En Proceedings of LREC 2010. Paris: ELRA, 2141-2148.
- Swerts, M., & Zerbian, S. (2010). Intonational differences between L1 and L2 English in South Africa. *Phonetica*, 67, 127-146.

Parte III. Fonética clínica

Velocidad de articulación y estructuración del espacio vocálico en función del acento en la afasia: un estudio preliminar

BAQUÉ, LORRAINE

1 Introducción

La afasia es una patología adquirida, que resulta de una lesión cerebral en las áreas del lenguaje, habitualmente en el hemisferio izquierdo, y que se caracteriza por un déficit del procesamiento lingüístico que afecta la expresión y/o la comprensión lingüística. Aunque la tipología de las afasias es objeto de controversia (Ardila, 2010), desde el punto de vista del componente fonéticofonológico se distinguen tradicionalmente las afasias por lesión anterior (p. ej. Broca) y posterior (p. ej. Wernicke o de Conducción). Las primeras se caracterizan por una habla no fluente, con numerosas distorsiones fonéticas que resultarían de un déficit a nivel "motor", mientras que las últimas, llamadas "fluentes", están asociadas a un déficit a nivel "premotor" (MacNeilage, 1982) que implica un déficit de la selección y/o ordenación fonémica. Con todo, numerosos estudios han mostrado que dichos déficits no son mutuamente excluyentes, y que, por ejemplo, se observan errores fonológicos (p. ej. metátesis) en las producciones de afásicos de Broca y un "subtle phonetic deficit" en los afásicos de Conducción (Vijayan & Gandour, 1995). Además la mayoría de trabajos sobre los errores fonético-fonológicos en la afasia se basan en el análisis de producciones consonánticas, mientras que el patrón acentual y el sistema vocálico han sido menos estudiados por considerarse globalmente preservados en la afasia. Sin embargo, estos dos aspectos son actualmente objeto de nuevos estudios.

Así, en lo tocante a la estructura acentual, distintos autores han descrito la ausencia de errores de patrón acentual en la afasia (p. ej. Nickels & Howard, 1999) coincidiendo con estudios previos (Dell, Schwartz, Martin, Saffran & Gagnon, 1997; Ellis, Miller & Sin, 1983; etc.) que consideraban que, incluso en las parafasias fonémicas o en los neologismos, los afásicos tienden a preservar la estructura métrica (esto es el patrón acentual y el número de sílabas). Con todo, se ha descrito la existencia de errores acentuales y se ha llegado a postular un déficit específico de la representación acentual en la afasia (Cappa, Nespor, Ielasi & Miozzo, 1997). Diversos de estos trabajos (para una revisión, cf. Howard & Smith, 2002) observan que dichos errores aparecen en palabras con patrones irregulares (o menos frecuentes) en la lengua considerada, con una eventual interacción con la estructura silábica. Para dar cuenta de dichas observaciones, la mayoría de modelos psicolingüísticos (Butterworth, 1992; Levelt, Roelofs & Meyer, 1999; Levelt, 1992) consideran que se accede por separado a la forma acentual y fonemática del léxico y que ambos aspectos se unen finalmente para formar palabras fonológicas. El debate radica en si la estructura acentual del léxico está almacenada en todos los casos (Butterworth, 1992; Laganaro, Vacheresse & Frauenfelder, 2002) o solo para los patrones acentuales irregulares mientras que sería calculada en el caso de los patrones más frecuentes (Levelt, 1992). Cabe señalar además que distintos autores (Howard & Smith, 2002; Nickels & Howard, 1999; Niemi, Koivuselkä-Sallinen & Hänninen, 1985; Pate, Saffran & Martin, 1987) han puesto de manifiesto un efecto del acento léxico sobre la integridad segmental de las producciones. Así, la probabilidad de omisión de sílabas y de errores fonemáticos aumenta significativamente en las sílabas átonas.

Por lo que respecta a las vocales, más allá de los estudios que concluyen a una relativa preservación del sistema vocálico (p. ej. Jacks, Mathes & Marquardt, 2010; Kent & Rosenbek, 1983; Ryalls, 1986), distintos autores han informado de la existencia de diferencias entre las producciones vocálicas de pacientes afásicos comparadas con las de sujetos control. Así, se han observado, en las afasias anteriores, alargamientos excesivos (Adem, 2006; Baum, Blumstein, Naeser & Palumbo, 1990), una reducción de la extensión del eje del primer y segundo formantes (Haley, Ohde & Wertz, 2001; Keller, 1978; Ryalls, 1981), una posteriorización vocálica y una reducción del área vocálica (Sugathan & Rajasudhakar, 2011; Whiteside, Grobler, Windsor & Varley, 2010). Asimismo, se ha constatado una mayor variabilidad intrafonémica en ambos tipos de afasias (Adem, 2006; Haley et al., 2001; Ryalls, 1986). Además, la duración vocálica se adapta a la velocidad de elocución (Baum, 1993) en el caso de las afasias fluentes vs no-fluentes, se observan menos errores vocálicos en sílaba acentuada que en sílaba átona (Semenza et al., 2007), y las realizaciones vocálicas se ven afectadas por factores como el contexto consonántico, la estructura silábica, el número de sílabas, etc. (Collins, Rosenbek & Wertz, 1983; Gelfer, Bell-Berti & Boyle, 2004; Semenza et al., 2007). Estos estudios apuntan a la existencia de un déficit fonético en las producciones vocálicas de las afasias anteriores y posteriores en el marco de una organización fonémica del sistema vocálico relativamente preservada. Sin embargo cabe añadir que se ha avanzado recientemente la hipótesis de la existencia de dos niveles de representación distintos para las consonantes y las vocales (Knobel & Caramazza, 2007) basándose en la descripción de una doble disociación entre sujetos afásicos con déficit en el procesamiento de las vocales y de las

consonantes. Si bien la descripción propuesta de dos sujetos afásicos con parafasias fonémicas esencialmente vocálicas pone en duda la necesaria preservación fonológica del sistema vocálico en esa patología (cf. también Semenza *et al.*, 2007), la hipótesis avanzada por Knobel y Caramazza es todavía muy controvertida (Monaghan & Shillcock, 2007).

Con este estudio preliminar pretendemos llevar a cabo un análisis acústico de producciones vocálicas en español que permita determinar en qué medida el sistema vocálico de los pacientes afásicos (fluentes y no fluentes) se distingue o no del de los sujetos sanos, y si el carácter acentuado o átono de la sílaba incide en la precisión vocálica. El objetivo es el de verificar si los dos grupos de afásicos presentan una reducción de su capacidad de producir gestos articulatorios distintivos para los diversos fonemas vocálicos. Para ello, se usan tradicionalmente índices como la ratio de extensión del eje del primer y segundo formantes (F1RR y F2RR), que se definen como índices de la capacidad de distinguir las vocales por su apertura y su lugar de articulación más o menos anterior o posterior respectivamente, o el área del espacio vocálico (VSA) en un espacio bi-dimensional F1-F2. Se ha demostrado en distintas patologías del habla que existe una correlación entre el tamaño del área vocálica y la inteligibilidad (p. ej. Weismer, Jeng, Laures, Kent, & Kent, 2001). El índice de la ratio de centralización formántica FCR (Sapir, Ramig, Spielman & Fox, 2010) da cuenta de la tendencia a "neutralizar" o "centralizar" las producciones vocálicas observada en distintas patologías del habla o del lenguaje. Recientemente se han creado nuevos índices para dar cuenta no solo de la distribución de los valores formánticos medios de las distintas vocales sino también de la distribución de la variabilidad intra- e interformántica (Huet & Harmegnies, 2000): los índices de dispersión intra- e interfonémica (CMintra y CMinter, respectivamente) y el "índice de organización del sistema" Phi. Diversos estudios (Audibert & Fougeron, 2012; Ciocca & Whitehill, 2013) han evidenciado el interés de llevar a cabo análisis comparativos de estos distintos índices para ver cuáles son aquellos que permiten un mejor diagnóstico diferencial de las patologías del habla entre sí.

2 Metodología

Han participado en este experimento tres grupos de sujetos: 4 afásicos de Broca, 4 afásicos de Conducción y 4 sujetos control, diestros y apareados en edad (medias de 56,5, 51,8 y 54,8 años respectivamente) y nivel socio-educativo (estudios secundarios).

Se ha sometido a todos los sujetos a una tarea de repetición de 240 palabras aisladas extraídas del corpus COGNIFON-español (Baqué *et al.*, 2006) en las que las vocales aparecen combinadas con las consonantes oclusivas, fricativas y africada del español en distintas estructuras silábicas y en sílaba átona y tónica.

En primer lugar tres fonetistas habituadas al habla afásica han evaluado todas las producciones y se han excluido todas aquellas que presentan un error de número de sílabas, de posición acentual, de sustitución vocálica o de pausa intraléxica según al menos una de las evaluadoras. En segundo lugar se ha procedido a analizar mediante el programa Praat 5.3.77 (Boersma & Weenink, 2014) las duraciones vocálicas y silábicas, así como el valor de los dos primeros formantes en su punto medio. Se han calculado posteriormente la velocidad de articulación (en sílabas por segundo) y, para cada sujeto, siete índices de estructuración del espacio vocálico a partir de los valores formánticos transformados en Bark (Audibert & Fougeron, 2012): la ratio de extensión del primer y del segundo formantes (F1RR y F2RR), el índice de centralización

formántica (cFCR), el área del pentágono vocálico (pVSA), la dispersión intra- e intercategorial (CMintra y CMinter) y el índice Phi de organización vocálica (Huet & Harmegnies, 2000).

A partir de los datos obtenidos, se han llevado a cabo dos tipos de análisis estadísticos para determinar el efecto del grupo (Broca, Conducción, Control) y del acento sobre los distintos parámetros considerados: para los aspectos temporales, regresiones lineales con efectos mixtos (con los sujetos y los ítems como variables aleatorias) y, para los valores de los índices del espacio vocálico, ecuaciones de estimación generalizada (con los sujetos como variable aleatoria).

3 Resultados

3.1 Descripción de los casos excluidos

Como se ha indicado supra, se han excluido del análisis las realizaciones que pueden introducir variables parásitas para el objetivo del estudio, esto es aquellas en que no coinciden el número de sílabas, el patrón acentual o alguna de las vocales, o en que aparece una pausa intraléxica.

Tipo de error	Número mínimo-máximo de ítems excluidos por sujeto (/240)			
	Broca Conducción		Control	
Pausas intra-ítem	5-11	1-44	0	
Ítem truncado	0-4	0-4	0	
+ 1 sílaba	0-4	0-1	0	
Patrón acentual	0-1	0	0	
Doble acentuación	0-1	0-1	0-11	
Sustitución	0-2	0-1	0-0	
vocálica				

Tabla 1. Distribución de las realizaciones excluidas en función del tipo de error y del grupo de sujetos

Los datos presentados en la tabla 1 permiten observar que los errores de patrón acentual o de sustitución vocálica son casi inexistentes, lo que parece indicar que los pacientes objeto de este estudio no presentan déficit en lo que respecta al acceso a la estructura métrica o a los fonemas vocálicos. En el resto del trabajo, presentaremos un análisis de las características de las realizaciones vocálicas de los distintos grupos de sujetos en función del carácter acentuado o átono de la sílaba.

3.2 Velocidad de articulación y duración vocálica

Los resultados obtenidos muestran un efecto simple del grupo sobre la velocidad de articulación (F(2, 9) = 11.794, p = .003), y los tests post-hoc muestran que los sujetos control presentan una velocidad media superior (3.92 síl/s.) a la de los afásicos de Conducción (3.10) y a de Broca (2.59), sin que exista diferencia estadísticamente significativa entre los dos grupos de afásicos.

¹ Cabe añadir un caso de parafasia verbal, que también ha sido excluída : "blanquita" → "platita".

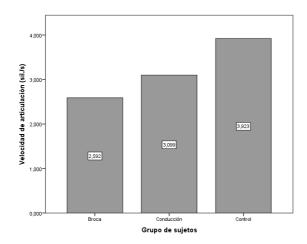


Figura 1. Velocidad de articulación (en sílabas por segundo) en función del grupo de sujetos

Se ha llevado a cabo un análisis separado de las duraciones vocálicas según si se hallan o no en sílaba prepausal.

Las duraciones vocálicas de sílabas no prepausales varían en función del grupo de sujetos (F(2, 9,0) = 8,731, p = ,008) y del carácter acentuado o átono de la sílaba (F(1, 479,6) = 53,136, p < .000). Existe asimismo un efecto de interacción del grupo y el acento (F(2, 4226,0) = 26,296, p < ,000). Los tests post-hoc muestran que la duración vocálica media es menor en los sujetos control que en los dos grupos de afásicos (0,098 vs. Conducción 0.154 y Broca 0.224 s), en sílaba átona que en sílaba tónica (0,149 vs. 0,165 s) y que el alargamiento en sílaba tónica es mayor en los sujetos control que en los dos grupos patológicos (35,37% vs. Conducción 9,59% y Broca 1,35%).

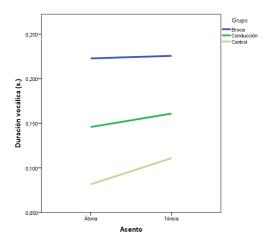


Figura 2. Duración vocálica de sílabas no prepausales en función del grupo de sujetos y del acento

Respecto a las duraciones vocálicas de sílabas prepausales se observa de nuevo un efecto simple del grupo de sujetos (F(2, 14,08) = 22,674, p < 000), del acento (F(1, 1731,78) = 45,308, p < 000) y un efecto de interacción entre las dos variables (F(2, 1756,42) = 20,109, p < 000). Los resultados de los tests post-hoc muestran de nuevo una duración superior en los sujetos patológicos que en el grupo control (Broca: 0,271 m Conducción: 0,244 vs. Control: 0,170 s), pero mientras que los dos grupos de afásicos alargan significativamente (p < 0.05) la sílaba tónica respecto de la átona (Broca: +43,49%, Conducción: +26.75%), se

observa la tendencia inversa (-20,59 %, diferencia no significativa) en los sujetos sanos.

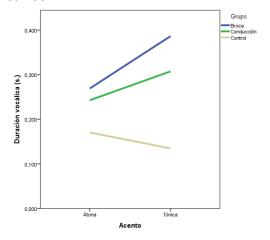


Figura 3. Duración vocálica de sílabas prepausales en función del grupo de sujetos y del acento

3.3 Índices de estructuración del espacio vocálico

El índice de extensión del eje del primer formante F1RR se ve afectado por el carácter acentuado o no de la sílaba (Wald- $X^2(1) = 37,507$, p <,000), lo que se caracteriza por una ampliación de la ratio de extensión del eje en sílaba tónica (1,84 vs. 1,60). Se observa asimismo una tendencia (Wald- $X^2(2) = 5,924$, p =,052) a la reducción del índice F1RR en los afásicos de Conducción (1,60) respecto a los afásicos de Broca y a los sujetos control (1,79 y 1,78 respectivamente), sin efecto de interacción entre el grupo y el acento.

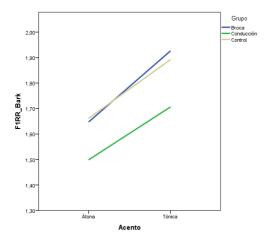


Figura 4. Índice de ratio de extensión del primer formante F1RR en función del grupo de sujetos y del acento

Respecto a la extensión del eje del segundo formante, F2RR, solo se ve afectado por el acento (Wald- $X^2(1) = 23,591$, p <,000): el índice F2RR es mayor en sílaba tónica (1,46) que en sílaba átona (1,34).

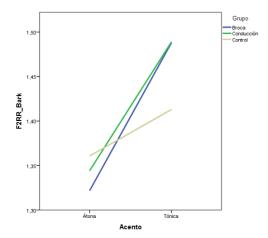


Figura 5. Índice de ratio de extensión del eje del segundo formante F2RR en función del grupo de sujetos y del acento

Del mismo modo, el único efecto significativo (Wald- $X^2(1) = 41,058$, p <,000) sobre el índice de centralización formántica cFCR y sobre el área del pentágono vocálico pVSA es el acento. Así se observa, para los tres grupos de sujetos, que las sílabas tónicas se caracterizan por una menor centralización (0,78 vs. 0,88) y por una mayor área vocálica (7,50 vs. 4,33).

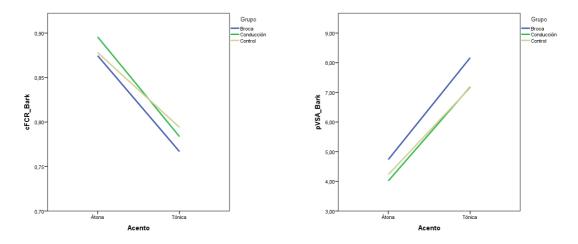


Figura 6. Índices de centralización formántica (cFCR, izquierda) y del área del pentágono vocálico (pVSA, derecha) en función del grupo de sujetos y del acento

Respecto al índice de dispersión intracategorial CMintra, no se obtiene ningún efecto significativo a p <,05, pero sí se observa una tendencia (p <,100) de un efecto de interacción entre el grupo y el acento (Wald- $X^2(2) = 5,221$, p = ,073), que se caracteriza por un valor significativamente inferior en sílaba tónica que en sílaba átona en los afásicos de Broca (101,43 vs. 106,56), contrariamente a lo que sucede (aunque sin diferencias significativas) en los sujetos sanos (102,79 vs. 99,87) y en los afásicos de Conducción (106,76 vs. 104,82).

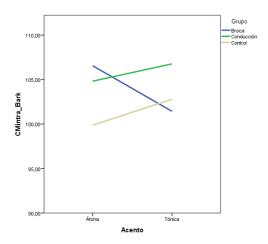


Figura 7. Índice de dispersión intracategorial CMintra en función del grupo de sujetos y del acento

En cuanto a la dispersión intercategorial CMinter, no existe ningún efecto simple pero sí un efecto de interacción entre el grupo y el acento (Wald- X^2 (2) = 7,261, p =,027), que se manifiesta por valores superiores en los afásicos de Broca que en los de Conducción o en los sujetos control (428,71 vs. 328,54 y 360,25 respectivamente), sin diferencias significativas en función del acento, excepto para los afásicos de Conducción para los que, al igual que los sujetos control, la sílaba acentuada está asociada a un mayor índice de dispersión intercategorial (353,40 vs. 303,67).

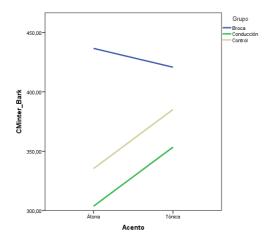


Figura 8. Índice de dispersión intercategorial CMinter en función del grupo de sujetos y del acento

Por último, no se observa ningún efecto significativo del grupo o del acento sobre los valores del índice de organización vocálica Phi.

4 Conclusiones

Los resultados obtenidos confirman en primer lugar que la velocidad de articulación aparece reducida en la afasia de Broca como era de esperar, pero también en la de Conducción, como mínimo en repetición de palabras aisladas. Este resultado va en el sentido de otros (Bird & Franklin, 1996) que observaron una reducción de la velocidad de articulación también en afasias fluentes.

Por otra parte, si bien ninguno de los dos grupos de afásicos comete errores significativos de patrón acentual, los pacientes alargan la vocal tónica respecto de la átona de manera diferente a como lo hacen los sujetos sanos. Así, en las

sílabas no seguidas de pausa la vocal tónica se distingue de la átona por un alargamiento importante en las producciones del grupo control, mientras que la diferencia es muy limitada en ambos grupos de afásicos. En sílabas prepausales el alargamiento de la vocal tónica es muy marcado en los dos grupos de pacientes contrariamente a lo que sucede en los sujetos sanos.

Respecto a los índices relacionados con el área vocálica, los resultados obtenidos muestran que, comparadas con las del grupo control, las producciones de los afásicos de Conducción presentan una reducción de la ratio de extensión del primer formante, una mayor dispersión intrafonémica y una menor distinción interfonémica global. Estas marcas de limitación de los grados de apertura, de imprecisión articulatoria y de hipo-discriminación fonémica respectivamente se reducen en sílaba acentuada aunque la diferencia respecto de lo observado en el grupo control se mantiene.

En cuanto al sistema vocálico de los afásicos de Broca, también se distingue del de los sujetos sanos por una mayor dispersión intrafonémica, pero también por una mayor distinción interfonémica. Cabe señalar que las diferencias respecto de las producciones del grupo control se reducen significativamente en sílaba tónica. Por consiguiente, se aprecia en sus producciones una imprecisión articulatoria en sílaba átona que desaparece en sílaba tónica, pero la diferencia radica en que los afásicos de Broca discriminan más que los normohablantes los fonemas vocálicos entre sí.

Estos resultados son globalmente congruentes con los de estudios previos (Adem, 2006; Haley et al., 2001; Ryalls, 1986) que concluyen a la existencia de un déficit fonético que afectaría el sistema vocálico no solo como era de esperar en los afásicos de Broca, sino también en los de Conducción, lo que es compatible con la idea de que dichos pacientes sufren de un cierto "subtle phonetic deficit" (Vijayan & Gandour, 1995), y que no pueden ser interpretados como el resultado de una mayor velocidad articulatoria (cf. resultados supra). Se aportan asimismo dos datos complementarios. El primero consiste en el efecto del acento sobre la estructura del sistema acentual de los dos grupos de sujetos: en sílaba acentuada, se observa una mejor organización intra- e interfonémica en los afásicos de Conducción (que mantiene sin embargo las diferencias con los sujetos control) y una organización intra- e interfonémica más parecida a la de los sujetos normohablantes en los afásicos de Broca. El segundo está relacionado con el hecho que el carácter desviante del sistema vocálico de los afásicos anteriores radica en su hiperdiscriminación fonémica. Sin embargo será necesario completar o precisar las conclusiones de este estudio preliminar con futuras investigaciones ampliando el número de sujetos y el tipo de tareas.

5 Agradecimientos

Este estudio ha sido financiado por el proyecto Cognipros (FFI2013-40419-P) del Ministerio de Economía y Competitividad. La autora agradece la colaboración de Anna Marczyk y Alba Rosas en la obtención de los corpus.

6 Bibliografía

Adem, H. (2006). Neurolinguistic Analysis:
Aspects of Language and Speech
Deviations in Palestinian Arabic
Aphasics. Tesis doctoral no
publicada. Johannes GutenbergUniversität Mainz.

Ardila, A. (2010). A proposed reinterpretation and reclassification

of aphasic syndromes. Aphasiology, 24(3), 363-394.

Audibert, N., & Fougeron, C. (2012).

Distorsions de l'espace vocalique:
quelles mesures? Application à la
dysarthrie. Actes Des 29èmes
Journées d'Etudes Sur La

- Parole/JEP-TALN-RECITAL 2012, 217-224.
- Baqué, L., Estrada, M., Le Besnerais, M., Marczyk, A., & Nespoulous, J. L. (2006). Corpus COGNIFON-Español para el análisis de producciones consonánticas en la afasia. Documento científico de uso restringido. Proyecto HUM2006-06582. Universitat Autònoma de Barcelona, 1-27.
- Baum, S. R. (1993). An acoustic analysis of rate of speech effects on vowel production in aphasia. Brain and Language, 44(4), 414-430.
- Baum, S. R., Blumstein, S. E., Naeser, M. A., & Palumbo, C. L. (1990). Temporal Dimensions of Consonant and Vowel Production: An Acoustic and CT Scan Analysis of Aphasic Speech. Brain and Language, 39, 33-56.
- Bird, H., & Franklin, S. (1996). Cinderella revisited: a comparison of fluent and non-fluent aphasic speech. Journal of Neurolinguistics, 9(3), 187-206.
- Boersma, P., & Weenink, D. (2014). Praat: doing phonetics by computer. Recuperado el 20 de 4 de 2014, de http://www.praat.org/
- Butterworth, B. (1992). Disorders of phonological encoding. Cognition, 42, 261-286.
- Cappa, S. F., Nespor, M., Ielasi, W., & Miozzo, A. (1997). The representation of stress: evidence from an aphasic patient. Cognition, 65(1), 1-13.
- Ciocca, V., & Whitehill, T. L. (2013). The acoustic measurements of Vowels. En Handbook of Vowels and Vowel Disorders, 113-137.
- Collins, M., Rosenbek, J. C., & Wertz, R. T. (1983). Spectrographic analysis of vowel and word duration in apraxia of speech. Journal of Speech, Language, and Hearing Research, 26, 224-230.
- Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. Psychological Review, 104(4), 801-838.
- Ellis, A. W., Miller, D., & Sin, G. (1983). Wernicke's aphasia and normal language processing: A case study in cognitive neuropsychology. Cognition, 15(1-3), 111-144.
- Gelfer, C. E., Bell-Berti, F., & Boyle, M. (2004). Are vowel errors influenced by consonantal context in the speech of persons with aphasia? Journal of

- the Acoustical Society of America, 115(5), 2424.
- Haley, K. L., Ohde, R. N., & Wertz, R. T. (2001). Vowel quality in aphasia and apraxia of speech: Phonetic transcription and formant analyses. Aphasiology, 15(12), 1107-1123.
- Howard, D., & Smith, K. (2002). The effects of lexical stress in aphasic word production. Aphasiology, 16(1-2), 198-237.
- Huet, K., & Harmegnies, B. (2000).

 Contribution à la quantification du degré d'organisation des systèmes vocaliques. En 23èmes Journées d'Etudes sur la Parole-JEP'2000, 225-228
- Jacks, A., Mathes, K. A., & Marquardt, T. P. (2010). Vowel Acoustics in Adults With Apraxia of Speech. Journal of Speech, Language, and Hearing Research, 53, 61-75.
- Keller, E. (1978). Parameters for vowel substitutions in Broca's aphasia. Brain and Language, 5(3), 265-285.
- Kent, R. D., & Rosenbek, J. C. (1983). Acoustic patterns of apraxia of speech. Journal of Speech and Hearing Research, 26, 231-249.
- Knobel, M., & Caramazza, A. (2007). Evaluating computational models in cognitive neuropsychology: The case from the consonant/vowel distinction. Brain and Language, 100, 95-100.
- Laganaro, M., Vacheresse, F., & Frauenfelder, U. H. (2002). Selective Impairment of Lexical Stress Assignment in an Italian-Speaking Aphasic Patient. Brain and Language, 81(1-3), 601-609.
- Levelt, W. J. (1992). Accessing words in speech production: stages, processes and representations. Cognition, 42, 1-
- Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. The Behavioral and Brain Sciences, 22(1), 1-38; discussion, 38-75.
- MacNeilage, P. F. (1982). Speech production mechanisms in aphasia. Speech motor control, 43-60.
- Monaghan, P., & Shillcock, R. (2007). Levels of description in consonant/vowel processing: Reply to Knobel and Caramazza. Brain and Language, 100(1), 101-108.
- Nickels, L., & Howard, D. (1999). Effects of lexical stress on aphasic word

- production. Clinical Linguistics & Phonetics, 13(4).
- Niemi, J., Koivuselkä-Sallinen, P., & Hänninen, R. (1985). Phoneme errors in Broca's aphasia: Three Finnish cases. Brain and Language, 26(1), 28-48.
- Pate, D. S., Saffran, E. M., & Martin, N. (1987). Specifying the Nature of the Production Impairment in a Conduction Aphasic: A Case Study. Cognitive Processes, 2(1), 43-84.
- Ryalls, J. H. (1981). Motor aphasia: acoustic correlates of phonetic disintegration in vowels. Neuropsychologia, 19(3), 365-374.
- Ryalls, J. H. (1986). An Acoustic Study of Vowel Production in Aphasia. Brain and Language, 29(1), 48-67.
- Sapir, S., Ramig, L. O., Spielman, J. L., & Fox, C. (2010). Formant Centralization Ratio: A proposal for a new acoustic measure of dysarthric speech. Journal of Speech and Hearing Research, 53(1), 114-126.
- Semenza, C., Bencini, G. M. L., Bertella, L., Mori, I., Pignatti, R., Ceriani, F., ...

- Caldognetto, E. M. (2007). A dedicated neural mechanism for vowel selection: a case of relative vowel deficit sparing the number lexicon. Neuropsychologia, 45(2), 425-430.
- Sugathan, N. & Rajasudhakar, R. (2011). Acoustic vowel space in individuals with Broca's aphasia. Frontiers of Research on Speech and Music, 1-5.
- Vijayan, A., & Gandour, J. (1995). On the Notion of a "Subtle Phonetic Deficit" in Fluent/Posterior Aphasia. Brain and Language, 48, 106-119.
- Weismer, G., Jeng, J.-Y., Laures, J. S., Kent, R. D., & Kent, J. F. (2001). Acoustic and intelligibility characteristics of sentence production in neurogenic speech disorders. Folia Phoniatrica et Logopaedica, 53(1), 1-18.
- Whiteside, S. P., Grobler, S., Windsor, F., & Varley, R. (2010). An acoustic study of vowels and coarticulation as a function of utterance type: A case of acquired apraxia of speech. Journal of Neurolinguistics, 23, 145-161.

La lingüística clínica a la vista de la fonética: adquisición y pérdidas fonéticas

TORDERA YLLESCAS, JUAN CARLOS

1 Jakobson y la adquisición de los fonemas en el desarrollo del lenguaje y la pérdida de fonemas. La revolución jakobsoniana

En su obra *Lenguaje infantil y afasia* Jakobson aborda un fascinante estudio sobre la adquisición fonológica en el niño así como la pérdida fonológica en los casos de sujetos con afasia. Respecto a la adquisición fonológica, Jakobson (1969 [1974]: 32-36) defiende la existencia de dos etapas en el desarrollo fonológico del niño, a saber, la etapa prelingüística y la etapa lingüística. La etapa prelingüística sería la etapa de balbuceo en la que el niño experimenta con todos los sonidos posibles, pero, según el maestro estructuralista, no tendría repercusiones fonológicas. Fisiológicamente, se trataría de una mera etapa exploratoria. Es la etapa lingüística (aproximadamente a partir del primer año de vida) la que tiene una relevancia en la adquisición de los fonemas. Solo cuando tiene una verdadera intención de significar/comunicar, aparecen los "sonidos de la lengua" (esto es, los fonemas).

La autonomía inicial de las numerosas sensaciones sonoras dispersas es reemplazada en el niño por una distribución conceptual de los sonidos articulados –paralela en cierto sentido a la de los colores [sic]. En lugar de la abundancia fonética del balbuceo se instala la austeridad

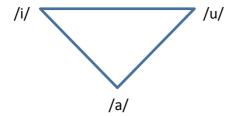
fonemática de los primeros escalones del lenguaje; una especie de deflación transforma los "sonidos salvajes" del balbuceo en valores lingüísticos (Jakobson, 1969 [1974]: 36-37).

Para el desarrollo propiamente fonológico, Jakobson (1969 [1974]) asume que el orden en la adquisición de los fonemas obedece a unos principios universales que son independientes de la lengua materna de partida.

Una vez alcanzada la primera etapa del lenguaje propiamente dicho, en cuanto se inicia la selección de los sonidos y la construcción de un sistema fonemático, se observa un orden de sucesión estrictamente regulado y universalmente válido (Jakobson, 1969 [1974]: 41).

Así por ejemplo, Jakobson (1969 [1974]: 69-72 y 106-108) defiende que, en cualquier lengua, los primeros fonemas en ser aprendidos son la vocal /a/y la consonante /p/, o bien, la consonante /m/. Desde un punto de vista fisiológico, la vocal /a/y es aquella que menos obstrucción ofrece de entre todas las vocales, es decir, es la vocal por excelencia. Las consonantes /p/y/m/, en cambio, ofrecen una oclusión total al paso del aire y, en el caso de la /p/, al ser sorda, no vibran las cuerdas vocales. La /p/y es, por tanto, la antítesis de la vocal /a/y/y o, dicho de otro modo, es la consonante "más consonántica".

Jakobson (1969 [1974]) considera que la adquisición se va realizando a partir de contrastes máximos entre los diferentes fonemas que componen el sistema fonológico de cada lengua. Por ejemplo, Jakobson (1969 [1974]: 69-72 y 106-108), siguiendo la propuesta de Stumpf, asume que la siguiente vocal en ser adquirida tiende a ser la vocal /i/, ya que dicha vocal contrasta por el grado de abertura respecto a la vocal /a/; mientras que la vocal /a/ es una vocal abierta, la vocal /i/ es una vocal cerrada. Adquiridas estas dos vocales, la siguiente vocal que puede ser adquirida es la vocal /u/. Esta vocal contrasta con la vocal /i/ en ser una vocal posterior o vocal velar. De este modo, se adquiere el sistema vocálico triangular.



Según Jakobson (1969 [1974]), todas las vocales de las diferentes lenguas se han de situar dentro de este triángulo articulatorio. Fisiológicamente, es comprensible que se dé este triángulo, puesto que la vocal /a/ es la vocal más abierta, mientras que las vocales /i/ y /u/ son las más cerradas. Y la vocal /i/ es la más anterior, mientras que la /u/ la más posterior. Así pues, toda vocal se ha de ubicar dentro de este espectro articulatorio. En cualquier caso, el mismo Jakobson (1969 [1974]: 107) advierte que existen casos de un sistema vocálico lineal, esto es, se aprende la /a/, luego la /i/ y posteriormente la /e/. De este modo, se prioriza los grados de abertura de las vocales en detrimento del punto de articulación (esto es, anterior/posterior). Solo en una segunda etapa, se adquieren las vocales /u/ y, finalmente, la vocal /o/. Tanto en el sistema vocálico triangular como lineal la vocal /o/ es la última en ser aprendida, según la teoría jakobsoniana.

De igual forma, se opera en el subsistema consonántico. Tras la adquisición de la /p/ (una vocal bilabial y, por tanto, anterior) se adquiere la consonante /t/, una consonante dental (Jakobson, 1969 [1974]: 73) y, posteriormente, la consonante /k/, una consonante velar o posterior. De este modo, se obtiene también un triángulo consonántico que abarcaría un amplio abanico de fonemas de las distintas lenguas.

Jakobson (1969 [1974]) da un paso más allá y ofrece algunas implicaciones tipológicas derivadas de su teoría. Por ejemplo, Jakobson (1969 [1974]: 70) asume que no pueden existir lenguas sin consonantes labiales (p. ej.: /p/o/m/). Aquellas lenguas que aparentemente muestren un inventario de este tipo se han de explicar mediante causas extralingüísticas (p. ej.: perforaciones labiales o malformaciones derivadas de la cultura de la lengua). Asimismo, según Jakobson (1969 [1974]: 77), se pueden establecer relaciones implicativas en la adquisición de las diferentes consonantes. Por ejemplo, la adquisición de las consonantes velares o posteriores, como es el caso de la /k/, de la /g/, o de la /x/, implica la adquisición de las consonantes dentoalveolares o anteriores, como es el caso de la /t/, de la /d/, o de la /s/, respectivamente. Igualmente, según Jakobson (1969 [1974]: 68), es un hecho igualmente universal que la adquisición de las dentales precede a las palatales.

Sin lugar a dudas, una de las relaciones más interesantes y, quizás, de las más cuestionables, es aquella que establece el maestro estructuralista referente a los fonemas tardíos en la adquisición del lenguaje y su frecuencia tipológica. En concreto, Jakobson (1969 [1974]: 83-84) llega a afirmar que aquellos fonemas que más tardíamente se aprenden son aquellos que son más raros tipológicamente hablando. Por ejemplo, en español, los fonemas $/\theta/$, /x/ y /r/ son de los últimos en ser adquiridos. Curiosamente, los fonemas $/\theta/$ y /x/ son fonemas desconocidos para el resto de las lenguas románicas, mientras que el fonema vibrante múltiple /r/ es desconocido en algunas lenguas como el chino. Sin embargo, en todas estas lenguas, los fonemas /p/ y /m/, que son las primeras en ser aprendidas, sí pertenecen a sus respectivos inventarios fonológicas.

Respecto al estudio afasiológico, Jakobson (1969 [1974]: 178) llegó a considerar estos trastornos del lenguaje como una oportunidad única para observar el proceso inverso al de la adquisición del lenguaje; de hecho, llega a calificar a este tipo de trastorno, como una regresión del lenguaje.

Según Jakobson (1969 [1974]: 45-56), en el caso de los afásicos, se asiste a casos en los que el hablante ha perdido el sistema de oposiciones fonológicas propias de una lengua dada ([±sonora], [±anterior], [±obstrucción], [±nasal]...) y que fue consolidando durante su infancia.

[Al igual que en el caso de la adquisición del sistema fonológico] lo mismo ocurre con su pérdida en el afásico; lo que importa no es el empobrecimiento de los sonidos articulables o audibles, sino el de los sonidos con función distintiva. En ello reside el verdadero origen de los trastornos de la producción o la comprensión de sonidos. A veces el afásico es capaz de emitir los sonidos que le faltan, los cuales persisten a menudo en sus gestos vocales, pero no impide que exista una pérdida de valor distintivo (fonemático) en el interior de los "signos arbitrarios del lenguaje". En ciertos casos esa pérdida engendra en el paciente fusiones e intricaciones de los sonidos entre sí, y en otros su pérdida irreversible: el sonido desaparece sin que nada lo reemplace (Jakobson, 1969 [1974]: 47-48).

Y al igual que se puede establecer de antemano un orden en la adquisición de los fonemas, en el caso de las afasias, también existe un orden predecible de los fonemas que se pueden perder en la afasia; en estos casos, el sujeto afásico se verá obligado a crear nuevos sistemas fonológicos.

Los elementos aislados del sistema fonemático desaparecen en la afasia fónica siguiendo un orden determinado. Mientras subsisten partes del sistema, estas forman un sistema bien ordenado, aunque empobrecido. Igual que el niño no se contenta con reducir pura y simplemente el modelo lingüístico, sino que atribuye nuevos valores al sistema simplificado, el afásico no se limita a ofrecer un desmantelamiento de su sistema antes más rico, sino que incluso a veces ofrece una reestructuración (Jakobson, 1969 [1974]: 48-49).

2 Críticas a la propuesta jakobsoniana

Las críticas que ha recibido la teoría de Jakobson han sido diversas. En primer lugar, se ha cuestionado la poca relevancia que Jakobson concedía a la etapa de balbuceo. En cambio, los estudios psicolingüísticos revelan que la etapa de balbuceo sí tiene una relevancia lingüística en la adquisición posterior del sistema fonológico. No existe, por tanto, una revolución entre la etapa de balbuceo y la etapa lingüística, sino que es un cambio gradual. Los carencias en una fase, por tanto, puede tener claras repercusiones en futuras etapas. (vid. Serra, Serrat, Solé, Bel & Aparici, 2000: 193 o Curtin & Werker, 2007: 589). Además, no es cierto que el niño produzca en esta etapa todos los sonidos de las lenguas. Existen sonidos que el niño no llega a pronunciar durante esta etapa. En estudios como los de Laura Bosch, Serra, Serrat, Solé, Bel & Aparici (2000: 193), se llega a afirmar que "no existe un repertorio universal de sonidos compartidos por todas las lenguas, a partir del cual se fijarán los del entorno y se eliminarán los no percibidos".

Respecto a la adquisición propiamente fonológica, el orden de adquisición ha sido, sin lugar a dudas, una de las cuestiones que más atención ha recibido en los estudios psicolingüísticos. Así por ejemplo, según se recoge en Serra, Serrat, Solé, Bel & Aparici (2000: 205-207), para el inglés y otras lenguas, las primeras vocales que se adquieren son /a/, /i/, /u/ y, posteriormente, las vocales /e/, /o/ (aunque admiten que existen estudios disonantes a este orden). Normalmente, se consolida antes el sistema vocálico que el consonántico. Respecto a las consonantes, a partir de diversos estudios psicolingüísticos, concluyen que, en inglés, se adquieren antes las oclusivas bilabiales (primero la sorda /p/) y las nasales /m/ y /n/. Posteriormente, se adquieren las oclusivas velares y las correspondientes sonoras de las oclusivas. También indica que la posición del fonema puede ser determinante. En concreto, en posición final, las sonoras como la /1/ y /d/ son las más tardías en aparecer.

Este orden en la adquisición se ajusta relativamente a la propuesta jakobsiana, pero se observan algunos hechos no esperados. Las consonantes velares se adquieren antes que las dentales, frente al modelo de Jakobson. En el fondo, de cumplirse esta tendencia, sería más razonable incluso que la misma propuesta de Jakobson. Si entre las vocales primero se adquieren las vocales más extremas (la /a/y la /i/) y luego se adquieren las vocales intermedias (como la /e/y la /u/), lo esperable es que con las consonantes ocurriera exactamente lo mismo; primero se deberían adquirir las consonantes más anteriores (como la bilabial /p/) y la más posterior (como la velar /k/) y, posteriormente, las consonantes intermedias (como la dental /t/).

Frente a estos estudios, existen otros que parecen dinamitar la esencia misma de la teoría jakobsiana. Por ejemplo, en el estudio de Gómez Fernández (1993), en el que hace una revisión de dicha teoría para el caso del español, el autor encuentra algunos contraejemplos. Para dicha revisión, parte del estudio hecho previamente por Hernández Pina realizado con un único sujeto (el hijo de Hernández Pina). Como el propio autor afirma, quizás un solo caso no sirva para construir una teoría de la adquisición fonológica, pero sí basta un único caso para deconstruir una teoría que se pretende universalista. En este estudio, el niño muestra el siguiente orden de adquisición vocálica: |a| > |e| > |i| >/o/ > /u/ (Gómez Fernández, 1993: 16 y 20). Es decir, el niño adquiere la vocal /e/ sin haber adquirido previamente la vocal /i/ y la vocal /o/ sin haber adquirido la vocal /u/. Ni el sistema vocálico triangular ni el sistema vocálico lineal de Jakobson podría prever este hecho. Respecto a las consonantes, el niño estudiado adquiere primero el fonema dental /t/, luego la /p/ posteriormente, la /m/ (Gómez Fernández, 1993: 22-23). Es decir, frente a la teoría jakobsoniana, el niño ha adquirido antes el rasgo dental que el bilabial. Poco después, el niño adquiere el fonema /b/, y a los 21 meses, el fonema /d/ y el fonema /g/. Aunque, como afirma Gómez Fernández (1993: 17), el desarrollo del rasgo sonoro es algo impreciso y vago en la teoría jakobsoniana, lo que se aprecia es lo esperable según la teoría de Jakobson: que primero se adquieran las labiales sonoras y luego el resto. Sin embargo, lo cierto es que, en este caso concreto, no parece razonable que el niño pronuncie la /d/ antes que la /b/. Dado que el primer fonema consonántico adquirido había sido el fonema /t/, hubiera sido esperable que el primer fonema sonoro fuera el correlato sordo de este fonema. Por otro lado, el niño adquiere antes el fonema /x/ que la /f/. Y junto a la /x/, adquiere la /k/ y la /s/. En este caso, se vuelve a cumplir que el fonema fricativo labial es adquirido tras la adquisición de las fricativas posteriores sordas. Por tanto, se observa que son más las ocasiones en las que no se cumple la teoría jakobsoniana que las ocasiones en que sí.

En el caso del inglés, los hechos son parejos a los apuntados. Tampoco la teoría jakobsoniana se cumple punto por punto. Por ejemplo, Owens (1996 [2003]: 258) señala que el niño muestra predilección por algunos sonidos ("sonidos favoritos" los llama) y estos son los que suele pronunciar antes. Estos sonidos son las vocales /a/, /e/, /i/ y /o/ y las consonantes /p/, /b/, /t/, /d/, /k/. Las consonantes adquiridas son las predichas por la teoría de Jakobson. En cambio, los sonidos vocálicos adquiridos son /a/, /e/, /i/ y /o/, es decir, se adquiere la $\langle o \rangle$ sin haber adquirido previamente la vocal $\langle u \rangle$. Asimismo, en Menn & Stoel-Gammon (1995: 348), a partir de otros estudios psicolingüísticos, se recoge que las primeras vocales adquiridas para el inglés son (en orden de adquisición): /a/, /i/, /u/, /o/ y $/\Lambda/$. Es decir, se adquiere la vocal /o/ sin haber adquirido la /e/ tan siquiera. Y en Kent y Miolo (1995: 313-314), a partir de los estudios de Kent y Bauer, se apunta que, en los niños de un año, las primeras vocales en ser adquiridas son $/\Lambda/$, /a/ y $/\epsilon/$. En este caso, el triángulo vocálico es mucho más reducido que el previsto por Jakobson. Las siguientes vocales en ser adquiridas varían mucho según diferentes factores. Entre estos factores, está la estructura silábica; por ejemplo, en la estructura V o VV, aparecen vocales tales como, $/ \frac{\omega}{v}$, $/ \frac{e}{v}$, $/ \frac{e}{v}$ o $/ \frac{v}{v}$; en el caso de CV, aparecen vocales como $/ \theta /$, / i / y, posteriormente, / I /.

Por todo ello, Menn & Stoel-Gammon (1995: 347) no dudan en poner en tela de juicio la universalidad de la adquisición. Además, advierten de que no es lo mismo pronunciar un sonido que haber adquirido una oposición fonológica del

lenguaje. La pronunciación de un sonido puede variar según su posición y la longitud misma de la palabra enunciada. Se han de tener en cuenta la variabilidad diatópica. Nosotros añadiríamos, incluso, la variabilidad diastrástica.

Respecto a la afasiología, se ha de señalar que no todas las afasias reproducen los mismos síntomas fonológicos, hecho del que el mismo Jakobson (1969 [1974]: 45-56) se percató. Caplan (1987 [1992]: 122-124) reconoce que la teoría de Jakobson es interesante aplicada al campo de la neurolingüística. Sin embargo, en Caplan (1987 [1992]: 257-260), se indica que los afásicos de Broca (afasia motora) producen solo las oclusivas sonoras en detrimento de las sordas; esto no ocurre con los de Wernicke (afasia sensitiva). Esto es un hecho interesante, dado que si la afasia es un trastorno que implica una involución de la adquisición del lenguaje, como pretende Jakobson, no se explica que sean los fonemas sonoros los que se mantengan, cuando, en la adquisición fonológica infantil, primero aparecen los fonemas oclusivos sordos y, posteriormente, los sonoros.

Asimismo, Caplan (1987 [1992]: 263), apoyándose en el estudio de Nespoulous, recoge que los afásicos de Broca, al tratar de repetir una palabra aislada, solo cambian un rasgo, bien sea el de sonoridad o bien el lugar de articulación, hecho que no ocurre con los de conducción. Si se acepta que la rigidez y la inflexibilidad de la teoría jakosoniana, no se puede aceptar que un fonema (p. ej.: el fonema /k/) cambie por otro al cambiar su punto de articulación (p. ej.: el fonema /t/) o bien su sonoridad (el fonema /g/). Debería existir un orden establecido (p. ej.: primero se pierde el rasgo de sonoridad y después existen confusiones articulatorias).

Por último, nos gustaría señalar que las relaciones implicativas que establece Jakobson en el campo tipológico no siempre se cumplen. Por ejemplo, la teoría jakobsoniana señala que una lengua no puede presentar en su inventario fonemas oclusivos posteriores (p. ej.: /t/ o /k/) sin tener fonemas anteriores o bilabiales (esto es, la /p/). Sin embargo, dichas lenguas sí existen, como es el caso del árabe.

A tenor de los abundantes contraejemplos apuntados, creemos que solo caben dos opciones: o bien se desecha en su totalidad la teoría jakobsoniana o bien se modifica parcialmente. Nuestro camino irá dirigido a rectificar los puntos más conflictivos de dicha teoría.

3 Revisión a la teoría jakobsoniana

Si se parte de un estudio exhaustivo como el de Vivar & León (2009), en el que se tiene en cuenta diferentes variables (como la posición de los fonemas estudiados) en una muestra amplia de sujetos, desde los tres años hasta prácticamente los seis, se puede observar que la teoría jakobsoniana tiende a cumplirse (primero se adquieren las bilabiales, luego las dentales y, por último, las velares; primero se adquieren las sonoras y luego las sordas; las líquidas se adquieren en los últimos estadios...).

Sin embargo, por los estudios presentados en el anterior epígrafe, se colige que a la teoría jakobsoniana se le ha de hacer algunas puntualizaciones. En primer lugar, la adquisición lingüística no obedece a reglas inexorables que no se puedan quebrantar, sino más bien a tendencias. En segundo lugar, es un acierto indudable hacer hincapié en el hecho de que el niño aprende por contrastes máximos, pero no las consecuencias que se han derivado de ello. Por ejemplo, Jakobson parte de que los primeros fonemas adquiridos son /a/ y/p/ por el contraste máximo que se da entre estos fonemas. Esto es una verdad a

medias. Como afirma Gómez Fernández (1993: 26), en el establecimiento del inventario fonológico de un niño, hemos de basarnos en criterios que permitan establecer "la determinación del orden y constitución de los fonemas a partir del orden de adquisición de las oposiciones pertinentes, entendiendo que un fonema está adquirido cuando lo están todos y cada uno de los rasgos pertinentes que lo constituyen esa lengua" (la cursiva es nuestra). Según el autor, en la oposición /a/y/p/, el niño adquiere solo la oposición vocal/consonante. En una posible oposición posterior, /páxa/ y /máxa/, el niño aprende solo la oposición consonante oral y consonante nasal. En la oposición /páxa/ y /káxa/, el niño aprende solo la oposición consonante bilabial oral y consonante no bilabial oral. En la oposición /páxa/ y /báxa/, el niño aprende solo la oposición consonante bilabial sorda oral y consonante bilabial sonora oral. Y, finalmente, en la oposición /páxa/ y /fáxa/, el niño aprende la oposición consonante oclusiva bilabial sorda oral y consonante fricativa labio(dental) sorda oral. Es entonces cuando se puede decir que el niño ha adquirido el fonema /p/.

Siendo esto así, hemos de concluir que, al principio, el niño no adquiere exactamente los fonemas /a/ y /p/. Desde una perspectiva adultocéntrica pero quizás más acertada, se ha de afirmar que lo que el niño ha aprendido son dos archifonemas de un reducido inventario fonológico: un archifonema vocálico frente a un fonema consonántico. Para el adulto, este archifonema infantil bien puede ser el fonema /p/, o el fonema /m/, el fonema /b/ o incluso cualquier consonante. Conforme el niño va aprendiendo archifonemas/fonemas, el niño va ajustando su sistema fonológico hasta asemejarse al del sistema fonológico adulto. En este punto, cesan las oposiciones por contraste máximo. De este modo, se explica por qué un niño puede producir /póka/ en lugar de /bóka/ o de /fóka/. Y, del mismo modo, se explica por qué el niño español presenta diversas confusiones con los fonemas /l/, /r/, /r/, que exigen un alto grado de oposición (basado en un contraste mínimo en la oposición primitiva vocálico/consonántico) y por qué los adultos japoneses y chinos no discriminan estos fonemas. Un hablante de japonés solo ha adquirido el archifonema /R/, mientras que un chino, el archifonema L, cuyas realizaciones fonéticas, en ambos casos, son [1], [\mathfrak{c}] y [\mathfrak{r}].

Una tercera cuestión es qué rasgos distintivos se consolidan antes. Jakobson asume un orden estricto e inexorable a este respecto: la primera oposición es vocal/consonante y, posiblemente, sea este el único universal que se cumpla. En las vocales, se establecen las oposiciones abierta/cerrada y luego anterior/posterior. En las consonantes, las oposiciones establecidas seguirían el siguiente orden: oral/nasal; labiales/dentales; labiales y dentales/palatales y velares (esto es, anteriores/posteriores); oclusivas/fricativas; africadas/oclusivas y fricativas y, finalmente, líquidas/consonantes y vocales.

No existe ninguna razón de peso para justificar que primero se hayan de consolidar las oposiciones del modo de articulación y luego las referentes al punto de articulación o viceversa ni respecto a qué elemento se ha de establecer esta nueva oposición. Jakobson vio parcialmente este hecho con el sistema vocálico al proponer dos tipos de sistemas vocálicos: el triangular y el lineal. Ante la oposición a/i, el niño puede establecer una nueva oposición basada en la anterioridad/posterioridad (sistema vocálico triangular) o bien basada en el grado de abertura (sistema vocálico lineal). Así pues, lo mismo habría de ocurrir en el sistema consonántico. Por ello, no es de extrañar que, tras adquirir el primer fonema consonántico, como puede ser el fonema /p/, el siguiente fonema en ser adquirido sea bien /m/, bien /t/, bien /k/ o bien /b/. Todo depende de qué nuevo rasgo distintivo adquiera el niño dentro de su particular

sistema fonológico. Y, a partir de esa nueva oposición, pueden establecerse otras que, evidentemente, no han de seguir necesariamente el orden impuesto por la teoría jakobsoniana. Por tanto, no existe prioridad en el orden de adquisición de los rasgos fonológicos ni fonemas que induzcan a unas oposiciones prioritarias¹. Este hecho explica la disparidad de datos recogidos en el anterior punto. Por ejemplo, Owens (1996 [2003]: 258) señalaba casos en los que el niño había aprendido las vocales /a, e, i, o/ sin haber aprendido la vocal /u/. Es razonable pensar que el niño adquiriera en primer lugar la oposición /a/, /i/ y, posteriormente, la vocal /e/. En este caso, tiene dos opciones: o bien desarrolla la oposición anterior/posterior respecto a la /i/ o bien respecto a la /e/. Lo que hace el sujeto es esta segunda opción, que no es ni mejor ni peor que la primera².

Diversos autores han señalado que los niños producen sustituciones, metátesis, asimilaciones, omisiones... en sus emisiones (Owens, 1996 [2003]: 259-260, Menn & Stoel-Gammon, 1995: 343; Serra, Serrat, Solé, Bel & Aparici, 2000: 200-202; Curtin & Werker, 2007: 589) e idénticos fenómenos se producen con los pacientes afásicos (Caplan, 1987 [1992]: 243-277). A nuestro juicio, la razón de estos hechos se podría explicar fácilmente a través de las teorías de Zipf y de la Teoría de la Optimalidad (OT, desde ahora) de Prince y Smolensky (1993).

Según se apunta en Van de Walle, J. & W. Willens (2006), Zipf apuntaba por los años treinta del siglo pasado que la comunicación está siempre dirigida por el principio del mínimo esfuerzo, que consta de dos principios dinámicos y contrarios, a saber, la fuerza de la unificación y la fuerza a la diversidad. La fuerza de la unificación nos induce a realizar el mínimo esfuerzo articulatorio, es decir, las oposiciones fonológicas son mínimas. La fuerza a la diversidad nos induce a los oyentes a interpretar los enunciados de forma inambigua y con el mínimo esfuerzo. En la adquisición del lenguaje (y, posiblemente, en la diacronía de cualquier lengua), estos dos principios están siempre en tensión, hasta conseguir un sistema "relativamente" estable, es decir, el adulto. Así pues, cuando el niño pronuncia /láta/ en lugar de /ráta/, será la fuerza a la diversidad la que invite al niño a seguir estableciendo más oposiciones (pues su mensaje continúa siendo ambiguo para su receptor)3.

Según Prince & Smolensky (1993 [2004]), en la generación de cualquier elemento lingüístico a partir de una forma subyacente intervienen tres componentes universales, a saber, GEN, CON y EVAL. GEN toma el input lingüístico observado y devuelve TODAS las formas posibles (p. ej.: GEN(/párke/ = [párke], [páke], ...), CONT evalúa dichas formas de acuerdo a las restricciones que son violadas y la importancia de estas. Finalmente, EVAL elige el candidato óptimo entre todos los candidatos. A partir de estos principios, China (2004) trata de explicar algunos de los casos apuntados. Para esta autora, por ejemplo, existen dos restricciones universales, NO CODA y

Dado que Jakobson es dado a establecer paralelismos entre la adquisición fonológica del lenguaje y la tipología lingüística, mutas mutandis, si existiera tal orden, posiblemente los resultados fonológicos de las lenguas romances desde el latín hubieran sido idénticos en todas ellas. Más aún, posiblemente, todas las lenguas del mundo tendrían el mismo inventario fonológico.

²Respecto al inventario fonológico de Kent y Bauer, $/\Lambda$, /a, y, $/\epsilon$, hemos de reconocer que es contraintuitivo, dado que se establece un triángulo articulatorio mínimo. No obstante, piénsese que estos son los fonemas que identifica el sujeto adulto. Para el infante, $/\Lambda$, puede ser ese archifonema vocálico posterior que recoja los fonemas /a, /a, /a, mientras que el archifonema $/\epsilon$, puede ser ese archifonema vocálico anterior que abarque los fonemas /a, /a, /a, al inglés. Lo interesante sería comprobar qué repercusiones posteriores tiene cuando en los primeros estadios se establece este tipo de oposiciones. ¿Sigue el ritmo habitual o se enlentece/acelera la adquisición fonológica?

³ Seria interesante corroborar empíricamente si lenguas como el inglés o el catalán, que disponen de un inventario fonológico más rico que el castellano, tienden a presentar palabras más cortas (¿monosilábicas?) que aquellas lenguas en las que el inventario es más reducido.

MAX. La restricción NO CODA determina que el mejor candidato será aquel que presenta una sílaba sin coda. La restricción MAX trata de ajustarse al input. Según la autora, para la gramática emergente del niño, la restricción MAX presenta un peso menor a otras como NO CODA. Así pues, el mejor candidato en algunos niños será [páke], tal como se muestra en China (2004: 29).

GEN (/párke/) =	NO CODA	MAX
a. [párke]	*!	
b. 🎾 [páke]		*

China (2004), en su detallado y concienzudo estudio, persigue explicar todos los casos señalados respecto a las modificaciones que se realizan en la etapa infantil (p. ej.: fricativización –[fása] en lugar de [gása]–, oclusivización –[méta] por [mésa]-, etc.). A nuestro entender, este estudio incurre en una pequeña inadecuación teórica al tratar los casos de asimilaciones o cambios en los puntos de articulación de un sonido dado. No solo debe aceptar que las restricciones impuestas así como su importancia puede variar de sujeto a sujeto (algo que la misma teoría de Prince & Smolensky, 1993 [2004] prediría, pues cada niño parte de gramáticas distintas), sino que, incluso, debería aceptar que dichas restricciones y su importantica varían dentro de la gramática del mismo sujeto. Por ejemplo, no es inhabitual que un mismo niño pronuncie /ráta/ como [láta] o como [dáta]. Y no menos importante es señalar que los distintos fenómenos apuntados por China (frontalización, oclusivización...) parecen responder a un único fenómeno general y no a restricciones ad hoc. Por tanto, sería conveniente reducir el conjunto de restricciones propuesto por esta autora (aunque, otros, como el de NO CODA, parecen necesarios para explicar otros fenómenos como el señalado).

El hecho de que el niño que trate de pronunciar /ráta/ difícilmente pronunciará /káta/ o /páta/ se debe a que el fonema /l/ comparte más rasgos opositivos con /r/ o con /d/. El fonema /r/ y /l/ son fonemas líquidos sonoros; el fonema sonoro /d/, articulado como aproximante [δ], se aproxima igualmente a la articulación líquida. De hecho, rehaciendo la escala de sonoridad que aparece en Prince & Smolensky (1993 [2004]: 12), nosotros propondríamos la siguiente⁴:

Escala de sonoridad

| Vocal baja | > | Vocal alta | > | Líquida | > | Nasal | > | *Aproximante* | > | Fricativa sonora | > | Fricativa sorda | > | Oclusiva sorda |

En los últimos años (Keller, 2002 o Jäger, 2007), dentro de la OT, se ha asumido que una restricción no se ha de cumplir o incumplir totalmente, sino que se puede incumplir en cierto grado. Por tanto, la violación de las restricciones no es un todo o nada, sino que se viola solo en un tanto por ciento. Siendo esto así y basado en la misma escala de sonoridad, se puede decir que si el niño parte del input /ráta/, el output [ráta] respeta el principio MAX un 100 %; el outpout [láta] posiblemente respete algo menos este principio (¿un 90 %?) así como el outpout [dáta] (¿un 80%?), pero lo respeta mucho más que un

⁴ En la diacronía de la lengua, se puede observar cómo las aproximantes en posición de coda tienen a vocalizar: CAPITALEM: [kaβδál] >> [kawδál]; RAPIDUM: [raβδo] >> [ráwdo]; REGULAM [réγ.la] >> [réjla]. Incluso, puntualmente, la líquida puede pasar a una aproximante: LAXARE >> /dexár/, pudiendo pronunciarse [δexár], si no se pronuncia tras pausa o nasal, es decir, en amplios contextos de pronunciación.

posible outpout [báta] (¿un 60 %?) o un outpout como [páta] (¿un 40 %?). Así pues, el niño, ante la imposibilidad de articular correctamente un determinado fonema, tratará de pronunciar el que más se le acerque. Consecuentemente, un deseo de la investigación futura es determinar los algoritmos adecuados que determinen con precisión en qué grado un fonema se acerca a otro, dado un sistema fonológico dado. Pues, evidentemente, si el sistema fonológico cambia, los porcentajes también lo harán.

GEN (/ráta/) =	MAX
a. [ráta]	100 %
b. [láta]	*:80?
c. [dáta]	*:80?
d. [páta]	**20 %

Los casos de afasia se podrían tratar de forma igualitaria. No obstante, cabría señalar que Jakobson partió de un error de partida. Jakobson vio en las afasias el proceso inverso al de la adquisición del lenguaje y esto no es exactamente correcto. En las afasias no degenerativas (como fruto de un accidente cerebrovascular, como un ictus, o de una lesión), los sujetos no tienden a perder más fonemas. De hecho, con el tiempo, ocurre exactamente lo contrario. El sujeto tiende a recuperarlos, aunque solo sea por recuperación espontánea. Así pues, si se quiere cumplir con el objetivo trazado por Jakobson, nuestra atención debería recaer en los síndromes afásicos fruto de una enfermedad neurodegenerativa, como son el alzhéimer o el párkinson. Solo en este tipo de afasias, se puede observar si se cumple esa regresión o "involución" fonológica. Creemos que, con los principios esbozados, dicho objetivo puede encararse.

En conclusión, hemos de afirmar que, pese a los años, la obra de Jakobson sigue siendo una obra imprescindible para abordar la adquisición y pérdida fonológicas, y no pocas de sus ideas mantienen actualmente plena vigencia. No obstante, algunos puntos merecen ser matizados y otros pueden ser enriquecidos a la luz de las nuevas aportaciones de la lingüística contemporánea. Creemos que puede resultar un camino muy fructífero si se aplican los principios teóricos de la OT a este campo de investigación y que existe todavía todo un mundo por delante por reexplorar.

4 Bibliografía

Buckingham, Hugh W., & Sarah S. Christman (2008). Disorders of Phonetics and Phonology. En B. Stemmer & H. A. Whitaker, Handbook of the neuroscience of language. Amsterdam y otros: Elsevier, 127-136.

Caplan, David (1987 [1992]). Introducción a la neurolingüística y al estudio de los trastornos del lenguaje. Madrid: Visor.

China, Nancy (2004): La sonoridad y la marcación en los errores comunes del habla infantil. *Revista Argentina de Neuropsicología*, 4, 23-37.

Curtin, S., & Werker, J. F. (2007). The perceptual foundations of phonological development. En M.

Garret Gaskell, *The Oxford handbook of psycholinguistics*. Oxford-Nueva York: Oxford University press.

Fernald, A. & Marchman, V. A. (2006). Language learning in infancy. En M. J. Traxler & M. A. Gernsbacher, *Handbook of psycholinguistics*. Amsterdam y otros: Elsevier, 1027-1071.

Gómez Fernández, D. (1993). La teoría universalista de Jakobson y el orden de adquisición de los fonemas en la lengua española. *Cauce*, 16, 7-30.

Jakobson, R. (1969 [1974]): Lenguaje infantil y afasia. Madrid: Editorial Ayuso.

Jäger, G. (2007). Maximum Entropy models and Stochastic Optimality Theory. En A. Zaenen, J. Simpson, T. H. King, J.

- Grimshaw, J. Maling & C. Manning (eds.), *Architectures, rules and preferences: variations on themes by Joan W. Bresnan.* Stanford: CSLI Publications, 467-479.
- Keller, F. (2002). Probabilistic Algorith learning and Optimality Theory. En *Linguistic Inquiry* 33:2, 225-244.
- Kent, Ray D., & Miolo G. (1995). Phonetic abilities in the first year of life. En P. Fletcher & B. MacWhinney, *The handbook of child language*. Malden-Oxford-Carlton: Blackwell publishing, 303-334.
- Menn, L., & Stoel-Gammon C. (1995).

 Phonological Development. En P.
 Fletcher & B. MacWhinney, *The handbook of child language*. Malden-Oxford-Carlton: Blackwell publishing, 335-359.

- Owens, R. E. (1996 [2003]). *Desarrollo del lenguaje*. Madrid: Pearson Educación/Prentince Hall.
- Prince & Smolensky (1993). *Optimality Theory*. Malden-Oxford: Blackwell.
- Saussure, Ferdinand De (1916 [1980]). *Curso de lingüística general*. Traducción al español de A. Alonso. Buenos Aires: Losada.
- Serra, M., Serrat, E., Solé, R., Bel A., & Aparici, M. (2000): *La adquisición del lenguaje*. Barcelona: Ariel Psicología.
- Van de Walle, J., & Willens, W. (2006). Zipf, George Kingsley (1902-1950). En K. Brown, *Encyclopedia of language and linguistics*. Amsterdam y otros: Elsevier. Volumen 14, 756-757.
- Vivar, P. & León H. (2009). Desarrollo fonológico-fonético en un grupo de niños entre 3 y 5,11 años. En de la revista CEFAC, Abril-Junio, 11(2), 190-198.

Parte IV. Fonética forense

Análisis semiautomático y automático de vocales con fines forenses en habla de micro y de móvil

DORTA, JOSEFA; JORGE TRUJILLO, CAROLINA; MARTÍN GÓMEZ, JOSÉ ANTONIO

1 Introducción

El objetivo del presente estudio es doble. En primer lugar, realizamos un análisis semiautomático de la estructura formántica de las vocales en una situación controlada que simula la que podría darse en un caso forense. Seguidamente, dado que se ha cuestionado la utilidad de la información proporcionada por los formantes en el análisis forense de voces, nos hemos planteado como segundo objetivo realizar un análisis biométrico automático de la fonación con el programa BioMet®foreProf².

En relación con el primer objetivo, se ha comprobado desde el punto de vista acústico que el F1 y el F2 son decisivos para el reconocimiento y percepción de cada vocal, mientras que los formantes superiores, normalmente por encima de los 3000 Hz, se relacionan más con características individuales puesto que varían en cada sujeto dependiendo de su configuración fisiológica laríngeobucal (Quilis, 1993; Martínez Celdrán, 1995; Martínez Celdrán & Fernández Planas, 2007; Albalá *et al.*, 2008). De ahí que en los estudios de tipo forense se dé gran importancia al análisis de la estructura formántica y a los campos de dispersión vocálicos para aportar información fonética relevante sobre el peso de la prueba que se presenta ante el tribunal en un caso determinado.

¹*En este trabajo se ha utilizado corpus del proyecto Estudio comparativo de la entonación y del acento en zonas fronterizas del 52716-P). Proyecto de I+D del programa estatal de fomento de la investigación científica y técnica de excelencia, subprograma estatal de generación del conocimiento (convocatoria 2014).

² Este programa está distribuido por BioMetro Soft S.L. y ha sido desarrollado por el grupo de informática aplicada al procesado de señal e imagen (GIAPSI) de la Universidad Politécnica de Madrid. Utilizamos la versión 4.0 (9.0), julio 2014.

Ahora bien, en los análisis de voz con fines forenses encontramos registros de voz grabados en circunstancias diversas, algunas de las cuales implican efectos importantes respecto de las grabaciones realizadas en las situaciones normales que suele manejar el experto. Es el caso del habla grabada a través de la telefonía móvil, circunstancia muy frecuente en las pruebas periciales de voz donde, además de los efectos derivados del ruido (mayor que en telefonía fija), el filtrado de frecuencias que se produce provoca una pérdida de la información por encima de los 3200 Hz, aproximadamente, lo que repercute en el análisis de ciertos segmentos, como las vocales objeto de interés en el presente trabajo.

En efecto, cuando el habla grabada se ha transmitido a través de telefonía fija y más aún de móvil se cortan las frecuencias altas por lo que no podemos contar, por una parte, con la información de formantes superiores a las frecuencias de filtrado y, por otra, se ha visto cierta alteración en los valores de los dos primeros formantes respecto del habla grabada a través de micrófono. Ello ha determinado que en estos casos se haya cuestionado la utilidad de la información proporcionada por la estructura formántica en el análisis forense de voces. Así, en uno de los trabajos pioneros sobre el estudio de formantes en registros de telefonía fija, Künzel advierte desde el mismo título ("Beware of the 'Telephone effect': The influence of telephone transmission on the measurement of formant frequencies", 2001) sobre los efectos de la telefonía digital debido a que por debajo y por encima de los 400 y 3400 Hz, respectivamente, disminuye progresivamente la amplitud de la señal. Debido a ello, en su experimento sobre la estructura formántica de las vocales del alemán, este autor utiliza un filtro que sitúa el rango de frecuencias entre 400 y 3400 Hz, lo que determina que en las vocales con formantes bajos, inferiores a 400 Hz, se incremente la intensidad de los armónicos de mayor frecuencia y, como consecuencia, el aumento de que habla Künzel en el F1 esté, según cálculo de Nolan (2002), en torno a un 6 %, en el F1 de las vocales medias y, especialmente, de las cerradas, tanto de voz masculina (5,13 %) como femenina (5,75 %), debido al desplazamiento del centro de gravedad del formante (mayor concentración de energía). Por tal razón, Künzel cuestiona la capacidad de este primer formante para identificar voces en casos forenses. Igualmente, en las vocales que presentan formantes en frecuencias superiores a 3400 Hz, los armónicos que se encuentran en frecuencias inferiores se intensifican, lo que se traduce en una disminución del valor del formante. En cambio, no se encuentran diferencias significativas en la variación del F2, que se mantiene mucho más estable. Nolan (2002), sin embargo, considera que los datos de Künzel no constituyen una razón suficiente para excluir el análisis formántico en casos de habla telefónica y aduce las razones por las que el F1 y, sobre todo, el F2, pueden ser de gran utilidad en este campo. Livijn (2004), en cambio, estima que las frecuencias de los tres primeros formantes de las vocales se modifican en el habla grabada en teléfonos móviles, por lo que ratifica que el análisis de la estructura formántica no es válido en casos forenses. Byrne y Foulkes (2004), por su parte, encuentran diferencias entre los valores de los formantes de las vocales tónicas, fundamentalmente las cerradas, dependiendo de si han sido grabadas en móvil o a través del micrófono. Observaron que, en general, el efecto de la telefonía móvil se hace notar fundamentalmente en el F1, puesto que las frecuencias de este formante aumentan respecto de las grabaciones en directo en un porcentaje más elevado que el que muestran los datos de Künzel (2001), esto es, un 29%. En cambio, en la misma línea que este último autor, concluyen que las medias del F2 no se alteraron de manera significativa.

En un estudio para el español, Jiménez (2011) analiza la estructura formántica (F1 y F2) de las vocales tónicas y su campo de dispersión en dos enunciados leídos en frase marco y grabados con teléfono móvil a cuarenta hombres cuya variedad es el español centropeninsular, con edades que oscilan entre los 35 y 55 años y con un nivel sociocultural medio-alto. Al comparar sus datos con los obtenidos por Ramón et al. (1979), Quilis & Esgueva (1983) y Martínez Celdrán (1995), concluye que los valores medios de estos autores son más extremos que los suyos, por lo se podría deducir que la telefonía móvil afecta a la estructura formántica haciendo que los dos primeros formantes tiendan a la centralización en el espacio vocálico. No obstante, al comparar sus valores con los de Albalá et al., observa que se da una gran coincidencia, por lo que concluye que las diferencias encontradas en relación con los tres primeros autores se deben más a efectos del tipo de corpus empleado y a las diferencias en la recogida de las muestras que al efecto del canal de transmisión; por su parte, atribuye las similitudes con Albalá et al. a la utilización de la misma herramienta de extracción de los formantes (un script de Praat).

Para el análisis biométrico automático de la fonación, como dijimos, utilizamos el programa BioMet®foreProf. Se trata de un paquete de software que puede registrar voz o habla o seleccionar cualquier segmento de voz previamente grabada para su parametrización con el objetivo de evaluar el perfil biométrico del locutor¹. El software considera un total de 68 parámetros que son evaluados de forma sincrónica con el ciclo de fonación: incluye el perfil de la fuente glótica, parámetros de distorsión o inestabilidad de la fonación que ayudan a caracterizar la disfonía (jitter, shimmer, relación armónico-ruido, etc.), parámetros cepstrales evaluados sobre la fuente glótica, singularidades espectrales de esta, estimados biomecánicos del pliegue vocal (masas y tensiones de los pliegues), parámetros temporales asociados al cierre y apertura de los pliegues, a defecto del cierre de estos y al temblor en la voz. El programa permite contrastar o cotejar los datos de voces dubitadas e indubitadas con fines forenses de manera que los patrones generados las aproximan o no y, a su vez, las relacionan o apartan de otras que se consideran normativas, entendiendo por estas aquellas que corresponden a locutores con fonación regular. En definitiva, este software permite grabar y analizar la voz y cotejarla a tres bandas (modelo regular o "normativo", voz indubitada y voz dubitada).

2 Metodología

Para el análisis formántico, a diferencia de Jiménez (2011) que compara los valores de su estudio con los de otros autores a pesar de que cambian los informantes, las condiciones de recogida de las muestras, etc., partimos de tres tipos de corpus: experimental no leído², espontáneo (en el que dos sujetos hablan libremente sobre temas no preestablecidos) y lectura de un texto emitidos por una misma informante y grabados a través de dos canales diferentes: micrófono (Shure SM48) y móvil. En este último caso, la señal procedía de un LG P970 Optimus Black³ y fue recogida en un teléfono GSM Nexus 4 de LG con el programa Call Recorder de Clever Mobile⁴ que comprime la señal en formato MP4; posteriormente, esta señal fue convertida a .wav para su análisis acústico. Los textos en ambas situaciones fueron emitidos por una mujer de 27 años, con nivel sociocultural alto y natural de Tenerife (Islas

¹ Este tipo de análisis está concebido para el reconocimiento automático de humanos basándose en rasgos conductales o físicos.

² En este caso, se trata de frases del tipo La cítara se toca con paciencia o El saxofón se toca con pánico, procedentes del corpus fijo que manejamos en AMPER (Atlas Multimedia de Prosodia del Espacio Románico). Véase la configuración del corpus en nuestra web (http://ampercan.webs.ull.es/).

³ Obviamente, en un caso forense este dato no suele conocerse.

⁴ Hemos elegido este método a pesar de que la voz se graba con ruido porque emulamos lo que se haría normalmente al hacer grabaciones de llamadas de móvil que serían entregadas como prueba en un caso forense.

Canarias). El propósito es comprobar si existe modificación entre los F1, F2 y F3 de las vocales grabadas en ambos medios con el objetivo de determinar un porcentaje de variación en los valores que sea imputable a la diferencia del canal de grabación y que pueda ser tenido en cuenta en casos forenses reales; comprobamos, además, el peso estadístico de las variables con el objetivo de corroborar si existen diferencias significativas que permitan discriminar entre las dos voces analizadas.

Para el análisis de los formantes tuvimos en cuenta las vocales procedentes de los tres tipos de corpus mencionados sin separar tónicas y átonas debido a que simulamos, como hemos dicho, una situación real en la que podemos suponer que el corpus grabado a través de móvil no sería demasiado extenso y, en consecuencia, el de vocales tampoco lo sería. Teniendo en cuenta lo dicho, el corpus final para el análisis formántico es de 481 vocales, cuya distribución se muestra en la tabla 1.

	Micro	Móvil
/a/	90	80
/e/	60	79
/i/	31	31
/0/	50	75
/u/	5	8
Total	208	273

Tabla 1. Corpus de vocales para el análisis de formantes

El análisis acústico semiautomático se realizó con el programa Praat⁵ utilizando para la extracción de los formantes el script de Mietta Lennes⁶ que halla los valores a partir de un fichero sonoro previamente etiquetado en un text grid. Los valores formánticos se extrajeron a partir de la medición de cada vocal en su punto medio más estable. Una vez obtenidos los valores, realizamos una carta de formantes, aunque teniendo en cuenta solo el F1 y F2, con un nuevo script de Kristine Yu 7. Finalmente, hemos realizado un análisis discriminante en SPSS para comparar estadísticamente los datos obtenidos en habla de móvil y de micrófono8.

Para el segundo objetivo, esto es, el análisis automático con el programa BioMet®foreProf, hemos elegido realizaciones átonas y tónicas de la vocal /a/ con una duración superior a 50 ms teniendo en cuenta que el programa está diseñado para funcionar mejor con vocales abiertas y medias. El corpus indubitado (de control) fue de 21 vocales y el dubitado de 23. El análisis biométrico de cada uno de los corpus, dubitado e indubitado, como se ha dicho, se basa en 68 parámetros que se ordenan estadísticamente de mayor a menor relevancia. Posteriormente se toman los más relevantes que permiten representar, por una parte, la relación entre la voz indubitada o de control y la voz dubitada y, por otra, la relación entre estas y los modelos ofreciendo el

⁵ Praat para Windows (5.3.63 de 24 de enero de 2014).

⁶ Collect_formant_data_from_files.praat (Copyright 4.7.2003).

⁷ Draw_formant_plot_from_table.praat (11.9.2010).

 $^{^8}$ En el análisis estadístico hemos desechado la $/\mathrm{u}/$ puesto que, como puede verse en la tabla 1, contábamos solo con 5 y 8 vocales grabadas en voz de micro y de móvil, respectivamente. Por tanto, el bajo número de datos formánticos no es suficiente para un análisis discriminante. Por otra parte, como puede verse en la misma tabla, de la vocal /i/ solo contamos con 31 casos en micro y en móvil, razón por la cual el modelo discriminante se generó a partir de ese número en todas las vocales.

resultado final de acuerdo con la escala de referencia ENFSI (European Network of Forensic Science Institutes):

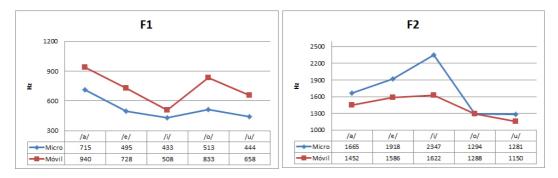
```
ILV<2,30: Prevalece Hip. Defensa
2,30<ILV<4,60: Prevalece DÉBILMENTE Hip. Fiscal
4,60<ILV<6,90: Prevalece MODERADAMENTE Hip. Fiscal
6,90<ILV<9,20: Prevalece FUERTEMENTE Hip. Fiscal
ILV>9,20: Prevalece MUY FUERTEMENTE Hip. Fiscal
```

Tabla 2. Escala de referencia ENFSI

3 Resultados

3.1 Análisis formántico semiautomático y estadística

Los resultados evidencian que la telefonía móvil afecta a la estructura formántica de las vocales. En efecto, en habla de móvil el F1 aumenta su frecuencia respecto de la de micrófono (figura 1), mientras que la de los formantes superiores F2 y F3 (figuras 2 y 3, respectivamente) desciende. Por tanto, los valores de los tres primeros formantes tienden a aparecer más centrados en habla de móvil que en la de micrófono, lo cual debe atribuirse en nuestro caso, exclusivamente, a la diferencia de canal de transmisión pues las demás circunstancias son idénticas en ambos casos.



Figuras 1 y 2. Gráficos a partir de los valores medios, en Hz, del F1 y F2

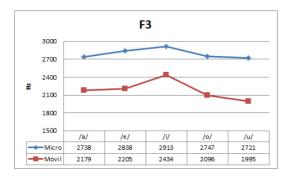


Figura 3. Gráfico a partir de los valores medios, en Hz, del F3

La diferencia entre telefonía móvil y micrófono en cada una de las vocales indica, como se puede observar en la tabla 3, que los porcentajes de diferencia son mayores en el F1 y F3, por este orden, que en el F2, sobre todo en /e/, /o/ y /u/; en el caso de /i/ no sucede lo mismo puesto que el F2 registra el mayor porcentaje de diferencia en esta vocal, lo que evidencia que se ve particularmente afectado en habla de móvil por el filtrado de las frecuencias altas que determina que su valor disminuya más que en otras vocales respecto del mismo formante en habla de micro.

	Porcentajes de diferencia										
	F	1	%	F2		F2		%	F	3	
	Micro	Móvil		Micro	Móvil		Micro	Móvil	%		
/a/	715	940	23,94	1665	1452	12,79	2738	2179	20,42		
/e/	495	728	32,01	1918	1586	17,31	2838	2205	22,30		
/i/	433	508	14,76	2347	1622	30,89	2913	2434	16,44		
/o/	513	833	38,42	1294	1288	0,46	2747	2096	23,70		
/u/	444	658	32,52	1281	1150	10,23	2721	1995	26,68		

Tabla 3. Diferencias, en porcentajes, entre los valores formánticos en habla de móvil y de micrófono

En la figura 4 se ilustran los triángulos acústicos obtenidos a partir de las medias de los dos primeros formantes de cada vocal (tabla 3); las figuras 5 y 6 muestran los campos de dispersión vocálicos obtenidos a partir de los valores de F1 y F2 individuales de cada vocal.

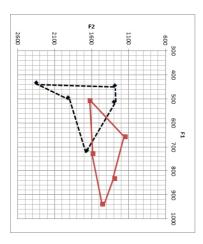


Figura 4. Triángulos acústicos en micro (línea discontinua) y móvil (línea continua). Valores en Hz

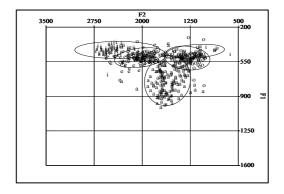


Figura 5. Campos de dispersión en habla de micrófono. Valores en Hz

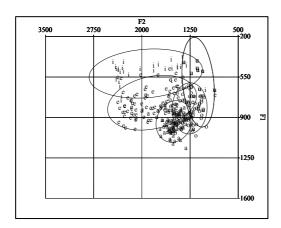


Figura 6. Campos de dispersión en habla de móvil. Valores en Hz

En las figuras 4, 5 y 6 se aprecia que los resultados del habla de micro y de móvil no aparecen muy próximos entre sí, con lo cual se podría interpretar que las dos voces pudieran haber sido emitidas por locutores diferentes. No obstante, sabiendo que no es así en este trabajo, es obvio que en los estudios forenses habrá de tenerse muy en cuenta el porcentaje de diferencia que se da entre la estructura formántica de habla de móvil y la de micro si se considera el estudio de los formantes en la identificación del locutor.

A partir de los valores individuales (no promediados) de los formantes vocálicos obtenidos en el análisis semiautomático, hemos realizado un análisis discriminante con el programa SPSS9 teniendo en cuenta los tres primeros formantes. La finalidad de este análisis es comprobar si existen diferencias significativas que permitan discriminar correctamente, con cierto porcentaje de error, entre los grupos (las dos voces analizadas) en función de las variables consideradas.

El modelo halló una función discriminante significativa [χ 2(12) = 167,72, p <0,05] (tabla 4) que separa los grupos (habla de móvil y habla de micrófono) de manera significativa. El valor de la correlación canónica (tabla 5) elevada al cuadrado (94 %) nos informa de una correlación muy alta entre las variables utilizadas para discriminar ambos medios y la función aislada. Dicho de otra manera, estas variables nos permiten separar los dos medios de grabación.

Contraste de	Lambda			
las funciones	de Wilks	Chi-cuadrado	gl	Sig.
1	,045	167,172	12	,000

Tabla 4. Lambda de Wilks

				Correlación
Función	Autovalor	% de varianza	% acumulado	canónica
1	21,104(a)	100,0	100,0	,977

Tabla 5. Correlación canónica

 $^{^{\}rm 9}$ Como dijimos en la nota 8, en el análisis estadístico hemos desechado la vocal /u/.

La separación entre medios que se consigue con esta función se ilustra en la figura 7, donde puede observarse que micro está en valores negativos de la función discriminante que ha aislado el programa.

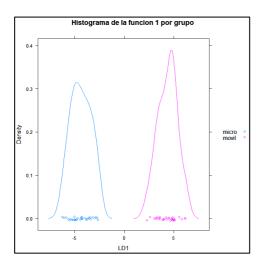


Figura 7. Separación entre micro y móvil según la función discriminante

Los coeficientes de la matriz de estructura (tabla 6) muestran los pesos (importancia) de las variables a la hora de separar a los grupos. Las variables discriminantes más efectivas para separar los grupos son oF1, eF3, aF3, oF3, iF3.

Formantes	Función
	1
oF1	-,309
eF3	,297
aF3	,294
oF3	,282
iF3	,272
aF1	-,214
iF2	,212
eF1	-,172
aF2	,127
eF2	,108
iF1	-,088
oF2	,010

Tabla 6. Matriz de estructura. Variables ordenadas por el tamaño de la correlación con la función

Del análisis estadístico se concluye, pues, que las variables se comportan de forma muy distinta en micro y en móvil y que, salvo en el caso de oF1, el F3 es el formante con mayor peso discriminante para separar los dos tipos de grabación analizados (micro y móvil).

3.2 Análisis biométrico automático de la fonación

Como se dijo anteriormente, hemos elegido realizaciones de la vocal /a/ para llevar a cabo el análisis biométrico de la fonación con el programa BioMet®foreProf. Dado que este programa se centra en la onda glótica, elimina la variabilidad interlocutor que se deriva de la configuración de los resonadores y que, por tanto, afecta a la estructura formántica. Como se dijo, BioMet®foreProf considera un total de 68 parámetros que permiten aproximar o alejar estadísticamente las voces contrastadas, entre sí y respecto de los modelos considerados normativos. El programa determina qué parámetros son los que muestran mayor capacidad de distinción entre voces, de modo que proyecta en el cotejo los más significativos.

De los 68 parámetros, nuestro análisis se ha basado en 14, de los cuales se han señalado como estadísticamente más relevantes 4 de perfil espectral. Este tipo de parámetros forma parte de la firma biométrica del locutor y de su comportamiento (normativo o no) y son muy importantes en la identificación y verificación del locutor y en la determinación de disfonías. Asimismo, se mostró altamente relevante uno de los parámetros cepstrales considerados (que también forman parte de la firma biométrica del locutor) y uno de los parámetros biomecánicos que se relacionan con el funcionamiento mecánico de la glotis, esto es, el que hace referencia a la masa dinámica de la cubierta del pliegue vocal.

De entre la gran cantidad de datos y gráficos derivada de este análisis nos interesa señalar, por su capacidad ilustrativa, la figura 8, donde se aprecia claramente la distribución de las medias de los tres parámetros más relevantes correspondientes a las voces dubitada (cuadrados en azul) e indubitada (rombos en rojo), así como al grupo de modelos (círculos en verde): se trata de parámetros de perfil espectral (23, 24 y 25). Si bien algunos puntos se entremezclan, queda clara la distribución de cada grupo analizado y la mayor proximidad estadística entre las voces dubitada e indubitada frente a los modelos normativos.

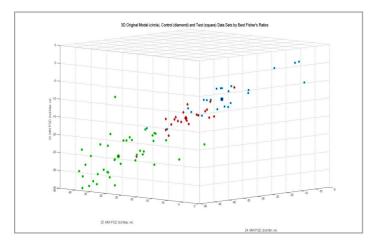


Figura 8. Distribución de las medias de los parámetros 1, 2 y 3 más relevantes

En la figura puede verse que cada grupo analizado presenta un punto de mayor tamaño. Se trata del centro de gravedad, o centroide, de cada uno de ellos. La distancia respecto de los centroides nos indica la variabilidad intralocutor y permite proyectar el llamado "Modelo del balancín" (figura 9).

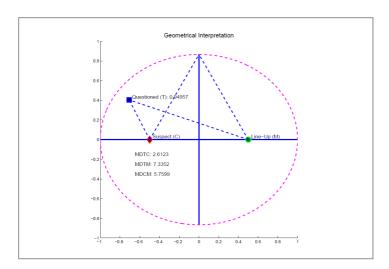


Figura 9. Diagrama de razonamiento balanceado

Esta figura 9 refleja la probabilidad estadística de semejanza entre la voz dubitada y la indubitada, la probabilidad de que la muestra dubitada se parezca al modelo de referencia y la probabilidad de que la indubitada se parezca a dicho modelo. Los resultados del cotejo presentan una amplia probabilidad de que las voces dubitada e indubitada sean la misma, puesto que se acercan entre sí y se alejan del grupo de control ya que, según la escala de referencia ENFSI (tabla 2), cuando el valor es mayor de 9,20 prevalece "muy fuertemente" la hipótesis del fiscal, esto es, se trata de la misma voz, y en nuestro caso el valor resultante es muy superior (23,49). Por ello, el informe final ofrecido por el programa indica que "El grado de convicción derivado de la evidencia acústica procesada SOPORTA MŬY FUERTEMENTE la hipótesis del fiscal frente a la hipótesis de la defensa".

Conclusiones

4.1 Análisis formántico semiautomático

- 1) Nuestros resultados ponen de manifiesto que la telefonía móvil afecta a la estructura de las vocales al provocar una centralización de los tres primeros formantes. Coincidiendo con los resultados que apuntaban Byrne & Foulkes (2004) comprobamos que en habla de móvil, debido al corte de frecuencias, sube el F1 en gran medida y baja el F2 en menor grado, salvo en /i/ en que desciende de forma muy destacada. Se comprueba, además, que el descenso del F3, exceptuando de nuevo el caso de /i/, es más acusado que el del F2 por lo que se puede concluir, como parece apuntar Nolan (2002), que este último formante es más válido en la identificación del locutor en tanto que en los dos tipos de habla experimenta menos cambios (con la excepción ya hecha de /i/).
- 2) Si se considera el análisis formántico en estudios forenses, es fundamental tener en cuenta el porcentaje de diferencia que se da entre los formantes de habla de móvil y la de micrófono (tabla 3). Para ello, en estudios posteriores habría que comprobar si se ratifican los porcentajes obtenidos en el presente trabajo.

4.2 Análisis estadístico de los formantes

- 1) El análisis discriminante de los tres primeros formantes reveló que el F3 es un formante de gran peso en todas las vocales a la hora de discriminar los dos medios (micro y móvil).
- 2) Como se vio en el análisis acústico, el F2 experimenta poca variación al comparar habla de micro y de móvil; de ahí que estadísticamente tenga menor peso que el F3 para discriminar los dos medios considerados en nuestro estudio.

Como conclusión general del análisis acústico semiautomático y del análisis estadístico podemos decir que cuando una de las muestras de habla procede de la telefonía móvil, el análisis formántico no se revela por sí mismo muy productivo en la discriminación de voces con fines forenses, ya que no podemos decidir con certeza si estamos ante la voz de un mismo locutor o ante dos voces diferentes, excepto si se determina un porcentaje fiable de variación entre ambos medios que permita subsanar los efectos de la telefonía móvil.

4.3 Análisis biométrico automático

- 1) Los resultados del cotejo de las voces dubitada e indubitada con el programa BioMet®foreProf arrojan una gran probabilidad de semejanza estadística entre ambas, por una parte, y de distancia, por otra, frente al grupo de control.
- 2) El programa ha demostrado, pues, un altísimo índice de seguridad al clasificar como muy similares dos voces que, dadas las características de nuestro estudio, sabemos que corresponden a una misma persona.

5 Bibliografía

- Albalá, M. J., Battaner, E., Carranza, M., Gil, J., Llisterri, J., Machuca, M. J., Madrigal, N., Marquina, M., Marrero, V., de la Mota, C., Riera, M., & Ríos, A. (2008). VILE: Nuevos datos acústicos sobre vocales del español. Language Design. Journal of Theoretical and Experimental Linguistics. Special Issue 1: New Trends in Experimental Phonetics: Selected Papers From the IV International Conference on Experimental Phonetics (Granada, 11-14 February 2008), 1, 1-14.
- Gómez Vilda, P., & Nieto Lluis, V. (2014). Documento técnico del programa BioMet®foreProf. Versión 1.0.
- Grupo de investigación Profondis. En http://ampercan.webs.ull.es/
- Jiménez Gómez, J. J. (2011). Estructura formántica y campo de dispersión de las vocales del español en telefonía móvil. Estudios Fónicos / Cuadernos de Trabajo, 1, 39-58. Recuperado el 3 de 10 de 2014, de <a href="http://www.estudiosfonicos.cchs.csic.es/fonetica/cuadernos/index.php/estfon/article/view/7" \t "_blank"

- Martínez Celdrán, E. (1995). En torno a las vocales del español: análisis y reconocimiento. *Estudios de Fonética Experimental*, VII, 195-218.
- Martínez Celdrán, E., & Fernández Planas, A. M.ª (2007). *Manual de fonética española*. Barcelona: Ariel.
- Künzel, H. J. (2001). Beware of the 'Telephone effect': The influence of telephone transmission on the measurement of formant frequencies. *Forensic Linguistics*, 8(1), 80-99.
- Nolan, F. (2002). The 'telephone effect' on formants: a response. *Journal: Forensic Linguistics*, vol. 9, n.° 1, 74-82.
- Livijn, P. (2004). A Comparison Between Four Common Ways of Recording and storing speech: Implications for forensic phonetics. *Proceedings, FONETIK,* Dept. of Linguistics, Stockholm University, Stockholm, 104-107.
- Quilis, A., & Esgueva, M. (1983). Realización de los fonemas vocálicos españoles en posición fonética normal. En M. Esgueva, & M. Cantarero (eds.), Estudios de fonética, vol. 1. Madrid: CSIC, 159-252.

El ritmo lingüístico en la caracterización del locutor. Percepción y aplicaciones judiciales

MARRERO AGUIAR, VICTORIA; LÓPEZ BASCUAS, LUIS ENRIQUE; MARTÍN FERNÁNDEZ, JOSÉ LUIS

1 Resumen

Presentamos en este trabajo los primeros resultados de un análisis sobre la percepción de las diferencias rítmicas y sus variaciones inter- e intralocutor. Hemos analizado, en función de las ocho métricas del ritmo más conocidas (%V, ΔV, ΔC, VarcoC, VarcoV, rPVI, nPVI-V y nPVI-C; Ramus, Nespor & Mehler, 1999; Low, Grabe & Nolan, 2000; White & Mattys, 2007, entre otros) tres sesiones de lectura de un mismo texto en 12 locutores de la base de datos AHUMADA (Ortega-García, González-Rodríguez & Marrero-Aguiar, 2000); se han seleccionado contrastes máximos entre hablantes y entre emisiones de un mismo hablante, controlando estrictamente el contenido segmental, las variaciones dialectales y la tasa de habla. Los estímulos han sido manipulados para privarlos de cualquier clave distinta a la estrictamente temporal (deslexicalización y asimilación de frecuencias e intensidades). Han sido presentados en una prueba de discriminación (paradigma 2IAX) a 10 jueces. Según nuestros resultados, las diferencias en métricas básicas como %V permitirían diferenciar emisiones de diferentes locutores, y no presentan variabilidad perceptible entre las distintas emisiones de un mismo locutor.

¹ Este trabajo se ha desarrollado en el marco del proyecto Estudio acústico y perceptivo de la variación prosódica inter e intralocutor en español- VILE-P (FFI2010-21690-C02-02) del Ministerio de Ciencia e Innovación.

2 Introducción

El estudio sobre las características individuales de la voz y sus aplicaciones forenses tradicionalmente no ha abordado el análisis de rasgos del dominio temporal, entre los que se cuenta el ritmo lingüístico, en comparación con la atención recibida por los pertenecientes al dominio de las frecuencias; entre los primeros, además, los elementos más asiduamente mencionados en la bibliografía han sido los relacionados con la tasa de habla y la proporción entre tiempo de habla y silencio (Battaner *et al.*, 2003). Sin embargo, muy recientemente (Leeman, Kolly & Dellwo, 2014) se ha empezado a considerar también el papel del ritmo en la caracterización individual del locutor; el ritmo lingüístico podría ofrecer la misma estabilidad intraindividual que otras conductas motoras, como los movimientos de los dedos o las piernas, y por lo tanto convertirse en un rasgo de gran interés en el ámbito judicial (Eriksson & Wretling, 1997, 1046).

2.1 El concepto de ritmo lingüístico

En este trabajo consideramos el ritmo lingüístico como la percepción de una sucesión regular de eventos con una estructura reconocible (en términos de duración o prominencia). Aunque esa regularidad física subyacente pueda ser más o menos mensurable, la percepción del ritmo no se relaciona directa o automáticamente con ningún evento acústico, sino que es una construcción mental resultado de una interpretación subjetiva de la señal: "es la mente humana la que percibe que ciertas claves físicas constituyen patrones rítmicos" (Auer, Couper-Kuhlen & Müller, 1999: 23).

En lingüística, tradicionalmente se ha identificado la ritmicidad lingüística con la isocronía; en principio, no necesariamente deben ir unidos ambos conceptos (en música existen ritmos anisócronos, los sincopados), pero lo cierto es que la investigación sobre ritmo ha estado condicionada por la búsqueda de la isocronía (cf. 2.2). En ocasiones se observa en la bibliografía otro paralelismo que, este sí, es indudablemente erróneo: el ritmo es un fenómeno diferente al de la tasa de habla o velocidad de elocución; no obstante, la segunda afecta al primero (Dellwo & Wagner, 2003), por lo que la velocidad de elocución ha de ser una variable controlada en las investigaciones sobre ritmo.

2.2 Los estudios sobre el ritmo lingüístico

Los estudios sobre el ritmo en las lenguas han seguido, durante la segunda mitad del siglo XX, la estela de la metáfora propuesta por Llyod James, en 1940, al diferenciar el ritmo tipo morse, propio del inglés y otras lenguas germánicas, del ritmo tipo ametralladora que observaba en las lenguas románicas; en cada grupo habría elementos diferentes que aparecen a intervalos regulares: el acento en el primero y las sílabas en el segundo. Pike, en 1945, reformula una propuesta (ampliada por muchos autores posteriores, como Abercrombie (1967), que la generalizó a las lenguas del mundo, o Ladefoged) en la que se establecen, pues, dos grandes categorías para el ritmo lingüístico: las lenguas de ritmo silábico (syllable-timed), caracterizadas por mostrar sílabas isócronas (típicas, se indicaba, de las lenguas románicas); y las de ritmo acentual (stress-timed), en las que la isocronía se da solo entre sílabas tónicas, de modo que la unidad rítmica es el pie; se consideró una características de las lenguas

germánicas. Posteriormente se propuso una tercera categoría, la de las lenguas de ritmo moraico, en las que la regularidad de intervalos se da entre moras (así ocurriría en japonés y tamil). Paralelamente se observaba una relación entre la complejidad de la estructura silábica y el tipo de ritmo: las reglas fonotácticas de las lenguas de ritmo acentual típicamente permiten más complejidad en ataques y codas silábicas; además tienen las vocales tónicas significativamente más largas que las átonas, típicamente reducidas; las lenguas de ritmo silábico tienden más a la estructura CV y tienen preponderancia de sílabas abiertas; la percepción del ritmo, por lo tanto, podría emerger de contrastes en la estructura silábica como mecanismo general.

A medida que aumentaban los estudios sobre emisiones reales en diversas lenguas, empezaron a surgir críticas respecto a esta visión; así, Dauer, por ejemplo, en 1983, observó que ni las sílabas mantenían una duración estable en lenguas de supuesto ritmo silábico, ni los intervalos interacentuales en las de ritmo acentual. También Borzone de Manrique y Signorini (1983) encontraron que en el español, considerado de ritmo silábico, la duración de las sílabas no era constante, y en cambio los intervalos interacentuales sí tendían a presentar una duración similar. Tampoco Wenk & Wiolland (1982) encontraron sílabas isócronas en francés, por lo que propusieron la existencia de unidades rítmicas más largas (la frase fonológica). Las moras, por último, tienen en japonés diferentes duraciones (Otake *et al.*, 1993). En definitiva, aunque se demostró que la hipótesis de la isocronía no encuentra soporte acústico, la idea de una tipología rítmica para las lenguas ha mantenido su vigencia.

A final de siglo, la psicolingüística evolutiva también hizo suya la propuesta dicotómica, y le dio un nuevo soporte: Mehler, Dupoux, Nazzi & Dehaene-Lambertz (1996) se basaron en ella para explicar cómo los niños pueden aprender aspectos de la fonología de la lengua materna, puesto que el ritmo les permite segmentar el continuo fónico, atendiendo a diferentes indicios, en un tipo de lengua u otro.

Pero el trabajo que marcó una nueva etapa en los estudios sobre ritmo lingüístico fue el de Ramus, Nespor & Mehler (1999): en él se defiende que "a simple segmentation of speech into consonants and vowels can:

- Account for the standard stress- / syllable-timing dichotomy and investigate the possibility of other types of rhythm;
- account for language discrimination behaviors observed in infants;
- clarify how rhythm might be extracted from the speech signal" (Ramus, Nespor & Mehler, 1999: 270-271).

Las fórmulas propuestas son las siguientes:

- %V: proporción de intervalos vocálicos sobre la duración total de un texto; se relaciona con la reducción vocálica en sílabas átonas y la complejidad de ataques y codas silábicas; por eso es más baja en las lenguas germánicas (cuyas vocales átonas se reducen, y cuya estructura silábica contiene más consonantes) y más alta en las romances (donde las vocales átonas apenas se acortan y cuyas sílabas son más sencillas).
- ΔV: desviación estándar de la duración de los intervalos vocálicos en un texto; se ve afectada por la existencia de oposición fonológica entre vocales largas y breves y por los fenómenos de alargamiento vocálico contextual.
- ΔC: desviación estándar de la duración de los intervalos consonánticos en un texto.

Una nueva variable aparece cuando Ramus (2002) introduce una tasa de normalización de la velocidad de habla¹: VarcoC, la desviación estándar de la duración de cada intervalo consonántico dividida entre la media de la duración

 $^{^{1}}$ Barry, Adreeva, Russo, Dimitrova & Kostadinova (2003) demostraron que ΔV y ΔC son inversamente proporcionales a la tasa de habla.

de todos ellos. Dellwo (2006) encontró que VarcoC resultó mejor que Δ C para distinguir inglés y alemán del francés, en todas las tasas de habla. El correlato de VarcoC para los intervalos vocálicos es VarcoV.

Low, Grabe & Nolan (2000) desarrollan otro argumento en este ámbito: el elemento esencial para capturar el ritmo es su naturaleza secuencial; para poder cuantificarla proponen usar el índice normalizado de variabilidad emparejada (*Normalised Pairwise Variability Index*) PVI, dividiendo la diferencia entre pares de intervalos vocálicos entre la suma de los intervalos. Luego Grabe & Low (2002) propusieron "a non-rate normalised PVI" para los intervalos consonánticos, porque la normalización podría enmascarar una variación rítmicamente relevante en la estructura de ataques y codas silábicos.

En los años siguientes aparecen nuevas fórmulas con matices destinados a captar la complejidad del ritmo lingüístico. Entre ellas, el *Compensation and Control Index* (CCI), de Bertinetto & Bertini (2008), definido como "un índice fonológicamente motivado": toma como base el rPVI y lo divide entre el número de segmentos fonológicos del texto (de la definición de estos últimos surge el carácter fonológico de la propuesta). Otra de las fórmulas recientes es %VO (Dellwo, Fourcin & Abberton, 2007), el porcentaje de segmentos sonoros de un texto, que considera prioritario el rasgo de sonoridad sobre el carácter consonántico o vocálico de un sonido.

En conclusión, aunque no podemos descartar la presencia de otras variables importantes en la codificación del ritmo en las lenguas (Prieto *et al.* [2012] resaltan el papel de la prominencia de frase, y las fronteras prosódicas), las diferencias de duración han demostrado jugar un papel importante, y las llamadas *métricas* del ritmo se han convertido en un instrumento generalizado para analizarlas. Considerando que los rasgos temporales son más resistentes que los frecuenciales ante las distorsiones generadas por ruido ambiente o por el filtrado telefónico (condiciones muy frecuentes en las grabaciones que se suelen presentar en el ámbito judicial), su aplicabilidad en el entorno de la llamada *fonética forense* constituirá el objetivo de los dos apartados siguientes.

2.3 ¿Es el ritmo un rasgo individual, caracterizador del hablante?

Existen numerosas fuentes de variabilidad rítmica. La mayor parte de la bibliografía se ha centrado en la que tiene su origen en la diversidad tipológica de las lenguas, pero también contamos con trabajos que demuestran las diferencias de ritmo entre diversas variedades geolectales de algunas de ellas (principalmente el inglés; para dialectos del español cf. O'Rourke 2008; Toledo, 2010). Se ha mencionado ya la influencia de la tasa de habla en la variabilidad del ritmo (Dellwo & Wagner, 2003); factores estilísticos también provocan fluctuaciones (por ejemplo, Payne *et al.*, 2010 encuentran diferencias rítmicas entre el habla entre adultos y la dirigida a niños); e incluso el contenido de la muestra: Widget *et al.* (2010) encuentran que las desviaciones entre diferentes locutores son menores que las causadas por las distintas oraciones del texto analizado.

Otro tipo de variabilidad más relevante, en el contexto de este trabajo, es el que tiene su origen en las diferencias entre el habla de distintos locutores (variabilidad interlocutor) y las emisiones de una misma persona (variabilidad intralocutor). Yoon (2010) encuentra que las PVIs son las métricas más estables entre varios hablantes (por lo tanto, las que menos interés revisten para

aplicaciones judiciales, donde lo que se busca son parámetros que caractericen al individuo). Mairano (2011: 106) encuentra, por el contrario, gran variabilidad en PVI entre locutores, mientras que el CCI resultaría más estable. En el trabajo de Loukina *et al.* (2011) los locutores tenían la posibilidad de repetir la lectura de los textos que se les presentaban: un 15 % de ellos fue repetido, y al comparar ambas emisiones se observó que la repetición "clearly did not affect the values of rhyhtm metrics" (Loukina *et al.*, 2011: 3262). Este resultado apunta, por tanto, a una baja variabilidad intralocutor en el ritmo lingüístico, lo cual es el objetivo de los rasgos válidos para la caracterización individual con fines judiciales.

2.4 Estudios de ritmo con finalidad forense

La lógica subyacente a la aplicabilidad del ritmo lingüístico en el ámbito forense está perfectamente expuesta en la siguiente cita de Dellwo, Leeman & Kolly (2012: 1584):

Humans have highly individual ways in which they move their legs when walking and that identification of individuals based on these movements is well possible. [...] We argue that speech is similar to walking in that it is a highly complex brain operated series of muscle movements which may be carried out to a considerable degree in individual ways [...] There is increasing evidence revealing that rhythm measures like %V, deltaV or the PVI may vary significantly between speakers.

En Leeman, Kolly y Dellwo (2014) se analizaron las *métricas* del ritmo anteriormente mencionadas, además de los intervalos entre picos silábicos (*syllable-peak-to-syllable-peak interval durations*, una medida basada en el concepto de *centros perceptivos*, Fowler 1979), en tres tipos de corpus, grabados con los mismos 16 sujetos (8 hombres y 8 mujeres, aunque no se encontraron efectos de esta diferencia): habla espontánea por micrófono, por teléfono y lectura por micrófono. Según sus resultados, %V y %VO permitieron distinguir entre locutores, tanto en habla espontánea como en lectura, sin presentar diferencias entre ambos estilos ni entre diferentes emisiones del mismo locutor, tal como requiere la comparación judicial de voces (Leeman, Kolly & Dellwo, 2014: 65).

Entre los temas que requieren, según los autores, investigación adicional, se menciona la necesidad de llevar a cabo estudios perceptivos: si los rasgos que caracterizan el ritmo lingüístico de una persona no resultan claramente perceptibles tampoco podrán ser fácilmente disimulados. Con el fin de contribuir al avance de conocimientos en este sentido hemos desarrollado este trabajo

3 Objetivos e hipótesis

El objetivo de este estudio es valorar la aplicabilidad de la caracterización rítmica del locutor en el ámbito judicial seleccionando las métricas más sensibles a la variación interlocutor y menos sensibles a la variación intralocutor, mediante un experimento diseñado con el mayor control posible de la variabilidad rítmica, excepto la dependiente del locutor; con ese fin se ha controlado la variabilidad lingüística y geolectal (todas las muestras pertenecen a la variedad centro-peninsular del español), el estilo de habla (lectura), el contenido segmental (idéntico para todos los locutores, y además deslexicalizado), la velocidad de habla (solo se consideran muestras con la

misma duración), la intensidad y la frecuencia fundamental (manipuladas para igualarlas).

Nuestras hipótesis principales son dos:

- Los oyentes serán capaces de discriminar dos emisiones basándose exclusivamente en diferencias de ritmo entre ellas.
- Los estímulos que se diferencian por duración vocálica serán los que obtengan mejores tasas de discriminación entre locutores y los menos afectados por la variación intralocutor.

4 Metodología

- a) Los parámetros analizados corresponden a ocho de las *métricas* del ritmo más utilizadas en la bibliografía:
 - %V: proporción de intervalos vocálicos sobre la duración total de la secuencia
 - ΔC: desviación estándar de las duraciones de todos los intervalos consonánticos en la secuencia
 - ΔV : desviación estándar de las duraciones de todos los intervalos vocálicos en la secuencia
 - VarcoC: cambios relativos en ΔC considerando la tasa de habla
 - VarcoV: cambios relativos en ΔV considerando la tasa de habla
 - rPVI-C (Pairwise variability index, non-rate normalised: diferencias de duración entre un intervalo consonántico y el siguiente, dividido por el número de intervalos de la secuencia
 - nPVI-C: similar a rPVI pero considerando la tasa de habla
 - nPVI-V: similar a nPVI-C pero considerando intervalos vocálicos y no consonánticos

Estímulos. Se han extraído de la base de datos AHUMADA (Ortega, González & Marrero, 2000), especialmente diseñada con fines forenses, datos de un texto equilibrado fonéticamente, leído por 12 locutores masculinos y grabados en tres sesiones de cada uno (por tanto, 36 emisiones de cada estructura silábica, 216 realizaciones). Aunque disponemos de secuencias de habla espontánea etiquetadas, para esta primera etapa se ha optado por la tarea de lectura de texto, que permitía controlar de forma precisa la variable contenido segmental. Se han seleccionado dos secuencias, una representante de la estructura silábica más simple, CV ("te toca la cara"), y la otra con mayor complejidad estructural, con predominio de CVC ("buscando los campos"). Se han controlado la aparición de rasgos geolectales en las secuencias y la tasa de habla, para evitar que las diferencias percibidas puedan residir en una mayor o menor velocidad de elocución (diferencias menores a las mínimas perceptibles, según Martín Fernández, López Bascuas & Marrero Aguiar, 2014). Los pares seleccionados fueron manipulados mediante un script para Praat elaborado Lahoz (2012) y resintetizados mediante Mbrola (Dutoit et al., 1996), con el fin de convertir todas las consonantes en /s/ y todas las vocales en /a/; en principio, este método iguala las diferencias en F0 e intensidad, pero como medida de control adicional, utilizamos una aplicación de Praat para trasladar el perfil tonal de un miembro de la pareja al otro, y garantizar así que el f0 no intervenía como variable. De los datos acústicos (extraídos durante una etapa previa del proyecto VILE-P), se han seleccionado las parejas que ofrecían mayor contraste en cada una de las métricas entre dos locutores diferentes y entre dos sesiones de un mismo locutor. Ha habido ocasiones en que el mayor contraste en una pareja coincidía entre diferentes métricas; así ocurrió, por ejemplo, con ΔC, VarcoC, rPvi y nPviC: en los cuatro casos, fue el locutor núm. 13, en sus

grabaciones 1 y 3, el que presentó los mayores contrastes. En la tabla 1 pueden verse todos los casos en que se dio esta circunstancia.

Tarea: finalmente la prueba contenía 20 parejas de estímulos (*a*, *b*) de sílabas tipo *sasasa*, en cuatro agrupaciones (*a-a*, *a-b*, *b-a*, *b-b*), que se presentaron repetidas ocho veces en orden aleatorio (pero en dos bloques separados, uno para la secuencia CV y otro para CVC), y por tanto constaba de 640 *ensayos*. Los sujetos debían indicar si se trataba de estímulos iguales o diferentes (tarea de discriminación con elección forzada). De acuerdo con la Teoría de Detección de Señales (Green & Swets, 1966), cuando los estímulos presentados son iguales, y el sujeto responde que lo son, nos encontramos ante un *rechazo correcto*; cuando en esta condición el sujeto responde que se trata de estímulos distintos, aparece una *falsa alarma*; por el contrario, cuando los estímulos son diferentes, se obtiene un *acierto* cuando el sujeto responde que lo son, y un *error* cuando los considera estímulos iguales.

Sujetos: 10 estudiantes universitarios, cuatro hombres y seis mujeres, sin problemas auditivos conocidos, sin conocimientos musicales avanzados y monolingües. Los sujetos no tenían experiencia previa en la tarea, pero contaron con una pequeña fase de entrenamiento. La prueba duró unos 45-50 minutos, y los sujetos recibieron una pequeña compensación económica por su participación.

5 Resultados

Se han obtenido 6368 respuestas (3184 corresponden a parejas iguales y 3184 a parejas distintas). La tasa global de identificación correcta (sumando aciertos y rechazos correctos) fue del 69 %; cuando los estímulos eran iguales, los rechazos correctos alcanzaron el 84 % de las presentaciones, pero cuando eran diferentes, solo un 54 % de las veces se identificaron como tales.

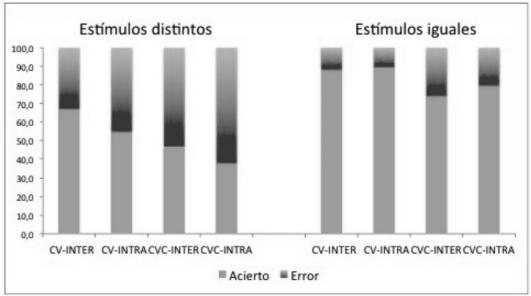


Figura 1. Resultados de estímulos CV y CVC en parejas iguales y diferentes de un mismo locutor (INTRA) y dos locutores distintos (INTER)

Como se comprueba en la figura 1, los resultados con secuencias CV son siempre mejores que con CVC, tanto en estímulos iguales como diferentes, y tanto en parejas de estímulos procedentes del mismo locutor como de distintos locutores; esta estructura presenta los resultados esperados: tasas de discriminación cercanas o superiores al 70 % para estímulos distintos procedentes de distintos locutores, y cercanas al 50 % (acierto por azar) para las que proceden del mismo locutor. En cambio, los estímulos obtenidos con

secuencias CVC presentan tasas de acierto inferiores al 50 % en todos los casos, e incluso una mejor discriminación de las diferencias intralocutor que las interlocutor en estímulos iguales. La prueba T de Student indica que cuando los estímulos son iguales las diferencias inter- e intralocutor no son significativas ni con estímulos CV (p = 0.27) ni con estímulos CVC (p = 0.08); por el contrario, ante estímulos distintos los resultados en emisiones de un mismo locutor y de dos hablantes diferentes resultaron estadísticamente significativas, tanto en estímulos CV (p = 0.033) como CVC (p = 0.027).

Para determinar cuáles son las *métricas* más adecuadas en el ámbito judicial nos centraremos solo en los resultados con estímulos diferentes, considerando tanto los aciertos en parejas de locutores distintos (INTER) como en diferentes emisiones del mismo hablante (INTRA).

		CV	
INTERLOCUTOR	%	INTRALOCUTOR	%
%V	93,75	%V	26,25
ΔC	41,25	ΔC=VarcoC=rPvi=nPviC	66,25
ΔV	84,38	ΔV	55,13
rPvi	51,88		
nPvi	39,10	nPvi	58,13
VarcoV	75,00	VarcoV	66,88
VarcoC=nPviC	81,88		
		CVC	
%V	61,25	%V	45,00
ΔC=VarcoC	59,38	ΔC=VarcoC	39,74
ΔV=VarcoV=nPvi=nPvi C	36,54	ΔV=VarcoV=nPvi	39,38
rPvi	29,38	rPvi=nPviC	26,25

Tabla 1. Porcentajes de acierto y medida de sensibilidad (d') en estímulos distintos según las diferentes métricas utilizadas

Las pruebas estadísticas de significatividad de las diferencias (ANOVA) en estímulos CV interlocutor indican que no hay diferencias significativas entre %V, Δ V, VarcoC=nPviC y VarcoV; ni entre rPVI, Δ C y nPVI, aunque sí entre ambos grupos (p <0.001). En estímulos CV intralocutor solo %V se diferencia significativamente de todas las demás *métricas*. (p<0.001). En cuanto a los estímulos CVC, en emisiones de locutores diferentes se observan diferencias significativas entre %V- Δ C=VarcoC y las demás métricas; en emisiones de un mismo locator no hay diferencias en ninguna de las métricas.

Las diferencias en %V son las que mejor responden a las hipótesis planteadas: resultan claramente perceptibles cuando corresponden a diferencias entre locutores y muy difíciles de discriminar cuando proceden del mismo locutor. El resto de las parejas presentan una perceptibilidad baja en estímulos CVC, aunque algunas (ΔV, VarcoV y VarcoC-nPVI-C) se distinguieron bien en estímulos CV de diferentes locutores. Las demás métricas presentan una baja discriminabilidad tanto cuando los locutores son diferentes como cuando es el mismo en diferentes emisiones.

Por lo tanto, nuestros resultados conducen a afirmar que las diferencias de ritmo en condiciones de máximo control de los estímulos son percibidas de forma consistente por sujetos hispanohablantes cuando la estructura silábica es simple (CV), con tasas de discriminación correcta superiores al $80\,\%$ en métricas como %V, Δ V o VarcoC-nPVI-C, pero fue difícil con estructuras más complejas (en estímulos CVC esas tasas no subieron del $66\,\%$). En general, las diferencias entre locutores fueron mejor percibidas que los contrastes intralocutor, aunque la diferencia entre las dos estructuras silábicas condiciona este resultado.

6 Conclusiones

Si retomamos las hipótesis planteadas al comienzo, podemos concluir que algunas de las diferencias de ritmo estudiadas han resultado claramente perceptibles contando única y exclusivamente con información temporal en la señal (se confirma, por tanto, nuestra primera hipótesis). En concreto, %V, una medida básica de duración vocálica, ha presentado el comportamiento esperado: muy alta discriminabilidad entre locutores diferentes y muy baja entre distintas emisiones de un mismo locutor (se confirmaría así también nuestra segunda hipótesis).

Algunos de los demás resultados obtenidos eran esperables: las métricas formuladas para compensar las diferencias en tasa de habla (VarcoC/V y nPviC/V) no resultan eficaces con nuestra muestra, cuya tasa de habla había sido controlada previamente.

No obstante, algunos resultados han sido inesperados. Esencialmente, el comportamiento de los estímulos construidos con predominio de secuencias CVC, contrario al previsto: una discriminación inferior al 50 % (acierto por azar) cuando los estímulos eran diferentes; y mayor discriminación para las diferencias intralocutor que para las interlocutor en algunas condiciones. Creemos que la explicación puede residir en el tratamiento experimental que han sufrido los segmentos consonánticos sonoros de la secuencia elegida ("buscando los campos"). Con la manipulación descrita en la metodología (la habitual en estudios de percepción del ritmo), se han convertido en [s], igual que las consonantes sordas. Se pierde así el rasgo de sonoridad, que según se ha demostrado en investigaciones anteriores (Leeman, Kolly & Dellwo, 2014; Dellwo, Leeman & Kolly, 2012) puede jugar un papel muy relevante en la codificación del ritmo lingüístico.

Por lo tanto, entre los próximos pasos de nuestra investigación se incluirá %VO entre las *métricas* analizadas, y se modificará la manipulación de los estímulos para su deslexicalización, con el fin de mantener el rasgo sonoro de las consonantes que lo presenten. Además, se corroborará en los datos acústicos correspondientes al texto completo de los 24 locutores analizados en VILE-P que la variabilidad rítmica en %V y %VO es baja entre distintas emisiones de un mismo hablante y alta entre emisiones de hablantes diferentes, y se realizará un nuevo experimento con mayor número de sujetos.

7 Bibliografía

Auer, P., Couper-Kuhlen, E., & Müller, F. (1999). Language in time: The rhythm and tempo of spoken interaction. New York: Oxford University Press.

Battaner, E., Gil, J., Marrero, V., Llisterri, J., Carbó, C., Machuca, MJ, de-la Mota & Ríos, A. (2003). VILE: Estudio acústico de la variación inter- e intralocutor en español. En SEAF 2003. Actas del II Congreso de la

Sociedad Española de Acústica Forense Barcelona, SEAF, 59-70.

Bertinetto, P. M., & Bertini, C. (2008). On modeling the rhythm of natural languages. In *Proceedings of the Fourth International Conference on Speech Prosody*. Recuperado el 12 del 11 de 2014, de http://linguistica.sns.it/QLL/QLL0

- 7_08/Bertinetto_Bertini_campinas.P
- Borzone de Manrique, A.M. y Signorini, A. (1983): Segmental duration and rhythm in Spanish. *Journal of Phonetics*, 11 (2), 117-128.
- Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of phonetics*, 11, 51-62.
- Dellwo, V., & Wagner, P. (2003).

 Relationships between rhythm and speech rate. 15th ICPhS, Barcelona.

 Recuperado el 11 del 12 de 2014 de http://pub.uni-bielefeld.de/luur/download?func=downloadFile&recordOId=1785384&fileOId=2352642
- Dellwo, V., Fourcin, A. & E. Abberton (2007). Rhythmical classification based on voice parameters, En: *International Conference of Phonetic Sciences (ICPhS)*, Saarbrücken, 1129-1132.
- Dellwo, V., Leemann, A., & Kolly, M. J. (2012). Speaker idiosyncratic rhythmic features in the speech signal. In *INTERSPEECH* 2012. Recuperado el 12 del 11 de 2014, de http://20.21019352.unknown.qala.com.sg/archive/archive_papers/interspeech_2012/i12_1584.pdf
- Dutoit, T., Pagel, V., Pierret, N., Bataille, F., & Van der Vrecken, O. (1996). The MBROLA project: Towards a set of high quality speech synthesizers free of use for non commercial purposes. In Spoken Language, 1996. ICSLP 96. Proceedings., Fourth International Conference on (Vol. 3, pp. 1393-1396). IEEE.
- Dutoit, T.; Pagel, V.; Pierret, N.; Bataille, F., & Van der Vrecken, O. (1996). The MBROLA project: Towards a set of high-quality speech synthesizers free of use for non-commercial purposes. Proceedings of the Fourth International Conference on Spoken Language (ICSLP'96). Philadelphia. IEEE: 1393-1396.
- Eriksson, A., & P. Wretling (1997): How Flexible is the Human Voice? A Case Study of Mimicry. Proceedings of the European Conference on Speech Technology, Eurospeech 1997, 1043-1046.
- Fowler, A. (1979). Perceptual centers in speech production and perception, *Perception and Psychophysics* 25, 375-388.

- Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. *Papers in laboratory phonology*, *7* (515-546).
- Green, D. M., & Swets, J. A. (1966): *Signal detection theory and psychophysics*. Oxford, Robert E. Krieger. Reimpreso en 1974.
- Lahoz, J. M.^a (2012). Syllable, accent, rhythm: typological and methodological considerations for teaching Spanish as a foreign language. *Revista Internacional de Lenguas Extranjeras*, (1), 129-150. Recuperado el 12 del 11 de 2014, de http://www.publicacionsurv.cat/ojs/revistes/index.php/rile/article/viewFile/13/15
- Lahoz, J. M.^a (2012). La enseñanza de la entonación, el ritmo y el tempo. En J. Gil Fernández, (ed.). *Aproximación a la enseñanza de la pronunciación en el aula de español*. Madrid. Edinumen, 93-132.
- Leemann, A., Kolly, M. J., & Dellwo, V. (2014). Speaker-individuality in suprasegmental temporal features: Implications for forensic voice comparison. Forensic science international, 238, 59-67.
- Ling, L. E., Grabe, E., & Nolan, F. (2000). Quantitative Characterizations of Speech Rhythm: Syllable-Timing in Singapore English. *Language and* speech, 43(4), 377-401.
- Loukina, A., Kochanski, G., Rosner, B., Keane, E., & Shih, C. (2011). Rhythm measures and dimensions of durational variation in speech. *The Journal of the Acoustical Society of America*, 129 (5), 3258-3270.
- Low, L. E., Grabe, E., & Nolan, F. (2000). Quantitative Characterizations of Speech Rhythm: Syllable-Timing in Singapore English. *Language and Speech*, 43(4), 377-401.
- Mairano, P. (2011). Rhythm typology: acoustic and perceptive studies (Doctoral dissertation, Università degli studi di Torino). Recuperado el 12 del 11 de 2014, de http://halshs.archives-ouvertes.fr/docs/00/65/42/61/PDF/MairanoPaolo_These.pdf
- Martín Fernández, J. L., López Bascuas, L. E., & Marrero Aguiar, V. (2014). Percepción del tempo en estímulos lingüísticos de diferente complejidad. VI CIFE, Valencia (comunicación).
- O'Rourke, E. 2008. Correlating speech rhythm in Spanish: Evidence from

- two Peruvian dialects. *Selected Proceedings of the 10th Hispanic Linguistics Symposium*. J. Bruhn de Garavito y E. Valenzuela. Somerville (eds.). Massachusetts: Cascadilla Proceedings Project. 276-287
- Ortega-Garcia, J., Gonzalez-Rodriguez, J., & Marrero-Aguiar, V. (2000). AHUMADA: A large speech corpus in Spanish for speaker characterization and identification. *Speech communication*, 31(2), 255-264.
- Otake, T., Hatano, G., Cutler, A., & Mehler, J. (1993). Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language*, 32(2), 258-278.
- Payne, E., Post, B., Astruc, L., Prieto, P., & Vanrell, M. D. M. (2010). A crosslinguistic study of prosodic lengthening in child-directed speech. Prieto *et al.*, 2012
- Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, 73(3), 265-292.
- Roach, P. (1982). On the distinction between 'stress-timed' and 'syllable-timed' languages. *Linguistic controversies*, 73-79.
- Toledo, G. (2010). Métricas rítmicas en tres dialectos Amper-España. *Estudios Filológicos* 45, 93-110.
- White, L., & Mattys, S. L. (2007). Calibrating rhythm: First language and second language studies. *Journal of Phonetics*, 35(4), 501-522.
- Wiget, L., White, L., Schuppler, B., Grenon, I., Rauch, O., & Mattys, S. L. (2010). How stable are acoustic metrics of contrastive speech rhythm?. *The Journal of the Acoustical Society of America*, 127(3), 1559-1569.
- Wretling, P., & Eriksson, A. (1998). Is articulatory timing speaker specific? Evidence from imitated voices. En *Proc. FONETIK* (vol. 98, 48-52). Recuperado el 12 del 11 de 2014, de http://www.ling.gu.se/~anders/papers/imit98.pdf
- Yoon, T. J. (2010). Capturing inter-speaker invariance using statistical measures of rhythm. In *Proceedings of speech prosody*. Recuperado el 12 del 11 de 2014, de http://20.21019352.unknown.qala.com.sg/archive/sp2010/papers/sp10_201.pdf

Efectos del cambio de lengua en la comparación de voces mediante LTAS

ROSEANO, PAOLO; FERNÁNDEZ PLANAS, ANA MA.; ELVIRA-GARCÍA, WENDY; MARTÍNEZ CELDRÁN, EUGENIO

1 Introducción

A lo largo de las últimas décadas se ha ido definiendo un conjunto de parámetros acústicos que se consideran aptos para ser utilizados en la comparación de voces con finalidades judiciales (v. Gil, 2014). Entre ellos se encuentra el Espectro Medio a Largo Plazo (o LTAS) que, como es sabido, es una representación del espectro promediado de la amplitud en una gama de frecuencias dada calculado para un intervalo de tiempo.

Desde los primeros intentos de utilizar el LTAS para el reconocimiento automático del locutor (Pruzansky, 1963), se ha abierto un filón de investigación cuya finalidad es valorar la fiabilidad de esta técnica en la comparación de voces (Zalewski *et al.*, 1975; Pittam, 1987; Nolan, 1983, Molina de Figueiredo & Bernales, 1999; Cicres, 2011; i.a.). Uno de los aspectos que han recibido cierta atención es el de la efectividad de esta técnica en la comparación de muestras de voz del mismo locutor que difieren entre sí en cuanto a variaciones debidas a ajustes fonatorios o a efectos del canal de transmisión (Klingholz *et al.*, 1988, Nordenberg & Sundberg, 2003; i.a.). Una atención menor se ha dedicado a la aplicabilidad del LTAS en los casos donde las muestras de voz dubitada e indubitada son de lenguas distintas (a causa, por ejemplo, de que un sospechoso poco colaborativo se niegue a grabar en la misma lengua de la muestra dubitada).

Los trabajos que se centran en la aplicabilidad del LTAS en la comparación de muestras de voz producidas en lenguas distintas por hablantes bilingües, además de ser escasos, no coinciden en sus conclusiones. Mientras Harmegnies et al. (1991), por ejemplo, concluyen que las características de cada lengua no influyen sobre el LTAS, estudios como los de Bruyninckx et al. (1994) y Manwa et al. (2012) abogan que la lengua hablada tiene efectos sobre los LTAS. Cabe destacar también que ninguno de estos trabajos está enfocado a la fonética judicial.

2 Objetivos

Este trabajo tiene dos objetivos, de los cuales uno es de tipo teórico y el otro tiene una caracterización más bien aplicada. Desde una perspectiva más teórica, se pretende retomar el estudio de los efectos del cambio de lengua sobre el LTAS, filón de investigación que en los últimos 20 años ha sido objeto de una atención marginal.

El segundo objetivo, de carácter metodológico y aplicado, es desarrollar una propuesta de rutina de análisis del LTAS con fines judiciales. La rutina en cuestión, preparada en el entorno de Praat (Boersma & Weenink, 2013), permite medir las distancias entre muestras de habla basándose en el LTAS (Majewski & Hollien, 1974; Zalewski et al., 1975). La automatización de la comparación entre los LTAS de dos muestras de voz no tiene como objetivo el de desautorizar al experto o de reducir la importancia de su papel, sino el de obviar algunos de los límites reconocidos de los métodos más tradicionales de comparación de muestras de habla, es decir su posible subjetividad y su alto coste en términos de tiempo necesario (Künzel, 2011: 39). Además, es importante subrayar como la literatura científica de los últimos años sugiere que la comparación automática del habla podría representar una ventaja justamente en los casos reales en los que se disponga de dos muestras de habla en lenguas diferentes una de las cuales sea poco conocida por el experto (Künzel, 2011: 41-42).

3 Metodología

3.1 Corpus

El corpus utilizado para este trabajo se compone de las grabaciones de 12 hablantes bilingües de Italiano regional de Friul y del dialecto friulano del Bajo Valle de Gorto, 8 mujeres y 4 hombres, leyendo el texto de *El viento Norte y el sol* en ambas lenguas¹. Las grabaciones se efectuaron en una sola sesión para cada hablante, en casa del mismo, en una habitación silenciosa, mediante una grabadora digital Marantz modelo PMD620 que llevaba conectado un micrófono direccional Shure SM58. Los archivos de sonido se grabaron en formato .wav mono con una frecuencia de muestreo de 44.100 Hz.

La decisión de utilizar voces masculinas y femeninas, así como la de trabajar con grabaciones de alta calidad de habla de laboratorio realizadas en una sola sesión (todas condiciones diferentes de las que se darían en un caso judicial real) responden a la necesidad de testar, en esta primera fase, el funcionamiento del método automático de comparación de voces en una situación próxima a la ideal. Si el resultado fuera satisfactorio en estas condiciones, se podría pasar a otras fases de experimentación que consistirían, por ejemplo, en utilizar voces

Los materiales se grabaron en diciembre de 2013 en el marco de un proyecto de investigación sobre las características rítmicas de lenguas en contacto (friulano e italiano) que el Laboratori de Fonètica de la Universitat de Barcelona está llevando a cabo en colaboración con la Universität Hamburg.

de hablantes del mismo sexo y con grados diferentes de alteración por efecto del canal de transmisión.

Los sistemas fonéticos del italiano y del friulano del Bajo Valle de Gorto presentan un grado de divergencia que podríamos definir moderado. En cuanto al sistema vocálico, ambas lenguas presentan siete timbres para las vocales tónicas [i, e, ϵ , a, δ , o, u], que en posición átona pasan a ser cinco en italiano [i, e, a, o, u] y seis en friulano [i, e, a, v, o, u]. En relación con los glides, ambas lenguas muestran los mismos, es decir [j, w]. Con respecto al sistema consonántico, las dos lenguas comparten una parte importante de elementos, es decir [m, n, n, n, p, b, t, d, k, g, ts, dz, tf, d3, f, v, s, z, f, r, l]. A esos, el italiano regional añade [Λ]; el friulano del Valle de Gorto [3, c, η]. En cuanto a la presencia de elementos geminados, cabe recordar que en friulano, a diferencia del italiano, existen también vocales tónicas fonológicamente largas [i:, e:, ϵ ; a:, δ ; o:, u:]. Además, mientras que el italiano presenta un repertorio muy amplio de consonantes geminadas, en friulano solamente se pueden producir en algunos casos por fonética sintáctica.

3.2 Script

La comparación entre los LTAS de las diferentes muestras se ha llevado a cabo mediante un script de Praat (Elvira-García *et al.*, 2014) que efectúa dos operaciones: 1) calcula el LTAS de cada muestra y, a continuación, 2) calcula las distancias entre ellos y las guarda en un documento de texto que puede ser gestionado mediante Excel.

3.2.1 Cálculo del LTAS

Las versiones más recientes de Praat permiten calcular, además del LTAS tradicional, también el pitch-corrected LTAS, que es una versión del LTAS que intenta corregir los efectos de las variaciones de F0 (Boersma y Kovačić, 2006) y que es la que se ha utilizado para este trabajo. El script que se ha predispuesto calcula el LTAS para un rango frecuencial que llega hasta los 8000 Hz, divididos en pasos de 100 Hz. Como resultado, el LTAS de cada muestra resulta ser una secuencia de 80 valores numéricos (figura 1), ordenados según la frecuencia a la cual corresponden. Los 80 valores numéricos en cuestión se guardan en una base de datos desde la cual se recuperarán en la fase siguiente del análisis.

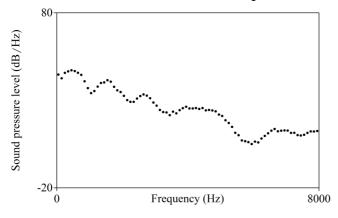


Figura 1. LTAS de un archivo que contiene la grabación de El viento norte y el sol

3.2.2 Cálculo del LTAS

Una vez determinado el LTAS de todas las muestras que se quieren comparar, el paso siguiente es el cómputo de las diferencias entre cada una de ellas y las demás. El cálculo de la diferencia entre dos secuencias de 80 números, que llamamos LTAS₁ y LTAS₂, es una operación matemática sencilla. La diferencia en cuestión, que denominamos Δ_{LTAS} , es la suma de los valores absolutos de las diferencias entre cada pareja de números que ocupan la misma posición en la secuencia dividida por el número de observaciones. En términos más formales:

$$\begin{split} LTAS_1 &= \{x_1, \ x_2, \ x_3, \ x_4, \ x_5... \ x_{80}\} \\ LTAS_2 &= \{y_1, \ y_2, \ y_3, \ y_4, \ y_5... \ y_{80}\} \\ \Delta_{LTAS} &= \sum (|x_1 - y_1| + |x_2 - y_2| + |x_3 - y_3| + |x_4 - y_4| ... + |x_{80} - y_{80}|)/80 \end{split}$$

Cuanto más grande es Δ_{LTAS} , mayor es la diferencia entre los LTAS de las dos muestras. Si fuera cierto que, en hablantes bilingües, el cambio de lengua no produce ningún efecto apreciable sobre el LTAS (que podemos considerar nuestra hipótesis de trabajo o Hp), el Δ_{LTAS} entre las dos muestras del mismo sujeto hablando lenguas distintas debería ser inferior al Δ_{LTAS} entre cada una de esas muestras y cualquier otra muestra de sujetos distintos. En términos más formales, si definimos como Δ_{id} la distancia entre los LTAS de la muestra indubitada y la dubitada y como D el conjunto de las distancias Δ entre todas las muestras del corpus (con exclusión, obviamente, de Δ_{id}):

$$Hp \Leftrightarrow \forall \Delta \in D : \Delta_{id} < \Delta$$

No obstante, antes de verificar esta hipótesis a partir de los datos de nuestro corpus, es preciso efectuar algunas correcciones al modelo matemático – demasiado sencillo- que se acaba de presentar. Estas correcciones son requeridas por la naturaleza misma del LTAS. Recordamos, de hecho, que el LTAS es función, en primer lugar, de la intensidad de la señal en cada banda de frecuencia. Esto puede tener consecuencias nefastas a la hora de comparar los LTAS de dos muestras, como se demuestra en el ejemplo a continuación. Imaginemos que se graben dos sujetos (el hablante A y el hablante B) leyendo el mismo texto en castellano (El viento norte y el sol). El hablante A es grabado con una grabadora conectada con un micrófono direccional puesto a 30 cm de su boca. La muestra así obtenida se llamará α_{30} . El hablante B es grabado con dos grabadoras idénticas conectadas con dos micrófonos del mismo modelo, puestos a 30 y a 70 cm de su boca. La muestras así obtenidas se llamarán β_{30} y β_{70} (figura 2).

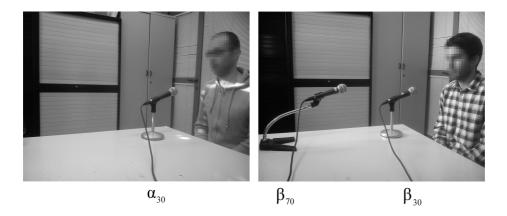


Figura 2. Condiciones de grabación del hablante A (panel de la izquierda) y del hablante B (panel de la derecha)

Los LTAS de las tres muestras, calculados con el método expuesto anteriormente, son las que se representan en la figura 3. Las dos líneas discontinuas representan la muestra β_{30} (línea de rayas, la más alta de las tres), la β_{70} (línea de puntos, la más baja) y la α_{30} (línea continua).

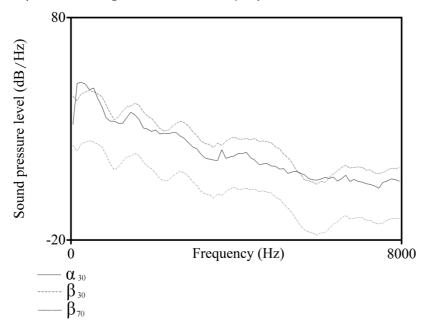


Figura 3. LTAS de las muestras α_{30} , β_{30} y β_{70}

De acuerdo con las fórmulas para el cálculo de las distancia que se han presentado anteriormente, la diferencia entre el LTAS de β_{30} y α_{30} sería mucho menor de la que hay entre los LTAS de β_{30} y β_{70} , a pesar de ser estas dos grabaciones de la misma emisión de voz. Por lo tanto, a partir del resultado de la fórmula en cuestión en un contexto judicial se tendría que rechazar la posibilidad de afirmar que β_{30} y β_{70} sean emisiones realizadas por el mismo hablante, ya que son las más diferentes del grupo. Esta conclusión es contraria a lo que sabemos que es cierto (y el ojo de un experto efectivamente reconoce que los LTAS de β_{30} y β_{70} tienen el mismo perfil, a pesar de encontrarse a alturas diferentes). El error deriva de la incapacidad de la fórmula antedicha para comparar eficazmente los perfiles de dos LTAS que, a pesar de tener un perfil

idéntico, se sitúan a alturas diferentes debido a la mayor/menor intensidad de grabación de la emisión.

La solución matemática que permite obviar errores tan evidentes consiste en calcular la media aritmética (\bar{x}) de los LTAS que se comparan y utilizarla para neutralizar las diferencias globales de intensidad entre las muestras, mientras las diferencias locales, es decir los picos y valles del LTAS, se mantienen. Por ejemplo, si comparamos los LTAS de β_{70} y β_{30} , observamos que \bar{x} (LTAS β_{70}) = 26 y \bar{x} (LTAS β_{30}) = 4, por lo tanto la diferencia entre las dos es $\Delta_{\bar{x}} = \bar{x}$ (LTAS β_{70}) - \bar{x} (LTAS β_{30}) = 22. Si, a continuación, sumamos a cada valor de la segunda secuencia del LTAS (en este caso β_{30}) el valor de la de Δ_x neutralizamos el efecto de la diferencia global de intensidad. Si se quiere visualizar gráficamente esta corrección, en el caso de las curvas de los LTAS β_{30} y β_{70} de la figura 3 se puede decir que se desliza hacia arriba la línea discontinua inferior (β_{30}) de 22 unidades, hasta sobreponerla a la superior (β_{70}), con el resultado que se observa en la figura 4.

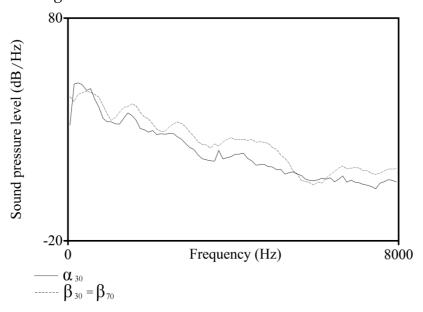


Figura 4. LTAS de las muestras α_{30} , β_{30} y β_{70} después de corregir los efectos de la intensidad global de la grabación

Resultados

Una vez calculadas las distancias entre el LTAS de cada muestra del corpus y los de todas las demás, se puede proceder a valorar si los resultados permiten confirmar nuestra hipótesis de trabajo, según la cual si es cierto que en hablantes bilingües el cambio de lengua no produce ningún efecto apreciable sobre el LTAS (o, en otras palabras, que el LTAS es un parámetros resistente al cambio de lengua), el ALTAS entre las muestras del mismo sujeto hablando dos lenguas distintas debe ser inferior al Δ LTAS entre cada una de las dos muestras del sujeto en cuestión y cualquier otra muestra de sujetos distintos. Esta hipótesis se cumple solo para 6 de los 12 informantes de nuestro corpus (figura 5).

N distancia S1S2L1L2	N distancia S1S2L1L2	N distancia S1S2L1L2	N distancia S1S2L1L2	N distancia S1S2L1L2
1 1,823782 02 02 ita fri	57 3,748049 05 04 ita ita	113 4,5573988 14 11 ita ita	169 5,2742606 14 11 fri fri	225 6,0719864 04 01 fri ita
2 1,829295 06 06 ita fri	58 3,7695174 03 02 ita ita	114 4,5583423 13 12 fri ita	170 5,2836707 11 05 fri ita	226 6,0981992 04 01 fri fri
3 1,998899 12 12 ita fri	59 3,7785162 11 02 ita ita	115 4,5674981 05 02 fri fri	171 5,2898241 15 12 fri ita	227 6,108689 10 04 fri fri
4 2,202308 11 11 ita fri	60 3,7796159 13 12 ita fri	116 4,5675297 05 03 ita fri	172 5,2947657 11 04 ita ita	228 6,118825 10 02 fri ita
5 2,281556 15 14 fri ita	61 3,7902638 13 01 ita ita	117 4,5868785 02 01 fri fri	173 5,3025714 15 06 ita ita	229 6,1305954 13 01 fri fri
6 2,324756 14 14 ita fri	62 3,8588683 03 01 fri fri	118 4,6189917 06 03 fri fri	174 5,3144265 10 02 fri fri	230 6,1346357 14 13 ita ita
7 2,358865 15 15 ita fri	63 3,876015 11 01 fri ita	119 4,6273547 04 03 fri fri	175 5,3175968 12 05 fri ita	231 6,1434437 14 11 fri ita
8 2,409567 15 14 ita ita	64 3,894169 13 02 fri ita	120 4,6798393 11 10 ita fri	176 5,3238184 15 02 ita ita	232 6,157609 13 05 ita ita
9 2,48638 15 10 ita ita	65 3,9381838 03 02 fri ita	121 4,6876054 12 10 fri fri	177 5,3287122 05 01 fri ita	233 6,1883119 12 04 fri ita
10 2,550965 12 01 ita ita	66 3,9459201 11 03 ita ita	122 4,7051348 13 06 fri ita	178 5,3508723 06 03 ita ita	234 6,2777669 10 06 ita fri
11 2,637762 12 11 ita ita	67 3,9538445 05 04 fri ita	123 4,7066011 02 01 ita ita	179 5,3514761 14 03 fri fri	235 6,3324038 10 04 ita ita
12 2,646978 13 13 ita fri	68 3,9541922 12 01 ita fri	124 4,7144144 13 06 ita ita	180 5,3650327 12 05 ita fri	236 6,3656479 14 06 ita fri
13 2,657574 12 11 fri fri	69 3,9772263 12 10 fri ita	125 4,7221318 10 03 fri fri	181 5,3705186 02 01 ita fri	237 6,3716675 06 01 fri fri
14 2,695544 10 10 ita fri	70 3,9932746 10 01 ita ita	126 4,7364495 12 10 ita fri	182 5,3821096 12 06 fri fri	238 6,3921832 15 05 ita fri
15 2,732518 05 05 ita fri	71 4,0098588 13 02 fri fri	127 4,7386446 13 11 fri fri	183 5,4372312 15 11 fri ita	239 6,3953489 14 05 ita ita
16 2,737132 15 14 fri fri	72 4,0135312 11 10 fri fri	128 4,7415498 06 02 ita ita	184 5,4404972 13 01 ita fri	240 6,4003101 13 10 ita fri
17 2,788951 12 01 fri ita	73 4,0487825 06 03 ita fri	129 4,7462712 12 06 ita ita	185 5,4556826 04 03 fri ita	241 6,4324102 14 03 fri ita
18 2,885154 12 11 ita fri	74 4,0502473 14 01 ita fri	130 4,7482411 11 04 ita fri	186 5,4958825 10 06 ita ita	242 6,4594522 15 04 ita ita
19 2,968571 15 10 ita fri	75 4,0560455 13 02 ita ita	131 4,7630066 10 02 ita ita	187 5,510206 14 01 fri ita	243 6,5027541 14 04 ita ita
20 2,973785 11 03 fri fri	76 4,0697353 14 11 ita fri	132 4,7645328 03 01 ita fri	188 5,5205824 15 03 fri ita	244 6,5073828 13 05 fri ita
21 2,992057 12 03 fri fri	77 4,0969278 10 02 ita fri	133 4,7655514 10 03 fri ita	189 5,5288655 13 06 fri fri	245 6,5381692 15 06 fri ita
22 2,995343 06 04 fri fri	78 4,1044663 14 03 ita fri	134 4,776113 05 04 fri fri	190 5,5495002 06 02 fri fri	246 6,5466059 13 04 fri fri
23 3,010459 13 12 ita ita	79 4,1055801 13 11 ita fri	135 4,7783661 06 02 ita fri	191 5,6081708 06 05 fri ita	247 6,565659 10 06 fri fri
24 3,069423 01 01 ita fri	80 4,1115489 11 04 fri fri	136 4,7856381 15 11 fri fri	192 5,6510237 05 01 fri fri	248 6,589456 13 04 ita fri
25 3,082765 04 04 ita fri	81 4,1254016 11 01 fri fri	137 4,8074629 12 02 fri ita	193 5,6658194 15 04 ita fri	249 6,5901065 04 01 ita fri
26 3,122116 03 01 fri ita	82 4,1510002 15 03 ita ita	138 4,8185952 14 03 ita ita	194 5,6717351 06 01 ita fri	250 6,6038721 13 04 ita ita
27 3,14968 11 03 ita fri	83 4,1578268 11 02 fri fri	139 4,8428453 11 05 ita fri	195 5,6772526 04 02 ita ita	251 6,6046825 15 04 fri fri
28 3,156878 15 11 ita fri	84 4,1585951 06 04 ita ita	140 4,845419 15 03 fri fri	196 5,6809868 14 12 fri ita	252 6,6143523 15 13 fri ita
29 3,189692 12 11 fri ita	85 4,1984762 11 06 ita ita	141 4,8472784 13 12 fri fri	197 5,6921174 12 04 fri fri	253 6,6620877 14 02 ita ita
30 3,201318 14 10 ita ita	86 4,2481547 13 02 ita fri	142 4,8493973 05 01 ita ita	198 5,7133359 06 02 fri ita	254 6,8420181 15 06 fri fri
31 3,279953 10 03 ita ita	87 4,2492439 12 02 ita fri	143 4,8513194 15 13 ita ita	199 5,7278332 13 05 ita fri	255 6,9078272 13 04 fri ita
32 3,293079 12 03 ita fri	88 4,2530035 12 03 ita ita	144 4,8581636 01 14 fri fri	200 5,7533805 04 02 fri fri	256 6,9456793 10 05 fri ita
33 3,344809 06 04 ita fri	89 4,2792913 10 01 ita fri	145 4,8657634 06 01 ita ita	201 5,7859903 14 02 ita fri	257 7,0356706 14 05 ita fri
34 3,372913 14 10 ita fri	90 4,2841625 10 01 fri fri	146 4,9019799 13 10 ita ita	202 5,7957967 12 05 fri fri	258 7,0739466 10 05 fri fri
35 3,389067 15 03 ita fri	91 4,3236761 11 01 ita fri	147 4,9145105 05 02 ita ita	203 5,7969752 15 13 ita fri	259 7,0763516 14 04 fri fri
36 3,445617 11 10 ita ita	92 4,324982 02 01 fri ita	148 4,9212734 06 05 ita ita	204 5,7970767 13 05 fri fri	260 7,1269199 15 02 fri ita
37 3,462443 15 01 ita ita	93 4,3397487 13 11 fri ita	149 4,9226157 12 03 fri ita	205 5,8249066 04 02 fri ita	261 7,1434355 14 13 ita fri
38 3,462651 15 12 ita fri	94 4,340485 14 01 ita ita	150 4,9342251 15 01 fri ita	206 5,826769 10 05 ita ita	262 7,1764371 13 10 fri fri
39 3,466536 03 02 ita fri	95 4,3701948 14 12 ita fri	151 4,9496505 13 03 fri fri	207 5,831728 13 03 fri ita	263 7,1788596 15 05 fri ita
40 3,477945 15 10 fri fri	96 4,3801404 15 01 fri fri	152 4,9549843 11 05 ita ita	208 5,8573992 14 04 ita fri	264 7,1790553 10 04 fri ita
41 3,4784 03 02 fri fri	97 4,3855506 12 02 ita ita	153 4,9656834 05 02 ita fri	209 5,8896757 13 06 ita fri	265 7,1796349 14 02 fri fri
42 3,495718 11 10 fri ita	98 4,4006929 06 04 fri ita	154 4,9690808 04 03 ita fri	210 5,8940702 12 06 ita fri	266 7,2294999 14 06 fri ita
43 3,512876 15 01 ita fri	99 4,4015751 05 02 fri ita	155 4,973564 05 03 fri fri	211 5,8951734 12 04 ita fri	267 7,3960997 15 13 fri fri
44 3,514671 12 10 ita ita	100 4,4024957 11 03 fri ita	156 5,0390311 04 03 ita ita	212 5,8982711 15 06 ita fri	268 7,448317 14 06 fri fri
45 3,517787 15 12 ita ita	101 4,4119675 03 01 ita ita	157 5,0427889 05 04 ita fri	213 5,9046595 10 04 ita fri	269 7,5160139 14 13 fri ita
46 3,553667 11 01 ita ita	102 4,4176631 14 12 ita ita	158 5,1060044 10 01 fri ita	214 5,9436431 04 01 ita ita	270 7,5248048 15 04 fri ita
47 3,571969 03 03 ita fri	103 4,4256535 12 02 fri fri	159 5,1211069 15 12 fri fri	215 5.9446123 04 02 ita fri	271 7,6451361 15 05 fri fri
48 3,600046 10 03 ita fri	104 4,4332542 13 03 ita fri	160 5,1819977 12 05 ita ita	216 5,9671681 13 10 fri ita	272 7,8957202 14 05 fri ita
49 3,617791 15 14 ita fri	105 4,4344999 15 02 ita fri	161 5,1918572 11 06 ita fri	217 5,9759363 10 05 ita fri	273 7,9040266 14 04 fri ita
50 3,630145 13 11 ita ita	106 4,436219 12 06 fri ita	162 5,2044056 13 01 fri ita	218 5.9791037 06 01 fri ita	274 8,1378226 14 02 fri ita
51 3,697113 11 02 ita fri	107 4,4473393 05 03 ita ita	163 5,2070669 13 03 ita ita	219 6,0119316 12 04 ita ita	275 8,5572492 14 13 fri fri
52 3,727361 15 11 ita ita	108 4,4528449 11 06 fri fri	164 5,230911 05 01 ita fri	220 6,0123772 15 02 fri fri	276 8,6101968 14 05 fri fri
53 3,738023 14 10 fri fri	109 4,4903786 06 05 ita fri	165 5,2343936 06 05 fri fri	221 6,0162025 14 06 ita ita	
54 3,738744 15 10 fri ita	110 4,4921758 11 02 fri ita	166 5,2581762 11 04 fri ita	222 6,034821 15 05 ita ita	
55 3,738908 11 06 fri ita	111 4,5103657 14 10 fri ita	167 5,2615851 11 05 fri fri	223 6,0429351 10 06 fri ita	
56 3,747244 12 01 fri fri	112 4,5144998 05 03 fri ita	168 5,2691976 14 12 fri fri	224 6,0695399 06 03 fri ita	

Figura 5. Distancias entre LTAS de las muestras. Se han evidenciado las líneas que contienen los resultados de la comparación de dos muestras del mismo hablante

Esto lleva a la conclusión que el LTAS, por sí solo y de acuerdo con el método de análisis descrito, no es un parámetro que permita comparar de forma segura muestras en lenguas distintas, por lo menos en el sentido *tradicional* de afirmar que dos muestras proceden de un mismo sujeto.

A pesar de esta limitación, el método que se ha presentado podría tener alguna utilidad práctica. Si se observa la tabla de la figura 5, en la cual los resultados de la comparación de los LTAS aparecen ordenados por distancia decreciente (es decir, desde los dos LTAS más parecidos hasta los dos más diferentes), se observa que –en general– los LTAS de muestras producidas por el mismo informante tienden a aparecer en posiciones altas del *ranking*. Concretamente, todos ellos se encuentran dentro del primer 20% de datos (primeros dos deciles) y la absoluta mayoría en el primer decil. Esta observación no permite afirmar que si la distancia entre los LTAS de dos muestras recae en este sector es razonable esperar que hayan sido producidas por el mismo hablante, pero seguramente sería posible afirmar lo contrario: si la distancia entre los LTAS de dos muestras el lenguas distintas recae en los últimos ocho deciles, se podría concluir que el análisis del LTAS depone en

contra de la hipótesis de la fiscalía (es decir que la muestra dubitada y la indubitada provienen del mismo hablante).

5 Conclusiones

Este trabajo, tal como se ha expresado en el apartado 2, tenía dos objetivos, de los cuales uno de carácter más teórico y el otro de tipo metodológicoaplicado. En cuanto a la pregunta de investigación teórica, se ha tenido que rechazar la hipótesis de trabajo, es decir que la diferencia entre los LTAS de las dos muestras del mismo sujeto hablando lenguas distintas debería ser inferior a la diferencia entre cada una de las esas muestras y cualquier otra muestra de sujetos distintos. Ya que esta hipótesis se ha rechazado, se ha tenido que concluir que el LTAS no parece un parámetro determinante en la comparación de muestras en lenguas distintas a la hora de avalar la hipótesis de la fiscalía, ni siquiera en condiciones experimentales ideales como las que se han utilizado (grabaciones de alta calidad, efectuadas en una sola sesión, de un número no especialmente grande de hablantes). No obstante, se ha observado como los datos proporcionados por la comparación automática de los LTAS, a pesar de no poder constituir un elemento a favor de la hipótesis de la fiscalía, podrían restarle credibilidad.

En cuanto al objetivo metodológico-aplicado, se ha podido definir un protocolo de análisis basado en un script de Praat que permite calcular automáticamente las distancias entre los LTAS de varias muestras. El script en cuestión, además de ser aplicable al caso de estudio concreto que se ha presentado, se puede utilizar en cualquier otra situación en que se quiera utilizar el LTAS como parámetro en la comparación de voces.

Determinar la fuerza de esta evidencia contraria (es decir cuánto en contra de la hipótesis de la fiscalía están los datos del LTAS) y determinar la manera más adecuada para expresar esta conclusión constituyen un tema que transciende las finalidades de este trabajo (v. Gil (2014) para una síntesis del debate sobre la presentación de los resultados de las comparaciones judiciales de muestras de habla).

6 Bibliografía

- Boersma, P., & Kovačić, G. (2006). Spectral characteristics of three styles of Croatian folk singing. Journal of the Acoustical Society of America, 119, 1805-1816.
- Boersma, P., & Weenink, D. (2013). Praat: doing phonetics by compute. Version 5.3.57. Descargado el 27 de octubre de 2013, de http://www.praat.org/
- Bruyninckx, M., Harmegnies, B., Llisterri, J., & Poch, D. (1994). Language-induced voice quality variability in bilinguals. Journal of Phonetics, 22, 19-31.
- Cicres, J. (2011). Los sonidos fricativos sordos y sus implicaciones forenses. Estudios Filológicos, 48, 33-48.
- Elvira García, W., Roseano, P., & Fernández Planas, A. M. (2014). Ltas differences. Praat script. Descargado el 23 de setiembre de 2013, de http://stel.ub.edu/labfon/en
- Gil, J. (2014). Más allá del "efecto CSI": avances y metas en fonética judicial. En Y. Congosto, M. L. Montero Curiel, & A. Salvador (eds.), Fonética Experimental, Educación Superior e Investigación. Madrid: Arco Libros, vol. I, 63-112.
- Harmegnies, B., Bruyninckx, M., Llisterri, J., & Poch, D. (1991). Effects of language change in voice quality in bilingual speakers. Corpus content effects. En Eurospeech 1991. Proceedings of the 2nd European conference onspeech communication and technology. Genova, Italy. 24-26 September, 1991. Vol. 1, 165-8.
- Klingholz, F., Penning, R., & Liebhart, E. (1988). Recognition of low-level of alcohol intoxication from speech signal. Journal of the Acoustical Society of America, 84(3), 929-935.
- Künzel, H. (2011). La prueba de voz en la investigación criminalística. Ciencia

- Forense, INACIPE-Academia Iberoamericana de Criminalística y Estudios Forenses, 1(1), 37-50.
- Majewski, W., & Hollien, H. (1974). Euclidean distance between longterm speech spectra as a criterion for speaker identification. Speech Communication Seminar, 1-3.
- Manwa, L. Ng, Chen, Y., & Chan, E. Y. K. (2010). Cantonese and English Produced by Proficient Cantonese-English Bilingual Speakers A Long-Term Average Spectral Analysis. Journal of Voice, 26(4), 171-176.
- Molina de Figueiredo, R., & Bernales, M. (1999). Reconocimiento de hablantes basado en el espectro a largo tiempo. En Actas del VI Simposio Internacional de Comunicación Social. Santiago de Cuba: Centro de Lingüística Aplicada, 1372-1378.
- Nolan, F. J. (1983). The Phonetic Bases of Speaker Recognition. Cambridge: Cambridge University Press.
- Nordenberg, M., & Sundberg, J. (2003). Effect on LTAS of vocal loudness variation. TMH Quarterly Progress Status Report, 1/2001.
- Pittam, J. (1987). The long term spectral measurement of voice quality as a social and personality marker: A review. Language and Speech, 30, 1-12
- Pruzansky, S. (1963). Pattern-matching Procedure for Automatic Talker Recognition. Journal of the Acoustical Society of America, 35, 354-358.
- Zalewski, J., Majewski, W., & Hollien, H. (1975). Cross-correlation between Long-Term Speech Spectra as a criterion for speaker identification. Acoustics, 34, 20-24.

7 Agradecimientos

Queremos expresar nuestro agradecimiento al Institut für Romanistik de la Universität Hamburg, que ha puesto a disposición las grabaciones, y la Dra. Machuca de la Universitat Autònoma de Barcelona, puesto que gracias a una conversación mantenida con ella en el marco del posgrado en Estudios Fónicos del CSIC surgió la primera idea para este trabajo.

Parte V. Fonética normativa

La fonètica en la gramàtica d'una institució normativitzadora (GNV, 2006)

SARAGOSSÀ, ABELARD

1 Introducció

Les gramàtiques de les institucions normativitzadores haurien d'estar elaborades pensant en el públic general, ja que la seua finalitat és contribuir a la formació d'una llengua pública adequada per a la societat a què l'obra va destinada. La gramàtica de l'Acadèmia Valenciana de la Llengua (*Gramàtica Normativa Valenciana*, 2006, (*GNV*)) està dins d'eixe marc, vist que comunica que "pot interessar al conjunt de ciutadans valencians, tant aquells que la usen habitualment i procuren expressar-se de la millor manera possible com aquells que desitgen aprendre-la" (*GNV*, 2006, §1.1).

Els destinataris de les obres que comentem condicionen la manera d'elaborar-les. Si una gramàtica vol que el públic puga entendre-la, haurà de procurar ser rigorosa, exigència general que podem concretar en quatre mesures: 1) Usar tan poca terminologia teòrica com siga possible. 2) Recórrer a un sol terme per a cada concepte. 3) Explicar als lectors cada terme teòric usat per a estudiar la llengua. 4) Evitar les anomalies metodològiques.

L'article té dos finalitats. La primera és mostrar que, en unes poques pàgines, es poden explicar els conceptes fonamentals de la fonètica: fonètica, fonema, fonologia i entonació; síl·laba: vocal i consonants; llengües amb tres, cinc o set vocals; consonants bàsiques (amb soroll): oclusives, fricatives i africades; sonores; consonants secundàries (sense soroll; necessàriament sonores): nasals, líquides; diftong i hiat; al·lòfon.

La segona finalitat del treball és demostrar que els lectors de la *GNV* (2006) no poden entendre la fonètica d'eixa gramàtica perquè el manual actua d'una manera poc ordenada, incorre sovint en anomalies metodològiques i, en els

pocs conceptes lingüístics que caracteritza, fa propostes poc coherents. Seria poc adequat presentar en un congrés de lingüística l'anàlisi d'una gramàtica poc significativa. Però la *GNV* (2006) hauria de representar el valencià i hauria d'aspirar a ser ben rebuda en moltes cases valencianes. Justament per la responsabilitat enorme que té l'Acadèmia Valenciana de la Llengua, opine que els lingüistes valencians tenim l'obligació de mostrar les limitacions de la *GNV* (2006) a fi que la institució les supere en una obra posterior.

2 Fonaments de la fonètica (i de l'ortografia)

2.1 Fonètica, fonema, fonologia i entonació. Ortografia

El tractament següent dels conceptes bàsics de la fonètica està elaborat a partir d'una gramàtica anterior a la GNV (Saragossà, 2003, §1.6.3, §5, §7.1.2) i de dos articles posteriors (Saragossà, 2009, 2010). La fonètica és una part de la teoria lingüística. Estudia els mitjans sonors que permeten exterioritzar l'ús del llenguatge, de manera que el pensament de l'emissor arribe al receptor. Quan parlem, es produïx una tira fonètica, que podem segmentar, allargar i acurtar (baixar / bai + xar / b + a + i + x + a + r; acurtar: baixà; allargar: $baixar\acute{e}$). Cada segment mínim de la tira és un fonema. Els fonemes són sons no segmentables que, en una llengua determinada, aprofiten per a expressar paraules i poden diferenciar-ne ($\{poc \neq pot\}$, $\{poc \neq puc\}$, $\{poc \neq foc\}$).

L'estudi dels fonemes constituïx una part de la fonètica, la fonologia. Tracta la formació dels fonemes (e, b, s), com els combina l'emissor quan parla i com els percep el receptor (b + e + s = bes). Quan parlem, no diem els fonemes en la mateixa intensitat, agudesa i duració (cf. les dos as de anà). L'entonació (o prosòdia) és la part de la fonètica que estudia el contrast entre els fonemes de la tira fonètica des de les tres propietats dites. El fonema és la unitat de la fonologia (producció i combinació) i de l'entonació (contrast).

L'ortografia és la representació gràfica de la fonètica. Conté dos classes de signes: els que representen els fonemes (les lletres) i els que indiquen l'entonació (signes de puntuació).

La fonètica anava al final de les gramàtiques (Sempere, 1560; Ballot, 1813; Amengual, 1835; Bofarull & Blanch, 1887; etc.). En canvi, en el segle XX sol anar al principi (Fabra, 1912, 1918; Marvà, 1932; etc.). L'actuació històrica és coherent, ja que la definició dels conceptes fonètics bàsics demana nocions de les altres parts de la lingüística (cf. *fonema*).

2.2 Síl·laba: vocal i consonants. Vocals: obertura i sonoritat

Quan parlem, unim els fonemes en síl·labes, que és la construcció que ordena els fonemes per a poder expressar les paraules. En cada síl·laba, hi ha necessàriament una vocal (o nucli de la síl·laba), i pot aparéixer una consonant o més d'una. La vocal és el fonema més perceptible de la síl·laba; tots els altres són consonants. Així, el fonema o de to és molt més perceptible que el fonema t. La perceptibilitat alta justifica que la vocal siga obligatòria. Les consonants són opcionals, ja que modifiquen la percepció de la vocal, com podem comprovar comparant paraules amb la mateixa vocal: to / do / bo / so. La centralitat de la vocal justifica la impressió de la síl·laba com a "colp de veu". Hi han paraules que només tenen una vocal (a, o, i); una vocal i una consonant (pi / os); dos

consonants (nas, gra, arc); tres (groc, pont). Alguna en té quatre (forns) i cinc (troncs).

Quan una llengua té tres, cinc o set vocals, són fàcils d'explicar. Són tan perceptibles perquè sempre vibren les cordes vocals (que són un múscul que tenim en la gola). Els lingüistes anomenen eixa característica com a sonoritat. Modulem el so que resulta amb l'obertura de la boca i la posició de la llengua. La vocal a és la més oberta; la llengua està plana. Si tanquem molt la boca i avancem la llengua, obtenim la i. Si la retrocedim, arribem a la u. Entre la a i la i, hi han llengües que tenen una vocal (com el castellà, la e) i n'hi han que en tenen dos (la e i la e, com el valencià i les llengües romàniques en general). Si una llengua té tres vocals anteriors (e, e i e), en sol tindre tres posteriors (e, e i e). En conjunt, totes les vocals tenen un grau d'obertura (entre la e i les vocals e i e). Per contra, totes les consonants són tancades (també en un grau variable, com comprovarem tot seguit). En resum, les propietats bàsiques de les vocals són l'obertura i la sonoritat.

2.3 Consonants bàsiques (*soroll*) i secundàries (acostament a les vocals)

Les consonants partixen dels valors contraris a les vocals. Les cordes vocals no vibren, de manera que les consonants resulten del *soroll* que fa l'aire quan topa amb un obstacle en la boca. El terme *soroll* és propi de la lingüística històrica (com ara, apareix en el DGLC, *vocal* i *consonant*). Els dos obstacles bàsics són previsibles: tancar la boca i acostar dos parts. Quan tanquem la boca en algun punt, acumulem aire i l'obrim, es produïx una explosió. Podem tancar la boca amb els llavis (*pa*), amb la punta de la llengua i les dents superiors (*tu*) i amb la part posterior de la llengua i el final del paladar (*ca*, 'gos' en balear). El nom d'eixes consonants és *oclusives* (de *cloure*, tancar). Poden ser bilabials (*pa*), dentals (*tu*) i velars (*ca*). Les oclusives són momentànies (i, per tant, no allargables).

Si acostem dos parts de la boca, l'aire xiula quan passa per la zona estreta. Si toquem amb les dents superiors el llavi inferior, obtenim fi. Si acostem la punta de la llengua a les genives, ix so. Si és la part anterior de la llengua i la part anterior del paladar, resulta ix. Com que hi ha un fregament o fricció, usem el nom consonants fricatives. Poden ser labiodentals (fi), alveolars (del nom de les genives en grec, so) i palatals (ix). Les consonants fricatives són duratives (i, per tant, allargables, ixxx...).

Hi ha una tercera classe d'obstacle, que és una combinació dels dos anteriors. Tanquem en algun punt el conducte bucal, però quan l'obrim no hi ha una explosió, sinó una fricció momentània (no allargable). Això passa en la consonant final de *despatx*; o en el plural de *tot* (*tots*). Les consonants africades poden ser palatals (*despatx*) i alveolars (*tots*).

Les consonants descrites són les bàsiques: les més separades de les vocals. Des d'eixa posició, hi ha un acostament progressiu a les característiques de les vocals. La primera fase consistix en afegir sonoritat a les consonants descrites, propietat que fa passar de to a do, de pes a bes, de cas a gas. Un exemple de les fricatives és fi contra vi. En les africades, tots contrasta amb dotze.

Sabem que, en les consonants bàsiques, hi ha un soroll. Però, si tanquem la boca i l'aire ix pel nas, l'obstacle ja no produïx soroll: *mà*. En eixes consonants (que es diuen *nasals*), només hi ha la sonoritat de les cordes vocals i una modulació de la sonoritat (com en les vocals). Per la falta de soroll, totes les

consonants secundàries són sonores. Les nasals poden ser bilabial (ham), alveolar (son) i palatal (any).

Les consonants que queden són les *líquides* (*l, ll, r, rr*). Es diuen així perquè ja estan molt prop de les vocals: el so que fan ix (o *fluïx*) com si fóra un *líquid*. Si posem la punta de la llengua en les genives i l'aire ix pels costats de la llengua, obtenim la paraula *el*. Si toquem la part anterior del paladar, ix *ell*. Finalment, tenim les consonants de *car* (la punta de la llengua s'esvara ràpidament per les genives) i de *carro* (repetim el moviment tres o quatre voltes). El nom històric és *vibrant simple* (*car*) i *vibrant múltiple* (*carro*). Les consonants líquides són les úniques que poden aparéixer entre una consonant i una vocal (*plou, prou*). Es poden vocalitzar (*alba* > *auba* en balear i en provençal).

2.4 Els conceptes diftong i hiat. El concepte al·lòfon

Podríem dir que la vocal més vocàlica és la que té més sonoritat: la que és més oberta (la *a*). En canvi, a mesura que es reduïx l'obertura de la boca ens acostem cap a les consonants. Això explica que les vocals més tancades (la *i* i la *u*) poden actuar com a nucli de síl·laba i com a consonants. Comprovarem l'afirmació anterior si comparem *veí* (dos síl·labes) i *mai* (una síl·laba). En *veí*, el fonema *i* actua com a vocal; en canvi, en *mai* funciona com a consonant. El mateix resultat obtindrem si comparem les sabates de llana de les criatures (*peücs*: dos síl·labes) i el nom del final de les cames (*peus*: una síl·laba).

Les dualitats descrites permeten definir els conceptes diftong i hiat. Tenim un hiat quan estan en contacte dos vocals una de les quals podria funcionar com a consonant (veí, peüc). Correlativament, parlem de diftong quan en una síl·laba estan en contacte una vocal i un fonema que, encara que pot funcionar com a vocal, en eixa síl·laba funciona com a consonant (mai, peu). Si hi ha un fonema davant i u darrere (aguaitar), parlem de triftong (en l'exemple, uai). Hi han llengües en què les consonants més pròximes a les vocals (les líquides) també poden actuar com a vocals. Així, la paraula dit és prst en xec.

Convé saber que els fonemes no els pronunciem sempre igual. U dels factors més influents és el fonema que va a continuació. Així, si un fonetiste diu aïllada la consonant nasal de *pren* i la de *prenc* comprovarem que sonen d'una manera prou diferent (la nasal de *prenc* no la pronunciem en les genives, sinó al final del paladar, com la c). Això no obstant, usem la mateixa lletra en els dos casos perquè la variant de *prenc* no aprofita per a diferenciar paraules. Les variacions d'un fonema s'anomenen *al·lòfons*. Els parlants no solem ser conscients de les diferències entre els al·lòfons d'una fonema, encara que siguen acusades (com en *pren* i *prenc*). En el cas de les vocals més tancades, no cal dir que en *veí* i *rei* hi ha sempre el mateix fonema, *i*. Els fonetistes diuen que hi han diferències en l'obertura (la *i* vocàlica seria més oberta que la *i* consonàntica), però els parlants no som conscients d'eixa particularitat: només sabem que en *veí* diem dos síl·labes; en *rei*, una.

Hi han llengües que no tenen síl·laba tònica (com ara el llatí clàssic o el basc històric). Les que en tenen, pronuncien una síl·laba d'una paraula en més intensitat, agudesa i duració que les altres (compareu pensàrem amb pensarem). La finalitat de la síl·laba tònica és facilitar al receptor que trobe les paraules que li diu l'emissor (sempre que sent una síl·laba tònica, sap que al voltant hi ha una paraula). Els estudiants aprenen a trobar la síl·laba tònica prou fàcilment, encara que hi haja un diftong (des-mai, hie-na, euro-peu, qua-tre). No és tampoc difícil separar el hiat quan una síl·laba és tònica (ve-í, pe-üc).

3 Anàlisi de la fonètica de la GNV (2006)

3.1 Fonètica i ortografia: separació i alhora mescla

Després de mostrar que en un espai curt podem explicar els conceptes fonètics necessaris en una obra divulgativa, analitzarem com actua la GNV (2006). Com molts manuals, tracta primer la fonètica ("ortologia", capítol 1, p. 20-35: 15 pàgines; 27%) i després l'ortografia (capítols 2-7, p. 36-77: 41 pàgines; 73%), ja que els conceptes ortogràfics bàsics es definixen a partir dels fonètics. Contràriament a Sanchis Guarner (1950), la GNV dedica prou més extensió a l'ortografia (el 73%; Sanchis: el 40%). Per la raó ja dita (l'ortografia és una representació de la fonètica), la fonètica és més significativa.

Contradictòriament amb la separació entre la fonètica I l'ortografia, la GNV dedica el segon capítol de l'ortografia a dos conceptes fonètics, la síl·laba i el diftong (1c):

- 1. El segon capítol de l'ortografia tracta conceptes fonètics (*c*)
 - a. Capítols 2-7: "ortografia" (p. 36-77)
 - b. Capítol 2: "L'alfabet" (p. 36-37)
 - c. Capítol 3: "L'agrupament dels sons" (p. 38-39).
 - i. "La síl·laba (§3.1). Conceptes generals (§3.1.1)"
 - ii. "Els diftongs (§3.2). Conceptes bàsics (§3.2.1)"
 - d. Capítol 4: "Les lletres" (p. 40-46).

Tot seguit, analitzarem la fonètica, amb inclusió del capítol posat en l'ortografia (1c).

3.2 Vocals: estructures i sistemes sense cap constituent definit

El manual tracta "l'estructura del vocalisme tònic" (2a). Tanmateix, els lectors no troben cap "estructura" per la raó que el manual no definix les vocals. Només diu quantes vocals té el valencià (set), les representa amb l'alfabet fonètic internacional i les exemplifica:

- 2. Contingut de la fonètica de la *GNV* (2006): les vocals
 - a. "Les vocals (§1.2) Les vocals tòniques (§1.2.1) Estructura del vocalisme tònic (§1.2.1.1)"
 - i. "El valencià té set vocals en posició tònica:
 - ii. "[símbols de l'alfabet fonètic internacional i exemples]"
 - b. "Les vocals àtones. Estructura del vocalisme àton (§1.2.2.1)"
 - i. El sistema vocàlic àton del valencià es reduïx a cinc vocals:
 - ii. [símbols de l'alfabet fonètic internacional i exemples]
 - iii. Les vocals àtones [i: mirar, oli] i [u: unflar, ímpetu] tenen com a correlat no vocàlic les semivocals [j] i [w], respectivament, que formen diftong amb la vocal precedent o següent: [j] hiena [...] [w] pasqua [...]"

Si passem a les vocals àtones (2b), tornem a trobar una "estructura", reforçada ara per un "sistema" (2b.i). Però, a pesar dels títols, la GNV seguix sense definir cap vocal. La terminologia usada ("estructures, sistemes") impulsa a pensar en un treball científic, encara que la sensació no va acompanyada

d'una propietat inherent de la ciència: comprendre l'objecte d'estudi (ara, entendre quina estructura formen les vocals).

En 2b, també hi han anomalies metodològiques. Parlant sobre les vocals àtones, apareix el concepte *semivocal*. Això no obstant, el manual estudiarà eixe concepte dos capítols més avant (en l'ortografia). Eixa irregularitat teòrica té efectes negatius en els lectors: no poder comprendre l'afirmació que "les vocals àtones [i] i [u]" tindrien "com a correlat no vocàlic les semivocals [j] i [w]" (2b.iii). Certament, la *GNV* usa símbols diferents ("[i] i [u]" contra "[j] i [w]"); però el manual no ha explicat quin valor representa cada símbol. L'absència de teoria fa que els símbols arriben als ulls dels lectors, però no a l'enteniment.

Per una altra banda, si els parlants no perceben la diferència entre dos al·lòfons tan diferents com són els de *pren* i *prenc*, no notaran tampoc la diferència entre els dos al·lòfons pròxim de *veí* i *rei*, de *peüc* i *peu* (§2.4). Ens trobem davant d'un altre factor que fa pensar als lectors en la ciència (ara, els símbols "[i], [j]), [u], [w]"), a pesar que de ciència o teoria no n'hem trobat. La veritat és que, si la GNV (2006) haguera tingut més en compte la gramàtica de Sanchis Guarner (1950, §38), potser no hauria usat l'hipotètic concepte *semivocal*, que Sanchis criticà obertament: "La diferència entre vocals i consonants no és específica sinó de grau, i així, a vegades, els mateixos fonemes vocals tenen funció de consonants i alguns, amb poca precisió, els anomenen semiconsonants".

Com que la tonicitat es referix a la síl·laba, resultarà que l'expressió *vocal tònica* (2a) comporta el concepte *síl·laba*, que la *GNV* tracta posteriorment (en l'ortografia).

Resumim. La GNV no definix el concepte *vocal*. Tampoc explica els conceptes *fonètica, fonologia* i *entonació*. Podríem pensar que l'objectiu és no tractar conceptes sobre els quals els lectors tenen alguna idea. Però, si això fóra cert, no s'entendria que la *GNV* usara térmens que fan pensar en un manual teòric (*estructura, sistema*); o que use símbols de l'alfabet fonètic internacional (sovint innecessàriament: [rej], [kwatre]); o que faça afirmacions que els lectors no poden comprendre (i, probablement, tampoc els membres de l'Acadèmia Valenciana de la Llengua), com ara que en *rei* o en *seu* hi hauria una "semivocal". A més, l'espai que el manual dedica a la teoria no és escàs, com comprovarem en l'estudi de les consonants (§3.3) i de la síl·laba (§3.4).

Seguint avant en l'anàlisi de la *GNV*, hem de constatar que el manual no usa el concepte *fonema*. Eixa noció és cabdal no solament per a comprendre el sistema fonètic d'una llengua, sinó també per a entendre què és una ortografia regular. Així, correspon una lletra als fonemes, no als al·lòfons. Eixa propietat justifica que grafiem igual *pren* i *prenc*; o que siguen regulars les grafies *rei* i *seu* i, correlativament, irregulars les de *rey* i *sew*. L'observació feta explica els dubtes que generen els signes gràfics *y* i *w* en castellà, en francés o en anglés (més que *lletres*, podríem dir que són *al·lògrafs*). Cal dir que Sanchis Guarner (1950, §33) usava el terme *fonema*, el diferenciava de *so* i el definí: "SONS I FONEMES. Aquells sons dels quals, en articular-los, el subjecte té consciència de llur valor específic, comprensiu pel que els escolta, s'anomenen 'fonemes'". La substitució de *fonema* per *so* en la *GNV* (2006) no deu ser un error, sinó una conseqüència de l'actuació de la *Gramàtica del català contemporani*, que també recorre sempre al *so* ("Els sons del català. Els sons vocàlics. Els sons consonàtics", etc., Julià, 2003, §1, §1.2.1, §1.2.2).

3.3 Consonants: explicació poc ordenada, poc profunda i sense exemples

La GNV no definix *vocal* ni *consonant*. En canvi, l'absència d'explicar l'articulació de les vocals contrasta amb la caracterització de de les consonants:

- 3. GNV (2006: §1.3.1): "Les consonants. Estructura del consonantisme"
 - a. "Les consonants es poden classificar a partir dels trets relacionats amb el lloc i el mode d'articulació. El lloc d'articulació és el punt del tracte vocal on es produïx l'obstrucció o la constricció de l'eixida de l'aire expirat, i permet classificar les consonants en bilabials, labiodentals, dentals, alveolars, palatals i velars.
 - b. OBSERVACIÓ: Les consonants s'articulen per mitjà del contacte o l'acostament de dos òrgans articuladors: els dos llavis (bilabials); el llavi inferior i les dents incisives superiors (labiodentals); l'àpex de la llengua i la cara interior de les dents incisives superiors (dentals); l'àpex o la làmina de la llengua i els alvèols (alveolars); el predors o la làmina de la llengua i el paladar dur (palatals), i el dors de la llengua i el vel del paladar (velars).
 - c. D'altra banda, el mode d'articulació permet classificar les consonants segons el tipus d'obstacle que troba l'aire expirat i segons la manera com l'aire passa per la glotis:
 - d. Atenent el tipus d'obstacle que troba l'aire expirat, les consonants es poden classificar en *oclusives*, *fricatives*, *africades*, *ròtiques* (que, al seu torn, poden ser *bategants* i *vibrants*), *laterals* i *nasals*.
 - e. OBSERVACIÓ: Les oclusives s'articulen amb una obstrucció total de l'aire seguida d'una explosió; les *fricatives*, en canvi, amb una constricció provocada per l'acostament de dos òrgans articuladors; les africades, combinant un primer moment oclusiu i un segon de fricatiu; les ròtiques, per mitjà d'un contacte (bategants) o diferents contactes (vibrants) de l'àpex i els alvèols; les laterals, per mitjà del contacte de la part central de la llengua i la cavitat oral superior, que crea un o dos canals laterals per on pot eixir l'aire, i les nasals, d'una banda, per mitjà d'un tancament total del conducte oral i, d'una altra, d'un abaixament del vel del paladar, que provoca que el passatge cap al nas quede obert, de manera que l'aire pot eixir pel nas, però no per la boca.
 - f. Atenent la manera com l'aire passa per la glotis, les consonants poden ser *sordes* o *sonores*: les *sordes* s'articulen sense vibració de les cordes vocals, i les *sonores*, amb vibració.
 - g. Així, la combinació dels dos trets permet una classificació global del conjunt de consonants, tal com queden reflectides en la taula següent: [...]" [en el quadre, apareixen unes "variants contextuals" de les consonants oclusives sonores, noció que la part teòrica no ha tractat].

L'absència de definir què la vocal i les consonants té efectes negatius en l'explicació de les consonants. Així, l'oposició amb les vocals (sonoritat i obertura, sense cap obstacle) impulsa a tractar les consonants per la classe d'obstacle, el lloc de la boca en què es produïx i quin soroll efectua. En canvi, els lectors de 3 es troben de sobte amb un "lloc i mode d'articulació" sense saber per què apareix eixa noció (3a). L'estructuració teòrica escassa facilita que la *GNV* tracte primer els punts d'articulació (bilabial, etc., 3a) i en acabant la manera d'articular les consonants (3c-d). Eixa actuació fa descriure els punts d'articulació sense poder posar exemples (3b). En canvi, si comencem per la classe d'obstacle exemplificarem cas a cas, des del principi (com hem fet en §2.3).

Convé afegir que no hi han tampoc exemples quan el manual explica les classes d'obstacle (3e). Eixa absència fa pensar en un interés pedagògic poc alt, i també en un rigor no fort. Un treball científic deu exemplificar cada propietat

per dos raons, una científica i l'altra pedagògica. La primera finalitat de la teoria és explicar dades empíriques. A més, l'aplicació a exemples mostra que la teoria és bona; i la verificació facilita l'assimilació dels lectors: hem arribat a la pedagogia. Els escrits que parlen i parlen sense exemplificar (com el de 3) solen tindre deficiències teòriques. En (3g), n'he indicat una: la *GNV* posa en el quadre final unes "variants contextuals" de les consonants oclusives sonores, noció que la part teòrica no ha tractat. El fet de pronunciar *cabre* amb una consonant bilabial sonora contínua (sense oclusió) és una de les manifestacions del concepte *al·lòfon*. Hem arribat a una anomalia metodològica doble: no explicar un concepte (*al·lòfon*) i donar una manifestació sense explicar-la tampoc (i sense exemplificar-la); la GNV es limita a posar en el quadre el símbol que la *b* contínua té en l'alfabet fonètic internacional.

En la caracterització de les consonants fricatives i les africades, falta la propietat que les separa: les fricatives són prolongables i les africades són com les oclusives, momentànies (3e). Les africades tenen propietats comunes amb les oclusives i les fricatives; però també propietats diferencials. No són com les oclusives perquè no tenen explosió; però tampoc són com les fricatives perquè són momentànies. En l'explicació de l'obstacle (3a-c), la GNV (2006) no recorre al soroll, que separa les consonants bàsiques de les secundàries. Eixa noció és pròpia de la lingüística històrica: "consonant adj. [...] | | f. Fonema que consisteix en un soroll produït a la boca (o a la gola) per una oclusió o fort estrenyiment del canal bucal, tot sol o acompanyat de la veu produïda en la laringe" (DGLC). En la GNV, l'efecte de no usar el soroll és agrupar les consonants bàsiques (les oclusives, les fricatives i les africades) amb les secundàries (les nasals i les líquides, que la GNV posa en l'orde contrari, 3a-c). Eixa actuació no mostra com ens acostem a les propietats de les vocals i impedix explicar que les nasals i les líquides han de ser sonores. De fet, la descripció de la sonoritat (3f) fa pensar que totes les consonants podrien ser sordes i sonores.

Ací, deixarem sense analitzar anomalies menors (com ara no usar el nom consonant líquida, que té conseqüències en la síl·laba: plou, prou; o substituir les expressions vibrant simple i vibrant múltiple per ròtic, bategant i vibrant, 3e).

3.4 Conceptes fonètics dins de l'ortografia (§3.1, 1b)

La *GNV* (2006) separa la fonètica i l'ortografia, però estudia tres conceptes fonètics (síl·laba, diftong i hiat) dins de l'ortografia. La síl·laba és bàsica per a caracteritzar les dos nocions que la *GNV* usa en la fonètica (vocal i consonant). Això significa que estem davant d'una altra anomalia metodològica (tractar en l'ortografia un concepte indispensable per a entendre els dos conceptes considerats en la fonètica, *vocal* i *consonant*):

- 4. Capítol tercer de l'ortografia de la GNV (2006): "L'agrupament de sons (§3). La síl·laba (§3.1). Conceptes generals (§3.1.1)"
 - a. "Les vocals i les consonants s'agrupen en síl·labes, susceptibles d'articular-se en una sola emissió fònica.
 - b. Tota síl·laba està formada necessàriament per una vocal, que funciona com a nucli, i opcionalment per una o diverses consonants.
 - c. Les síl·labes poden ser lliures o travades: les lliures acaben en vocal (com les de la paraula *a-me-ri-ca-na*) i les travades en consonant (com les de la paraula *per-mís*)".

La definició "una sola emissió fònica" és una variant de la definició intuïtiva "un colp de veu", que hem intentat justificar més amunt (§2.2). Convé notar que

Moll (1937, §14), Sanchis Guarner (1950) i el DCVB havien actuat amb l'enfocament adequat: unint *síl·laba*, *vocal* i *consonant*, i definint les vocals i les consonants pel grau de sonoritat:

5. "En l'articulació de les vocals, la disposició dels òrgans forma una obertura prou gran per tal que l'aire ixca sense obstacle, resultant el so de gran perceptibilitat i capaç de servir de nucli a la síl·laba. És menor, en canvi, la caixa de ressonància que es forma en articular les consonants, les quals resulten menys perceptibles i han d'anar subordinades a les vocals en la formació de les síl·labes." (Sanchis Guarner, 1950, §38 i §39)

Tornant a la *GNV*, u es pregunta quins "conceptes generals" (en plural) hi han en 4. L'envoltori del manual fa pensar en la teoria; però, en el contingut de 4, no n'hi ha.

En una llengua, és tan important la combinació dels fonemes en la síl·laba com l'inventari de fonemes. El valencià té una síl·laba molt consonàntica, la qual la llengua viva (sobretot entre Sagunt, Llíria i Xàtiva) manté molt bé. Tenint en compte això, ens hem de preguntar què diu la *GNV* sobre l'estructura sil·làbica del valencià. No he trobat res, en contrast amb la dedicació al tema que vaig fer en Saragossà (2003, §5.3.8: L'estructura sil·làbica del valencià; §5.3.9: Contrast entre les síl·labes del valencià i del castellà).

El segon concepte fonètic que la *GNV* (2006) tracta en l'ortografia és *diftong*. Eixa noció és un efecte de l'acostament progressiu entre les vocals i les consonants (les dos vocals més tancades estan prop de les consonants i, com a conseqüència, és factible usar-les com a consonants). Però, com que no hem trobat cap d'eixos trets en el tractament de les vocals i de les consonants, serà ben difícil que la GNV vaja ben orientada:

- 6. Diftong en la GNV (2006): "Els diftongs (§3.2), Conceptes bàsics (§3.2.1)"
 - a. "S'anomena diftong l'agrupament d'una vocal i una semivocal pertanyents a una mateixa síl·laba, com en ai-re, io-gurt, cau-re i qua-tre. Es parla de triftong en aquells casos en què una vocal està precedida i seguida d'una semivocal dins d'una mateixa síl·laba, com en cre-ieu i guai-tar. Per contra, la successió de dos vocals pertanyents a síl·labes diferents constituïx un hiat, com en lli-bre-ri-a i ma-nu-al."

És fàcil constatar que la GNV vol definir un concepte (diftong) recorren a un concepte lingüístic que no ha definit (semivocal). L'efecte inevitable d'eixa anomalia metodològica és que els lectors no poden comprendre la definició. Si la GNV haguera tingut en compte el tractament de Sanchis, no solament hauria definit bé la síl·laba, sinó que hauria evitat un terme ("semivocal") tan corrent com poc definit (d'una manera aplicable).

4 Conclusions

Els conceptes de la fonètica no són difícils d'explicar si tenim en compte els predecessors (Moll, 1937; Sanchis Guarner, 1950; DGLC, DCVB) i apliquem el sentit comú que tenim com a parlants. La gramàtica de l'Acadèmia Valenciana de la Llengua hauria de ser un manual clar, que descriguera les estructures lingüístiques del valencià a fi de mostrar la part bàsica de la llengua i, alhora, donar confiança als valencians en la seua forma de parlar. Perquè, contra l'opinió predominant, en la fonètica i en la flexió, en la semàntica i en la sintaxi, el valencià popular aplica a la comunicació la immensa major part de les estructures lingüístiques que el caracteritzen. Certament, la institució normativitzadora també deu tractar les vacil·lacions i els problemes. Però la

incorporació d'eixos dos camps hauria d'anar al costat del bàsic, que és explicar les estructures lingüístiques del valencià.

Per a poder desplegar les operacions dites, una gramàtica necessita usar terminologia teòrica, que deu explicar als lectors a fi que entenguen el manual. En la part de la teoria lingüística que hem tractat, la GNV s'encara molt poc a eixe objectiu (només en les consonants, que tracta d'una manera sensiblement deficient, §3.3, 3). El resultat és que una gran part de la terminologia teòrica no està definida (fonètica, fonema o "so", vocal, consonant, "semivocal"...). En altres treballs (embastats o en curs de publicació), mostraré que la teoria és un bé escàs en la GNV (2006). Per contra, el desorde i les anomalies metodològiques són frequents (n'hem trobat en cada apartat, §3.1, §3.2, §3.3, §3.4). A més, els títols, la terminologia i els símbols usats fan pensar en un manual teòric: "estructura del vocalisme tònic", "el sistema vocalic aton", "La síl·laba. Conceptes generals", "Els diftongs. Conceptes bàsics", ús de símbols fonètics com si equivalgueren a teoria. La unió d'eixos tres factors (poca teoria, quantitat remarcable d'irregularitats metodològiques i aparença teòrica) fa que per a un lector valencià siga difícil transitar pel manual, en contra de la voluntat de l'Acadèmica Valenciana de la Llengua (§1).

També cal destacar la poca vinculació que hem trobat als predecessors (Fabra, Moll i sobretot Sanchis Guarner, §3.2, §3.3, §3.4). Una de les obligacions de l'Acadèmia Valencina de la Llengua hauria de ser estudiar bé els gramàtics valencians i procurar incorporar a la seua gramàtica totes les aportacions que han fet. La tendència que hem trobat és més aïna la contrària a l'arrelament històric. Així, el terme *lletra* no en necessita més. En canvi, la part de la fonètica de la GNV (2006) conté tres térmens diferents en un espai reduït (p. 21-24): grafema (§1.2.1.3), lletra (§1.2.1.3) i grafia (§1.2.2.2), fet que va contra el principi metodològic d'usar un sol terme per a expressar un concepte. Els lectors veuen una certa predisposició a la novetat terminològica (sense constatar-la ni, per tant, justificar-la). Com ara, les consonants vibrants (simples o múltiples) esdevenen "ròtiques bategants i ròtiques vibrants" (§3.3, 3d). Això es completa amb un cert gust per canviar paraules populars clares per cultismes. La boca esdevé tracte vocal (§3.3, 3a); enfront dels noms tancament i acostament, trobem obstrucció i constricció (§3.3, 3a); la punta de la llengua es torna l'àpex de la llengua (§3.3, 3b); les genives no apareixen mai, sempre els alvèols (§3.3, 3a); el rigor es manifesta en descripcions a on diria que no és necessari (en compte de la punta de la llengua toca les dents superiors, 9 paraules, la GNV n'escriu 14: "l'àpex de la llengua toca la cara interior de les dents incisives superiors"). Tot plegat no fa pensar en les persones planes que han mantingut la vitalitat del valencià i l'usen en la comunicació pública, sinó en comportaments que apunten cap a l'aparença, cap a l'elitisme, actituds que lliguen poc amb el caràcter dels valencians.

Convindria completar este treball amb un altre que mire si la *GNV* (2006) descriu quines són les propietats més pertinents del sistema fonològic del valencià i quins són els problemes que cal solucionar per a arribar a un model fonètic identificador, assimilable i practicable per a la majoria de valencians. En relació a eixe contingut, he apuntat que no he trobat l'estudi de l'estructura sil·làbica, que dóna tanta personalitat al valencià.

També convindria estudiar si l'extensió enorme dedicada a l'ortografia (el 73%, §3.1) aprofita per a ajudar als valencians a superar la por davant de les irregularitats ortogràfiques pròpies de les llengües romàniques; o si, per contra, la GNV presenta l'ortografia valenciana més difícil de com és en realitat. La consulta del principi del primer capítol dedicat a les lletres (*GNV*, 2006, §4.1) és útil per a deduir si el manual va de les qüestions bàsiques a les secundàries,

dels usos regulars als irregulars, i si avança d'una manera ordenada i metòdica. No cal dir que eixa actuació contribuïx a formar els lectors, els fa saber què és important i què és secundari i, en definitiva, els impulsa a tindre confiança en la seua llengua.

5 Bibliografia

- Amengual, J. J. (1835). *Gramática de la Lengua Mallorquina*. Palma: P.J. Gelabert.
- Ballot, J. P. (1813). *Gramàtica y apologia de la llengua cathalana*. Barcelona: Piferrer.
- Bofarull, A., & Blanch, A. (1867). *Gramática de la lengua catalana*. Barcelona: Espasa Hermanos.
- DCVB (1926-1968). *Diccionari Català-Valencià-Balear*. Palma de Mallorca: Ed. Moll.
- DGLC: Fabra, P. (1932). Diccionari General de la Llengua Catalana. Barcelona: Edhasa.
- Fabra, P. (1912). *Gramática de la lengua catalana*. Barcelona: L'Avenç.
- Fabra, P. (1918). *Gramàtica catalana*. Barcelona: Aqua, 1981.
- GCC (2002). *Gramàtica del Català Contemporani*. Solà, J., Llorent, M. R., Mascaró, J., & Pérez Saldanya, M. (directors). Barcelona: Empúries.
- GNV (2006). *Gramàtica normativa valenciana*. València: Acadèmia Valenciana de la Llengua.
- Gómez Font, X. (1997). Andreu Sempere (1510-1572) i la seua Prima Grammaticae Latinae Institutio. Alcoi: Ajuntament d'Alcoi.
- Julià i Muné, J. (2002). Els sons del català, *GCC* (2002), volum 1, 39-90.
- Marvà, J. (1932). Curs superior de gramàtica catalana. Barcelona: Barcino, 1968.
- Moll, F. de B. (1937). Rudiments de gramàtica preceptiva per ús dels escriptors baleàrics. Palma de Mallorca: Les Illes d'Or.
- Sanchis Guarner, M. (1950). *Gramàtica* valenciana. València: Torre.
- Saragossà, A. (2003). Gramàtica valenciana raonada i popular. València: Tabarca, 2005.
- Saragossà, A. (2009). Els sis conceptes bàsics de la fonètica en quatre diccionaris (DGLC, DCVB, GDLC i DIEC), Llengua & Literatura. Revista anual de la Societat Catalana de Llengua i Literatura, 20, 257-298.
- Saragossà, A. (2010). Els conceptes síl·laba, vocal i consonant en els diccionaris DGLC, DCVB, GDLC i DIEC,

- Quaderns de Filologia. Estudis lingüístics, XV, 199-218
- Sempere, Andreu (1560). *Prima Grammaticae Latinae Institutio*. Reeditat en Gómez
 Font (1997).

Parte VI. Fonética vocálica y consonántica

Las estrategias compensatorias de los hablantes chilenos: una ampliación acústica de la investigación de /s/implosiva

BOLYANATZ, MARIŠKA A.

1 Introducción

La investigación de la variación de la /s/ en posición de coda, plenamente estudiada en los campos de la dialectología y la sociolingüística, necesita llegar a modernizarse a través de la investigación acústica. Esta actualización se puede realizar gracias a los avances en la calidad y disponibilidad del equipo de grabación, así como por el uso de los programas de análisis fonético como Praat (Boersma & Weenink, 2013) y Audacity (Audacity Team, 2013). En el presente trabajo, usando estos medios, se analizan la duración y la calidad vocálicas en el habla de ocho chilenos que debilitan de manera variada la /s/ implosiva en la lectura de dos cuentos. A través de este estudio, se pretende contribuir a los parámetros de análisis de /s/ implosiva. De esta manera, es decir, con estándares de análisis más precisos, podremos acercarnos más al estudio de la lengua usada diariamente entre los hablantes.

1.1 Sobre la /s/ codal

La /s/ en posición implosiva se debilita en aproximadamente la mitad de América. Este debilitamiento corresponde con áreas dialectales que tradicionalmente se han denominado zonas de "tierras bajas" (Henríquez Ureña, 1921; Menéndez Pidal, 1962; Parodi, Claudia, 2011; Rosenblat, 1933), como las costas pacíficas de Centro- y Suramérica, y también las islas del Caribe. La posición implosiva es una posición tipológicamente marcada en todas las lenguas. Es decir, los hablantes de los idiomas tienden a preferir una sílaba que consista de una consonante más una vocal, la estructura silábica universalmente más común (Macneilage, Davis, Kinney & Matyear, 2000). El debilitamiento de /s/ implosiva, por lo tanto, es otro indicador más de la tendencia hacia la sílaba perfecta, o abierta.

Los investigadores de sociolingüística han demostrado que la manifestación de la /s/ se condiciona por factores internos (o lingüísticos) y externos o sociales (Brown & Cacoullos, 2002; Carvalho, 2006; Lipski, John, 1986; Lynch, 2009; Poplack, 1980a, 1980, 1986; Terrell, 1975, 1977, 1978). Por ejemplo, el segmento que le sigue a la /s/ puede afectar su uso. Varios autores demuestran que el debilitamiento de /s/ a [h] y hasta $[\emptyset]$ se condiciona por la consonante siguiente. Es decir, la presencia de [s] sibilante es muy poco común frente a una consonante en estos dialectos.

Otros factores lingüísticos que condicionan la manifestación de /s/implosiva son la tonicidad de la sílaba en la que aparece la /s/, la tonicidad de la sílaba siguiente, la posición de /s/ dentro de la palabra, la frecuencia léxica de los ítems, su estatus morfémico, y el número de sílabas que contiene la palabra.

De mucha importancia para el presente estudio, hay varios factores sociales que condicionan la manifestación de la /s/ en las tierras bajas. Por ejemplo, la clase socioeconómica a la que pertenece una persona puede afectar su pronunciación de /s/: las personas educadas suelen ser más conscientes de la prescriptividad de [s], y por ende la utilizan más. Además, el género de la persona puede afectar el uso de /s/: las mujeres tienden a usar la variante más conservadora, mientras los hombres utilizan la variante innovadora. Asimismo, la edad afecta el uso de /s/: las personas mayores tienden a conservar el fonema /s/ (o como [s] o como [h]), mientras los hablantes menores suelen usar un sistema de elisión más avanzada que la de sus precursores. Por último, y de mucha importancia para el presente estudio, es el estilo el cual afecta la pronunciación de /s/. En los estilos más formales, se suele usar el alófono con más prestigio dentro de la comunidad de habla (sea [s] o [h], mientras en los estilos menos formales, se tiende a reducir la /s/ a una variante menos prestigiosa.

1.2 Los recursos compensatorios teorizados

En el español, se marca la pluralidad de manera redundante. Por lo tanto, todos los componentes de una frase nominal concuerdan en género y número. Sin embargo, la elisión de toda /s/ en posición final de palabra (como por ejemplo, dentro de un sintagma nominal) puede llegar a neutralizar el contraste entre singular y plural.

Se han mencionado varios recursos compensatorios que pueden utilizar los hablantes de los dialectos radicales en los que se debilitan las /s/ finales, como sucede con los hablantes del español chileno. Como se ha propuesto, tienden a

aparecer vocales cerradas en sílabas abiertas (como por ejemplo [kome]), mientras en sílabas trabadas tienden a aparecer vocales abiertas (como por ejemplo [komes]). Por ejemplo, Navarro Tomás (1966) propone que cuando los puertorriqueños eliden la /s/ final, no queda ninguna otra huella de su pluralidad menos la calidad abierta de la vocal que la precede (73). De igual modo, Oroz (1966) señala que en Chile, la aspiración y la elisión de la /s/ implosiva ocasionan una apertura de la vocal precedente, y también un alargamiento de duración de la misma. Pereira y Soto (2011) indican que la duración vocálica en el español, como no es fonológicamente distintiva, puede estar vinculada a la variación pragmática (155), como la que tratamos aquí.

Resnick y Hammond (1975) fueron pioneros al comenzar a analizar las estrategias compensatorias de manera experimental, a través de una investigación de producción y percepción entre hablantes cubanos que vivían en Miami. Intentaron verificar de manera empírica si existía alguna diferencia fonética o fonológica entre las vocales que precedían el fonema /s/ que se manifestaba como la sibilante [s] y el cero fonético [Ø]. Es decir, pretendían averiguar si los oyentes podían percibir una diferencia entre frases y palabras que terminaban en /s/ elidida, y frases y palabras que terminaban sin /s/ al final. Por ejemplo, les presentaron a los participantes frases como (1):

(1)

- a. Hay varios problemas que no comprendes / komprendes /
- b. Hay varios problemas que no comprende / komprende /

Además, seleccionaron varias palabras en posición final de enunciado con esta misma diferencia entre $/s/\sim/\emptyset/$, y algunas palabras aisladas que contrastaban en posición media de palabra, como *pescado* /peskado/ versus *pecado* /pekado/. Cuatro hablantes de Miami produjeron estas frases que luego se les presentaron a 20 participantes. Los oyentes solo pudieron discriminar entre palabras en los casos de contraste de /s/ en medio de palabra (es decir, *pescado vs. pecado*). Resulta que las vocales que precedían las /s/ elididas en estos casos se alargaron un promedio de 33-40 %. Sin embargo, no encontraron ninguna diferencia significativa ni regularizada en la duración y la calidad de las vocales previas a la /s/ elidida a final de palabra o de enunciado.

Figueroa (2000) intentó reproducir los resultados de Resnick y Hammond entre hablantes puertorriqueños, citando la 'condición distintiva' de Kiparsky (1982), quien plantea que si se elide la /s/ implosiva, como tiene una función morfológica tan importante, debe haber alguna clase de evidencia compensatoria, que sea fonológica, morfológica, o sintáctica. Los resultados de Figueroa apoyan la hipótesis de Resnick y Hammond: sus participantes solo podían distinguir las palabras con contraste de /s/ en medio de la palabra.

Sin embargo, en ninguna de las dos investigaciones queda clara la forma en la que controlaban la producción de la /s/ implosiva de los participantes. Se puede suponer que a los participantes les dieron instrucciones específicas sobre la producción de la /s/. Por lo tanto, se requiere una investigación de la calidad vocálica en la que la producción de la /s/ implosiva esté menos controlada.

1.3 Las preguntas de investigación

A diferencia del estudio de Resnick y Hammond, cuyo estudio se enfocaba en las vocales previas a $/s/versus/\emptyset/$, aquí nos interesa investigar la vocal previa a un fonema /s/ a final de sílaba. La pregunta de ellos es de gran interés,

pero por el momento aquí nos limitaremos a un ambiente más controlado: la variación vocálica frente a /s/. Por ende, las siguientes preguntas guiaron la investigación del presente estudio sobre el comportamiento de las vocales frente al fonema /s/ en posición final de sílaba:

- 1. ¿Hay diferencias de *duración* vocálica entre las vocales que preceden [s] y [h] y las que preceden un cero fonético $[\emptyset]$?
- 2. ¿Hay diferencias de *calidad* vocálica entre las vocales que preceden [s] y [h] y las que preceden un cero fonético [Ø]?

Al igual que Resnick & Hammond y Figueroa, propongo que sí hay una diferencia de duración de las vocales frente a los varios alófonos de /s/ en medio de la palabra. En cuanto a la /s/ implosiva a final de palabra, propongo la hipótesis nula. Es decir, propongo que no hay ninguna diferencia de duración ni de calidad vocálica ante los alófonos [s] y [h], y $[\emptyset]$ (cero fonético).

2 Metodología

En el presente estudio, los participantes fueron 4 hombres y 4 mujeres de origen chileno, quienes viven actualmente en Los Ángeles. Todos llegaron a Estados Unidos, y en particular, a Los Ángeles, después de los 25 años de edad. En Bolyanatz (Bolyanatz, 2014) aclaro que la población de chilenos en Los Ángeles es bastante homogénea. Es decir, los chilenos que llegan a esta zona suelen pertenecer a una clase sociocultural media alta, y casi todos se formaron en la universidad en Chile. Por lo tanto, a diferencia de otros estudios sociolingüísticos sobre la variación de /s/, clase social o socioeconómica no entra al análisis del presente estudio.

Los participantes contribuyeron datos en tres tareas: primero, realizaron una entrevista sociolingüística con la investigadora, luego leyeron 3 narrativas breves, y para terminar, leyeron unas 100 palabras en una frase portadora *_ por favor*). En el presente estudio, se presentan los resultados de la lectura de dos de las narrativas, originalmente utilizadas por Valdivieso y Magaña (1988) en su investigación sociolingüística de la /s/ implosiva en Concepción. La primera narrativa es de estilo formal, y se trata del Quijote. Contiene muchas palabras de baja frecuencia, con un total de 159 palabras. Dentro de este conteo, hay 50 casos de una /s/ de coda silábica (los cuales forman parte de la 'envoltura de la delimitación,' ing. envelope of variation; es decir, entran al análisis). La otra narrativa es de un estilo mucho menos formal, sobre un partido de fútbol al que va un grupo de amigos. En esta narrativa, hay 78 incidencias de /s/ codal. En los dos textos, la envoltura de la delimitación consiste en vocales monoptongadas que no aparecen frente a los fonemas /s/ ni /x/. Por ejemplo, no forma parte del análisis la /s/ final de la palabra *llegamos* en la frase *llegamos justo*, pero sí entra al análisis la /s/ en posición de coda de /xusto/. Se analiza el fonema /s/ en distintos contextos, como por ejemplo a final de palabra pero antes de una consonante o vocal con la que comienza la siguiente palabra, e incluso se incluye la /s/ al final del enunciado.

La grabación de las muestras se realizó usando una grabadora digital de MIC Olympus (LS-14). Luego se pasó la señal a la computadora con una frecuencia de muestreo de 44,1 kHz y una cuantización de 16 bits. Se analizó la señal con el programa Praat (Boersma, Paul & Weenink, David, 1999), segmentando manualmente las vocales. El corpus consiste en 996 vocales monoptongadas previas al fonema /s/, de medio de palabra o al final de palabra.

Como bien sabemos, la calidad vocálica depende (entre otras cosas) del tamaño del tracto vocal, y que esto mismo puede variar mucho entre hombres y

mujeres. Por ende, las vocales se normalizaron a través del NormForm (Thomas & Kendall, 2014), utilizando el método de ANAE de Labov (quien lo utilizó en su proyecto del Atlas del inglés de Norteamérica [2006]).

Como propone Pérez (2007), hay muchas variantes intermedias de los alófonos del fonema /s/, como por ejemplo [s] y [z], [h] y [ĥ]. Además, existen diferentes niveles de aspiración de la vocal precedente, algunos que se pueden considerar como la coarticulación de la vocal y la [h]. Por lo tanto, es imprescindible establecer bien los criterios de segmentación. Para fines de comparación con investigaciones anteriores, catalogamos cada alófono en una de tres categorías, que denominamos [s], [h], y [Ø] (siguiendo las definiciones de Pérez):

[s], o sibilante	presencia de ruido o fricción, concentrado en sectores altos de	
	frecuencia (sobre los 3-4 kHz)	
[h], o aspirada	presencia de ruido de fricción de baja intensidad, distribuido a	
_	lo largo de todo el rango de frecuencias visibles en el espectro	
[Ø], cero fonético	no se observa ruido de fricción de ningún tipo; sin embargo, la	
	vocal anterior puede tener un nivel de aspiración, utilizando el	
	segundo formante como marcador de coarticulación frente a un	
	segmento separado	

Tabla 1. Parámetros de categorización de /s/

En los siguientes cuadros, se puede observar la variación que cabe dentro de cada categoría en el habla de los participantes. Por lo tanto, propongo una ampliación de estos criterios, para así poder estandarizar la segmentación y el análisis de esta variante tan decisiva.

En la primera figura, vemos una [s] canónica americana (es decir, predorso dental) a final de frase.

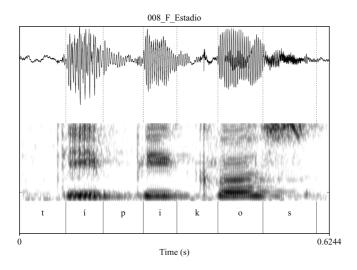


Figura 1. [típikos] de típicos; [s] canónica (008.F)1

En esta primera figura, se nota la división entre la periodicidad de la vocal /o/ que precede la [s], y la energía aperiódica de la [s] que acompaña el ruido en el sector sobre los 4.5 kilohercios.

¹ En todos los espectrogramas, el rango de hercios es de 0-7000.

En la figura 2, se ve una [s] debilitada en el enunciado de uno de los participantes. Esta [s] aparece comúnmente frente a [t], como en este mismo caso.

001M samplereduced s short

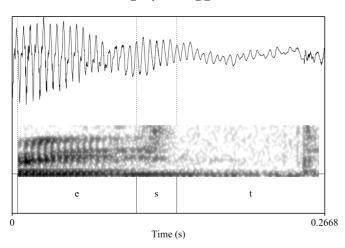


Figura 2: [est] de restablecer; [s] debilitada (001.M)

Esta [s] debilitada, con su duración reducida y los formantes visibles, puede volver a plantearnos la pregunta de por qué una fricativa predorso dental sorda se debilita a una fricativa glotal sorda. Según el modelo de la fuente y el filtro (source-filter theory), la [s] tiene dos fuentes de sonido: una glotal (o sea el aire pasa por los pliegues vocales abiertos), y una oral (el aire pasa por la aproximación creada por la lengua que se acerca a los alvéolos; (Johnson, 2011: 152-159). Este proceso se denomina desbucalización, y es un proceso de lenición en la que una consonante pierde su obstrucción dentro del tracto vocal, siendo el sonido resultante una consonante laríngea (como [h], [ĥ], o [ʔ]; (O'Brien, 2012). Propongo que esta variante alofónica de [s] debilitada necesita incorporarse a la categorización de la descripción de /s/ implosiva.

En las figuras 3 y 4 a continuación, vemos dos ejemplos de lo que categoricé como variantes de [h]. Como anteriormente se ha mencionado, la segmentación de la [h] versus la [Ø] es complicada. Por lo tanto, es necesario recurrir a estrategias establecidas, y ser consistente al segmentar. Por ejemplo, Ertan (2013), al estudiar la /h/ del turco, propone tres variantes distintas: una [h] con formantes, una [h] con fricación pero sin formantes, con energía en las bajas frecuencias, y una [h] con muy poca energía. Como parámetro de segmentación, Ertan (2013: 91) señala que el inicio de la [h] se acompaña por "una disminución considerable de la amplitud de la onda". Garellek (c. p.) recomienda también enfocarse en los formantes, sobre todo el segundo (el F2). Si el segundo formante de la vocal disminuye abruptamente, frente a un periodo de fricación (aunque sea sonora), lo más probable es que le siga una [h]. Si los formantes disminuyen gradualmente, lo más probable es que no le siga una [h], sino una vocal con aspiración.

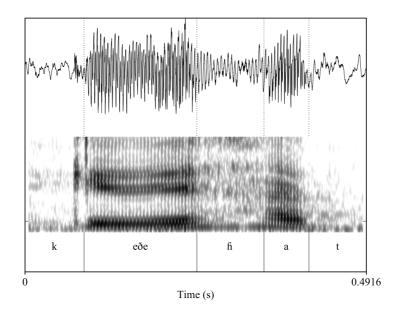


Figura 3. [kéðehat] de quedes atrás, [h] (006.F)

En la figura 3, se observa una fricativa glotal sonora [fi] entre dos vocales átonas. Se perciben ligeramente los formantes que actúan como puente entre las dos vocales (como propone Keating (1988), y se puede observar la barra de sonoridad en las frecuencias bajas.

En la figura 4 (abajo), se puede ver que el segundo formante de la vocal [e] desaparece de repente, a la vez que aumenta la fricación en todas las frecuencias (aunque concentrándose en la zona de los formantes). Por lo tanto, ítems como éste se categorizaron como aspirados.

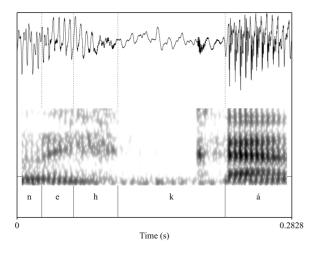


Figura 4. [nehká] de un escándalo, [h] (001.M)

En la figura 5 abajo, tenemos 2 incidencias de elisión de /s/. La primera es una /s/ verbal, del verbo es. Los tres formantes disminuyen gradualmente, hasta el cierre velar de la [k] siguiente.

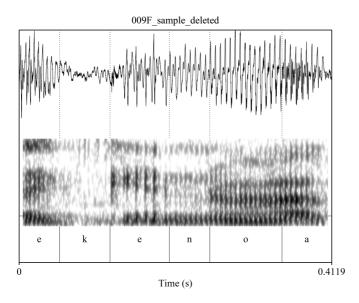


Figura 5. [ekenoa] de es que nos acomodamos, [Ø] (009.F)

De igual modo, entre las vocales átonas [o] y [a], del pronombre *nos* y la primera sílaba de *acomodamos*, no hay ruido o fricción que represente una [s] o una [h]; más bien la vocal átona [o] pierde un poco de energía pero se mantiene bastante clara, y directamente comienza la [a] siguiente.

En esta sección, se han planteado algunas ampliaciones en la descripción de la segmentación y análisis de la /s/ en posición de coda. La segmentación de /s/ implosiva tiene un carácter subjetivo, pero esperamos que esta investigación ayude a disminuir esta subjetividad.

3 Resultados

3.1 Resultados estadísticos descriptivos

Primero, se darán a conocer los resultados estadísticos descriptivos. En la figura 6, se observa una gráfica de los promedios de frecuencia de uso de cada alófono de /s/ en el estudio actual.

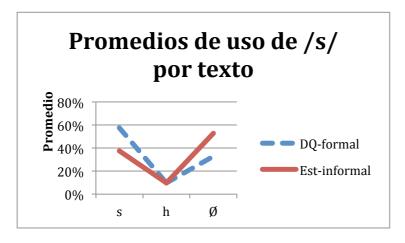


Figura 6. Promedios de usos de alófonos de /s/ por cuento (estudio actual)

Como se puede observar, [h] es el alófono menos usado de los tres; hay casi el mismo promedio de uso de [h] en cada texto. Se nota también que la sibilante [s], como la variante de mayor prestigio, es más frecuente en la narrativa más formal, mientras el cero fonético es más común en la narrativa más informal. Esta figura se puede comparar con la figura 6 (a continuación), donde se representan los promedios encontrados para las mismas lecturas en Concepción por Valdivieso & Magaña (1988).

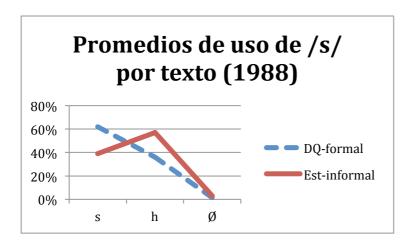


Figura 7. Promedios de usos de alófonos de /s/ por cuento (de Valdivieso & Magaña 1988)

Como se puede notar, la frecuencia del cero fonético del estudio actual es mucho más alta que en el estudio llevado a cabo hace unos 25 años. De igual modo, la frecuencia de la aspirada en la presente investigación es mucho menor, mientras la frecuencia de la [s] es bastante constante. La frecuencia regular de la [s] se podría atribuir al estilo de habla de la tarea (o sea, lectura de narrativa). Sin embargo, la variación (y casi intercambio de frecuencias de [h] y $[\emptyset]$ de los dos estudios) se podría explicar de dos maneras. Primero, puede haber diferencias de codificación de los alófonos de /s/ de los dos métodos (es decir, una metodología de análisis impresionista versus un análisis acústico). La segunda posibilidad es que la diferencia que observamos es que hay un cambio en marcha, o variación diacrónica de las variantes de /s/.

3.2 Los resultados de los análisis estadísticos

Los análisis estadísticos se realizaron en el ambiente y lenguaje de programación estadística R (R Core Team, 2014), utilizando modelos de regresión lineales con efectos mixtos. El factor de 'participante' fue considerado como efecto aleatorio, para así controlar por las diferencias de velocidades de habla de cada persona, otro factor que podría contribuir a la manifestación de la /s/ implosiva. Como efectos fijos, se incluyeron factores sociales y factores lingüísticos. Los factores sociales tomados en cuenta son el género del hablante y el nivel de formalidad del texto que leyeron. Los factores lingüísticos incluidos en el análisis son la identidad de la vocal ([a], [e] u [o]), el segmento fonológico siguiente, el acento de la sílaba, la posición de la /s/ (en medio de o a final de palabra), y el alófono de [s] (categorizado bajo los tres alófonos discutidos arriba). Dado el carácter de los ítems léxicos en los cuentos (y en la lengua en sí), hubo muy pocas incidencias de vocales [i] y [u] en las posiciones de interés, las cuales por lo tanto fueron excluidas del análisis. Del mismo modo, se excluyó el alófono [h], pues formó solo el 10 % del muestreo total.

Nos dirigiremos primero a la primera pregunta de investigación sobre la duración vocálica: ¿difiere la duración de la vocal previa a una [s] o una [h] de la duración de la vocal previa a un cero fonético? Un modelo de regresión lineal nos revela varios efectos principales y un efecto de interacción. Primero, se darán a conocer los resultados de los efectos fijos. Una vocal es más breve frente a las oclusivas sordas (como /p/, /t/, y /k/), mientras es más larga frente a una pausa (p <.001). Este efecto es de esperar, por la influencia de alargamiento vocálico a final de enunciado. Se encontró también un efecto de acortamiento vocálico dependiendo de la posición de la vocal. Es decir, las vocales a final de palabra son aproximadamente 9 milisegundos más cortas (p<.001). Este efecto es algo inesperado; se esperaría también un efecto de alargamiento de final de palabra. Además, hubo un efecto fijo de tonicidad: las vocales tónicas son aproximadamente 25 mseg más largas que las átonas (p <.001); esto, fonéticamente esperado (la duración de la vocal es un indicador de tonicidad). Por último, se encontró un efecto de vocal: la vocal [e] es aproximadamente 12 mseg más breve (p <.001), mientras no hay diferencias significativas de duración de [a] versus [o] (el cual se esperaría por el nivel de apertura de la boca al pronunciar la [a]).

Además de los efectos fijos, hubo un efecto de interacción, como se puede apreciar en la figura 8.

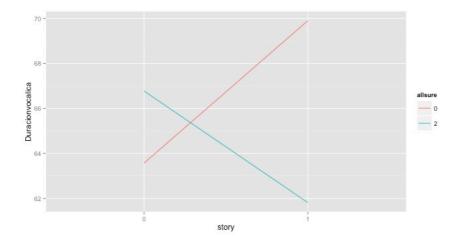


Figura 8. Interacción del factor de 'estilo' en el uso de [s] vs. [Ø] y la formalidad del cuento

En esta figura, se evidencia que la variable 'estilo' incide en la realización de las variantes de la variable /s/ (similarmente al estudio de Pérez). Es decir, en el cuento más formal, las vocales que le preceden a [s] son más largas que las vocales que le preceden a un cero fonético, pero esta diferencia no es significativa (es mínima la diferencia). Por otra parte, en el cuento informal, cuando una vocal precede un cero fonético, es aproximadamente 8 milisegundos más larga, y esta diferencia es significativa (p <.01). Es decir, en el estilo que más se aproxima al habla espontánea, hay una diferencia significativa de duración vocálica: la vocal previa a [Ø] es más larga que la vocal previa a [s]. Esta diferencia se tiene que considerar de manera conservadora: puede ser una estrategia compensatoria por la elisión del marcador de pluralidad manifiesto, pero también se podría explicar como parte de la tendencia hacia el alargamiento en la sílaba abierta (Hayes, 1989).

Para contestar la segunda pregunta de investigación sobre la calidad de las vocales frente a /s/, nos dirigiremos primero a los resultados del primer

formante. No hubo ningún efecto de interacción, pero sí se obtuvieron unos efectos fijos. A pesar de la normalización de las vocales, se encontró que los valores de F1 varían significativamente entre hombres y mujeres (p <.01). Es decir, el valor de hercios del primer formante de las mujeres son aproximadamente 32 hercios menos que el de los hombres, lo cual quiere decir que las vocales de las mujeres se producen con la lengua en una posición significativamente más alta en el espacio oral. Se encontró un segundo efecto fijo: el de identidad de la vocal. Este efecto es de esperar: el primer formante indica la altura de la lengua, entonces los valores de F1 de la vocal [a] son mucho más altos que los valores de F1 de las vocales [e] y [o], porque la [a] se produce con la lengua baja y la boca más abierta.

La segunda parte de esta pregunta de investigación trata de los valores del segundo formante. Se encontró un efecto fijo de la tonicidad de la vocal: cuando una vocal precedente a /s/ es tónica, es aproximadamente 45 hercios más baja que cuando precede una vocal átona (p <.02). Es decir, las vocales átonas se producen de una manera más adelantada en la cavidad oral. Hubo además dos efectos de interacción. Primero, la posición de la vocal (es decir, en medio de la palabra o a final de palabra precedente a /s/) afecta la realización de las variantes de la variable /s/: el valor del formante 2 es más bajo (es decir, la vocal se produce de una manera menos adelantada) cuando la vocal precede a un $[\varnothing]$ en medio de la palabra. A final de palabra, sin embargo, las vocales que le preceden a un cero fonético son más adelantadas.

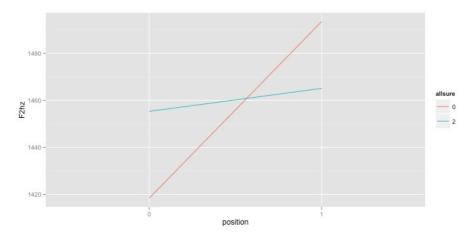


Figura 9. Efecto de interacción en los valores de F2: alófono vs. posición

Por último, también se encontró un efecto de interacción de la identidad de la vocal: la variable 'vocal' incide en las realizaciones de la variable /s/ en cuanto a los valores de F2. Ello tiene mucho sentido: las vocales, aparte de categorizarse por su altura, se clasifican por el adelantamiento de la lengua, el cual se mide por el valor de F2. Por lo tanto, no es sorprendente que las vocales tengan valores significativamente diferentes de F2.

4 Conclusiones

Como se ha planteado, el propósito de esta investigación ha sido contribuir con más parámetros a la metodología y análisis de /s/ implosiva, además de determinar si los participantes chilenos utilizan efectos o estrategias compensatorias para compensar la /s/ elidida.

En cuanto a la primera pregunta de investigación, si la duración vocálica puede ser un sitio de compensación, se ha descubierto que a final de palabra, en el estilo más informal, los hablantes alargan una vocal final frente a un cero fonético. Este alargamiento vocálico se podría considerar una estrategia compensatoria, pero no se debe descartar la posibilidad de que sea solo un alargamiento de la sílaba abierta. Para confirmar esta hipótesis, de modo similar a la investigación de Resnick y Hammond, se requiere un estudio que compare las vocales finales previas al fonema /s/y al fonema /O/w.

Además, se necesita considerar que la tasa de aspiración vocálica puede también ser una estrategia compensatoria: cuando se elide el fonema, queda una vocal aspirada como huella de lo elidido.

Por último, se requiere también descubrir alguna medida que pueda hacer de la segmentación y codificación de los alófonos de /s/ un proceso menos incierto. De esta manera, la investigación de /s/ implosiva puede convertirse en algo más riguroso, formalizado, y preciso, usando los métodos de la fonética acústica.

5 Bibliografía

- Audacity Team. (2013). Audacity (Version 2.0.3.0). In http://audacity.sourceforge.net/
- Boersma, P., & Weenink, D. (1999). Praat: doing phonetics by computer (Version 5.3.51). In http://www.praat.org/.
- Bolyanatz, M. (2014). Lo do idioma: A variationist examination of Chilean Spanish final /s in Los Angeles. Presented at the OSU Conference on Hispanic and Lusophone Linguistics (OSUCHILL), Ohio State University, Columbus, OH.
- Brown, E. L., & Cacoullos, R. T. (2002). ¿Qué le vamoh aher?: Taking the syllable out of Spanish/s/reduction. University of Pennsylvania Working Papers in Linguistics: Papers from NWAV, 30(8), 3.
- Carvalho, A. M. (2006). Spanish (s) aspiration as a prestige marker on the Uruguayan-Brazilian border. *Spanish in Context*, 3(1), 85-114. doi:10.1075/sic.3.1.07car
- Ertan, E. (2013). Spectrographic Characteristics of Turkish /h/. *Open Journal of Acoustics*, 03(04), 97-102. doi:10.4236/oja.2013.34015
- Figueroa, N. (2000). An acoustic and perceptual study of vowels preceding deleted post-nuclear /s/ in Puerto Rican Spanish. In H. Campos, et. al (Ed.), Hispanic Linguistics at the Turn of the Millenium (pp. 66-79). Somerville, MA: Cascadilla Press.
- Hayes, B. (1989). Compensatory lengthening in moraic phonology. *Linguistic Inquiry*, 253-306.

- Henríquez Ureña, P. (1921). Observaciones sobre el español de América. *Revista de Filología Española*, 8, 357-390.
- Johnson, K. (2011). *Acoustic and Auditory Phonetics* (3rd ed.). Malden, MA: Wiley-Blackwell.
- Keating, P. A. (1988). Underspecification in Phonetics. *Phonology*, 5(2), 275-292.
- Kiparsky, P. (1982). *Explanation in Phonology*. Dordrecht, The Netherlands: Foris Publications Holland.
- Labov, W., Ash, S., & Boberg, C. (2006). The Atlas of North American English: Phonetics, Phonology and Sound Change. Berlin: Mouton de Gruyter.
- Lipski, John. (1986). Reduction of Spanish word-final /s/ and /n/. *The Canadian Journal of Linguistics*, 31(2), 139-156.
- Lynch, A. (2009). A Sociolinguistic Analysis of Final /s/ in Miami Cuban Spanish. *Language Sciences*, 31, 767-790.
- Macneilage, P., Davis, B., Kinney, A., & Matyear, C. (2000). The motor core of speech: A comparison of serial organization patterns in infants and languages. *Child Development*, 71, 153-163.
- Menéndez Pidal, R. (1962). *Estructuralismo e historia: homenaje a André Martinet*. En D. Catalán, (3.ª ed., pp. 99-165). Universidad de La Laguna: La Laguna: Biblioteca Filológica.
- Navarro Tomás, T. (1966). El español en Puerto Rico; contribución a la geografía lingüística hispanoamericana. Río Piedras, Universidad de Puerto Rico.
- O'Brien, J. (2012). An experimental approach to debuccalization and supplementary

- *gestures*. University of California, Santa Cruz, Santa Cruz, CA.
- Oroz, Rodolfo. (1966). *La lengua castellana en Chile*. Santiago, Facultad de Filosofía y Educación: Universidad de Chile.
- Parodi, Claudia. (2011). El otro México: español chicano, koineización y diglosia en Los Angeles, California. En Villanueva, Rebeca & Butragueño, Pedro Martín (Eds.), (pp. 217-243). Presented at the Realismo en el análisis de corpus orales: Primer Coloquio de Cambio y Variación Linguística, México, D. F.: El Colegio de México.
- Pereira, D. I., & Barba, J. S. (2011). Duración absoluta de las vocales del español urbano y rural de la provincia de Ñuble. *Boletín de Filología*, 46(1), 153-161.
- Pérez, H. E. (2007). Estudio de la variación estilística del fonema/s/en posición implosiva en el habla de los noticieros de la televisión chilena. *RLA. Revista de Lingüística Teórica Y Aplicada*, 45(1), 101-115.
- Poplack, S. (1980a). Deletion and disambiguation in Puerto Rican Spanish. *Language*, 371-385.
- Poplack, S. (1980). The notion of the plural in Puerto Rican Spanish: competing constraints on/s/deletion. *Locating Language in Time and Space*, 1, 55.
- Poplack, S. (1986). Acondicionamiento gramatical de la variación fonológica en un dialecto puertorriqueño. In R. A. Núñez Cedeño, I. Páez Urdaneta, & J. M. Guitart, Estudios sobre la fonología del español del Caribe (pp. 96-107). Caracas: La Casa de Bello.

- R Core Team. (2014). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. In http://www.R-project.org/
- Resnick, M. C., & Hammond, R. M. (1975). The Status of Quality and Length in Spanish Vowels. *Linguistics*, 13(156). doi:10.1515/ling.1975.13.156.79
- Rosenblat, A. (1933). La lengua y la cultura en Hispanoamérica. Tendencias lingüísticas y culturales. In *Vom Leben un Wirken der Romanen* (vol. 1, p. 28). Universidad de Berlín: Spanische Reihe.
- Terrell, T. D. (1975). Functional Constraints on Deletion of Word Final /s/ in Cuban Spanish. *Proceedings of the First Annual Meeting of the Berkeley Linguistics Society*, 431-437.
- Terrell, T. D. (1977). Constraints on the aspiration and deletion of final/s/in Cuban and Puerto Rican Spanish. *Bilingual Review/La Revista Bilingüe*, 4(1/2), 35-51.
- Terrell, T. D. (1978). Sobre la aspiración y elisión de/s/implosiva y final en el españjol de Puerto Rico. *Nueva Revista de Filología Hispánica*, 27(1), 24-38.
- Thomas, E. R., & Kendall, T. (2014). Norm:
 The Vowel Normalization and
 Plotting Suite (Version 1.1). In
 http://lvc.uoregon.edu/norm/norm
 1.php
- Valdivieso, H., & Magaña, J. (1988). Variación lingüística: La /s/ implosiva en Concepción. Revista de Lingüística Teórica Y Aplicada, 26, 91-103.

Sons, cervells i usos lingüístics. Caracterització diafàsica de vocals tòniques del lleidatà en contextos multilingües

CARRERA-SABATÉ, JOSEFINA

1 Introducció

L'objectiu d'aquest article és presentar la variació acústica de vocals tòniques emeses per adolescents² de Lleida que segueixen l'ESO i el Batxillerat en centres d'educació secundària i que tenen el català i el romanès com a L1 (i, evidentment, són usuaris actius de castellà).

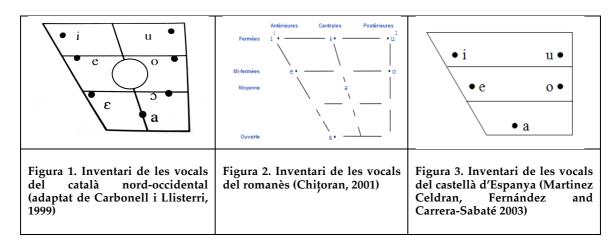
Lleida és el punt de referència escollit per diversos motius: d'una banda, Lleida ha tingut un creixement constant de població estrangera, com ha ocorregut de manera global en molts indrets del món; segons les dades de l'Idescat de l'any 2010 (http://www.idescat.cat), Lleida és una ciutat d'uns 138.000 habitants, uns 29.000 dels quals són de nacionalitat estrangera, i té un creixement migratori del 4,37‰ (a Catalunya el creixement és de l'1,34‰). D'una altra banda, Lleida compta amb col·lectius nombrosos agrupats per nacionalitats diferents; segons un informe de l'Ajuntament de Lleida (2004), els

¹ Aquest treball s'ha beneficiat de l'ajut a projectes d'investigació referència núm. FFI2010- 22181-C03-02 del MECI. Agraeixo la bona disponibilitat dels estudiants del Col·legi Episcopal i de l'Institut Joan Oró de Lleida, així com l'ajut de Celi Canyadell i Gemma Mirmi per poder contactar amb els informants. També vull agrair els comentaris de Jane Stuart-Smith i Tamara Rathcke en versions anteriors d'aquest treball.

² Els adolescents han estat àmpliament reconeguts com a part important en els processos de variació i canvi lingüístics. En comunitats lingüístiques amb parlants procedents de nacionalitats diverses, s'ha demostrat que la variació fonètica dels adolescents està relacionada amb l'estructura social de la comunitat, els grups socials i ètnics (Kirkham & Moore 2013).

col·lectius més nombrosos provenen del Marroc i Algèria, de Colòmbia i Equador i també de Romania.

El fet de comparar parlants de procedències lingüístiques diferents s'explica perquè, al marge d'aspectes demogràfics i socioculturals que acabo de descriure, els inventaris vocàlics de les llengües comparades són força diferenciats. A continuació presento els inventaris de referència per al català nord-occidental, romanès i espanyol d'Espanya (perquè forma part dels usos lingüístics dels enquestats-) (figures 1, 2 i 3), on es pot comprovar que el català és l'única llengua que compta amb les vocals $/\epsilon/i$ /5/.



Les vocals mitjanes obertes del català, preses aïlladament o conjuntament amb les altres vocals que conformen l'inventari del català, han estat analitzades des de la producció (amb estudis comparatius interdialectals -vegeu, per exemple, Carrera-Sabaté & Fernández-Planas (2005) i Recasens & Espinosa (2006)– i també des de la percepció (vegeu, per exemple, Mora, Keidel & Flege, 2011; Sebastián-Gallés & Díaz, 2012; Sebastián-Gallés, Echeverría, & Bosch, 2005). Aquestes aportacions coincideixen a assenyalar més inestabilitat en les vocals [ɛ] i [ɔ] que en les extremes (com ara [i a u]) tant des del punt de vista productiu (perquè, en general, en les primeres hi ha més variació en l'F1), com perceptiu (perquè són més difícils de percebre, especialment si es tenen en compte les parelles [e e] i [ɔ o] en relació amb d'altres com ara [i ε] o bé [o u]). A més a més, des d'un punt de vista més general i comparant sistemes vocàlics, la presència de dos graus d'obertura en les vocals mitjanes és menys frequent que la presència d'un grau d'obertura (Maddieson, 1984) i, tal com assenyala Amengual (2014), les vocals mitjanes del català són més marcades (tipològicament menys preferides) i, en definitiva, més vulnerables a seguir la influència d'altres llengües que només tenen un grau d'obertura (com és el cas del castellà i del romanès, en el cas que m'ocupa).

Al marge d'això, si es prenen en consideració processos d'aprenentatge de segones llengües, sembla que les dificultats perceptives i productives que s'observen a l'hora d'aprendre els sons d'una L2 no només vénen causades per les propietats fonètiques dels sons particulars d'una L1 sinó per l'estructura del sistema fonològic, de manera que són tots els trets fonètics de la L1 i les distàncies entre els sons que conformen el sistema, els que influeixen en la producció i percepció d'una L2 (Brown, 2000). Cal tenir en compte, endemés, que com més diferències hi ha entre un so de la L1 i un de la L2, més fàcilment

es distingiran (i es produiran) els trets específics de la L2 (Flege, 1995). A banda d'aquests aspectes, la recerca sobre l'aprenentatge fonològic i fonètic d'una L2 s'ha centrat en variables com l'edat dels parlants i també en variables socials com ara l'ús de la L1 i la L2, la instrucció i pràctica de la L2, el temps de residència, l'aptitud, la motivació i els aspectes cognitius dels parlants (vegeune una revisió a Major, 2008). En un altre ordre de coses, la sociofonètica també ha considerat recurrentment variables com l'edat o el gènere, que poden explicar tant processos de canvi en curs com de manteniment de variables. Tal com afirma Labov (2001: 367), "Women deviate less than men from linguistic norms when the deviations are overtly proscribed, but more than men when the deviations are not proscribed". També, tenint en compte el gènere dels informants, s'han trobat diferències en les mesures dels formants vocàlics que estan relacionades amb el gènere i amb l'edat dels parlants (Huber *et al.*, 1999).

Quant als usos lingüístics, la literatura variacionista ha proveït la comunitat lingüística d'exemples relacionats amb situacions comunicatives i de contacte lingüístic. D'una banda, Labov (1966) va definir un contínuum diafàsic mitjançant cinc contextos d'ús: conversa lliure (estil casual); conversa dirigida (estil acurat); lectura de text (estil formal), lectura de llistes de paraules (estil més formal); lectura de parells mínims (estil molt formal), contextos que han servit de canemàs per a molts estudis. Aquests contextos classifiquen els estils de parla segons el grau d'atenció al propi discurs dels parlants i, tenint en compte els tipus d'activitats cerebrals que comporten, també han estat objecte d'estudi en altres disciplines com la psicolingüística. D'aquesta manera, quan llegim, el cervell duu a terme tres tasques diferents: a) analitza formes escrites, b) connecta la llengua escrita amb la llengua parlada i c) busca el significat de les paraules. Sembla que aquestes tasques activen tres parts diferents del còrtex esquerre del cervell (Dehaene, 2007). Segons Tokimura et al. (1996), quan es llegeix en veu alta hi ha una excitabilitat de l'àrea motora de l'hemisferi esquerre, i, en parla espontània, hi ha una excitabilitat de les àrees motores dels dos hemisferis cerebrals. Se sap, també, que les estratègies de lectura d'una L2 vénen determinades per la llengua primera (Nakada et al., 2001). També sembla que en parlants de francès llengua estrangera que tenen una bena competència, l'elocució és menys fluida en la lectura en veu alta que en parla espontània (Alazard *et al.*, 2011).

Encara dins de l'estudi de la variació de la llengua, s'ha observat que les interaccions cara a cara juguen un paper molt important per tal que hi hagi acomodació lingüística i, en definitiva, canvis lingüístics (Trudgill, 1989); en aquest sentit, els mecanismes d'acomodació que els individus duen a terme en la construcció de les gramàtiques individuals poden explicar com es representa cognitivament el significat social de la variació lingüística (Celata & Calamai, 2014). I és que les gramàtiques no són fixes ni estàtiques, sinó que es veuen modelades per l'experiència de cada parlant-oient. D'acord amb aquest marc teòric, l'adquisició, estructuració, organització, variació i canvi lingüístic d'una llengua no són aspectes independents sinó facetes diferents d'un mateix sistema que és complex i que es va adaptant als usuaris, que interactuen els uns sobre els altres i que es comporten seguint mecanismes perceptius i també motivacions socials².

Si focalitzem l'atenció en l'inventari fonològic d'una llengua i d'una variant dialectal determinada, i partint del que acabo de descriure, es pot predir, a grans trets, que els sons d'un sistema lingüístic es van afaiçonant i modificant a

¹ Speech Learning Model (Flege, 1995).

² Segueixo els postulats del sistema adaptatiu complex (CAS, a Beckner et al., 2009).

partir de l'ús que en fan els parlants-oients. Quan l'inventari fonològic de referència és utilitzat per parlants de procedències diverses, amb altres inventaris fonològics com a L1, hi ha, necessàriament, transferències i interaccions entre tots els sistemes implicats que provoquen que les modificacions fòniques arribin a ser força evidents. Aquest podria ser, per exemple, el cas del català, que s'utilitza molt sovint en contextos multilingües.

La hipòtesi de partida d'aquest treball és que hi ha diferències en la realització de les vocals [ɛ] i [ɔ] en lleidatanoparlants que tenen diferents L1 i que aquestes diferències es veuen modificades per factors discursius (diferències entre lectura i entrevista semidirigida) i socials, com el gènere, els anys de residència a la ciutat de Lleida i les llengües d'ús amb els amics.

2 Metodologia

Els informants analitzats són 27 adolescents de 12 a 17 anys que viuen a Lleida, estudien en centres de secundària de la ciutat i parlen lleidatà amb fluïdesa; no tenen cap defecte de pronúncia ni presenten problemes auditius. Es troben distribuïts en dos grups segons la L1: 12 tenen el lleidatà com a L1 (i usen, normalment, aquesta varietat) i 15 tenen el romanès com a L1. En cada grup de parlants hi ha subagrupacions relacionades amb el gènere: 6 nois i 6 noies en el grup L1 català i 7 nois i 8 noies en el grup L1 romanès. Al marge d'això, i en relació amb els informants L1 romanès, hi ha diferències entre els enquestats que tenen a veure amb l'edat d'arribada a Lleida i, consegüentment, l'edat d'aprenentatge del català i, des d'una altra òptica, les llengües d'ús a casa, amb els amics i al centre d'educació secundària. A més a més, tots els enquestats L1 romanès coneixen perfectament el castellà i l'utilitzen al costat del català. En aquest sentit, es fa difícil establir si el català o el castellà són la L2 o la L3 dels enquestats, atès que les dues llengües s'aprenen de forma paral·lela i, molts cops, el castellà s'usa amb més freqüència que el català.

3 Estímuls

L'observació lingüística dels informants L1 català i L1 romanès s'ha dut a terme mitjançant diverses proves: 1) lectura de frases portadores amb logòtoms que tenen totes les vocals del català en contextos consonàntics simètrics, 2) lectura d'un conte breu, i 3) converses semidirigides. En aquest treball em centraré en les dades obtingudes de dues tasques: la lectura en veu alta d'un conte breu, que prové d'un recull de rondalles populars catalanes, i una part narrativa d'una conversa semidirigida. Les vocals que apareixen en tots dos fets de parla no presenten totes la mateixa freqüència d'ús, cosa que també ocorre en la llengua natural (vegeu Rafel, 1980).

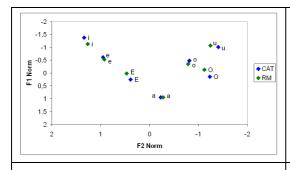
Les variables lingüístiques per analitzar són els valors de F0, F1, F2 i F3 de cada vocal, obtinguts mitjançant Praat. També el context consonàntic adjacent a les vocals. Per tal de poder comparar tots els enquestats, s'han normalitzat els formants seguint el mètode de Lobanov (Ferrari, 1980 i Recasens, 2008). Les variables socials preses en consideració per a l'estudi són la L1 de cada informant, el gènere, el temps de residència a la ciutat de Lleida, així com la llengua utilitzada a casa, al centre d'educació secundària i amb els amics.

L'anàlisi estadística s'ha dut a terme mitjançant els programes Statgraphics i SPSS, utilitzant ANOVA mulfactorial i d'un nivell i MDS.

4 Resultats i discussió

4.1 Generalitats

L'anàlisi de la carta formàntica de les freqüències amitjanades de totes les vocals emeses pels parlants permet observar, d'entrada, que les diferències més rellevants en totes les vocals es troben en les vocals mitjanes, tal com era previsible (vegeu les figures 4 i 5). De manera més concreta, s'observa que el camp de dispersió de les vocals mitjanes baixes dels parlants que tenen el romanès com a L1 no coincideix amb el dels parlants que tenen català com a L1 ni en les dades provinents de la lectura ni de l'entrevista semidirigida. En la lectura, les diferències són més rellevants que en l'entrevista semidirigida.



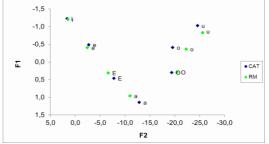


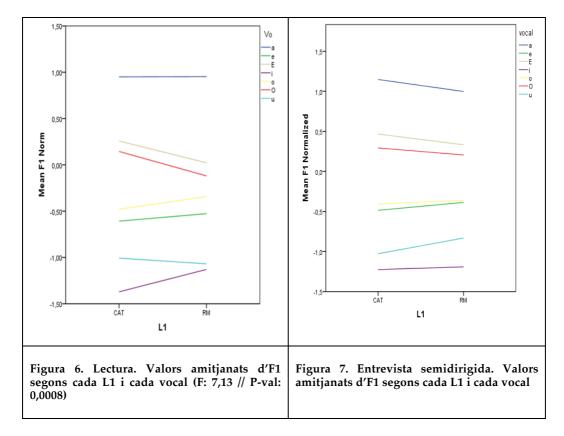
Figura 4. Carta formàntica corresponent a les freqüències amitjanades dels valors normalitzats de F1 i F2 per a cada grup de parlants: L1 català (CAT), L1 romanès (RM). Lectura en veu alta

Figura 5. Carta formàntica corresponent a les freqüències amitjanades dels valors normalitzats de F1 i F2 per a cada grup de parlants: L1 català (CAT), L1 romanès (RM). Entrevista semidirigida

4.2 F₁₃

De manera més concreta, les diferències més rellevants entre els parlants se situen en les vocals mitjanes obertes [ɛ] i [ɔ] perquè presenten l'F1 més elevat quan els parlants tenen L1 català. Cal tenir en compte, però, que les diferències només resulten significatives en les dades provinents de la lectura, per bé que els resultats de l'entrevista segueixen un patró similar (vegeu les figures 6 i 7). També és important destacar que els valors formàntics de F1 d'aquestes vocals en contacte amb contextos palatals es modifiquen significativament (són més elevats) i en contacte amb contextos velars presenten més dispersió.

³ Després d'analitzar les valors d'Fo, F1, F2 i F3, s'ha observat que només són significatius els resultats corresponents a F1.



Al marge d'aquest aspecte, i encara parlant del primer formant, per a les dades provinents de la lectura (aquesta tendència també es dóna en els resultats de l'entrevista), es constaten diferències significatives si es tenen en compte les distàncies entre els F1 amitjanats de les parelles de vocals següents: [ɛ e] i [ɔ o]. D'acord amb això, els parlants L1 català presenten distàncies més grans entre les vocals mitjanes obertes i tancades ([ɛ e] i [ɔ o]) que els L1 romanès (vegeu la figura 8).

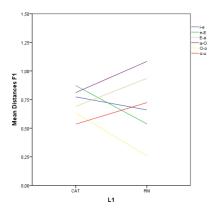


Figura 8. Lectura. Valors amitjanats de les distàncies vocàliques entre els valors normalitzats de F1 segons cada L1 i segons cada parella de vocals analitzada ([ε-e]: F=4,8; p-val=0,03; [ɔ-o]: F=10,45; p-val=0,0034)

Seguint en l'anàlisi de les distàncies entre les vocals, en cap dels grups de resultats obtinguts no s'observen diferències significatives entre les vocals extremes ([i a u]), de manera que, a grans trets, tant [i a] com [a u] són equidistants en tots dos grups. Aquesta observació concorda amb els resultats

de Recasens i Espinosa (2006: 664) en la comparació de les vocals de quatre grans zones dialectals del català, sobretot pel que fa a la parella [ia]. I és que, paral·lelament al que apunten en relació amb els dialectes del català, en lleidatà, les diferències en les distàncies de les vocals mitjanes entre els parlants es veuen compensades per la distància constant entre [ia] i també [au] que presenten. D'alguna manera, es pot parlar d'un sistema compensatori que ve donat per un rang global de F1 constant en tots els parlants i que delimita les diferències en les distàncies de les vocals mitjanes.

4.3 Context comunicatiu

La incidència del tipus de tasca (lectura i entrevista semidirigida) a què s'han sotmès els enquestats no ha resultat significativa per trobar diferències entre els F1 de les vocals analitzades, contràriament al que s'esperava. Tal com s'observa en la figura 9, la distribució dels F1 de les vocals dels enquestats segons els contextos comunicatius permet només distingir les emissions dels enquestats L1 romanès i català. S'observa que en les vocals mitjanes obertes [ɛ] i [ɔ] els punts dels L1 catalans se situen a la part més baixa del gràfic, mentre que en el cas de les vocals mitjanes tancades i també en les vocals tancades i obertes els punts es troben més barrejats, cosa que demostra, un cop més, que les diferències en les emissions dels enquestats es troben en les vocals mitjanes obertes.

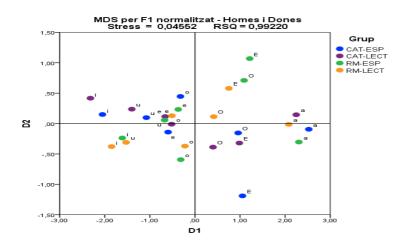
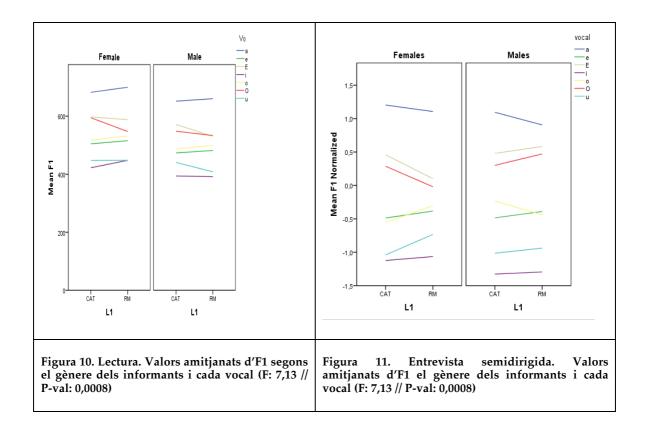


Figura 9. MDS de F1 de totes les vocals en relació amb el tipus de tasca (lectura/entrevista)

4.4 Gènere

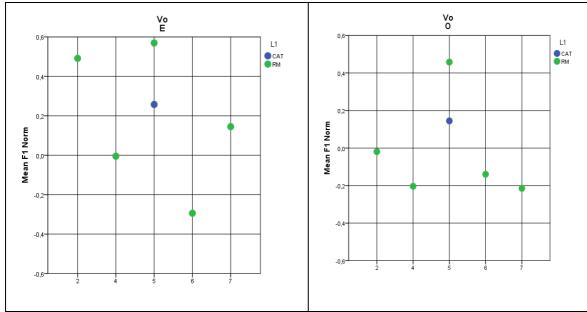
La connexió de la variable social gènere amb els valors del primer formant de les vocals [ɛ] i [ɔ] permet determinar que les noies amb L1 català presenten un F1 molt elevat, fins i tot més alt que el dels nois del mateix grup lingüístic d'origen, i són les que emeten les variants vocàliques més obertes de tots els informants. Al marge d'això, tal com es mostra en les figures 10 i 11, els nois amb L1 romanès i català presenten un grau similar d'obertura en les vocals mitjanes baixes.



4.5 Llengües d'ús

La llengua d'ús amb els amics ha resultat un factor important a l'hora de determinar les variacions en l'F1 de les vocals [ɛ] i [ɔ] dels grups de parlants L1 romanès i català. I és que quan el català és una de les llengües utilitzades en les relacions d'amistat, els valors d'F1 de les vocals [ɛ] i [ɔ] tendeixen a augmentar significativament (vegeu els gràfics 9 a 17). En alguns casos d'informants L1 romanès, els valors d'F1 de [ɛ] i [ɔ] són fins i tot més alts que els dels informants L1 català (vegeu les figures 12 a 15). En aquest sentit, doncs, es pot afirmar que en els informants L1 romanès que més utilitzen català hi ha un acostament a les realitzacions fòniques dels informants L1 català i, en aquest cas concret, els exemplars d'aquests parlants (en termes de Pierrehumbert, 2001) tendeixen a uniformitzar-se amb els que emeten els informants L1 català. Un cop més, l'ús de la llengua i la interacció cara a cara entre els parlants té un paper molt important en l'acomodació lingüística (Trudgill, 1989).

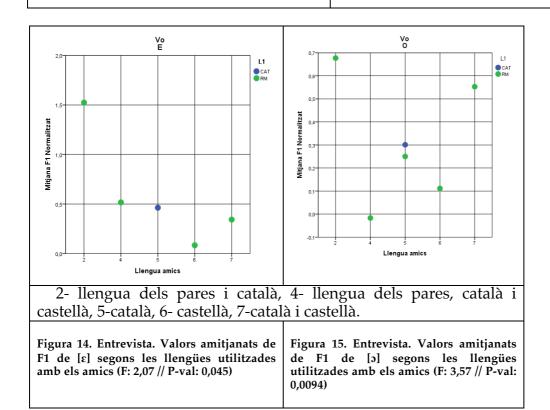
⁴El temps de residència a la ciutat de Lleida, contràriament al que s'esperava, no ha estat una variable significativa, paral·lelament al que s'ha observat en altres estudis (veg. Oller 2013).



2- llengua dels pares i català, 4- llengua dels pares, català i castellà, 5-català, 6-castellà, 7-català i castellà.

Figura 12. Lectura. Valors amitjanats de F1 de $[\epsilon]$ segons les llengües utilitzades amb els amics (F = 12.02 // p = 0.0000)

Figura 13. Lectura. Valors amitjanats de F1 de [\mathfrak{d}] segons les llengües utilitzades amb els amics (F = 11.94 // p = 0.0000)



5 Observacions generals

Les vocals del lleidatà emeses per parlants amb L1 català i romanès són força similars exceptuant les vocals mitjanes [ɛ] i [ɔ]. La diferència en la pronunciació

de [ɛ] i [ɔ] entre tots els grups de parlants es troba en el grau d'obertura d'aquestes vocals (F1), de manera que les vocals emeses per parlants amb L1 català són més obertes que les emeses per parlants L1 romanès. A més a més, les distàncies entre l'F1 de les parelles [ɛe] i [ɔo] dels parlants L1 català són més grans que les dels altres enquestats.

Les diferències en l'obertura de les vocals mitjanes es poden explicar a partir de diversos factors:

- 1) Factors intralingüístics: interdialectalment, les vocals mitjanes baixes del català són més variables que les vocals altes o que la vocal baixa (Carrera-Sabaté & Fernández-Planas, 2005; Recasens i Espinosa, 2006); a més, els contrastos entre [ɛe] i [ɔo] són menys productius que d'altres en què es combinen amb vocals altes (vegeu exemples aplicats a la percepció del català oriental a Mora *et al.*, 2010). Paral·lelament a aquestes observacions, les dades recollides en aquest estudi permeten observar variació en les vocals mitjanes de tots els enquestats. Cal tenir en compte, però, tal com es dóna en diferents dialectes del català (vegeu Recasens & Espinosa, 2006), que totes les vocals dels tres grups de parlants s'han emès dins d'unes distàncies similars en les vocals extremes ([i], [a] i [u]), cosa que confereix estabilitat al sistema vocàlic lleidatà, independentment de la llengua d'origen dels parlants.
- 2) Factors interlingüístics: les vocals del sistema fonològic del romanès són diferents del català perquè no tenen les vocals $/\epsilon/i$ / $\sigma/$. A més a més, cal tenir en compte que els usuaris d'aquestes llengües es troben en contacte permanent amb una tercera llengua: el castellà, que tampoc no compta amb els fonemes $/\epsilon/i$ / $\sigma/$. En aquest sentit, el tret més marcat de les vocals mitjanes del català respecte del romanès i castellà (vegeu Armengual, 2014) facilita que el F1 de $/\epsilon/i$ / $\sigma/$ no sigui tan elevat en informants L1 romanesos com en L1 catalans.
- 3) Factors situacionals i socials: no s'observen diferències significatives entre els valors de F1 obtinguts en la lectura d'un conte i en una entrevista semidirigida i tampoc han estat significatius els anys de residència a la ciutat de Lleida, contràriament al que es preveia inicialment. En canvi, les dones i els homes analitzats presenten actuacions diferenciades segons cada L1, seguint els postulats de la literatura variacionista. El contacte amb la llengua catalana en els àmbits dels amics aporta llum en les diferències d'F1 de les vocals mitjanes obertes: els adolescents que tenen L1 romanès i que usen el català amb els amics tendeixen a emetre les vocals [ɛ] i [ɔ] amb un F1 més elevat que els altres. En resum, el major contacte dels parlants L1 romanès amb la llengua catalana provoca que els seus exemplars es vagin modificant en la mateixa direcció que es troben les emissions dels L1 catalans. Aquesta és, doncs, una prova més de la importància de les interaccions dels parlants per a l'acomodació lingüística i també de la variabilitat de la llengua com a sistema adaptatiu complex als usuaris i als contextos d'ús.

6 Referències bibliogràfiques

Ajuntament de lleida (2005). *Informe sobre la immigració a Lleida durant l'any* 2004. Recuperat el 21 de 3 de 2014, de http://www.paeria.es/cat/ajuntame nt/noticies.asp?Detall=True&IdNotic ia=3025

Alazard, C., Astésano, C., & Billières, M. (2011). *Multiphonia (multimodal database of phonetics teaching methods in classroom interactions*). Recuperat el 21 de 3 de 2014, de http://sldr.org/sldr000780/fr

- Amengual, M. (2014). the perception and production of language-specific midvowel contrasts: shifting the focus to the bilingual individual in early language input conditions. *International journal of bilingualism* 1-20. Recuperat el 20 de 9 de 2014, de http://ijb.sagepub.com/content/early/2014/08/12/1367006914544988
- Beckner, C., Blythe, R., Bybee, J., Christiansen, M., Croft, W., Ellis, N., Holland, J., Ke, J., Larsen-Freeman, D. & Schoenemann, T. (2009). Language is a complex adaptive system. Language Learning 59, 1, 1-26.
- Brown, C. (2000). The interrelation between speech perception and phonological acquisition from infant to adult. En J. Archibald (ed.), Second language acquisition and linguistic therory. Oxford: Blackwell, 4-63.
- Carbonell, J., & Llisterri, J. (1999). Catalan. En *Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press, 62
- Carrera-Sabaté, J., & Fernández-Planas, A. M. (2005). *Vocals mitjanes tòniques del català. estudi contrastiu interdialectal.* Barcelona: Horsori.
- Celata, C., & Calamai, S. (2014). Sociophonetic perspectives on language variation. En *Advances in Sociophonetics*. John Benjamins. Amsterdam/Philadelphia, 1-13.
- Chiţoran, I. (2001). The Phonology of Romanian: A Constraint-Based Approach. Berlin & Nova York: Mouton de Gruyter.
- Dehaene, S. (2007). *Les neurones de la lecture*. Odile Jacob.
- Ferrari, S. (1980). Evaluation of vowel normalization procedures. *Journal of the Acoustical Society of America*, 67, 1, 253-261.
- Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. En W. Strange (ed.) Speech perception and linguistic experience: Issues in cross-linguistic research, Timonium, MD: York Press, 233-277.
- Huber, J. E., Stathopoulos, E. T., Curione, G. M., Ash, T. A. & Johnson, K. (1999). Formants of children, women, and men: The effects of vocal intensity variation. *Journal of the Acoustical Society of America*, 106, 3, 1532-1542.
- Kirkham, S., & moore, E. (2013). Adolescence. En: J.K. Chambers & Natalie Schilling (eds) *The Handbook*

- of Language Variation and Change. Oxford: Wiley-Blackwell, 277-296.
- Labov, W. (1966). *The social stratification of English in New York City*. Washington D. C.: Center for Applied Linguistics.
- Labov, W. (2001). Principles of Linguistic Change. Vol. 2: Social Factors. Oxford: Blackwell.
- Maddieson, I. (1984). *Patterns of sounds*. Cambridge: Cambridge University Press.
- Major, R. C. (2008). Transfer in second language phonology. En H. Edwards, G. Jette & M. L. Zampini (eds.) *Phonology and second language acquisition*. Amsterdam: John Benjamins, 63-94.
- Martinez-Celdrán, E., Fernández-Planas, A.M., & Carrera-Sabaté, J. (2003). Castilian Spanish. *JIPA* 33(2), 255-259
- Mora, J. C., Keidel, J. L., & Flege, J. E. (2011). Why are the Catalan contrasts between /e/-/ε/ and /o/-/ο/ so difficult for even early Spanish-Catalan bilinguals to perceive? En M. Wrembel, M. Kul, & K. Dziubalska-Kołaczyk, Achievements and perspectives in the acquisition of second language speech: New Sounds 2010. Volume II. Frankfurt am Main: Peter Lang, 183-193.
- Nakadaa, T.; Fujiia, Y., & Kweeb, I L. (2001).

 Brain strategies for reading in the second language are determined by the first Language. *Neuroscience Research*, 40, 4, 351-358
- Oller, J. (2013). Variables que incideixen en el coneixement de català i castellà de l'alumnat estranger a Catalunya: un estudi amb alumnat de sisè de primària. En F. X. Vila & E. Salvat (eds.) *Noves immigracions i llengües*. Barcelona: MRP & Universitat de Barcelona, 77-95.
- Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition, and contrast. En J. L. Bybee & P. Hopper (eds.), Frequency and the emergence of linguistic structure. Amsterdam: John Benjamins, 137-157
- Rafel, J. (1980). Dades sobre la freqüència de les unitats fonològiques del català. En Estudis universitaris catalans XXIV. Estudis de llengua i literatura catalanes oferts a R. Aramon i Serra en el seu setantè aniversari (2). Barcelona: Curial, 473-496.
- Recasens, D. (2008). Mètodes de normalització i de representació de

- dades acústiques i articulatòries. *Estudios de Fonética Experimental*, XVII, 331-341.
- Recasens, D., & Espinosa, A. (2006).

 Dispersion and variability of Catalan vowels. *Speech Communication*, 48, 645-666.
- Sebastián-Gallés, N., & Díaz, B. (2012). First and second language speech perception: Graded learning. Language Learning, 62, 131–147.
- Sebastián-Gallés, N., Echeverría, S., & Bosch, L. (2005). The influence of initial exposure on lexical representation: comparing early and simultaneous bilinguals. *Journal of Memory and Language*, 52, 240-255.
- Tokimura, H., Ridding, M. V., Tokimura, Y., Amassian, V. E., & Rothwell, J. C. (1996). Short latency facilitation between pairs of threshold magnetic stimuli applied to human motor cortex. *Electroencephalography and Clinical Neurophysiology*, 101, 263-272
- Trudgill, P. (1989). Contact and isolation in linguistic change. En L. Earl Breivik & E. H. Jahr (eds.) Language change: Contributions to the Study of Its Causes. Berlin: Mouton de Gruyter, 227-237.

Effects of place of articulation on Spanish coda /s/ perception

DEL SAZ, MARÍA

1 Introduction

Speech perception is a process affected by numerous factors that should be taken into account. Some of these factors derive from the listeners while others concern the stimuli used in the experiments. Listeners' attributes encompass the type of listener, their age and age of learning a second or foreign language, their first language (L1) and their experience with the second language (L2). Stimuli's characteristics include the psychoacustic salience of the sounds, the variability found in the signal, such as intra-speaker and inter-speaker variability, the specific characteristics of the listeners' L1 or L2 dialect, and the possible allophonic changes due to phonetic context.

Well-known is the study by Repp (1982) in which it was proven that /r/ and /1/ were perceived differently depending on whether the following sound was /d/ or /g/. Before the alveolar stop, listeners tended to perceive /1/, while they perceived /r/ before the velar stop. Along these lines, Mann & Repp

(1980) also studied the perception of /s/ and $/\int/$ before voiceless stops /t/ and /k/. They found that /s/ was best perceived before /k/, while $/\int/$ was best perceived before /t/, and viceversa: each voiceless stop was identified best depending on the preceding sibilant.

This also holds true for L2 speech perception, for example, in terms of allophonic changes due to phonetic context. L1 American English listeners in Strange, Bohn, Nishi, and Trent's study (2005) found L2 German [y] to be a poorer example of American English [u:] than German [u:] in citation-form /hVp/ context, while in sentence-embedded /bVp/, /dVp/ and /gVk/ contexts, both German sounds were seen as good exemplars of L1 [u:], most likely because American English back rounded vowels are fronted in these contexts and thus become more similar to German front rounded vowels.

Indeed, differences in temporal and spectral parameters found in place and manner of articulation or voice contrasts can affect the acoustic characteristics of the sounds and, thus, perception (Strange & Shafer, 2008).

1.1 Spanish coda /s/

The center of this study, Spanish coda /s/, encompasses two of the possible realizations of this coda: aspiration and sibilance. While sibilance is regarded as the mainstream variety, aspirating varieties are actually found in Spain and countries in South and Central America. This study focuses on two varieties of Peninsular Spanish: Castilian (sibilance) and Andalusian (aspiration).

While Spanish aspiration of coda /s/ has been described extensively, the perception of such phenomenon has been studied to a lesser extent. As an L2, Schmidt (2009) studied the perception of Dominican Spanish by listeners of L1 American English before and after a three-week study abroad program in the Dominican Republic. Even when improvement was shown after their stay abroad, the mainstream features of Spanish still rendered better identification results than the variant-specific characteristics, such as aspiration. As an L1, Schmidt (2011) showed that native speakers of an aspirating dialect identified Argentininan and Venezuelan aspiration of coda /s/ more accurately than native speakers of a non-aspirating dialect. In terms of Andalusian Spanish as L1, O'Neill (2005) demonstrated that native speakers of this variant were unable to determine whether and underlying /s/ was present in absolute sentence-final position, given to the neutralization of aspiration in this context.

In this respect, this present study analyzes the identification of L2 Spanish coda /s/ as a morphological marker in plural nouns and second-person verbs in the present tense, from stimuli recorded by Western Andalusian Spanish and Castilian Spanish speakers. One of the main differences between these two variants of Spanish is that the former aspirates syllable- and word-final /s/, while the latter maintains sibilance.

The most typical characteristic of Andalusian Spanish, also common to other varieties of Spanish, is the aspiration of coda /s/. Nevertheless, what seems to distinguish Andalusian Spanish further from other varieties in terms of aspiration is the effect that this aspiration of coda /s/ has on the following sounds. In particular, and relevant for this study, is its effect on voiceless and voiced stops. To begin with, Spanish voiced stops are uttered as such in initial position after a pause and after nasals /m, n/ (also after /l/ in the case of /d/). In any other context they become approximants (Quilis, 1993), usually of less duration and higher intensity than stops and fricatives (Martínez Celdrán,

2013). Thus, the effect of aspiration on these approximants makes them become fricative sounds (Alvar, 1996). In terms of voiceless stops, as pointed out by Torreira (2007), the main difference lies in a longer post-aspiration and shorter pre-aspiration and closure found in Andalusian Spanish (at least in comparison with Porteño and Puerto Rican Spanish). Therefore, Andalusian Spanish aspiration is not only a matter of $/s/\rightarrow$ [h], but rather a more complex phenomenon, summarized in Table 1 below.

	sibilance	aspiration
/b/	[β]	[v]-[f]
/d/	[ð]	[δ]-[θ]
/g/	[γ]	[x]-[h]
/p/	[p]	[p ^h]
/t/	[t]	[th]-[[ts]1
/k/	[k]	[k ^h]

Tabla 1. Spanish allophones of voiced and voiceless stops after sibilance and aspiration

Previously, Del Saz (2014) reported the identification of Western Andalusian Spanish aspiration before the three voiceless stops by L2 Spanish learners. In this case, aspiration is transferred to the following sounds, turning unaspirated voiceless stops [p, t, k] into aspirated ones [p^h, t^h, k^h] (Torreira, 2007). The author found that L1 American English listeners identified L2 target words when followed by the dental voiceless stop significantly better than before the velar and the bilabial stops. After an acoustic analysis, it was concluded that the length of the preceding vowel was a key factor to identification; the dental voiceless stop presented the longest duration and, therefore, the longest preaspiration.

In this current study, we report our findings for the same set of L2 listeners according to not only their identification of sibilance before the voiceless stops, but also their identification of aspiration and sibilance before the three voiced stops.

2 Methods

2.1 Participants

Participants were 84 General American English listeners ($M_{\rm age}=20.36$), learners of Spanish as a foreign language in US universities during the spring semester of 2013. They were unfamiliar with Spanish aspiration of coda /s/ and had not stayed in a Spanish-speaking country for over 3 months. They also reported no speech or hearing disorder.

2.2 Stimuli

Stimuli consisted of 80 sentences with embedded target words that were followed by ten different phonetic contexts. Target words were nouns and verbs, among which were those that ended in /s/, i.e. plural nouns and second-person verbs in the present simple, and their counterparts, i.e. singular nouns and third-person verbs ending in vowel. These words were followed by others

¹New affricate sound described in Moya Corral (2007) and Ruch (2010), not included in the identification task of this study.

beginning with one of the following sounds: /p, t, k, b, d, g, m, n, 1/ and vowel. For each phonetic context, two sentences were produced. See below examples of sentences with nouns and verbs ending in /s/, followed by /d/:

Digo **leches** demasiado lento Digo **bocas** demasiado lento **Comes** demasiado deprisa **Estás** decorando nuestra casa

For the purposes of this article, only results concerning target words followed by voiced and voiceless stops will be reported.

2.3 Procedure

The 80 sentences were recorded by two speakers of Castilian Spanish (one male, one female) and two speakers of Andalusian Spanish (one male, one female) at the Phonetics Laboratory of the University of Seville, under the presence of the experimenter. The recording was done at a 44.1 kHz and 16 bps sampling rate in a sound-proof booth, using a Marantz Professional PMD671 solid-state recorder and a Shure SM48 microphone. Sentences were then normalized to -6dB and set to 70dB with Adobe Audition CS5.5. Speakers were instructed to speak at a normal conversational rate, as if they were talking to a friend. The resulting 320 sentences (80 sentences x 4 speakers) were then included in a test specially designed for the study, mounted in the author's university webpage.

Experiments were run in computer rooms with individual stations and headphones. Participants first signed an informed consent form and filled out two background questionnaires. Subsequently, they received oral and written instructions in English that they would listen to sentences in Spanish and would need to select the one they heard from the two forced-choice written options given. Each participant listened to a different set of 60 sentences randomly chosen from the corpus, with no feedback provided. As a training method, the test played five sentences and showed the correct answer, so that they became familiar with the task and could adjust volume settings. Participants had to listen to each sentence twice in order to proceed to the next one, and were allowed unlimited time to complete the test, although the average time of completion was 20 minutes.

Following Strange and Shafer (2008), a justification for the use of this experimental design is provided. The use of an identification task, instead of categorization or discrimination tasks, ensures that the stimuli is "compared against internal representations of phonetic/phonological categories" (p. 183). The use of real words instead of nonsense items, despite the fact that lexical and phonetic effects might be confounded, ensued from the assertion that "experiments using real world materials more accurately reflect the receptive problems of L2 learners" (p. 163). Embedding target words in sentence contexts instead of using citation-form stimuli avoids the problem of the results not being easily generalizable to real-world situations of understanding continuous speech. The fact that stimuli were not conversational per se, but rather read in a conversational tone, still can render results that "may be generalized to some real-world situations, such as the language classroom" (p. 163).

2.3.1 Acoustic analysis

Acoustic analyses were carried out to determine stimuli characteristics that may have influenced the outcome of the identification rates according to place of articulation. For Andalusian stimuli, we measured the duration of preaspiration before voiceless stops and fricatives, as well as post-aspiration (Voice Onset Time, VOT) found in voiceless stops. Likewise, we determined the harmonics-to-noise ratio (HNR) to corroborate the amount of pre-aspiration found before these contexts. For Castilian stimuli, we measured the duration and intensity of the sibilants before voiceless stops and approximants, as well as spectral information such as center of gravity (COG), i.e., the mean concentration of energy in a sound, and dispersion, i.e., the distance with respect to the COG (standard deviation).

3 Results

Word-final /s/ is phonotactically possible both in Spanish and English, and also has a similar function as a morphological marker for plural nouns and verb person in the present. Word-final /h/ is not phonotactically possible in English and is not a possible realization of /s/, while the opposite is true for aspirating dialects of Spanish. Furthermore, the effects of aspiration on the following stops produce allophones that are actually phonemes in English, but again, never as an indication of an underlying /s/. This, together with a lack of familiarity with this particular feature, accounts for the mismatch in identification accuracy between Andalusian and Castilian Spanish stimuli.

Listeners identified Andalusian aspiration as coda /s/ with less than 15% of accuracy, while they identified Castilian sibilance with more than 90% of accuracy. Nevertheless, in both cases phonetic context played a role in identification, particularly as far as the (inter)dental context is concerned, as we will see in the following sections.

3.1 Aspiration and sibilance before voiceless stops

Despite the disparity in the overall identification accuracy of coda /s/ between the two Spanish variants, there seems to be some indication that the place of articulation of the following phonetic context plays a role in speech perception. As shown in Table 2, aspiration was significantly best identified as coda /s/ before the dental voiceless stop [tʰ] ($\chi^2(2) = 100.971$; p <.001), while the opposite was true for the identification of sibilance as coda /s/ ($\chi^2(2) = 178.043$; p <.001). In this case, /s/ rendered the highest results before the velar stop [k], and the lowest before the dental stop [t].

	aspiration		sibilance	
/p/	[p ^h]	8.55 (5.07)	[p]	93.02 (2.63)
/t/	[t ^h]	14.26 (5.42)	[t]	92.55 (3.60)
/k/	[k ^h]	11.83 (7.26)	[k]	96.76 (2.00)

Table 2. Mean values (SD) of identification of aspiration and sibilance as coda /s/ before voiceless stops

At this point, it is convenient to revisit Mann and Repp's study (1980) mentioned in the introduction section. Their findings showed that stop identification was affected by the preceding sibilant fricative at the same time as fricative sibilant perception was affected by the following stop, in a reciprocal

effect. As the authors point out, there exist coarticulatory effects the lead to a listener's "preference for combinations of fricatives and stops with opposite polarities on the place-of-articulation dimension" (p. 235).

3.2 Aspiration and sibilance before voiced stops

Before going deeper into this section, it is convenient to remember that, while phonologically we are dealing with voiced stops, phonetically these sounds are voiced approximants in Castilian Spanish and fricatives after Andalusian Spanish aspiration. The main difference between these approximants and fricatives lies in their duration, which is longer for the fricatives, and in their intensity, which is lower for the fricatives.

Table 3 below shows the mean identification rates and standard deviations of aspiration and sibilance as coda /s/ before the three voiced stops, with a similar outcome to that found in the context of voiceless stops. Aspiration was significantly best identified before the interdental fricative [δ] ($\chi^2(2) = 272.145$, p <.001), while sibilance was significantly less identified before the dental approximant [δ] ($\chi^2(2) = 185.136$, p <.001) than before the other two sounds. Once more, we see opposite tendencies that may be related to coarticulatory effects, especially in the context of (inter)dental sounds.

	aspiration		sibilance	
/b/	[v]	17.97 (5.65)	[ß̞]	97.31 (2.75)
/d/	[ð]	30.22 (11.06)	[ð]	85.80 (11.13)
/g/	[x]	9.08 (5.37)	[γ]	97.26 (3.99)

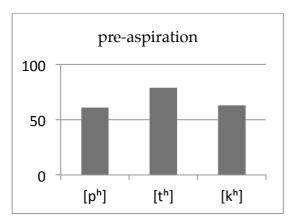
Table 3. Mean values (SD) of identification of aspiration and sibilance as coda /s/ before Andalusian fricatives and Castilian approximants

3.3 Acoustic characteristics

In an attempt to find acoustic characteristics of the stimuli that may implement these results, especially concerning the (inter)dental contexts, we conducted acoustic analyses of aspiration and sibilance in relation to the phonetic context that followed these codas.

3.3.1 Pre- and post-aspiration

Measurements of pre-aspiration values were carried out from the onset of the preceding vowel to the beginning of the stop closure, in the case of voiceless stops, and the beginning of the fricative, in the case of fricatized approximants. Figure 1 shows that pre-aspiration was significantly longer before [th] ($\chi^2(2) = 6.03$, p <.05) than before [kh] and [ph], and longer before [ð] than the two other fricative sounds, although not significantly so.



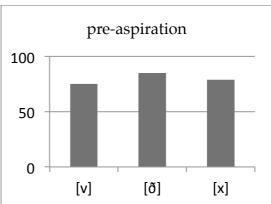


Figure 1. Pre-aspiration mean values before voiceless and voiced stops

To complement duration HNR was also calculated, i.e., the amount of preaspiration in relation to the preceding vowel. Table 4 shows that, indeed, HNR was lower before $[t^h]$ and $[\check{o}]$, indicating a higher presence of aspiration (noise) with respect to vowel (harmonics). Likewise, HRN was lower before stops than before fricative sounds.

[p ^h]	13.2 (1.05)	[v]	16.4 (0.81)
[t ^h]	7.3 (1.14)	[ð]	13.0 (3.12)
$[k^h]$	8.1 (1.07)	[x]	15.3 (0.62)

Table 4. HNR mean values (SD) before voiceless stops and fricative sounds

In correlation with the above said, post-aspiration, or VOT, was measured from the beginning of the stop burst to the beginning of the following vowel. Although post-aspiration in [th] was slightly shorter than that in the other stops, there were no statistically significant differences between the three sounds in this case (Figure 2).

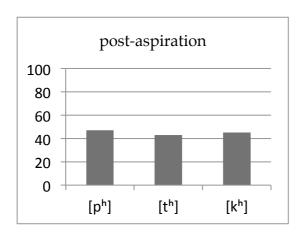


Figure 2. Post-aspiration (VOT) mean values in voiceless stops

3.3.2 Spectral information

To measure the acoustic cues of /s/, apart from accounting for intensity and duration, we also measured its center of gravity (COG) and dispersion. For this purpose, we used a script to be run by Praat (Boersma & Weenink, 2009), adapted by the technician at the Laboratory of Phonetics of the University of Seville. Table 5 shows the values for [s] before the three voiceless stops. No

statistically significant differences were found between the three phonetic contexts, except for COG, which was higher for [s] before the dental stop ($\chi^2(2)$ = 5.915, p = .05).

	[k]	[p]	[t]
intensity	68.836 (1.487)	66.493 (3.551)	67.000 (2.154)
duration	62.60 (14.40)	55.25 (9.46)	56.25 (11.76)
COG	5315.246 (487.152)	4959.341 (532.251)	5732.890 (781.833)
dispersion	1735.012 (255.352)	1908.128 (496.308)	1992.983 (192.913)

Table 5. Spectral information mean values (SD) of [s] before voiceless stops

In a similar manner, the same values were calculated for [z] before the three voiced approximants. Once more, the only statistically significant value found was COG, which was higher for [z] before the dental approximant ($\chi^2(2) = 6.695$, p < .05).

	[ß]	[ð]	[y]
intensity	66.693 (3.669)	69.908 (2.915)	69.475 (3.567)
duration	66.83 (17.34)	61.75 (22.70)	62.51 (26.170)
COG	4958.090 (1059.328)	5742.061 (1915.487)	4461.001 (761.350)
dispersion	1896.014 (376.511)	1991.298 (511.699)	1844.090 (273.432)

Table 6. Spectral information mean values (SD) of [z] before approximants

3.4 Discussion

The acoustic analyses above lead to two significant findings. One the one hand, pre-aspiration was longer before dental stop [tʰ] and interdental fricative [ð]. On the other hand, the COG of [s] and [z] was higher before the dental stop [t] and the dental approximant [ð], respectively, providing evidence that both voiced and voiceless sibilants become dental in such contexts. Given that our articulatory system tends to produce sounds as similar as possible, and that our perceptive system needs sounds to be as distinguishable as possible (Marrero, 2001), the opposite tendencies between Andalusian and Castilian Spanish seem to be in agreement with claims that the perception of speech involves the direct perception of articulatory gestures (Best & Tyler, 2007; Strange, 2011). The more similar the place of articulation, the less indistinguishable sounds are.

As Ohala (1993) pointed out, the listener plays a crucial role in sound change. In the context of dental sounds, this may be the reason for the development of a new affricate sound [t³] in Andalusian Spanish (Moya Corral, 2007; Ruch, 2010), as an avolution from pre-aspiration to post-aspiration before the dental stop. In fact, while most of the examples from Andalusian aspiration before the dental stop presented pre-aspiration and post-aspiration, on a few occasions we observed the appearance of this new affricate sound in non-target words. Figures 3, 4, and 5 below show waveforms for Castilian Spanish and Andalusian Spanish pronunciations of the word "está" (he/she/it is).

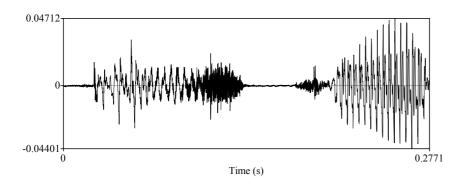


Figure 3. Waveform of Castilian Spanish "está" [es.tá]

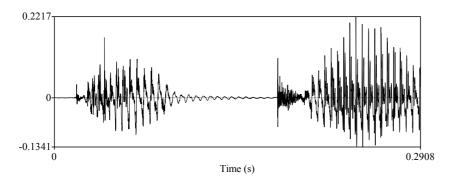


Figure 4. Waveform of Andalusian Spanish "está" [eh.thá]

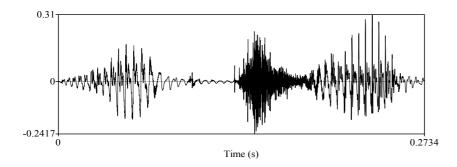


Figure 5. Waveform of Andalusian Spanish "está" [e.tsá]

Torreira (2007) defends a gestural reorganization in which the glottal opening for the aspirated /s/ and the supraglottal closure for the following stop overlap instead of being sequential. Certainly, Andalusian Spanish is evolving in given aspects connected to aspiration (Ruch & Harrington 2014), and the speaker as listener seems to have an important role in these changes.

4 Conclusions

In this study we investigated the effect of phonetic context on the identification of aspiration and of sibilance as Spanish coda /s/. The choice of L2 listeners unfamiliar with aspiration provided insight as to what articulatory and acoustic characteristics of the context may aid in the identification of this coda. A combination of both types of characteristics seems to be at play, particularly connected to (inter)dental contexts.

In the case of Andalusian Spanish, its evolution has led O'Neill (2009) to consider aspirated stops $[p^h, t^h, k^h]$ as part of a new set of phonemes working in

opposition to their unaspirated counterparts. Instead of being the result of an overlap of gestures, as proposed by Torreira (2007), these set of sounds would be phonetic categories in itself and not the result of coarticulatory gestures. While Parrell (2012) corroborates the post-aspiration phenomenon in Andalusian Spanish, he claims that it is still to soon to determine whether this process is phonetic or phonological.

As this is an ongoing process, more research is still needed to examine the diachronic and synchronic aspects of these developments, particularly in relation to their effect on speech perception.

5 References

- Alvar, M. (1996). Manual de Dialectología Hispánica. El Español de España. Barcelona: Ariel.
- Best, C. T. & Tyler, M. D. (2007). Nonnative and second-language speech perception. Commonalities and complementarities. In O-S. Bohn & M. J. Munro (Eds.), Language Experience in Second Language Speech Learning: In Honor of James Emil Flege, 31-52. Amsterdam, NLD: John Benjamins Publishing Company.
- Boersma, P. & Weenink, P. (2009). Praat: Doing Phonetics by Computer (Version 5.2.46). Retrieved from http://www.praat.org/
- Del Saz, M. The Perception of Western Andalusian Spanish Aspirated Stops by General American English Listeners. Paper presented at the International Workshop on Cross-Language Speech Perception (CEHUM 2014), Braga, Portugal, January 30-31, 2014.
- Mann, V. E., & Repp, B. H. (1980). Influence of preceding fricative on stop consonant perception. Journal of the Acoustical Society of America, 69(2), 548-558.
- Marrero, V. (2001). Fonética y Fonología. Addenda. Madrid: Universidad Nacional de Educación a Distancia.
- Martínez Celdrán, E. (2013). Caracterización Acústica de las Aproximantes Espirantes en Español. Estudios de Fonética Experimental, 11-35.
- Moya Corral, J. A. (2007). Noticia de un sonido emergente: La africada dental procedente del grupo -st- en Andalucía. Revista de Filología de La Universidad de La Laguna, 25, 457-465.
- Ohala, J. J. (1993). Sound change as nature's speech perception experiment. Speech Communication, 13, 155-161.

- O'Neill, P. (2005). Utterance final /s/ in Andalusian Spanish. The phonetic neutralization of a phonological contrast. Language Design: Journal of Theoretical and Experimental Linguistics, 7, 151-166.
- O'Neill, P. (2009). The effect of s-aspiration on occlusives in Andalusian Spanish. Oxford University Working Papers in Linguistics, Philology & Phonetics, 12, 73-86.
- Parrell, B. (2012). The role of gestural phasing in Western Andalusian Spanish aspiration. Journal of Phonetics, 40(1), 37-45.
- Quilis, A. (1993). Tratado de Fonología y Fonética Españolas. Madrid: Gredos.
- Repp, B. H. (1982). Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception. Psychological Bulletin, 92, 81-110.
- Ruch, H. (2010) Affrication of /st/-clusters in Western Andalusian Spanish: variation and change from a sociophonetic point of view. Proceedings of the Workshop "Sociophonetics, at the crossroads of speech variation, processing and communication". Pisa, December 14th-15th, 2010.
- Ruch, H., & Harrington, J. (2014). Synchronic and diachronic factors in the change from pre-aspiration to post-aspiration in Andalusian Spanish. Journal of Phonetics, 45, 12-25.
- Schmidt, L. B. (2009). The effect of dialect familiarity via a study abroad experience on L2 comprehension of Spanish. In J. Collentine et al. (Eds.), Selected Proceedings of the 11th Hispanic Linguistics Symposium, (pp. 143-154). Somerville, MA: Cascadilla Proceedings Project.
- Schmidt, L. B. (2011). Acquisition of Dialectal Variation in a Second

- Language: L2 perception of aspiration of Spanish /s/. (Unpublished doctoral dissertation). Indiana University, Bloomington.
- Strange, W. (2011). Automatic selective perception (ASP) of first and second language speech: A working model. Journal of Phonetics, 36(1), 141-157.
- Strange, W., Bohn, O.-S., Nishi, K., & Trent, S. A. (2005). Contextual variation in the acoustic and perceptual similarity of North German and American English vowels. Journal of the Acoustical Society of America, 118, 1751-1762.
- Strange, W. & Shafer, V. L. (2008) Speech perception in second language learners. In J. G. Hansen Edwards & M. L. Zampini (Eds.), Phonology and Second Language Acquisition, (pp. 153-191). John Benjamins: Amsterdam.
- Torreira, F. (2007). Pre- and post-aspirated stops in Andalusian Spanish. In P. Prieto, J. Mascaró & M. J. Solé (Eds.), Prosodic and Segmental Issues in Romance, 67-82. Amsterdam: John Benjamins.

Estudio del cambio lingüístico en el sistema vocálico del habla pasiega

MORENO CARRAL, MANUEL

1 Introducción

El propósito de este artículo es explorar el estado actual de la modalidad lingüística propia del área de la Pasieguería, en la zona centro oriental de Cantabria. Específicamente nos enfocaremos en el estado de conservación de su sistema vocálico, describiéndolo mediante el uso de la herramienta de análisis fonológico computerizado más habitual, Praat

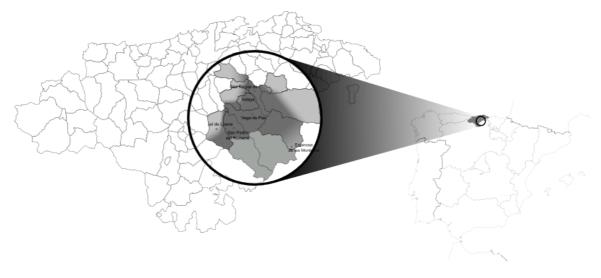


Figura 1. Localización de la pasieguería en Cantabria/Burgos y en España (mapa derecho)

2 Los pasiegos

Nuestro estudio se enfoca en la variedad vernácula del área conocida como "La Pasieguería", esta variedad recibe el nombre de pasiego (en el propio dialecto: pasiegu o pasiigu).

Esta zona se extiende por ambas vertientes de la zona más oriental de la cordillera cantábrica. En la vertiente cantábrica se extiende por un territorio comprenido por las cabeceras de los ríos Pas, Pisueña y Miera; mientras que en la vertiente burgalesa corresponde a la zona de las Machorras, entre los ríos Trueba y Lunada, que desaguan en el Ebro. Se trata de una zona eminentemente rural, mal comunicada y con una población que no supera los 5 000 habitantes. Estos, los pasiegos, se dedican fundamentalmente a la ganadería de vacuno para producción de leche. En este respecto es conocida la cultura pasiega por haber desarrollado desde el siglo XV uno de los primeros sistemas de explotación intensiva del terreno, *la muda*, basado en la transterminancia. Debido a que el ganado lácteo requiere de ordeños diarios, los pasiegos debían acompañar al ganado, mudándose hasta una docena de veces al año. Esto, junto con la complicada orografía del terreno, impedía en muchos casos la escolarización de los niños, contribuyendo hasta tiempos al aislamiento de este territorio y, sin duda, en lo conservador de su dialectología.

2.1 Repaso a la bibliografía

El pasiegu fue descrito por primera vez de manera sistemática en (Penny, 1970). Esta obra es la referencia para cualquier estudio posterior, pues analiza todos los aspectos del dialecto de una manera amplísima y está basado en una recogida de datos propia que incluye varios miles de palabras. Penny concluyó que el dialecto era "fundamental mente de tipo leonés", sin embargo, también apunta a un "influjo profundo del idioma oficial". Dicho esto, podemos dividir el resto de la bibliografía en dos grandes etapas, antes de los trabajos del profesor Penny, y después.

En el periodo previo existen numerosos autores de la literatura montañesa que dedicaron breves pasajes en sus obras a los pasiegos, ninguna de estas anotaciones son de demasiado interés, pues o bien carecían de rigor, o bien se trataba de juicios del autor², o bien son tan breves que no aportan gran cosa o bien se trataba de diccionarios. Menéndez Pidal fue quien desencadenó los más importantes estudios en este área, puesto que, según él relata, en una excursión por la zona escuchó lo que le parecía una metafonía. Esto le llevó a influir en que el ALPI fuese especialmente denso sobre el centro-oriente de la antigua provincia de Santander. Años después, Pidal dedicó un par de opúsculos al tema de la metafonía y al origen de los pasiegos³ en los que, acertadamente, relacionaba el origen de la metafonía del habla pasiega con los bables del centro de Asturias.

Además, García Lomás (1960) dedicó una obra de cierta extensión a las gentes de la Pasieguería, sin embargo, esta carece de rigor tanto en el ámbito lingüístico⁴ como en la sociología, donde se observa el ideario oficialista de la España franquista.

^{1.} Penny (1970). Conclusiones generales, pág. 382, §§518-20.

² Prejuicios, las más de las veces, debidos fundamentalmente al desconocimiento y al carácter reservado de estas gentes.

³ Véanse Menéndez Pidal (1954, 1959).

⁴ Penny refutó gran parte de sus afirmaciones, como una supuesta metafonía de adjetivos y sustantivos femeninos.

A posteriori de nuestra obra de referencia, existen cantidad de estudios teóricos sobre el sistema vocálico pasiego, todos basados en los datos ofrecidos por Penny, y que presentan ideas para justificar el porqué de los procesos vocálicos del dialecto, especialmente las armonías vocálicas. Baste citar a Hualde (1989) y a Picard (2011), aunque su obra no es relevante para lo que se expone en este artículo.

Además de los estudios derivados de la obra de Penny, en la zona se han realizado estudios originales de ámbito mayor, como F. Juncal (1998), que analizó la metafonía y el neutro de materia en todo el centro-oriente de Cantabria, encontrando formas vestigiales de metafonía en un área mucho más amplia de la que anteriormente se pensaba, que abarca todo el centro-oriente del territorio. Además, Holmquist (1982) llevó a cabo un estudio sociolingüístico en el occidente de Cantabria del que podemos extraer datos importantes, como la distribución del cambio lingüístico en una comunidad similar. Sin embargo, lo fundamental es que Holmquist usa equipo de analisis moderno, y ofrece resultados por formantes de los morfemas de género, esto nos permite cuantificar las diferencias ya detectadas por Penny y confirmar definitivamente la existencia de vocales centrales en la dialectología de Cantabria. A pesar de esto sus datos no extrapolan directamente, puesto que Ucieda, su lugar de estudio, no pertenece al mismo bloque dialectal que la Pasieguería, y el vocalismo átono final es bastante distinto más allá de algunas generalidades.

Por último, de los estudios de ámbito mucho mayor, destacan el Atlas Lingüístico y Etnográfico de Cantabria (Alvar López, 1995), más que por su aportación al conocimiento del dialecto, por reportar datos de campo sobre su lo crítico de su estado de conservación. En la misma dirección apuntan el COSER (Fernández Ordóñez, 2005) y sus entrevistas.

Como nota de interés, cabe mencionar que los entrevistadores a lo largo de los años han destacano la dificultad para poder obtener datos del habla cotidiana de los pasiegos por razones propias a su cultura, como la dispersión de la población y su carácter muy reservado, que hace que normalmente el cambio de registro para con el desconocido sea muy intenso, sin duda fruto de la fuerte diglosia que lleva generaciones establecida en esta zona.

2.2 El sistema vocálico pasiegos

Según el modelo expuesto por Penny, el vocalismo tónico pasiego es prácticamente igual que el castellano, aunque conserva ciertos aspectos propios de las hablas leonesas, como la la metafonía y la presencia de ciertas diptongaciones⁶.

2.2.1 El vocalismo tónico: Metafonía

Penny nos habla de un vocalismo tónico con 8 posibilidaes, las 5 vocales [a e i o u] comunes con el castellano, y otras 3 vocales que él denomina *mixtas* (centrales) y que son [ɐ ɨ ʉ]⁷. Estas aparecen en posición tónica solo si la vocal

⁵ Prescindiremos de describir el sistema de consonantes Para más información sobre este aspecto: Penny (1970). Evolución fonética. Págs 74-94, §59.

⁶ Prescindiremos de describir las diptongaciones. El lector interesado puede encontrar su descripción en Penny (1970)., Evolución fonética. Págs 59-60, §38.

⁷Usando la transcripción proporcionada por el propio autor en Penny (2009). En Penny (1970) se usa el mismo superconjunto del alfabeto fonético de la RFE que usa el ALPI.

final absoluta es una [#], es decir, el morfema nominal de masculino singular. A este proceso de armonización de la tónica con la final cerrada, la tradición lingüística asturleonesa lo denomina *metafonía vocálica*. Además de la metafonía desencadenada por [#] final, existe metafoná por la palatal cerrada [i], como en *isti* < *esti* (este), que no entraremos en describir en este artículo.

Como vemos en la tabla 1, [#] se opondría a [u], morfema de neutro⁸ de material y masculino plural, teniendo un uso equivalente al asturiano central - o-.

Cambio	Dialecto	Singular [+META]	Plural [-META]	Neutro [-META]
	Pasiegu	nilu [ˈni lu]	pelus [ˈpe.lus]	pelu [ˈpe.lu]
e > i	Asturianu	- pilu [ˈpɨ.lʉ]	pelos ['pe.los]	pelo [ˈpe.lo]
	Glosa	pelo	pelos	cabello
	Pasiegu	utru [ˈʉ.tɾʉ]	otrus ['o.trus]	otru [ˈo.tɾu]
o > u	Asturianu	<i>uiru</i> [u .u u]	otros ['o.tros]	otro [ˈo.tɾo]
	Glosa	otro	otros	otro
a > 129	Pasiegu	blancu [ˈblɐŋ.kʉ]	blancus	blancu
a > e9			[ˈblaŋ.kus]	[ˈblaŋ.ku]
a > e	Asturiano	blencu [ˈbleŋ.kʉ]		blanco
a > c	L.l/Ay	·	blancos ['blan.kos]	['blaŋ.ko]
a > 0	Asturiano Sb	bloncu [ˈbloŋ.kʉ]	,	[Diaij.KO]
	Glosa	blanco	blancos	blanco

Tabla 1. Patrones de metafonia en asturianu central y pasiegu.

2.2.2 Vocalismo átono

2.2.2.1 Cierres de vocales átonas finales

El pasiego solo tiene 4 posibilidades para sus vocales átonas finales [3 i u t].

- 1. Se articulan [3] las palabras con una /a/ en su silaba átona final. Por ejemplo, en la morfología nominal. En la morfología verbal, sin embargo, encontramos que si se añade un pronombre en enclisis: acáldalu [aˈkal.da.lu] (recógelo), la /a/ deja de ser final y se articula como [a]. Los clíticos solo se cierran si aparecen en enclisis, acáldala [aˈkal.da.lɜ] (recógela), pero no en próclisis: la acaldó [la a.ka.lˈdo] (la recogió).
- 2. Las vocales anteriores se articulan como [i] siempre que vayan en silaba átona final. *Monti* (monte) *puinti* (puente) *vieni* (viene). Hasta en los más recientes neologismos, como *tupper* ['ta.pir] (tupperware).
- 3. Las vocales posteriores se articulan como [u] siempre que vayan en silaba átona final. También en los neologismos recientes, como *parkinson* [ˈpar.kin.sun] (enfermedad de Parkinson).
- 4. La -*u* de las palabras con género masculino singular se pronuncian con [u].
 - Entre las consecuencias de los puntos 2 y 3 están la existencia de pares que solo se diferencian por su vocal media/cerrada entre los elementos átonos clíticos que pueden ocupar posición final o no (me/mi, te/ti). Estos pares también tienen su función en las posiciones pretónicas, como veremos a continuación.

8 Se llama neutro de materia a la distinción morfológica que hacen los dialectos asturleoneses centro-orientales entre los elementos contables e incontables, y que se manifiesta en la morfología de género los elementos del sintagman nominal.
9 Recientemente localicé en el ALPI la respuesta [kea.ru] en la pregunta 35 (carro) del punto 408, Bustantegua. Serán necesarios más datos para saber si es un vestigio de un modelo pasiego anterior similar al Asturianu L.l/Ay (de Llena y Ayer) (tabla 1).

2.2.2.2 Armonía de pretónicas

La vocal de la sílaba tónica condiciona la selección de las vocales pretónicas dentro del grupo clítico. Si la sílaba tónica contiene una vocal cerrada ([i u]) o un diptongo creciente ([jV wV])¹º, la vocal inmediatamente siguiente (siempre comenzando por el final) solo podrá ser cerrada o abierta (/a/), pero no media. Los ejemplos son abundantes en la conjugación verbal: bebi / bibía (presente y pretérito del verbo beber). En caso de añadir un elemento átono previo, se seleccionará teniendo en cuenta la vocal tónica del verbo: se bebi, si bibía (se bebe, se bebía). La única excepción la encontramos en la vocal abierta [a], que rompe la propagación de este fenómeno, como en dejaría (pero no *dijaría).

Un caso específico de lo anterior se produce si la vocal cerrada es además central, resultado de una metafonía, pues se limita aún más la selección de las pretónicas; en este caso solo se seleccionar de entre el conjunto de centrales [v i u], por ejemplo: rintiru [rin't i.ru] (rentero), afectando también a los elementos preclíticos: il pirru [il 'pi.ru] (el perro). De manera análoga al caso de las no centralizadas, la vocal abierta centralizada [v] detiene la propagación del cierre, pero no así de la centralización, pudiendo ser precedida por [v e i o u], como en el soldàu [el sol'dvu] (el soldado). Por alguna razón que no es bien comprendida, la /e/ no se centraliza en este caso.

2.2.3 Grupos de vocales

Cabe mencionar que el dialecto tiene aversión a los hiatos, que son reducidos en un gran número de ocasiones, bien mediante la inserción de una yod antihiática, como en *ideya* (idea) o *seya* (sea, verbo ser); bien mediante el cierre de una vocal media: *rial* (real), *pial* (pedal y peal); bien mediante un cambio de acento, *habiá* (había).

2.3 Objetivo del estudio

Una vez visto era sistema vocálico pasiego en los años 1960, nos interesa saber cómo ha cambiado. Específicamente este estudio tiene como fin comprobar cómo ha cambiado el uso del fenómeno de la metafonía vocálica, además de describirla de una manera precisa mediante el uso de herramientas de análisis fonético computerizado como Praat, que nos den una fotografía clara del sistema vocálico.

También nos interesa conocer de una manera más genérica la evolución del resto de fenómenos descritos, especialmente los relacionados con el vocalismo átono, que ya en la bibliografía se advierte como dificil de encajar con el modelo propuesto por Penny, especialmente en el terreno de los elementos proclíticos.

3 Metodología

3.1 Selección de participantes

Partimos del conocimiento de que el escenario ideal sería comparar los datos de Penny con nuevos datos obtenidos de las generaciones jóvenes del dialecto, para obtener un intervalo temporal lo más amplio posible, sin embargo, los

¹⁰ Existen en el dialecto 15 diptongos con [j] y [w] que se listan en (Moreno Carral, 2014: 9).

datos se Penny no están publicamente disponibles, por lo que solo podemos usarlos como marco teórico. De esta forma beberemos conseguir datos nuevos de dos estadios del dialecto se parados por el máximo tiempo posible, esto es, de los hablantes de mayor edad y de los jóvenes en edad escolar.

Para este fin seleccionamos dos grupos:

- 13 participantes de las edades más avanzadas, considerando el único criterio de que retuvieran en el mayor grado posible las características conocidas como dialectales. Para ello nos servimos de informantes locales de confianza, que nos ayudan a componer un grupo de 13 participantes, 8 hombres y 5 mujeres; con una edad media de 79,6 años (Máx.: 78 años, mín.: 69 años) y provenientes de toda el área geográfica.
- 2. 35 jóvenes en edad escolar de 4 centros educativos de la comarca, identificados por sus profesores como los más conservadores y de edades entre 8 y 18 años. También se intentó contar con informantes de todo el territorio.

3.2 El experimento

El experimento se realizó en forma de entrevista y cuestionario. Todo el proceso se desarrolló durante mayo de 2014. Los experimentos se grabaron con una grabadora digital Zoom H1, con las mejores especificaciones posibles11, generando más de 50 GB de archivos .wav de audio.

La primera parte consistía en una entrevista guiada, donde se intentaba que los participantes hablasen sobre temas que propiciasen la aparición de los fenómenos comentados12. La duración de las entrevistas fue muy variable ya que los jóvenes tuvieron un tiempo limitado a alrededor de 15 minutos por requerimiento de los centros escolares, mientras que los mayores no dispusieron de límite temporal. En el caso de los participantes mayores fueron de incontable ayuda los informantes que nos acompañaron en la conversación, ayudando a romper la diglosia.

En la segunda parte se realizaba un cuestionario fotográfico, donde se requería al participante nombrar el contenido de 29 fotografías sin ayuda. Este cuestionario alternaba fotografías de objetos, lugares y animales, intentando que las respuestas esperadas alternasen metafonía (final en -u) y ausencia de la misma (final en -3). Se dividía en 8 series, una por cada vocal tónica no metafonética (/a e i o u/) y tres diptongos crecientes (/je we ja/). A modo de ejemplo, la serie para /a/ era: caja, gato, vaca, caballo, esperando obtener [ˈka.xɜ], ['ge.tu], ['ba.k3], [ke'be. Λ u].

3.3 Procesado de datos

La primera parte del procesado de la conversación fue en vivo, anotándose en una hoja de cuestionario cada vez que el entrevistador registraba un fenómeno concreto. En un futuro se planea obtener resultados estadísticos sobre la frecuencia de uso de fenómenos en el habla de los entrevistados, lo que requerirá la transcripción de todo el material recogido.

En el cuestionario fotográfico se usa el software de edición de audio Audacity para recortar las entrevistas del cuestionario y limpiar las palabras, resultando en un archivo .wav de alrededor de dos minutos para cada uno de los 48 participantes. Este archivo contiene las respuestas aisladas. Seguidamente se

¹¹Stereo Sound, Bitrate: 320kbps, Sampling Frequency: 44.1kHz.

¹² Especialmente se buscaba la aparición del neutro de materia, hablando sobre la ganadería láctea (los productos lácteos son incontabes), o sobre el tiempo (viento, calor o frío son incontables).

procesa estos archivos en Praat, separando manualmente¹³ cada una de las vocales de cada una de las palabras, y obteniendo sus formantes primero y segundo mediante un script que los anota en dos *tiles* de un archivo .textgrid.

Tras esto y mediante el uso de un script Python se extraen de los formantes de los archivos textgrid, generando un archivo de valores separados por comas (.csv) que se puede importar a excel, donde calculamos la media y la desviación estandar de todas las vocales. Estos datos sin normalizar se representan en dos gráficas, una para cada grupo.

4 Resultados

4.1 Grupo 1 (mayores)

Los resultados de la conversación guiada muestran que la mayoría de los fenómenos del vocalismo átono se conservan con total regularidad en los hablantes de mayor edad. En este grupo todos los participantes realizaban el cierre de las vocales átonas finales y el de las pretónicas. Sin embargo, empezamos a notar una tendencia a considerar los elementos clíticos como palabras con una sola sílaba átona, por lo que se observan las versiones de estos con vocal cerrada en lugares donde Penny describe una vocal abierta, como en la frase *il gatu di casa* (el gato de la casa). Además, todos los participantes conservaban el neutro de materia en todas sus funciones.

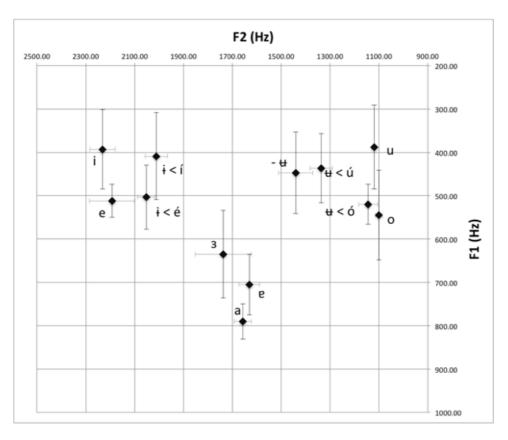


Figura 2. Gráfico de vocales obtenido de los datos del cuestionario lingüístico para el Grupo 1 (mayores), las barras de error corresponden a la desv. típica.

¹³ Por alguna razón, los scripts segmentadores del castellano no funcionaban bien, es posible que las diferencias en los sistemas vocálicos sean demasiado grandes.

En cuanto a los resultados de la encuesta¹⁴, son los reflejados en la figura 2, donde podemos ver que efectivamente existe una centralización de las vocales en el proceso de metafonía. Sin embargo, estas vocales son dificilmente diferenciables en el habla viva y no encajan con la descripción dada por Penny, que decía que se diferenciaban solo por el redondeamiento de los labios y que a veces incluso se confundían. Además de la metafonía, se observa el alófono cerrado de /a/ y la vocal velar, que procede del análisis de la tónica de palabras como *uña* o *pluma*, por lo que no aporta datos sobre el morfema de neutro.

La mayoría de los miembros del grupo reproducían sin problema ejemplos de discurso aplicando metafonía, por ejemplo, citando a sus abuelos o padres. Solo un dos realizaba el cambio de manera espontánea, aunque no demasiado frecuente.

4.2 Grupo 2 (jóvenes)₁₅

Los resultados de la conversación guiada muestran un gran deterioro de los rasgos dialectales en favor de la pronunciación netamente castellana. Se puede decir que el único rasgo que se mantiene con pujanza es el cierre de las átonas finales, especialmente en los plurales masculinos, que es tan o más común que el cierre el masculino singular. Se ve reforzada la tendencia a cerrar las vocales de los clíticos que se observa en los mayores, siendo casi regular el cierre de las preposiciones y los pronombres en todas las posiciones. Curiosamente el artículo masculino raramente se pronuncia diferente al castellano *el*. Solo es posible observar de manera regular el cierre de pretónicas en expresiones que los jóvenes reciben hechas (como topónimos o frases hechas), raramente en creaciones propias.

Los jóvenes desconocen la metafonía por completo, a excepción de dos participantes que las reconocieron, una de ellas fue capaz de proporcionar varias docenas de palabras que decía de uso habitual en su familia. como *budigu* (fresquera), *caldiru* (caldero), *curdiru* (cordero), etc.

¹⁴ Los datos numéricos pueden encontrase en el Apéndice B de (Moreno Carral, 2014).

¹⁵ Esta parte muestra los resultados para un subgrupo de 10 participantes escogidos aleatoriamente.

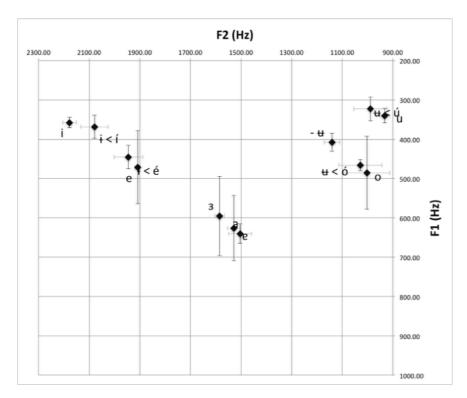


Figura 3. Gráfico de vocales obtenido de los datos del cuestionario lingüístico para el Grupo 2 (jóvenes), las barras de error corresponden a la desv. típica.

De los resultados del cuestionario fotográfico (figura 3) podemos ver que la centralización ya no existe, no siendo los antiguos fonemas sino un grupo de alófonos. Se observa la aproximación de lo que podemos llamar alófonos de /a/. Es reseñable que a pesar de la caída de la metafonía, el 80 % de los jóvenes mantienen el neutro de materia con vitalidad en todos sus usos excepto la morfología adjetival¹⁶. Es destacable el grado de separación que los jóvenes conservan entre la /u/ procedente de la tónica de *uña* y *pluma*, y la -u procedente del morfema de masculino singular.

5 Conclusiones

Es preocupante cuantificar cuán acusado es el deterioro del dialecto, aunque es lo esperable dado su nula protección o reconocimiento.

Los datos obtenidos confirman parcialmente lo obtenido por Fernández Juncal (1998). Será necesario esperar a los datos de la transcripció de las conversaciones para poder establecer comparaciones con la parte sociolingüística del trabajo de Holmquist, sin embargo, podemos decir que, a falta de describir el morfema de neutro, se observa un patrón similar en cuanto a la estructura de las vocales velares.

Para el futuro se plantean algunas opciones, como la realización de test de percepción con los que cuantificar el acercamiento de los morfemas de masculino y neutro. También sería posible extender el cuestionario lingüístico: (1) incluyendo palabras esdrújulas que nos permitirían analizar el comportamiento de las postónicas no finales, además de (2) otras vocales como la -i procedente del cierre de -e final átona y la (3) vocal media velar centralizada [θ] que solo aparece directamente ante [ɐ].

¹⁶ Es habitual en Cantabria la autocensura del uso del neutro en la morfología adjetival, pero raramente es la gente capaz de suprimirlo de los pronombres en anáfora donde son habituales expresiones como "me lo bebí todo [el agua]".

6 Bibliografía

- Alvar López, M. (1995). Atlas lingüístico y etnográfico de Cantabria. Arco-Libros.
- Fernández Juncal, C. (1998). Variación y prestigio: estudio sociolingüístico en el oriente de Cantabria. (Consejo Superior de Investigaciones Científicas, Ed.) Obtenido de http://gredos.usal.es/jspui/bitstrea m/10366/22434/1/978-84-7800-100-2.pdf
- Fernández Ordóñez, I. (2005). Corpus oral y sonoro del español rural. Recuperado el 2014, de Universidad Autónoma de Madrid: http://uam.es/coser
- García Lomas, A. (1960). *Los Pasiegos*. (E. L. Estvdio, Ed.) Santander, Cantabria, España.
- Holmquist, J. C. (1982). Language Loyalty and Linguistic Variation in Spanish Cantabria. Princeton University.
- Hualde, J. (1989). Autosegmental and metrical spreading in the vowel harmony systems of north-western Spain. Linguistics, 7 (5), 773-806.
- Menéndez Pidal, R. (1959). Algunas precisiones sobre la metafonía de

- Santander y de Asturias. Archivum: Revista de la facultad de Filología, 9, 236-248.
- Menéndez Pidal, R. (1954). Pasiegos y Vaqueiros: dos cuestiones de geografía lingüística. Universidad de Oviedo.
- Menéndez Pidal, R., & Navarro Tomás, T. (1947-54). Atlas Lingüístico de la Península Ibérica. Recuperado el 2014, de Atlas Lingüístico de la Península Ibérica, searchable database. : alpi.ca
- Moreno Carral, M. (2014). Diachronic Change of the Vowel Harmony in the Speech of Pas. Universitat Pompeu Fabra, Treballs de fi de màster, Barcelona.
- Penny, R. J. (1970). El habla pasíega: ensayo de dialectología montañesa. London, England: Tamesis Books.
- Penny, R. J. (2009). Vowel harmony and metaphony in Iberia, a revised analysis. Estudos de Lingüística Galega, 1, 113-124.
- Picard, M. (2001). Vowel harmony, centralization, and peripherality: the case of Pasiego. Linguistics, 39(1), 117-132.

Normalització de les vocals mitjanes anteriors del català en parla espontània

RIUS-ESCUDÉ, AGNÈS; TORRAS COMPTE, FRANCINA

1 Introducció

En aquest article presentem els resultats obtinguts en una investigació basada en l'anàlisi acústica de les vocals mitjanes anteriors normalitzades del català central en un context de parla espontània.

Hem utilitzat el *Corpus oral de parla espontània* de Font-Rotchés (2006), i el de Rius-Escudé (en premsa), que contenen 60 hores de material audiovisual, del qual hem seleccionat 177 vocals mitjanes anteriors, 98 d'altes i 79 de baixes, emeses per 67 informants catalans, homes i dones, de la varietat dialectal central, d'edats compreses entre els 18 i els 80 anys, de diferents professions que s'expressen en un context de parla espontània.

Per fer l'anàlisi acústica, hem utilitzat l'aplicació d'anàlisi i síntesi de veu PRAAT, que ens ha permès extreure els valors dels dos primers formants (F_1 i F_2) de cada so vocàlic, i de l'aplicació SPSS Statistics 17.0 per comprovar la fiabilitat dels resultats. Per calcular la desviació típica, realitzar els histogrames i els gràfics de caixes, hem utilitzat el programa d'anàlisi estadística GSTAT2.

Amb l'objectiu de reduir les diferències acústiques associades a les peculiaritats morfològiques de la cavitat oral dels informants, hem relativitzat les dades dels formants vocàlics seguint el procediment de normalització intrínseca del parlant S-centroid, proposat per Watt & Fabricius (2002). Així, hem obtingut els valors acústics normalitzats dels formants F_1 i F_2 de les vocals mitjanes anteriors del català en parla espontània.

2 Objectius

Els objectius inicials que ens vam proposar a l'hora d'estudiar les vocals mitjanes anteriors del català en parla espontània són els següents:

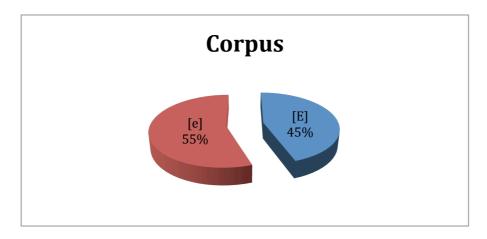
- 1. Descriure acústicament les vocals mitjanes anteriors normalitzades del català central en parla espontània.
- 2. Obtenir el camp de dispersió de les vocals mitjanes anteriors normalitzades i analitzar-ne la zona de confluència.
- 3. Comparar les dues vocals mitjanes anteriors normalitzades en parla espontània.

3 Corpus i metodologia

Per dur a terme aquesta investigació, hem utilitzat el *Corpus oral de parla espontània* de Font-Rotchés, (2006), i el de Rius-Escudé (en premsa), que contenen 60 hores de material audiovisual, de programes dels canals de televisió TV3, TV2, Canal 33, 8TV, UBTV i TOTTV, emesos els anys 1996, 1997, 1999, 2000 i 2012, amb formats molt diversos, com magazins (La Columna, Malalts de tele, Els matins de TV3, Divendres, 8 al dia), tertúlies (El rondo, Efectivament, Agora), entrevistes (Quarts de nou), concursos (A+A; Tic, tac, toe), debat (L'ou o la gallina), reportatges (Entre Línies, Catalunya Avui, Jutjats), enquestes al carrer, entre d'altres.

Él corpus de les vocals mitjanes anteriors consta de 177 paraules, que corresponen a 67 informants, d'edats compreses entre els 18 i 80 anys, de diferents professions i que tenen en comú que parlen la varietat dialectal central.

Del total de paraules que contenen una vocal mitjana anterior, 79 tenen una vocal mitjana anterior baixa [E], un 45%, i 98 una vocal mitjana anterior alta [e], un 55%. (Vegeu gràfic 1).



Gràfic 1. Percentatge de vocals mitjanes anteriors altes i baixes

Per fer l'anàlisi acústica de totes aquestes vocals, ens hem servit de l'aplicació d'anàlisi i síntesi de veu PRAAT (Boersma i Weenink, 1992-2012), la qual ens ha permès extreure els valors dels dos primers formants (F_1, F_2) de cada so vocàlic.

Hem analitzat els dos primers formants perquè, tal com diuen Martí (1984), Matas (1997), Recasens (1991), Badia i Cardús (2002) i Carrera-Sabaté (2010), la freqüència central dels formants F_1 i F_2 té un interès fonamental per identificar les vocals àtones i tòniques del català.

Hi ha, doncs, una relació entre els mecanismes articulatoris i la posició dels formants a l'espectrograma. Per a l'anàlisi de les vocals, segons Badia i Cardús (2002: 75), són suficients els dos primers formants, la diferent disposició dels quals determina el timbre peculiar de cada so vocàlic. Aquests dos formants coincideixen amb les zones principals de ressonància en els òrgans supraglòtics, és a dir, a la zona de la faringe i a la boca. El primer formant correspon a la zona de ressonància que va des de la laringe fins al punt d'articulació i el segon, des del punt d'articulació fins als llavis, que encara es poden tirar endavant i fer més gran aquest ressonador.

L'F₁, segons Recasens (1991: 39), depèn directament del grau d'obertura de la cavitat oral (com més obertura més alt és el formant i, com menys obertura, més baix) i del grau de constricció de la faringe (com més constricció més alt és el formant i, a l'inrevés).

L'F₂, segons Recasens, en el cas de les vocals no labialitzades, com les mitjanes anteriors, depèn de l'avançament lingual i del grau de constricció dorsopalatal (com més avançament i més constricció, més alt és el formant, i com menys avançament i menys constricció, més baix).

Martínez Celdrán (1994: 82) hi afegeix la incidència que té la posició de la llengua en l'F₂, baixa de freqüència a mesura que la llengua va retrocedint des del paladar dur.

Finalment, les dades que hem obtingut a partir de PRAAT han estat organitzades i classificades a l'aplicació SPSS Statistics 17.0 per poder-ne fer l'explotació i comprovar-ne la fiabilitat dels resultats. Amb aquesta aplicació, hem pogut analitzar totes les variables del nostre estudi i buscar-ne la variança d'un factor (ANOVA) per obtenir els resultats del valor acústic mitjà i de la desviació estàndard.

3.1 Normalització de les vocals

Amb l'objectiu de reduir les diferències acústiques associades a les peculiaritats morfològiques de la cavitat oral dels informants, com ja han fet altres investigadors (Herrero & Jiménez, 2012), hem relativitzat les dades dels formants vocàlics seguint el procediment de normalització intrínseca del parlant S-centroid, proposat per Watt & Fabricius (2002).

Els quatre objectius generals del procediment de normalització de vocals, segons Thomas & Kendall (2007), citant Disner (1980) i Thomas (2002), són:

- a. eliminar la variació causada per les diferències fisiològiques entre els parlants;
- b. preservar les diferències sociolingüístiques / dialectals en la qualitat vocal;
- c. preservar les distincions fonològiques entre les vocals;
- d. modelar els processos cognitius que permeten als oients normalitzar les vocals pronunciades per diferents parlants.

Per a Watt & Fabricius (2002), la tècnica de normalització de formants vocàlics permet la comparació visual i estadística directa de triangles de vocals per a diversos informants de diferent sexe, calculant, per a cadascun d'ells, un centre de gravetat S al pla F₁-F₂. El procediment conegut com a S-central de Watt & Fabricius funciona diferent del de Lobanov (1971) i Nearey (1977) perquè és com una eina d'investigació sociofonètica de tres maneres diferents. En primer lloc, redueix la variació en les relacions d'àrea d'espais de vocals (tracta d'igualar les àrees d'espai de vocals); en segon lloc, millora la superposició de polígons de vocals, i, en darrer lloc, reprodueix les posicions relatives de la vocal que significa dins de l'espai vocàlic.

Amb aquest mètode (Watt & Fabricius, 2002), els valors formatius de les vocals de cada individu es relativitzen en relació al punt central d'un triangle, els vèrtexs del qual estan determinats pels valors dels dos primers formants de la vocal anterior i, de la vocal baixa a i d'una vocal posterior u imaginària d'aquest individu. Els valors dels dos formants d'aquesta vocal imaginària u s'agafen del valor del primer formant de la vocal anterior i (vegeu figura 1) ; S representa el punt central del triangle.

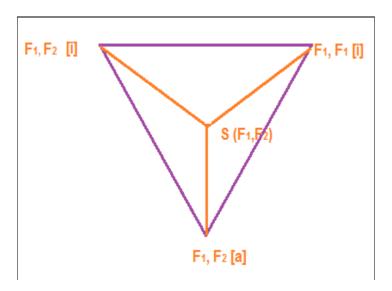


Figura 1. Representació del punt S, segons Watt & Fabricius (2002)

Per a cada informant, segons Watt & Fabricius (2002), cal fer el càlcul del punt S a partir de la fórmula següent :

$$S(Fn) = [i] Fn + [a] Fn + [i] Fn$$

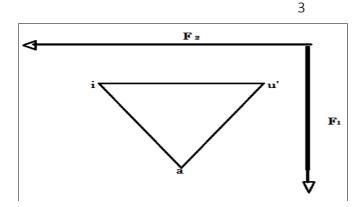


Figura 2. Representació esquemàtica del triangle vocàlic ideal per a calcular S. i=mínim F_1 , màxim F_2 ; u'=mínim F_1 , mínim F_2 , quan F_1 (u') i F_2 (u') = F_1 (i). Watt & Fabricius (2002:164)

Després d'obtenir els valors de S, un per a $l'F_1$ i un altre per a $l'F_2$, per normalitzar les dades, es divideixen els valors en Hz dels formants de cada vocal concreta, (F_1, F_2) , pel valor que correspon al punt S (S_1, S_2) i, d'aquesta manera, obtenim dos nous valors sense dimensió per a cada vocal perfectament comparables entre si:

 SF_1 (valor normalitzat) = F_1/S_1

 SF_2 (valor normalitzat) = F_2/S_2

Per il·lustrar-ho, agafarem un exemple concret del nostre corpus. Al gràfic 2 podem veure tots els valors de l' F_1 i de l' F_2 de les vocals [a] i [i] que hem obtingut d'un mateix informant (7 D) i busquem el valor mitjà de cadascun dels formants.

INFORMANT	[i]	[i]	[a]	[a]
7 D	$\mathbf{F_1}$	\mathbf{F}_{2}	$\mathbf{F_1}$	\mathbf{F}_{2}
	425	1343	894	1541
	395	2037	999	1762
	498	2048	915	1584
	598	2175		
MITJANA	479	1900	936	1629

Gràfic 2. Valors en Hz dels dos primers formants de les vocals [a] i [i] produïdes per un mateix parlant i la mitjana corresponent

Un cop obtinguts els valors mitjans de cada vocal, apliquem la fórmula de Watt & Fabricius per trobar la S_1 i la S_2 (vegeu gràfic 3).

SF ₁	479 (F ₁ [i])	+	936 (F ₁ [a])	+	479 (F ₁ [i])	:3	631
SF_2	1900,7 (F ₂ [i])	+	1629 (F ₂ [a])	+	479 (F ₁ [i])	:3	1336

Gràfic 3. Valors de la S_1 i la S_2 de l'informant 7 D un cop aplicada la fórmula de Watt & Fabricius (2002:164)

Amb els resultats de la S_1 i la S_2 , ja podem normalitzar els valors de cada formant de l'informant 7 D dividint el valor en Hz per SF_1 en el cas del primer formant, i per SF_2 en el cas del segon, tal com podem veure al gràfic 4.

INFORMANT	[i]	[i]	[i]	[i]	[a]	[a]	[a]	[a]
7 D	$\mathbf{F_1}$	\mathbf{F}_{2}	F _{1 NORM}	F _{2 NORM}	$\mathbf{F_1}$	\mathbf{F}_{2}	F _{1 NORM}	F _{2 NORM}
	425	1343	0,674	1,005	894	1541	1,417	1,153
	395	2037	0,626	1,525	999	1762	1,583	1,319
	498	2048	0,789	1,533	915	1584	1,450	1,186
	598	2175	0,948	1,628				

Gràfic 4. Valors de l'F1 i de l'F2 de les vocals [a] i [i] de l'informant 7 D en Hz i normalitzats

I, d'aquesta manera, seguim el mateix procediment amb la resta de vocals de l'informant 7 D (vegeu gràfic 5):

[E]	[E]	[E]	[E]	[e]	[e]	[e]	[e]
F ₁	\mathbf{F}_2	F ₁ NORM	F ₂ NORM	F ₁	\mathbf{F}_2	F ₁ NORM	F ₂ NORM
600	2488	0.951	1.862	512	2486	0.811	1.861
660	2095	1.046	1.568	563	1847	0.892	1.382
691	2035	1.095	1.523	566	2015	0.897	1.508

Gràfic 5. Resultats normalitzats de les vocals mitjanes anteriors

En utilitzar valors normalitzats, com ja hem dit abans, els resultats es poden comparar perfectament entre si. Ara bé, en contrapartida, es perd la referència que proporcionen els Hertz ja que tots els valors relatius dels formants passen a expressar-se en valors amb dimensió SF_1 i SF_2 , i no en Hertz (Hz).

4 Resultats

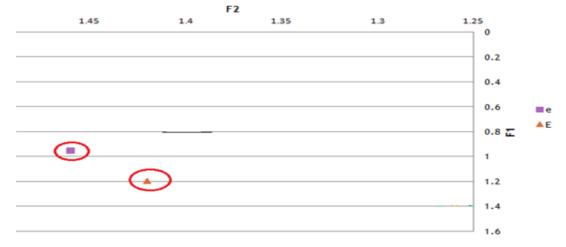
Per fer aquesta anàlisi, ens hem basat en un total de 177 vocals mitjanes anteriors normalitzades: 79 [E], un 45% del total del corpus, i 98 [e], un 55%, produïdes per 67 informants. A partir d'aquestes dades, hem obtingut els resultats següents:

μ	ANOVA	$\overline{\mathbf{F}}_{1}$	\mathbf{F}_2	ANOVA	\mathbf{F}_{2}
[e]	$\mathbf{F}_{1}[\mathbf{e}]/[\mathbf{E}]$	[E]	[e]	F ₂ [e]/[E]	[E]
0,95	0,000	1,19	1,46	0,102	1,42

Taula 1. F₁ i F₂: diferències entre la [e] i la [E] normalitzades

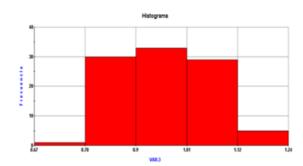
A la taula anterior, veiem que ambdós sons vocàlics presenten una diferència significativa en $l'F_1$, (p = 0.000), mentre que en $l'F_2$ no ho és gens (p = 0.102).

Si observem els resultats obtinguts en el gràfic 6, es marca clarament la distància entre la posició de la vocal anterior alta [e] i la de la baixa [E]. Visualment, doncs, cada so té el seu espai i la vocal anterior baixa és més centralitzada que l'alta.



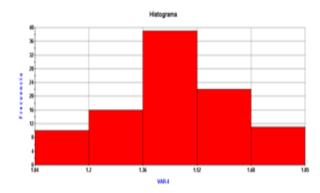
Gràfic 6. F1 i F2 de la [e] i la [E] normalitzades

Si analitzem els histogrames que corresponen a la vocal mitjana anterior alta normalitzada, podem comprovar que en un 95% dels casos el so en $1'F_1$ es produeix entre els valors 0.78 i 1.12 (vegeu gràfic 7).



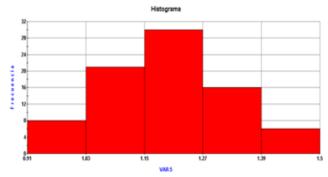
Gràfic 7. Histograma de l'F1 de [e]

En canvi, en l' F_2 , hi ha una amplitud que va de l'1,04 al 1,85 i destaca, amb un 40%, la franja del valor d'1,36 a 1,52 (vegeu gràfic 8).



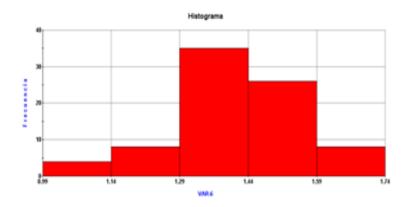
Gràfic 8. Histograma de l'F2 de [e]

Si ens fixem en els histogrames que corresponen a la vocal mitjana anterior baixa normalitzada, podem comprovar que en un 66% dels casos el so en l'F₁ es produeix entre els valors 1,03 i 1,39 (vegeu gràfic 9).



Gràfic 9. Histograma de l'F₁ de [E]

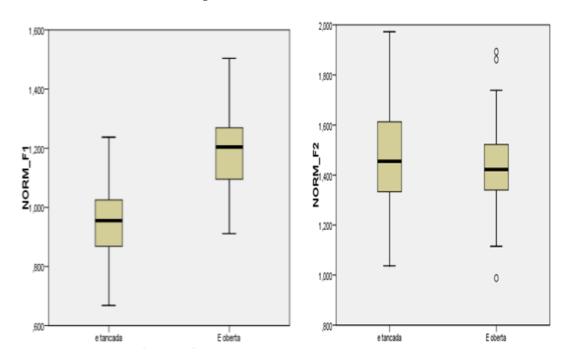
En $l'F_2$ de la vocal mitjana anterior baixa normalitzada els valors van des de l'1,29 al 1,59, un 60%, i ocupa una franja global que va des del 0,89 fins a l' 1,74 (vegeu gràfic 10).



Gràfic 10. Histograma de l'F2 de [E]

En resum, si comparem els histogrames 7 i 9, que corresponen a l'F₁ de les dues vocals anteriors, podem comprovar que tenen una zona més estable que la que correspon a l'F₂, vegeu gràfics 8 i 10, en els quals hi ha una àmplia dispersió.

A continuació, si ens fixem en els gràfics següents, 11 i 12, podem veure que en $1'F_1$ de les vocals mitjanes anteriors normalitzades el camp de dispersió és clar i el camp de la vocal mitjana anterior alta no coincideix gens amb el de la mitjana anterior baixa, és a dir, cada vocal té el seu propi espai i no se sobreposen en cap punt. En canvi, en $1'F_2$ la vocal mitjana anterior alta té un espai molt ampli que inclou l'espai de la mitjana anterior baixa, és a dir, hi ha una zona de confluència important.



Gràfic 11. F₁ de les vocals mitjanes anteriors

Gràfic 12. F₂ de les vocals mitjanes anteriors

5 Conclusions

Després d'analitzar acústicament les 177 vocals mitjanes anteriors normalitzades del català central en parla espontània (98 d'altes i 79 de baixes), hem obtingut els valors mitjans de cada so, tant del primer formant com del segon. Hem pogut comprovar que només hi ha diferències significatives (0.000)

entre ambdós en l'F1. En l'F2, aquesta diferència no és gens significativa (0.102) i, a més a més, el valor de la vocal mitjana anterior alta en aquest formant és més alt que el de la mitjana anterior baixa.

[e]: F₁ 0,95 F₂: 1,46 [E]: F₁ 1,19 F₂: 1,42

Quant al camp de dispersió i la zona de confluència entre les vocals mitjanes anteriors normalitzades del català central són clares en el primer formant i cadascuna té el seu propi espai, la qual cosa no passa en l'F₂, en què l'espai de la vocal mitjana anterior baixa s'inclou dins del de l'alta.

Aquests resultats ens animen a continuar investigant en l'anàlisi de les altres vocals del català normalitzades per comprovar si les tendències que hem observat es mantenen. Així i tot, cal tenir en compte aquests resultats en l'ensenyança i l'aprenentatge de la pronunciació de les vocals mitjanes anteriors del català com a L2, no només des del punt de vista de la producció, sinó també de la percepció.

6 Referències bibliogràfiques

- Badia i Cardús, M. (2002). *Introducció a la fonètica i a la fonologia catalanes*.

 Barcelona: Curial Edicions Catalanes i Publicacions de l'Abadia de Montserrat.
- Boersma, P. & Weenink, D. (1992-2012). PRAAT. Doing phonetics by computer, Institute of Phonetic Sciences. University of Amsterdam. http://www.praat.org [Consulta: 28 de 5 de 2014].
- Carrera-Sabaté, J. (2010). Vocals mitjanes anteriors del català i castellà extretes d'entrevistes radiofòniques: caracterització i comparació acústiques. En Creus, I., Puig, M., & Veny, J. R., XV Col·loqui de l'AILLC. Lleida, 7-11 setembre de 2009. Barcelona: PAM, vol. 1: 367-377.
- Disner, S. (1980). Evaluation of vowel normalization procedures. *Journal of the Acoustical Society of America*, 67, 253-261.
- Font-Rotchés, D. (2006). Corpus oral de parla espontània. Gràfics i arxius de veu, dins *Biblioteca Phonica*, 4. http://www.publicacions.ub.edu/revistes/phonica-biblioteca/ [Consulta: 27 de maig de 2014].
- Herrero, R., & Jiménez, J. (2012). De la coarticulación a la armonía vocálica en valenciano. http://www.academia.edu/5470361
 /Herrero_and_Jimenez_De_la_coarticulacion_a_la_armonia_vocalica_en_valenciano_[consulta_el_5_de_5_de_2014].

- Lobanov, B.M. (1971). Classification of Russian vowels spoken by different speakers. *Journal of the Acoustical Society of America*, 49 (2B), 606-608.
- Martí, J. (1984). Paràmetres vocàlics del català, *Folia Phonetica*, 1, Lleida, 23-44.
- Martínez Celdrán, E. (1994). *La fonètica*. Barcelona: Empúries.
- Matas, J. (1997). Validació de la carta vocàlica del català oriental central mitjançant la tècnica de l'escalfament multidimensional. Barcelona. EFE, VIII, 253-270.
- Nearey, T. (1977/8). Phonetic Feature

 Systems for Vowels. Indiana

 University Linguistics Club.

 Disponible a: http://

 www.ualberta.ca/~tnearey/Nearey1

 978_compressed. pdf
- Recasens, D. (1991). Fonètica descriptiva del català. Barcelona: Institut d'Estudis Catalans.
- Rius-Escudé, A. (en premsa). Corpus oral de parla espontània.
- Thomas, E. R. (2002). Instrumental phonetics. A J. K. Chambers, P. Turudgill & N. Schilling-Estes (ed.), *The handbook of language variation and change*. Oxford: Blackwell, 168-200.
- Thomas, E. R., & Kendall, T. (2007). NORM: The vowel normalization and plotting suite.
- Disponible a http://ncslaap.lip.ncsu.edu/tools/n orm/index.php (Consulta: 18 de 11 de 2013).

Watt, D. J. L., & Fabricius, A. H. (2002). Evaluation of a technique for improving the mapping of multiple speakers' vowel spaces in the F1-F2 plan. Leeds Working Papers in Linguistics and Phonetics, 9: 159-173.

Tratamiento perceptivo de las categorías / p,t,k / /b,d,g / . Esbozo del espacio acústico de distribución categorial a partir de los parámetros VOT, Explosión y Transición

ROSAS SALVANS, ALBA

1 Introducción

La oposición fonológica entre las series de sonidos consonánticos /p, t, k/ y /b, d, g/ ha sido y sigue siendo objeto de debate desde diversas perspectivas de la fonética experimental: desde la pertinencia del rasgo de sonoridad¹ como

 $^{^1}$ En este estudio, utilizaremos el término sonoridad como una etiqueta funcional que nos permita simplificar las referencias a la diferencia fonológica existente entre /p/y/b/, /t/y/d/, y/k/y/g/, aun coincidiendo con Pérez (2001) en el hecho de que este término resulta confuso por referirse, stricto sensu, a la presencia/ausencia de energía regular sobre las bajas frecuencias durante la realización del segmento consántico y, por lo tanto, directamente relacionado con el parámetro de VOT, el cual se ha demostrado no consituir el único parámetro explicativo de las diferencias fonológicas entre estas series de sonidos.

realidad psicolingüística en la representación mental del hablante, hasta los parámetros acústicos (o rasgos fonéticos) que permiten marcar tal oposición.

En efecto, junto a la clásica asociación de esta oposición fonológica con el rasgo de sonoridad, autores como Torreblanca (1979), Celdrán (1989, 2009), Celdrán & Fernández-Planas (2007) o Veiga (2002), señalan la conveniencia de considerar la tensión como elemento distintivo de estas series de sonidos.

En los estudios descriptivos de las características fonético-acústicas que permiten distinguir /p, t, k/ de sus correlatos /b, d, g/, el parámetro VOT es ciertamente el que aparece más documentado (Abramson & Lisker, 1964; Castañeda, 1986; Soto-Barba, 1999). Sin embargo, otros parámetros como la duración absoluta, el IREDUS, la intensidad de la onda periódica (Soto-Barba, 1999), la velocidad de transición (Quilis, 1981) y la tensión articulatoria (Celdran, 1984; Canalleda & Kuhlman, 1987) parecen también contribuir en la distinción acústica de estas series de sonidos.

Por otro lado, si el VOT no resulta ser un parámetro suficiente en la descripción fonético-acústica de las diferencias entre las realizaciones de /p, t, k/y/b, d, g/, tampoco parece constituir una pista suficiente para explicar la categorización perceptiva de estas series de sonidos. En efecto, diversos estudios demuestran como la percepción categorial de estos sonidos puede mantenerse aún cuando las marcas estrictamente asociadas a la sonoridad no están presentes en los estímulos (Celdrán, 1993; Soto-Barba, 1994).

En el ámbito de la fonética clínica, diversos estudios han podido constatar cómo, frente a un déficit motor o de la gestión motora con efectos sobre la gestión de las cuerdas vocales y/o coordinación de los articuladores inferiores con los superiores, algunos pacientes son capaces de mantener el contraste fonético entre sonidos sordos y sonidos sonoros, aun 1) cuando los valores de VOT de sus realizaciones de /b, d, g/ son inferiores a las medianas de los hablantes control e incluso, 2) en aquellas realizaciones de sonidos consonánticos sonoros que han sido percibidos como sordos por uno o varios examinadores especializados en el análisis del habla patológica (Baqué *et al.*, en prensa).

Nuestro estudio, inscrito en el marco del proyecto COGNIFONII¹ para la evaluación del habla patológica, defiende la hipótesis según la cual un déficit en la gestión de la actividad glótica (activación/coordinación con articuladores superiores) en la marcación de la oposición sordo/sonoro se puede compensar reforzando otras marcas fonéticas que intervienen también en la oposición fonológica del rasgo de sonoridad. En este sentido, el estudio que presentamos propone un análisis del tratamiento perceptivo de las marcas acústicas de tres de los parámetros asociados a la distinción de estas series de sonidos: el clásico VOT² (Abramson & Lisker, 1964; Castañeda, 1986; Soto-Barba, 1999), la explosión (Rosas *et al.*, 2012) y la transición (Quilis, 1981), con el objetivo de determinar en qué medida la combinación de los valores de estos parámetros puede tener un efecto en la distinción fonológica de estas series de sonidos y constituir, por consiguiente, una estrategia de compensación articulatoria pertinente en el habla patológica.

² Intervalo de tiempo transcurrido entre el inicio de la actividad glótica (con su correspondiente "barra de sonoridad" o energía regular sobre las frecuencias más bajas como correlato acústico) y el momento de la explosión como inicio de la actividad de los articuladores superiores (con su correspondiente "barra de explosión" como correlato acústico).

¹Este estudio se ha realizado en el marco del proyecto COGNIFON y COGNIFON-II (HUM2006-06582 y FFI2010-15320) "Evaluación de la producción y la percepción fónica de pacientes afectados de patologías del lenguaje : Descripción lingüística y cognitiva", financiados por el Ministerio de Ciencia e innovación Español. Los proyectos se han llevado a cabo en la "Universitat Autònoma de Barcelona", bajo la dirección de Lorraine Baqué y en colaboración con la Universidad Francesa de "Toulouse II-Le Mirail".

2 Metodología

Los estímulos de este estudio se han generado a partir de manipulaciones sobre las producciones /pa, ba, ta, da, ka, ga/ realizadas en habla natural por un locutor hispanohablante. Para crear los estímulos del test se han extraído los períodos de explosión y transición de dichas realizaciones, los cuales se han combinado con diferentes variaciones en los valores de duración del VOT. Los tres parámetros se han combinado sobre las realizaciones del locutor para obtener los estímulos siguientes:

- Una serie de 11 variaciones sucesivas de VOT por intervalos de 10 ms (desde -60 ms hasta 30 ms, incluyendo un estímulo con el VOT original) combinadas con el período de explosión extraído de la realización del correlato sordo y el período de transición extraído de la realización del correlato sordo.
- Una serie de 11 variaciones sucesivas de VOT por intervalos de 10 ms (desde -60 ms hasta 30 ms, incluyendo un estímulo con el VOT original) combinadas con el período de explosión extraído de la realización del correlato sonoro y el período de transición extraído de la realización del correlato sordo.
- Una serie de 11 variaciones sucesivas de VOT por intervalos de 10 ms (desde -60 ms hasta 30 ms, incluyendo un estímulo con el VOT original) combinadas con el período de explosión extraído de la realización del correlato sordo y el período de transición extraído de la realización del correlato sonoro.
- Una serie de 11 variaciones sucesivas de VOT por intervalos de 10ms (desde -60 ms hasta 30 ms, incluyendo un estímulo con el VOT original) combinadas con el período de explosión extraído de la realización del correlato sonoro y el período de transición extraído de la realización del correlato sonoro.

El resultado de estas manipulaciones permite obtener un total de 132 estímulos, de los cuales 44 corresponden al grupo articulatorio /pa_ba/, 44 al grupo /ta_da/ y 44 al grupo /ka_ga).

Los participantes realizaron dos tipos de tests complementarios en el estudio de la percepción categorial: un primer test de identificación dónde debían indicar si el estímulo que se les presentaba correspondía o no a X (siendo X = pa/, X = ba/, X = ta/, X = da/, X = ka/, X = ga/ según el test que realizaran), y un test de discriminación en el que se les presentaba dos estímulos correspondientes a la misma serie de variaciones, con una única diferencia de 20ms sobre los valores de VOT y se pedía a los participantes que indicaran si los dos estímulos presentados eran o no exactamente iguales.

Los tests se distribuyeron en tres sesiones de trabajo con el fin de controlar un posible efecto aprendizaje entre las respuestas de los tres puntos de articulación del estudio:

- Sesión 1: Identificación de /pa/, Identificación de /ba/ y Discriminación de estímulos del grupo /pa_ba/
- Sesión 2: Identificación de /ta/, Identificación de /da/ y Discriminación de estímulos del grupo /ta_da/
- Sesión 3: Identificación de /ka/, Identificación de /ga/ y Discriminación de estímulos del grupo /ka_ga/

3 Resultados

Un análisis descriptivo mediante tablas de contingencia nos ofrece una primera visión del efecto de la explosión sobre la percepción sorda/sonora de los estímulos, y nos permite observar un desplazamiento de la frontera que separa los espacios acústicos asociados a la identificación de /pa/, /ta/ y /ka/ de aquellos asociados a sus correlatos sonoros /ba/, /da/ y /ga/ sobre las 11

categorías de variación de VOT (figura 1). En efecto, el porcentaje de identificación del correlato sordo se reduce cuando la explosión del estímulo presenta las características acústicas del correlato sonoro, de la misma manera que se reduce el porcentaje de identificación del correlato sonoro cuando la explosión del estímulo presenta las características acústicas del correlato sordo.

No se observan diferencias en el porcentaje de respuestas obtenidos en función de la explosión cuando el VOT presenta valores inferiores a -50 ms y superiores a 10 ms en el grupo de los sonidos bilabiales, y con valores inferiores a -30 ms y superiores a 00 en el tratamiento de las velares. A partir de estos valores, se observan diferencias en las respuestas de identificación.

Para el grupo /ta_da/, las diferencias en las respuestas aparecen sobre todas las variaciones de VOT.

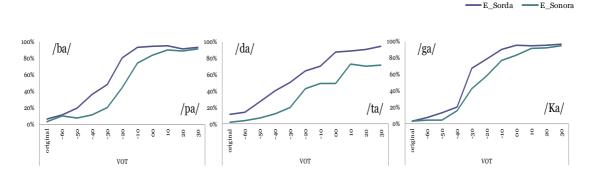
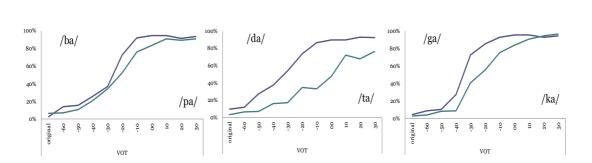


Figura 1. Efecto de la explosión sobre la identificación de los sonidos /pa_ba/, /ta_da y /ka_ga/

El mismo análisis realizado esta vez sobre el parámetro de transición permite constatar un efecto muy significativo de este parámetro (p=.000) sobre la categorización de los sonidos como sordos o sonoros. Si comparamos las líneas de categorización obtenidas en función de la variable transición para cada grupo articulatorio, obtenemos la siguientes observaciones: El grupo /ta_da/ presenta un desplazamiento de la línea de categorización más pronunciado que el de los sonidos velares y bilabiales, así como parece ser, de nuevo, el único grupo para el que el efecto de este parámetro influye incluso en el tratamiento perceptivo de los estímulos que conservan sus valores originales de VOT. Para el grupo /pa_ba/, el desplazamiento de la frontera por efecto de la transición resulta menos pronunciado y aparece principalmente en el tratamiento de la serie de estímulos comprendida entre las variaciones de VOT -20 ms y VOT 00 ms. Finalmente, para el grupo de las velares, el desplazamiento de la frontera influye en los valores de VOT comprendidos entre -30 ms y 00 ms. El efecto de la transición en la percepción de las categorías sorda/sonora en las tres series de correlatos es muy muy significativo (p = .000).



T sorda -

T Sonora

Figura 2. Efecto del período de transición sobre la identificación de los sonidos /pa_ba/, /ta_da y /ka_ga/

Finalmente, un análisis de regresión logística con efectos mixtos nos permite completar los resultados referentes a la contribución de cada una de las variables del estudio en la generación del efecto sordo/sonoro para la percepción del habla del oyente, con los valores de interacción que puedan establecerse entre sí. Para este fin, hemos seleccionado los tres parámetros fonéticos (*i. e.* VOT, Explosión y Transición) como variables explicativas y la respuesta de los participantes (variable binaria con las categorías sí/no) como variable dependiente. Se han eliminado progresivamente del modelo aquellas interacciones no significativas. Los resultados obtenidos son los siguientes:

Para el grupo /pa_ba/: Observamos un efecto simple muy significativo de los tres parámetros fonéticos: Explosión (Wald-X²(1) = 96.324 p =.000), Transición (wald-X²(1) = 23.311 p =.000), y VOT (Wald-X²(10) = 857.473 p =.000), así como un efecto de interacción entre el VOT y la Explosión (Wald-X²(10) = 28.274, p=.001) y la Explosión y la Transición (Wald-X²(1) = 20.805 p<.001).

Para el grupo /ta_da/: Observamos de nuevo un efecto simple muy significativo de los tres parámetros fonéticos: Explosión (Wald- X^2 (1) = 203.7523, p=.000) , Transición (Wald- X^2 (1) = 257.3532, p=.000), y VOT (Wald- X^2 (10) = 679.4225, p=.000), así como efectos de interacción entre el VOT y la transición (Wald- X^2 (10) = 28.0358, p=.002) y la Explosión y la Transición (wald- X^2 (1) = 5.2763, p=.021) . Observamos también un efecto significativo de interacción entre los tres parámetros, VOT, Explosión y transición (Wald- X^2 (10) = 19.3137, p=.036).

Para el grupo /ka_ga/: Observamos un efecto simple muy significativo de los tres parámetros fonéticos: Explosión (Wald-X²(1) = 56.3453 p =.000), Transición (Wald-X²(1) = 77.4523, p =.000), y VOT (wald-X²(10) = 866.1455, p =.000), así como un efecto de interacción entre el VOT y la Transición (Wald-X²(10) = 28.8342, p = 0.001) y entre la Explosión y la Transición (Wald-X²(1) = 7.9661, p = 0.005).

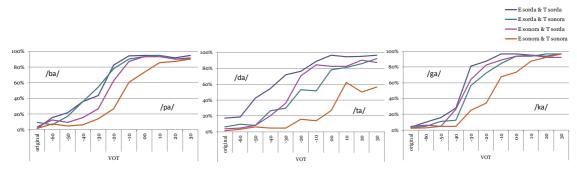


Figura 3. Respuestas obtenidas en cada una de las series de combinaciones de estímulos en el grupo /pa ba/ (izquierda), /ta da/ (centro) y /ka ga/ derecha.

La combinación de explosión y transición sorda y de explosión y transición sonora provoca un desplazamiento de la frontera de identificación categorial sobre les valores de VOT equivalente a 10ms en el grupo /pa_ba/, a 20 ms en el grupo /ka_ga/ y de hasta 50 ms en el grupo /ta_da/ (figura 4).

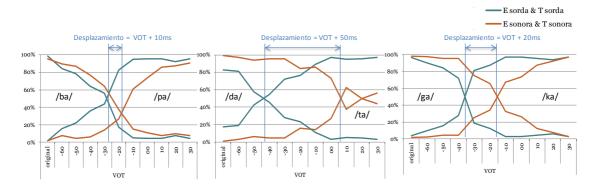


Figura 4. Desplazamiento de los puntos de intersección de las curvas de categorización sordo/sonoro en las series de estímulos con explosión y transición sorda (en verde), y las series con explosión y transición sonora (en naranja). Resultados del grupo /pa_ba/ (izquierda), /ta_da/ (centro) y /ka_ga/ (derecha)

Los gráficos de las respuestas obtenidos en el test de discriminación permiten observar muestra un desplazamiento de la zona de mayor discriminación de los estímulos desde los valores de VOT más negativos en el caso de las series de estímulos donde se combinan la explosión y la transición sordas, hacia los valores de VOT más positivos en las series de estímulos dónde la explosión y la transición son sonoras. Este desplazamiento es especialmente evidente en las respuestas obtenidas en los grupos /ta_da/ y /ka_ga/ (figura 5).

Aunque el desplazamiento de los picos de discriminación es evidente en las series de los 3 grupos articulatorios, es en la diferencia entre las series con explosión y transición sonora y las de explosión y transición sorda o explosión y transición con sonoridad combinada donde este aparece más destacado. Si comparamos las curvas de tratamiento obtenidas en cada serie de estímulos para cada grupo articulatorio, observamos lo siguiente:

En el grupo /pa_ba/ y /ka_ga/ el desplazamiento es progresivo en el orden en el que se presentan las series sobre la gráfica, *i. e.*, la zona de mayor discriminación se sitúa hacia los valores más negativos del VOT en la series que combinan explosión y transición sordas, y se desplaza progresivamente hacia los valores de VOT más positivos en las series con explosión sorda y transición sonora, seguidas de las de explosión sonora y transición sorda y, finalmente, de las que combinan explosión y transición sonoras. La serie de estímulos que

combinan explosión sorda con transición sonora en el grupo /ta_da/, a diferencia de todas las demás, no presenta ningún pico de discriminación que permita identificar una zona de discriminación diferenciada.

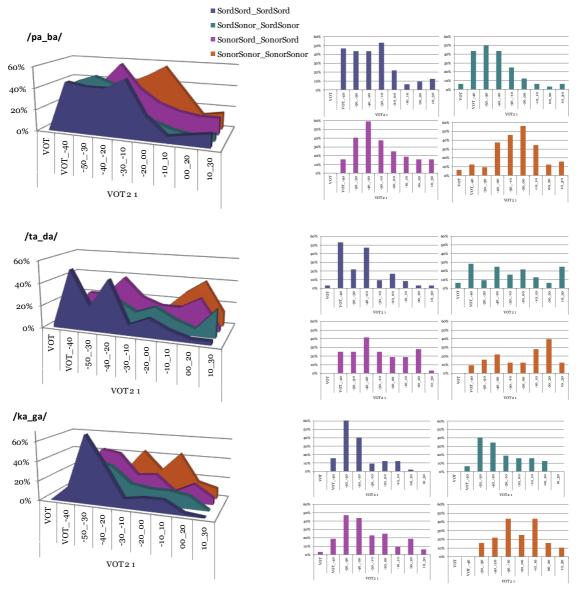


Figura 5. Porcentajes de respuestas "diferente" obtenidas en la tarea de discriminación igual/diferente. La diferencia entre los estímulos de las parejas presentadas consistía en todos los casos en una única variación de 20 ms en los valores de VOT. Las respuestas se muestran organizadas en función de la combinación de parámetros del estímulo: en índigo la serie de estímulos con explosión y transición sordas, en verde las series con Explosión sorda y transición sorda, en violeta las series de estímulos con Explosión sonora y transición sorda y en naranja las series con explosión sonora y transición sonora. Se muestran los resultados obtenidos en el grupo /pa_ba/ (arriba), /ta_da/ (en el centro), y /ka_ga/ (abajo)

El análisis estadístico realizado mediante la aplicación de un modelo logístico lineal con efectos mixtos sobre las respuestas obtenidas en los tests de discriminación, del que se han eliminado progresivamente aquellas interacciones no significativas, nos muestra los siguientes resultados:

El VOT aparece como un parámetro muy significativo en la discriminación de estímulos de los tres grupos articulatorios: $/pa_ba/(Wald-X^2(8) = 84.8782, p = .000)$, $/ta_da/(Wald-X^2(8) = 37.7363, p = .000)$ y $/ka_ga/(Wald-X^2(8) = 9.706, p = .000)$.

La Explosión y la Transición solo presentan efectos simples con valores muy significativos en el grupo /ka_ga/ (Wald-X² (1) = 20.807, p = 0.000) para el parámetro de explosión, y (Wald-X² (1) = 19.992, p =.000) para el parámetro de transición, mientras que sus efectos no son significativos para la identificación de los estímulos en los grupos /pa_ba/ y /ta_da/ (p >.1).

El efecto de interacción entre el VOT y la Explosión es significativo en los tres grupos articulatorios, con un valor de significatividad de (Wald-X² (8) = 37.6960 p =.000) en el tratamiento de los estímulos del grupo /pa_ba/, un valor de (Wald-X² (8) = 33.5265, p =.000) en /ta_da/ y un valor de (Wald-X² (8) = 24.052, p = 0.002) en /ka_ga/.

El efecto de interacción entre el VOT y la transición es muy significativo en el tratamiento de los estímulos del grupo /ta_da/ (Wald- X^2 (8) = 26.3180, p <0.001) y muy muy significativo en el tratamiento de /ka_ga/ (Wald- X^2 (8) = 34.625, p = .000).

Finalmente, el grupo / pa_ba / presenta un efecto de interacción significativa entre los tres parámetros del estudio (Wald- $X^2(8) = 21.1151$, p = 0.007).

4 Conclusiones

Las características acústicas de los parámetros fonéticos correspondientes a Explosión, Transición y VOT parecen contribuir de manera muy significativa en la identificación de los correlatos sordo/sonoro de las categorías fonológicas /pa, ta, ka/ /ba, da, ga/. Aunque en los valores de VOT extremos (VOT negativos a partir de -60 ms y VOT positivos a partir de 20 ms) el tratamiento perceptivo de los estímulos parece ser menos sensible a las características acústicas de la explosión y de la transición, el VOT no constituye un parámetro suficiente para explicar los procesos de categorización de los correlatos oclusivos /pa_ba/, /ta_da/ y /ka_ga/. En efecto, la distancia entre los puntos de intersección de líneas de categorización en función de la Explosión y la Transición muestra un desplazamiento de las fronteras categoriales asociadas a la sonoridad sobre el eje de valores del VOT de 10ms en el grupo /pa_ba/, 20 ms en el grupo /ka_ga/ y de hasta 50 ms en el grupo /ta_da/ (figura 4).

En los resultados de discriminación, si bien es verdad que la explosión y la transición solo presentan efectos simples significativos en el tratamiento de los estímulos del grupo articulatorio /ka_ga/, los valores de significatividad obtenidos en los análisis de interacción de estos parámetros insisten de nuevo en la pertinencia de los mismos en el tratamiento de estos estímulos.

Por otro lado, el desplazamiento de la zona de mayor percepción de las diferencias entre estímulos sobre los valores de VOT en función de la combinación de parámetros explosión y transición (más positivos cuando las características acústicas de la explosión y la transición corresponden a sonidos sonoros), parece también confirmar el efecto fonológico de estos parámetros (explosión y transición).

Los resultados obtenidos en este estudio nos permiten confirmar la pertinencia del uso de otros parámetros distintos al VOT en la distinción fonética de los series /p, t, k/ de /b, d, g/ que permitan contribuir a marcar la oposición de estas series de sonidos en casos de dificultad en la gestión de la actividad glótica.

Finalmente, la constatación de efectos evidentes de los tres parámetros analizados no excluye la posible participación de otros parámetros en la percepción categorial del continuo sordo-sonoro. En efecto, consideramos que algunas de las explicaciones propuestas para la marcación del rasgo de

sonoridad, como es el caso de la tensión, podrían estar relacionadas de forma latente con los resultados.

5 Bibliografía

- Abramson, A. S., & Lisker, L. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20, 384-422.
- Baqué, L., Marczyk, A., Rosas, A, Estrada, M. (en prensa). Handicap, repair strategies and communicative effectiveness at the phonic level: Evidence from a multiple-case study. En C. Astésano & M. Jucla (eds.), Neuropsycholinguistic Perspectives on Language Cognition. Routledge Taylor & Francis Group. Series: Explorations in Cognitive Psychology.
- Castañeda, M. L. (1986). El V.O.T. de las oclusivas sordas y sonoras españolas. *Estudios de fonética experimental, II,* 93-110.
- Martínez Celdrán, E. (1989). *Fonología* general y española. Barcelona: Teide.
- Martínez Celdrán, E. (1991b). Duración y tensión en las oclusivas no iniciales del español: un estudio perceptivo. *Revista Argentina de Lingüística*, vol. 7, 1, 51-71.
- Martínez Celdrán, E. (1993). La percepción categorial de /b-p/ en español basada en las diferencias de duración. Estudios de Fonética Experimental, V. Universitat de Barcelona, Laboratori de Fonètica, 223-239
- Martínez Celdrán, E. (2009). Sonorización de las oclusivas sordas en una hablante murciana. *Estudios de Fonética Experimental*, 18, 253-251
- Martínez Celdrán, E., & Fernández Planas, A. M. (2007). *Manual de fonética*

- española. Articulaciones y sonidos del español. Barcelona: Ariel.
- Nespoulous, J.-L., Baqué, L., Rosas, A., Marczyk, A., & Estrada, M. (2013). Aphasia, phonological and phonetic voicing within the consonantal system: preservation of phonological oppositions and compensatory strategies. *Language Sciences*, 39, 117-125.
- Soto-Barba, J. (1994). ¿Los fonemas /b/ y /p/ se diferencian por la sonoridad? Estudios filológicos, 29, 33-37.
- Soto-Barba, J., & Valdivieso, H. (1999). Caracterización fonético-acústica de la serie de consonantes /p-t-k/ vs. /b-d-g/. *Onomazein*, 4, 125-133
- Pérez, H. E. (2001): La noción de rasgo. El caso de las consonantes oclusivas del español. *Onomazein*, 6, 327-336
- Rosas A., Baqué L., & Estrada M. (2012). Consonant production by a Catalan native speaker suffering from anarthria: a case study. En *Man and Speech 50 years of SUVAG*. Scientific and professional monograph of the VIIth International Symposium of the Verbotonal System Zagreb: SUVAG Polyclinic, 673-688.
- Torreblanca, M. (1979). Un rasgo fonológico de la lengua española. *Hispanic Review*, 47, 1, 455-468.
- Veiga, A. (2002). Estudios de fonología funcional. A Coruña: Toxosoutos.
- Quilis, A. (1981). Fonética acústica de la lengua española. Madrid: Gredos.

Resilabificación de la aspiración de /-s/ ante oclusiva dental sorda. Parámetros acústicos y variación social

VIDA-CASTRO, MATILDE

1 Introducción

En este trabajo se presentan los primeros resultados obtenidos al analizar, en una muestra de informantes universitarios de la ciudad de Málaga, el proceso de aspiración de /-s/ implosiva ante consonante dental tensa /t/.

El estudio de este proceso pasa por definir sus fases, delimitar las principales variantes, describir los parámetros acústicos de cada una de ellas y, por último, rastrear cuál es la dimensión social del proceso.

La variedad lingüística empleada en la ciudad de Málaga pertenece al área dialectal del español meridional y uno de sus rasgos más destacados es el debilitamiento y pérdida de los sonidos distensivos, especialmente el fricativo alveolar sordo /-s/ que, tal y como se afirma tradicionalmente, adopta diferentes soluciones: la aspiración, la asimilación o la pérdida absoluta de dicho segmento. Así, una secuencia como /los gátos/ se puede realizar como [lohgátoh], [loxátoh] o [logáto].

Tal y como se ha documentado en trabajos anteriores (Vida, 2004), el proceso de debilitamiento y pérdida de /-s/ implosiva está muy avanzado en la ciudad de Málaga, ya que la pérdida de este segmento es prácticamente categórica en la mayor parte de los contextos. Sin embargo, ante consonante oclusiva tensa la /-s/ parece retenerse, ya que la variante mayoritaria es la aspiración.

Pero lo que más llama la atención es que en posición interior ante consonante dental tensa /t/, es cada vez más frecuente escuchar un sonido africado [ts] que aparece principalmente entre los jóvenes universitarios. Así, por ejemplo, se oye [pátsa] por *pasta*, [itsórja] por *historia* o [detsíno] por *destino*.

Esta realización, que se escucha con bastante frecuencia entre los jóvenes universitarios de la ciudad de Málaga, se ha documentado también como variante prestigiosa en otras zonas del andaluz occidental, como Sevilla (Ruch, 2008) o Antequera (Málaga) (Moya, 2007).

Como se puede observar en la figura 1, se trata de una realización que presenta ruido propio de una sibilante tras la consonante oclusiva. Se trata de un sonido africado [ts] que, tal y como se va a presentar a continuación, podría ser el resultado de una resilabificación del fonema /-s/.

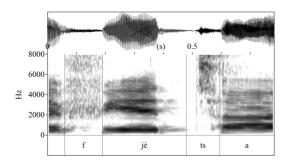
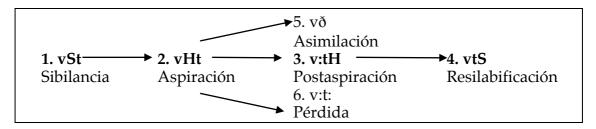


Figura 1: Espectrograma de la variante africada, [ts], de /-s/ en la secuencia fiesta

1. Descripción del fenómeno

La hipótesis que aquí se plantea es la siguiente: ante consonante dental tensa /t/, el fonema fricativo alveolar sordo /-s/ sufre un proceso de resilabificación que da como resultado una variante africada [ts]. Se trata de un proceso progresivo que se produce en diversas etapas (cfr. cuadro 1):



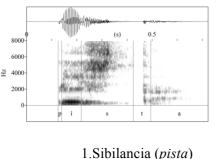
Cuadro 1: Etapas del proceso de resilabificación de /-s/ implosiva ante /t/ en la ciudad de Málaga

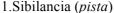
- En primer lugar se produce un debilitamiento del sonido fricativo alveolar sordo [s] en uno aspirado [h]. A partir de la realización aspirada [h] se pueden obtener diversas soluciones.
- La primera de ellas supone la transformación (5) de la oclusiva sorda siguiente en una aproximante [ð] (o, incluso, una fricativa [θ]): bastante realizado como [baðánte]. Esta variante es, en cualquier contexto y posición, minoritaria en la ciudad de Málaga.
- Otra posibilidad es que la /-s/ desaparezca completamente (6). Esta solución es muy poco frecuente en este contexto aunque sí es la variante preferida en Málaga en el resto de situaciones.

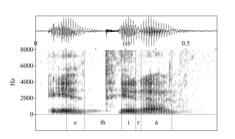
- Una de las soluciones mayoritarias en el contexto preoclusivo sordo es una variante que se podría denominar "postaspirada" (3) ya que la aspiración propiamente dicha se asocia a la consonante posterior, es decir, a la [t], que se percibe ahora como un sonido aspirado [t¹] (O'Neill, 2010; Torreira, 2012).
- La siguiente fase del proceso en este contexto particular es el sonido africado [ts] (punto 4): la /-s/ subvacente, que se había debilitado en posición implosiva, reaparece ahora en el ataque de la siguiente sílaba.

En la figura 2 se pueden observar los espectrogramas que diferenciarían las diferentes fases que se proponen para el proceso de resilabificación:

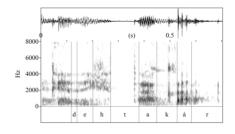
- 1. El primer espectrograma corresponde a la variante sibilante: se aprecia perfectamente el ruido de la fricativa alveolar sorda justo detrás de la vocal. Esta variante prácticamente no se escucha en la variedad de Málaga y, en este estudio, solo se documenta en los estilos de lectura.
- 2. El segundo espectrograma pertenece a una variante aspirada. En la imagen se observa cómo el ruido de la fricativa dura menos tiempo y se produce en frecuencias más bajas. Esta realización es muy poco frecuente en Málaga, tanto en los estilos informales como en los más formales de lectura.
- 3. La tercera imagen representa un caso en el que el ruido de la aspiración aparece ahora tras la barra de explosión de la oclusiva siguiente y da lugar a un VOT relativamente más largo que el que aparece en los casos de la sibilancia y la aspiración. Corresponde con lo que en este trabajo se denomina variante postaspirada.
- 4. El último espectrograma representa la última fase del proceso de resilabificación: el VOT de la oclusiva se llena de ruido fricativo propio de la alveolar sorda y da lugar a una variante africada.



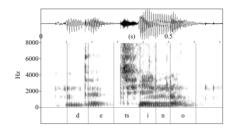




3. Postaspiración (estirar)



2. Aspiración (destacar)



4. Resilabificación (destino)

Figura 2: Espectrogramas de las diferentes fases del proceso de resilabificación de /-s/ ante dental tensa

2 Metodología

El estudio de las características acústicas del proceso de resilabificación de /-s/ en posición interior ante consonante dental tensa /t/ forma parte de una investigación más amplia cuyo objetivo último es el análisis, también acústico, de todo el proceso de debilitamiento y pérdida de /-s/, es decir, tanto en posición interior como final y en todos los contextos posibles.

Esta investigación particular se enmarca dentro del Proyecto para el Estudio Sociolingüístico del Español de España y de América (PRESEEA) de tal modo que aquí se ha asumido la propuesta metodológica de dicho proyecto.

Los datos se han obtenido a través de la grabación de entrevistas semidirigidas a una muestra representativa de hablantes estratificada según tres variables sociales básicas: sexo, edad y nivel de instrucción (cfr. tabla 1). En total se llevaron a cabo, durante el segundo semestre de 2013, 54 entrevistas realizadas a otros tantos hablantes procedentes de la ciudad de Málaga. Además de la entrevista, de unos 30 minutos de duración, cada informante leyó un texto, una lista de frases y una lista de palabras. En esta última, se incluyen palabras con /-s/ en posición interior ante diversos contextos consonánticos, palabras que no tienen /-s/ en la raíz, como pata o meto y un tercer grupo de palabras que tienen el fonema africado postalveolar /tʃ/, como techo o mucho. En la lectura del texto y de la lista de frases aparecen también casos de /-s/ en posición final. Los informantes tuvieron que rellenar además un cuestionario sociológico y de actitudes lingüísticas, que sirve para completar la información sociológica de los hablantes analizados.

Nivel de	Jóvenes (20-34 años)		Medianos (35-54 años)		Mayores (55 o más)	
instrucción	Hombres	Mujeres	Hombres	Mujeres	Hombres	Mujeres
Bajo	3	3	3	3	3	3
Medio	3	3	3	3	3	3
Superior	3	3	3	3	3	3

Tabla 1: Muestra estratificada de hablantes de la ciudad de Málaga

Los datos que aquí se presentan son todavía parciales, ya que corresponden al grupo de informantes de instrucción superior de entre 20 y 54 años de edad y, dado el carácter específico de este trabajo, se analiza sólo aquél grupo de palabras en los que el fonema /-s/ va seguido de oclusiva dental tensa /t/. El conjunto de datos analizados corresponde a los cuatro estilos recogidos: la entrevista, la lectura del texto, la lectura de frases y la lectura de palabras. De cada entrevista se analizaron unos 6 minutos aproximadamente: los 2 primeros corresponden a la parte inicial de la entrevista, los 2 siguientes a la parte central y los otros dos al fragmento final de la misma.

Los datos que aquí se presentan son todavía escasos para los objetivos del proyecto general, sin embargo son interesantes para observar las tendencias en el sentido de la hipótesis que se ha planteado.

En total se han analizado 900 palabras de las cuales:

- 741 son unidades que contienen distintas realizaciones de /-s/ subyacente en posición interior ante consonante dental tensa / t/ (pasta, estirar, historia, etc.).
- 57 son palabras que no presentan /-s/ subyacente ante /t/ (pata, meto, etc.). Se trata de un grupo de control que permite comparar las características acústicas de las distintas realizaciones de /vSt/ con las características acústicas de /vt/ cuando no existe ninguna /-s/ en la raíz.

Las 102 restantes son unidades léxicas que contienen el fonema africado postalveolar sordo /tʃ/ propio del español (cacho, techo, pichón, mocho y mucho). Se obtiene así un segundo grupo de control que facilita la comparación de las propiedades acústicas de la realización africada [ts] con las del fonema africado del español /tʃ/ (mosto, realizado como [mótso] con mocho).

Tras clasificar las palabras según el alófono de /-s/ que en ellas se percibe, se procedió a analizar las características acústicas de cada variante. Para ello se utilizó el programa *Praat* (Boersma/Weenik, 2013). Se analizaron, entre otros parámetros, los siguientes valores:

- La duración de la vocal que precede al segmento /-s/ (o a los segmentos /t/ o /tʃ/ en los casos de los grupos de control).
- En los casos en los que segmento /-s/ se realizaba como sibilante [s] o como aspiración [h], se medía la duración en ms. de dicho ruido y su centro de gravedad (COG), es decir, la concentración media de energía de un momento espectral de la consonante.
- En todos los casos se medía la duración total de la /t/, así como la duración del momento oclusivo y del VOT por separado.
- En los casos en los que el segmento /-s/ se realizaba como postaspirada [t^{*}] o como africada [ts] se analizó también el centro de gravedad y la intensidad del ruido que aparece tras la barra de explosión.

Para averiguar los valores referentes a la duración y a la intensidad de los segmentos analizados se utilizó el *script* elaborado por Wendy Elvira-García (2013): *Prosodic data extraction* en el Laboratorio de Fonética de la Universidad de Barcelona y para medir el centro de gravedad se empleó el *script* desarrollado por Christian DiCanio: *Spectral moments of fricative spectra* en los Laboratorios Haskins.

A continuación se llevó a cabo el análisis estadístico de los resultados con el programa SPSS© (*Statistical Package for the Social Sciences*) para *Mac* en su versión 21. Como se refleja en el punto 3, el análisis ha consistido, principalmente, en la realización de pruebas no paramétricas que reflejan las tendencias generales del comportamiento de los datos, todavía escasos. El objetivo es presentar las características acústicas de los cuatro alófonos de /-s/implicados en el proceso de resilabificación y compararlos entre sí.

Del mismo modo, se busca si hay diferencias significativas entre la realización de palabras que no presentan /-s/ en la raíz (pata) y los pocos casos en los que la /-s/ se elide en este contexto. Se trata de averiguar si la elisión de /-s/ se compensa de algún modo en el contexto.

También se comparan las características de la variante africada [ts] con [tʃ] para comprobar si entre ellas hay un margen amplio de distinción que hace posible la evolución del proceso.

3 Resultados

Los datos obtenidos en la ciudad de Málaga (tabla 2) reflejan que, en posición interior ante consonante dental sorda /t/, se escuchan principalmente dos realizaciones alofónicas de /-s/ implosiva: la postaspiración $[t^*]$, que aparece un 35,6 % de las veces y la variante africada [ts], que aparece un 31,4 % de las ocasiones. La pérdida es poco frecuente en este contexto (8,6%) y la preaspiración (3,9%) es prácticamente residual. La sibilancia aparece con más frecuencia (20,4%), pero, como se verá a continuación, es exclusiva de los estilos de lectura.

_

	N	%
Sibilancia [s]	151	20,4
Aspiración [h]	29	3,9
Postaspiración [th]	264	35,6
Africada [ts]	233	31,4
Pérdida [∅]	64	8,6
Total	741	100

Tabla 2: Distribución de las variantes de /-s/ ante dental tensa en doce informantes universitarios de la ciudad de Málaga

En el gráfico 1 se presentan las frecuencias relativas de aparición de cada variante de /-s/ según el estilo ordenado según el grado de atención: entrevista, lectura de un texto, lectura de frases y lista de palabras. Durante la entrevista, que es el estilo más espontáneo, las variantes mayoritarias son la postaspirada y la africada. La postaspiración es también la realización más frecuente durante la lectura del texto, seguida de la sibilancia que, como se puede observar, es la variante más frecuente en los dos estilos más cuidados: la lectura de frases y de la lista de palabras.

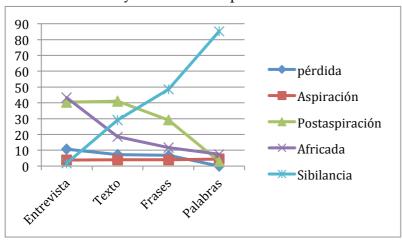


Gráfico 1: Aparición de las distintas variantes de /-s/ según el estilo

Estos datos, que se detallan en la tabla 3, indican que la variante más prestigiosa es la que corresponde al español estándar donde la /-s/ subyacente se asocia a la coda silábica. Esta variante, frecuente en los estilos gráficos, aparece con menor frecuencia conforme baja el nivel de atención que el informante presta a su forma de pronunciar y parece trasladarse a la posición inicial de sílaba dando lugar, en este contexto particular, a las dos variantes postaspiradas, la postaspiración propiamente dicha y la realización africada.

	Entrevista	Texto	Frases	Palabras	Total		
Pérdida [∅]	48	9	7	0	64		
	10,7 %	7,3 %	6,8 %	0,0 %	8,6 %		
Aspiración [h]	17	5	4	3	29		
	3,8 %	4,0 %	3,9 %	4,5 %	3,9 %		
Postaspiración	181	51	30	2	264		
[t ^h]	40,5 %	41,1 %	29,1 %	3,0 %	35,6 %		
Resilabificación	193	23	12	5	233		
[ts]	43,2 %	18,5 %	11,7 %	7,5 %	31,4 %		
Sibilancia [s]	8	36	50	57	151		
	1,8 %	29,0 %	48,5 %	85,1 %	20,4 %		
Total	447	124	103	67	741		
	100 %	100 %	100 %	100 %	100 %		
Coeficiente Chi cuadrado de Pearson: 343,39 (12). Sig.: ,000							
	V de Cramer: ,393. Sig.: ,000						

Tabla 3: Distribución de las variantes de /-s/ según el estilo. Frecuencias absolutas, relativas y significación estadística

Los datos presentados en la tabla 3 parecen indicar que, en la ciudad de Málaga, el proceso de debilitamiento del segmento /-s/ en posición interior ante oclusiva dental tensa da lugar a un proceso de resilabificación, por el cual dicho segmento pasa a ocupar la posición de ataque de la sílaba siguiente dando lugar a dos variantes, una postaspirada y otra africada.

En la tabla 4 se presentan los datos correspondientes a la distribución de las diferentes variantes de /-s/ realizadas durante la entrevista según los dos grupos de edad manejados en la muestra aquí analizada. Como se puede observar, son los hablantes más jóvenes los que propician el uso de la variante africada. La muestra aquí analizada es pequeña, por lo que todavía es pronto para sacar conclusiones definitivas de estos datos, pero la tendencia apunta a que la variante africada es relativamente innovadora dentro de la variedad utilizada en la ciudad de Málaga.

	20-34	35-54	Total		
Pérdida [Ø]	19	29	48		
	7,3 %	15,5 %	10,7 %		
Aspiración [h]	4	13	17		
	1,5 %	7,0 %	3,8 %		
Postaspiración [th]	102	79	181		
	39,2 %	42,2 %	40,5 %		
Resilabificación	134	59	193		
[ts]	51,5 %	31,6 %	43,2 %		
Sibilancia [s]	1	7	8		
	0,4 %	3,7 %	1,8 %		
Total	260	187	447		
	100,0 %	100,0 %	100,0 %		
Coeficiente Chi cuadrado de Pearson: 32,357 (4). Sig.: ,000					
	V de Cramer	: ,269. Sig.: ,000	-		

Tabla 4: Distribución de las variantes de /-s/ según grupos de edad en el estilo entrevista. Frecuencias absolutas, relativas y significación estadística

En la tabla 5 se presentan los valores medios resultantes de analizar, para cada una de las variantes implicadas en el proceso de resilabificación (sibilante,

aspirada, postaspirada y africada), la duración de los dos principales momentos de la oclusiva posterior: el cierre y el VOT y el centro de gravedad del ruido de la fricción (independientemente de si este se ha producido antes o después de la consonante oclusiva)

	Sibilancia [s]	Aspiración [h]	Postaspiración [tʰ]	Africada [ts]	Sig. (Kruskal-
			[-1		Wallis)
Número de casos	151	29	264	233	
Duración del cierre	0.07 ± 0.02	0.06 ± 0.02	0.05 ± 0.02	0.04 ± 0.02	,000
Duración del VOT	0.02 ± 0.01	0.03 ± 0.01	0.04 ± 0.01	0.06 ± 0.01	,000
Centro de gravedad	5824 ± 1437	3723 ± 1047	4031 ± 1397	5278 ± 1735	,000

Tabla 5: Comparación de los parámetros acústicos de /-s/ según su realización. Medias, desviación típica y significación estadística

Tal y como se puede observar, conforme se avanza en el proceso de resilabificación, el momento de cierre de la oclusión se va reduciendo mientras que el tiempo de emisión de la voz (VOT) se va haciendo más largo. Por otro lado, el centro de gravedad del ruido de la fricción es mayor en las dos variantes extremas del proceso (sibilante y africada) y disminuye en las variantes aspiradas.

En la tabla 6 se presentan los valores de la duración tras recodificar en dos grupos las variantes implicadas: sibilante [s] y aspirada [h] por un lado, y postaspirada [th] y africada [ts] por otro. Como se puede comprobar las diferencias entre los dos grupos son estadísticamente significativas: las variantes postaspiradas [th] y [ts] presentan un VOT más largo y un cierre más corto en la consonante posterior que la preaspirada y la sibilante.

	Sibilante y aspirada [s] y[h]	Postaspirada y africadas [tʰ] y [ts]	Significación (U de Mann-Whitney)
Número de casos	180	497	
Duración del	0.07 ± 0.02	0.04 ± 0.02	,000
cierre			
Duración del VOT	0.02 ± 0.01	0.05 ± 0.01	,000

Tabla 6: Comparación de los valores medios del cierre y de la duración del VOT, según el tipo de realización, distensiva o tensiva. Medias, desviación típica y significación estadística

Estos datos apoyan la idea de que la /-s/ subyacente, que en el español estándar aparece en la posición de coda silábica, se asocia en la variedad de Málaga y en este contexto particular, al ataque de la sílaba posterior, que presenta ahora una mayor duración del VOT dando lugar a variantes postaspiradas.

En cuanto al centro de gravedad de la fricción, la tabla 7 refleja cómo las diferencias entre los grupos oponen las dos variantes aspiradas frente a las dos sibilantes. Los valores del centro de gravedad de la variante sibilante y de la variante africada son significativamente superiores a los que se obtienen al medir las variantes aspiradas.

	Sibilante y africada [s] y [ts]	Aspirada y postaspirada [h] y [tʰ]	Significación (U de Mann-Whitney)
Número de casos	348	293	
Centro de gravedad	5492 ± 1645	4001 ± 1368	,000

Tabla 7: Comparación de los valores medios del COG según el tipo de realización. Medias, desviación típica y significación estadística

Este último dato refuerza la hipótesis de la resilabificación: las características acústicas del ruido de la variante sibilante son similares a la del ruido de la variante africada, ya que se trata de un desplazamiento de posición de una misma sustancia fónica.

La resilabificación es una solución óptima en el contexto preoclusivo dental, ya que el mantenimiento de toda la sustancia fónica en el ataque permite la diferenciación de posibles pares mínimos sin renunciar a una mayor naturalidad en la pronunciación, naturalidad que se pierde al realizar las variantes sibilante y aspirada.

La pérdida absoluta dificultaría la distinción de pares mínimos, ya que si se comparan aquellos casos en los que se realiza la variante elidida con el grupo de control en los que el hablante pronuncia una palabra sin /-s/ subyacente, se observa que las únicas diferencias significativas afectan a la duración de la vocal anterior, pero en el sentido contrario al esperado, es decir, que la vocal es más larga cuando no existe una /-s/ subyacente que cuando esta existe pero se ha eliminado completamente en la pronunciación (ver tabla 8).

	Pérdida [∅]	Palabra sin /-s/	Significación (U de Mann-Whitney)
Número de casos	64	57	
Duración de la vocal previa	$0,094 \pm 0,033$	$0,159 \pm 0,036$,000

Tabla 8: Duración de la vocal anterior a la elisión de /-s/ frente a la de la ausencia de /-s/. Medias, desviación típica y significación estadística

Los datos de la tabla 8 demuestran que no hay diferencias significativas que permitan distinguir la [t] resultante de una elisión de /-s/ frente a la [t] que se realiza cuando tal /-s/ no existe en la raíz. La mayor duración de la vocal previa en el caso de las palabras que no llevan /-s/ en su estructura lexemática es contraria a la hipótesis que considera que la pérdida de /-s/ se debe compensar con una mayor duración de la vocal anterior. Aunque es necesario el análisis de más casos para la explicación de este comportamiento particular, sí se puede pensar que los datos que aquí se presentan apoyan la idea de que la variante africada es una solución óptima dentro del proceso de debilitamiento de /-s/, ya que mantiene toda la sustancia fónica de la raíz en una posición más natural y no da lugar a la posibilidad de confundir significados.

Ahora bien, si la aparición de un sonido africado [ts], puede resultar una solución óptima que permite distinguir pares mínimos como *moto* y *mosto*, realizado este último elemento como [mótso], cabe preguntarse si la utilización de este alófono podría impedir la distinción de parejas como *mosto* y *mocho*, realizadas como [mótso] y [mótfo] respectivamente, o como *castillo* y *cachillo* [katsíʎo] y [katʃíʎo]. Para comprobar este extremo se han comparado las realizaciones de la variante africada de /-s/ con el grupo de control en el que

tabla 9.			
	Resilabificación [ts]	Africada [ʧ]	Significación (U de Mann-Whitney)
Número de casos	233	102	
Duración de la consonante	0.10 ± 0.02	0.12 ± 0.03	,000
Duración del cierre	0.04 ± 0.02	0.05 ± 0.02	,000
Duración del VOT	0.05 ± 0.01	0.07 ± 0.02	,000
Contro do grazzadad ruido	5278 ± 1735	4420 ± 1188	000

los hablantes pronunciaron palabras con / t J / Los resultados se presentan en la tabla 9.

Tabla 9: Valores acústicos de los dos sonidos africados en convivencia por el proceso de resilabificación. Medias, desviación típica y significación estadística

Como se puede comprobar en la tabla 9, las diferencias entre el sonido africado [ts], resultante del proceso de resilabificación de /-s/ y el fonema africado [tʃ] propio del sistema español son claras y estadísticamente significativas. Existe un margen amplio entre ambos sonidos que permite que el proceso de resilabificación avance sin peligro de que puedan confundirse pares mínimos como *castillo* y *cachillo*: el sonido africado [tʃ] es significativamente más largo y su centro de gravedad se sitúa en frecuencias medias más bajas que el sonido africado alveolar [ts].

4 Conclusiones

Aunque los resultados que aquí se presentan son todavía parciales, de los datos analizados y presentados en el punto anterior se pueden extraer las siguientes conclusiones:

- 1. Entre los hablantes de instrucción superior, la realización del fonema /-s/ en posición interior ante consonante dental tensa /t/ ha desaparecido ya prácticamente de la posición implosiva y se realiza en la tensión de la sílaba posterior, tal y como demuestra el hecho de que las variantes mayoritarias son la postaspirada y la africada.
- 2. La variante africada, considerada aquí como la fase más avanzada de todo el proceso de resilabificación, es favorecida por los hablantes más jóvenes.
- 3. Las características acústicas que diferencian a las dos variantes innovadoras de las conservadoras son la posición del ruido, el aumento progresivo de la duración del VOT de la dental y una disminución del tiempo que dura el momento de cierre.
- 4. En tercer lugar, la variante africada comparte con la sibilante el centro de gravedad del ruido de la fricción. Se asiste, por tanto, a un proceso de resilabificación en el que la sustancia fónica se traslada desde la posición implosiva hacia la tensión silábica.
- 5. La variante africada resultante de dicho proceso de resilabificación supone una solución óptima ya que permite la realización de una sílaba abierta sin perder sustancia fónica. Del mismo modo, esta variante impide la confusión de pares mínimos.
- 6. Por último, las diferencias acústicas que se registran entre la variante africada alveolar [ts] y la postalveolar [tʃ] son suficientemente amplias, de tal modo que el proceso de resilabificación no resulta limitado.

Cabe afirmar que en la ciudad de Málaga se asiste a un proceso de resilabificación del segmento /-s/ en posición interior ante dental tensa /t/. Dicho proceso da lugar a la aparición de un sonido africado alveolar [ts] que

implica la recuperación de toda la sustancia fónica de la raíz lexemática de las palabras implicadas. El material fónico de la raíz, que etimológicamente produce una sílaba cerrada, se traslada ahora al ataque de la sílaba siguiente dando como resultado una sílaba abierta: se mantiene, por tanto, toda la sustancia lexemática sin renunciar a la naturalidad en la pronunciación.

5 Bibliografía

- Moya Corral, J. A. (2007). Noticia de un sonido emergente: la africada dental procedente del grupo -st- en Andalucía. *Revista de Filología* 25, 457-465
- O'Neill, P. (2010). Variación y cambio en las consonantes oclusivas del español de Andalucía. Estudios de Fonética Experimental, 19, 11-41.
- Ruch, H. (2008). La variante [ts] en el español de la Ciudad de Sevilla: aspectos fonético-fonológicos y sociolingüísticos de un sonido innovador. Zürich. Romanisches Seminar der Universität Zürich.
- Torreira, F. (2012). Investigating the nature of aspirated stops in Western Andalusian Spanish. *Journal of the International Phonetic Association* 42/1, 49-63.
- Vida Castro, M. (2004). Estudio sociofonológico del español hablado en la ciudad de Málaga. Alicante, Universidad.

Parte VII. Tecnologías del habla

Una herramienta para la transcripción prosódica automática con etiquetas Sp_ToBI en Praat

ELVIRA-GARCÍA, WENDY; ROSEANO, PAOLO; FERNÁNDEZ PLANAS, ANA MA.; MARTÍNEZ CELDRÁN, EUGENIO

1 Introducción

Los estudios lingüísticos de las últimas décadas han resaltado la importancia de la prosodia y, en especial, la entonación, dentro del conjunto de los sistemas lingüísticos. Este hecho, junto con las posibilidades técnicas que ahora se abren para su estudio, ha posibilitado el aumento de los estudios sobre entonación. El carácter complejo de la entonación, donde hay que tener en cuenta simultáneamente cuestiones como la amplitud y la alineación de los movimientos tonales, conlleva que la manera idónea de estudiarla sea a través de transcripciones de la curva entonativa, para lo que se han venido creando durante el siglo XX diferentes modelos de entonación (para una reseña véase Prieto, 2002). Sin embargo, la transcripción prosódica exige una cantidad de recursos humanos y de pericia del investigador (Syrdal, Hirschberg, McGory & Beckman, 2001) que muchas veces no es rentable para el objetivo que persigue la investigación.

Esto ha hecho que desde la misma invención de los modelos de transcripción se haya intentado crear paralelamente instrumentos informáticos capaces de acelerar la labor del investigador (Lea, 1980; Rietveld, 1984). Por una parte, se han creado interfaces de usuario que ayudan a acelerar la transcripción manual (Syrdal *et al.*, 2001). Y por otra, transcriptores prosódicos automáticos que

realizan un análisis de la curva de F0 y la etiquetan de acuerdo con las convenciones del sistema transcripción para el que fueron creados (Hirst y Espesser, 1993; Mertens, 2004; Rosenberg, 2010).

La transcripción prosódica automática, como el resto de los componentes de las tecnologías del habla, se puede abordar desde dos perspectivas. Por un lado, es posible llevarla a cabo usando modelos estadísticos predictivos y técnicas de *machine learning*. Por otro, se puede realizar implementando el conocimiento lingüístico disponible en la actualidad sobre una lengua.

Si bien es cierto que, hasta el momento, los sistemas de reconocimiento y síntesis de habla basados en estadística han demostrado ser los más productivos y son, de hecho, los que se han implementado los sistemas que se comercializan hoy en día, el análisis basado únicamente en el conocimiento lingüístico es posible y, aunque hasta ahora había sido menos fiable, permite arrojar luz sobre los fenómenos lingüísticos que todavía desconocemos.

Este trabajo tiene dos objetivos. Por una parte, nos proponemos presentar Eti-ToBI, una herramienta lingüística para la transcripción automática de la entonación del español que usa las convenciones de notación Sp_ToBI.

Por otro lado, este trabajo pretende mostrar que un transcriptor basado en el conocimiento lingüístico es posible y, por tanto, aplicable al reconocimiento de voz. De esta manera, se quiere estimular el desarrollo de los sistemas de reconocimiento de voz basados en conocimiento de los sistemas lingüísticos, para revalorizar la labor del lingüista en el campo de las tecnologías del habla.

1.1 Los transcriptores prosódicos automáticos basados en sistemas ToBI

De los modelos de transcripción prosódica existentes en la actualidad, las convenciones Tones and Break Indices (ToBI) constituyen, seguramente, el modelo que goza de más popularidad. Los sistemas ToBI son un conjunto de normas de transcripción prosódica específicas para cada lengua que están basadas en el modelo métrico-autosegmental o AM (Pierrehumbert, 1980). Esta teoría entiende la entonación como una sucesión de tonos en la que estos son autosegmentos que se anclan a las posiciones prominentes de la frase entonativa (sílabas acentuadas y límites de frase). A partir de estos fundamentos teóricos, se han creado una serie de convenciones para la transcripción prosódica fonológica de corpus, en un primer momento, para el inglés, MAE_ToBI (Beckman y Elam, 1997), y, más tarde, para otras lenguas entre las que se encuentra el español, Sp_ToBI (Beckman, Díaz-Campos, McGory y Morgan, 2002; Estebas-Vilaplana y Prieto, 2008; Face y Prieto, 2007; Prieto y Roseano, 2010).

Los sistemas de anotación ToBI, en principio, constan de cuatro niveles (*tiers*) de transcripción. En el primer nivel, se detalla la sucesión de tonos de la frase; en el segundo, la transcripción segmental; en el tercero, la separación prosódica de los componentes de la frase (*Break Indices*); y en el cuarto se anotan posibles comentarios. No obstante, en algunos sistemas ToBI (Jun, Lee, Kim & Lee, 2010), se ha decidido incluir un nivel más para posibilitar una transcripción fonética de la curva entonativa. Para el español, la propuesta de cariz fonético convive con la fonológica desde 2002 (Dorta, 2013; Fernández Planas, Martínez Celdrán, Salcioli Guidi, Toledo & Castellví Vives, 2002; Martínez Celdrán & Fernández Planas, 2003; Roseano & Fernández Planas, 2013).

En Sp_ToBI, se etiquetan dos tipos de eventos tonales, los acentos tonales (pitch accents) y los tonos de frontera (boundary tones). Los acentos tonales están

relacionados con la sílaba tónica y en la notación se marca con un asterisco (*) el tono que coincide con la misma. Los tonos de frontera pueden estar relacionados con el límite de una frase intermedia (-) o de una frase entonativa (%). El sistema ToBI (como el modelo AM) está, en principio, basado en niveles, es decir, anota los tonos dependiendo de si estos son altos (H) o bajos (L). Para el español, además de los dos niveles básicos propuestos en las primeras versiones del sistema, se contempla un nivel medio (!H) en los tonos de frontera y uno extraalto (¡H) en los acentos tonales. Así, la curva entonativa se configura como una sucesión de tonos de diferentes niveles. En Sp_ToBI, se prevén dos tipos de acentos tonales: monotonales (T*) y bitonales (T*+T o T+T*) y para los tonos de frontera, tres tipos: monotonales (T%), bitonales (TT%) y tritonales (TTT%). El último acento tonal de la frase entonativa junto con el tono de frontera configuran el núcleo o configuración nuclear, mientras que todos los acentos anteriores a ese se clasifican como prenucleares.

Este sistema, debido a su generalización de uso y a su especificidad para cada lengua, es probablemente el sistema que más aplicaciones de transcripción automática tiene. La primera de ellas que conviene considerar es anterior al propio nacimiento del sistema. Pierrehumbert (Pierrehumbert, 1983) creó un sistema de reconocimiento para el modelo AM que buscaba las inflexiones de la curva entonativa a través de derivadas para después aplicar unas condiciones que atribuían a cada movimiento su acento tonal correspondiente.

A partir de ese momento los sistemas de transcripción basados en AM dejaron de basarse solamente en el conocimiento lingüístico de la lengua que describen y pasaron a basarse en modelos de predicción estadística. Estos modelos predictivos, basados en técnicas de regresión y machine learning exigen un etiquetaje manual previo de un gran número de datos que usan como entrenamiento para predecir la transcripción de nuevos datos. Pero, por grande que sea la cantidad de datos que manejan, esta siempre es finita, lo que hace que sea virtualmente imposible conseguir un modelo de predicción capaz de prever todos los casos con estas técnicas. Los modelos de predicción utilizados dependían en gran parte del avance del reconocimiento de habla para unidades segmentales. Y esto hizo que las técnicas de reconocimiento de la prosodia utilizadas coincidieran con aquellas que estaban en boga en el mundo del reconocimiento de habla de las unidades segmentales en cada momento.

Así, encontramos modelos basados en árboles de decisión (Black & Hunt, 1996; Lee, Kim & Lee, 2002), regresión lineal (Rietveld, 1984), modelos de Markov (Ross & Ostendorf, 1996; Wightman & Ostendorf, 1994), principio de máxima entropía (Sridhar, 2008) y la combinación de varios de éstos (Wagner, 2008). Pero probablemente el transcriptor prosódico automático MAE_ToBI que más éxito ha tenido es el creado por Rosenberg (2010), programado en Java, no se limita a la transcripción entonativa, sino que funciona de manera completamente autónoma, localiza fronteras, los posibles acentos tonales y los etiqueta.

Ādemás de para el inglés, ha habido intentos de crear transcriptores prosódicos para otras lenguas como el koreano (Lee *et al.*, 2002), el japonés (Campbell, 1996; Noguchi & Kiriyama, 1999), el italiano (Savino, Refice & Daleno, 2002) o el sueco (Frid, 1999), entre otros.

1.2 Los transcriptores prosódicos automáticos del español

En la actualidad, también existen transcriptores prosódicos automáticos para el español. De ellos, algunos están basados en otros modelos prosódicos como

el IPO (Garrido Almiñana, 2008) y, en un caso, el transcriptor efectúa un análisis con el sistema Sp_ToBI (Escudero-Mancebo, González-Ferreras, Vivaracho-Pascual & Cardeñoso-Payo, 2014). Este transcriptor, como la mayoría de los mencionados hasta ahora, se apoya en métodos estadísticos, concretamente en la lógica difusa, para realizar una propuesta de etiquetaje que proporciona con un nivel de certeza y en última instancia es el investigador quien decide si esa propuesta de tonos es la más adecuada para cada caso. El transcriptor está siendo utilizado actualmente para acelerar la transcripción del corpus del español Glissando, únicamente con frases declarativas.

2 Metodología

Para cumplir los objetivos que se han expuesto en el apartado 1, se ha creado un script de Praat (Boersma y Weenink, 2014) que etiqueta los tonos en tres niveles: uno superficial o fonético, uno profundo o fonológico y uno estandarizado. Puede etiquetar, en el nivel superficial, 13 tipos de acentos prenucleares, 15 nucleares y 10 tipos de tonos de frontera, lo que supone un total de 150 configuraciones nucleares. En el nivel profundo, distingue entre 9 acentos prenucleares, 8 nucleares y 10 tonos de frontera: un total de 80 configuraciones nucleares que son posteriormente estandarizadas. El script puede reconocer los patrones descritos en Prieto y Roseano (2010); reconoce, por tanto, los patrones de variedades de español habladas en Castilla, Cantabria, islas Canarias, República Dominicana, Puerto Rico, zona andina de Ecuador y Venezuela, Chile, Argentina y México.

El script trabaja a partir de las sílabas tónicas de la frase, por lo tanto, su uso requiere que los audios tengan un TextGrid asociado con una separación de sílabas en un *tier* de intervalos y, en ese *tier*, una marca (a escoger por el investigador) en los intervalos que correspondan a una sílaba tónica.

2.1 Formalización del sistema

El primer paso para el análisis automático es el tratamiento de la señal. En este caso para procesar la señal y facilitar la búsqueda de F0 se han tomado dos medidas cautelares. La primera consiste en pasar un filtro que permite mantener solo las frecuencias inferiores a 1000 Hz; de esta manera se palian las distorsiones que los sonidos de altas frecuencias como las fricativas puedan causar en la curva de F0. La segunda medida consiste en calcular el rango de la frase entonativa automáticamente para cada producción. Esta operación se ha efectuado con la técnica en dos pasos descrita por Hirst (2011) en la que, en el primer paso, se extrae la curva de F0 de la frase con un rango muy amplio para obtener el mínimo y el máximo de la frase. En un segundo paso se extrae un nuevo objeto de F0 donde el mínimo del rango se establece multiplicando por 0,75 el primer cuartil del rango anterior y el máximo multiplicando por 1,5 el tercer cuartil (De Looze, 2010).

Seguidamente, se procede al análisis de la curva entonativa obtenida. Este análisis es posible gracias al uso de una transcripción en tres niveles que permite efectuar una primera transcripción fonética objetiva. El paso del nivel 1 al 3 de análisis nos lleva a una aproximación de la representación fonológica progresiva. Pero hay momentos en los que solo un etiquetador humano puede realizar una transcripción realmente fonológica. Es el caso, por ejemplo, de los procesos de truncamiento (con palabras agudas) en los que el tono de frontera no se llega a realizar en un plano superficial.

Como se ha resaltado para otras lenguas (Lee et al., 2002), la transcripción prosódica automática en sistemas ToBI es solo posible si se parte de una transcripción prosódica fonética, que constituye una de las características básicas de las aproximaciones del Laboratori de Fonètica de la UB (Fernández Planas et al., 2002; Martínez Celdrán y Fernández Planas, 2003; Roseano y Fernández Planas, 2013). En este primer nivel de etiquetaje, se transcriben todos los movimientos que son susceptibles de ser percibidos de manera lingüística, esto es los movimientos que superan 1,5 st (Pamies, Fernández Planas, Martínez Celdrán, Ortega-Escandell y Amorós Céspedes, 2002). Los movimientos se pueden buscar en puntos fijos determinados *a priori* por el script aplicando los conocimientos sobre la prosodia del español (ej. entre inicio y final de tónica) o en puntos determinados por la misma curva, es decir, diferencias entre el pico y el valle de un movimiento. A partir de estos movimientos, se establece un etiquetaje con etiquetas tritonales que dan cuenta de los mismos. El resultado es una transcripción que podemos equiparar a la transcripción fonética estrecha en fonética segmental. Esta transcripción es independiente de la lengua, en el sentido en que cada movimiento se etiqueta usando símbolos que se corresponden con la información acústica de la curva.

El segundo nivel, el nivel profundo, es la equivalencia suprasegmental a una transcripción fonética ancha. En ella, los datos acústicos se interpretan de acuerdo a la fonología entonativa del español. Es en este nivel donde se tienen en cuenta las reglas de implementación fonológica específicas para el español. Estas reglas permiten simplificar las etiquetas tritonales que pudiera haber en el primer nivel y convertirlas en etiquetas bitonales relevantes para el español. En este nivel, además, se tienen en cuenta los movimientos descendentes de la curva y se comprueba si son fonológicos o causados por el fenómeno de declinación. Esto afecta a dos tipos de movimientos: en el prenúcleo, acentos tonales desacentuados, que fonéticamente se producen como H+L*; y en el núcleo, acentos descendentes fonéticos (H+L* L%) que en realidad no tienen una diana alta fonológica y son por tanto simplemente bajos (L* L%). Estas conversiones son posibles gracias a que en español un acento tonal H+L* consiste en un descenso dentro de la sílaba tónica donde el primer punto de la tónica está todavía en un nivel alto (Estebas-Vilaplana y Prieto, 2010; Prieto, 2014).

En el tercer nivel, el estandarizado, la transcripción del nivel profundo se normaliza para hacerla coincidir con el sistema actual de Sp_ToBI. Esta estandarización es necesaria únicamente para las configuraciones nucleares por: a) razones de convención del sistema (ya que el sistema Sp_ToBI no admite tonos en la cola en la configuración nuclear, así, algo como L*+H L% pasaría a ser L* HL%) o b) contrastes fonológicos que se dan solo para ciertas configuraciones, por ejemplo, !H% solo es fonológico cuando va detrás de L+H*.

3 Resultados

Para evaluar la actuación del script se ha comparado el etiquetaje proporcionado por este (en concreto, el etiquetaje del tercer tier o normalizado) con el realizado por un etiquetador humano. El análisis estadístico del nivel de acuerdo entre etiquetadores se ha llevado a cabo con la técnica de la *kappa* de Cohen, que sirve para medir el nivel de acuerdo entre dos etiquetadores descartando los acuerdos que hayan tenido lugar debido al azar (Cohen 1960;

La última versión del sistema de etiquetaje superficial se puede consultar en Roseano y Fernández Planas (2013).

Cohen 1968), y se ha calculado con una herramienta en línea (GraphPad 2014). Los valores del estadístico oscilan entre -1 (desacuerdo total) hasta 1 (acuerdo perfecto) donde 0 correspondería al nivel del azar. Normalmente, un kappa a partir de .60 se considera bueno.

Para llevar a cabo la comparación se etiquetó un corpus basado en la emisión de 1186 frases producidas por 4 hablantes de 4 variedades diferentes de español peninsular: español de Cantabria, español de Madrid, español de Sevilla y español de Barcelona. Las hablantes, todas mujeres, tenían unas edades comprendidas entre los 21 y 28 años y eran informantes representativas de su punto de encuesta. Los datos se grabaron a través de una grabadora Marantz PMD620 conectada a un micrófono direccional Shure SM58 y las grabaciones se realizaron en casa de los informantes.

La transcripción prosódica manual fue realizada por un único experto transcriptor humano. La decisión de realizar la comparación con un sólo individuo se debe al gran tamaño del corpus. Sin embargo, esto no supone un problema metodológico debido a que el transcriptor ha participado anteriormente en experimentos de nivel de acuerdo entre etiquetadores ToBI y sus resultados han sido coherentes con los del resto de transcriptores (Roseano, Fernández Planas, Elvira García, Cerdà Massó y Martínez Celdrán, aceptado).

La comparación de las transcripciones se realizó dividiendo los eventos tonales en: 1) acentos prenucleares, 2) acentos nucleares, 3) tonos de frontera, ya que el script usa diferentes fórmulas para cada uno de ellos.

Entre todos ellos hay un nivel de acuerdo >80 % y valores de kappa por encima de .7. Si además tenemos en cuenta los valores de kappa ponderados, el nivel de acuerdo sube hasta .8 lo que se considera un muy buen nivel de acuerdo (tabla 1).

Evento tonal	n	% Acuerdo	Kapp a	Kappa ponderado
Acentos prenucleares	1660	94.94 %	0.907	0.859
Acentos nucleares	1186	88.11 %	0.831	0.832
Tonos de frontera	1186	81.28 %	0.756	0.822

Tabla 1. Valores de acuerdo y kappa entre el etiquetador automático Eti-ToBI y el etiquetador humano

4 Discusión

Tal y como se ha visto, los resultados de fiabilidad del *script* son muy buenos, pero no alcanzan una fiabilidad del 100 %. Esto tiene dos causas. La primera de ellas es la limitación técnica que supone trabajar con algoritmos de predicción de F0 a partir de datos acústicos, y no trabajar a partir de los datos reales. Esto hace que, en final de emisión, donde hay ensordecimiento, el algoritmo de detección de pitch de Praat pueda interpretar erróneamente periodicidad donde no la hay y en esos casos el script coloca un tono en consecuencia.

La segunda es debida a un problema intrínseco del mismo sistema de transcripción, el Sp_ToBI. Es numerosa la bibliografía (tabla 2) que resalta que, aunque los sistemas ToBI son fiables, el grado de acuerdo entre diferentes etiquetadores nunca es perfecto. Por lo tanto, en un nivel fonológico, no así en el fonético, son abundantes los casos de ambigüedad donde existe más de una etiqueta posible. En estos casos, la etiqueta proporcionada por Eti-ToBI puede no coincidir con la etiqueta que darían los investigadores.

	Acentos tonales		Tonos de frontera		
Sistema	% acierto	Kappa	% acierto	Kappa	Fuente
Cat_ToBI	0.61	0.46	0.86	0.69	Escudero et al. 2012
					Syrdal y McGory
AM_ToBI	0.72	0.67	0.86	0.84	2000
E_ToBI	0.86	0.51	0.89	0.79	Yoon et al. 2004
G_ToBI	0.71	-	0.86	-	Grice et al. 1996
K_ToBI	0.52	-	0.82	-	Jun et al. 2000

Tabla 2. Acuerdo entre etiquetadores humanos para diferentes sistemas ToBI.

A este último problema también se enfrentan los transcriptores basados en estadística, que en estos casos suelen calcular cuál es la etiqueta más probable y colocarla o dar al investigador los datos para que él elija. En el caso de Eti_ToBI nuestro objetivo no es proporcionar al investigador una etiqueta probable, sino proveer una etiqueta acústicamente posible y fonológicamente descriptiva. En cualquier caso, el único modo de salvar este obstáculo es conseguir una descripción de la entonación del español que dé cuenta de manera detallada de la interacción fonético-fonológica y esto sólo es posible a través de la investigación.

5 Conclusiones

En este trabajo se ha presentado un transcriptor prosódico automático con etiquetas Sp_ToBI basado únicamente en los datos acústicos y el conocimiento fonológico de la entonación, un instrumento que no usa técnicas de predicción estadística. El sistema es válido para un amplio abanico de variedades del español, de las cuales reconoce una gran cantidad de contornos.

El trabajo prueba que, a partir de las formalizaciones del conocimiento lingüístico adecuadas, es posible llegar a un nivel de fiabilidad equiparable al de los transcriptores humanos y los transcriptores automáticos basados en predicción estadística. En este caso, la formalización que ha permitido llegar a ese nivel de fiabilidad es un análisis prosódico fonético que se ha probado como una transcripción objetiva, clara y universal basada únicamente en datos acústicos. Esta transcripción, por no estar condicionada por la lengua, sería aplicable a otros sistemas ToBI.

6 Bibliografía

Beckman, M., Díaz-Campos, M., McGory, J. T. & Morgan, T. A. (2002). Intonation across Spanish, in the Tones and Break Indices framework. *Probus*, *14*, 9-36. doi:10.1515/prbs.2002.008

Beckman, M., & Elam, G. A. (1997). *Guidelines for ToBI Labelling*. The Ohio
State University Research
Foundation.

Black, A. W., & Hunt, A. J. (1996).

Generating F 0 contours from ToBI labels using linear regression. In ICSLP 96. Fourth International Conference on Spoken Language Proceedings (pp. 1385-1388).

Philadelphia: IEEE. doi:10.1109/ICSLP.1996.607872

Boersma, P., & Weenink, D. (2014). *Praat:* doing phonetics by computer. En http://www.praat.org/

Campbell, N. (1996). Autolabelling Japanese ToBI. En ICSLP 96. Fourth International Congress on Conference on Language Processing Proceedings, vol. 4, 2399-2402. Philadelphia: IEEE. En http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=607292

De Looze, C. (2010). Analyse et interprétation de l'empan temporel des variations prosodiques en

- français et en anglais. Aix-en-Provence. En http://halshs.archivesouvertes.fr/tel-00470641/
- Dorta, J. (ed.). (2013). Estudio comparativo preliminar de la entonación de Canarias, Cuba y Venezuela. Madrid-Sta. Cruz de Tenerife: La Página ediciones.
- Escudero, D., Aguilar, L., Vanrell, M. del M., & Prieto, P. (2012). Analysis of inter-transcriber consistency in the Cat_ToBI prosodic labeling system. *Speech Communication*, *54* (4), 566-582. En http://www.sciencedirect.com/scie
 - http://www.sciencedirect.com/science/article/pii/S0167639311001749
- Escudero-Mancebo, D., González-Ferreras, C., Vivaracho-Pascual, C., & Cardeñoso-Payo, V. (2014). A fuzzy classifier to deal with similarity between labels on automatic prosodic labeling. En *Computer Speech y Language*, vol. 28, 326-341. doi:10.1016/j.csl.2013.08.001
- Estebas-Vilaplana, E., & Prieto, P. (2008). La notación prosódica del español: una revisión del Sp-ToBI. *Estudios de Fonética Experimental*, 17, 264-283. En http://www.raco.cat/index.php/EFE/article/view/140072/0
- Estebas-Vilaplana, E., & Prieto, P. (2010). Castilian Spanish intonation. Transcription of Intonation of the Spanish Language, Lincom Europa, München, 17-48.
- Face, T., & Prieto, P. (2007). Rising accents in Castilian Spanish: a revision of Sp-ToBI. *Journal of Portuguese Linguistics*, 6 (1), 117.
- Fernández Planas, A. M., Martínez Celdrán, E., Salcioli Guidi, V., Toledo, G., & Castellví Vives, J. (2002). Taxonomía autosegmental en la entonación del español peninsular. En *Actas del II Congreso de Fonética Experimental*,180-186. Sevilla.
- Frid, J. (1999). An environment for testing prosodic and phonetic transcriptions. En *Proceedings of ICPhS* 99, 2319-2322). San Francisco. En http://lup.lub.lu.se/record/529087/file/1624474.pdf
- Garrido Almiñana, J. M. (2008). Modelling Spanish Intonation for Text-to-Speech Applications. Universitat Autònoma de Barcelona. En http://www.tdx.cat/handle/10803/4885
- Grice, M., Reyelt, M., Benzmüller, R., Mayer, J., & Batliner, A. (1996). Consistency in transcription and

- labelling of German intonation with GToBI. In *Proceedings of the ICSLP*, 1716-1719). Philadelphia.
- Hirst, D. (2011). The analysis by synthesis of speech melody: from data to models. *Journal of Speech Sciences*, 1(1), 55-83. En http://www.journalofspeechsciences.org/index.php/journalofspeechsciences/article/viewArticle/21
- Hirst, D., & Espesser, R. (1993). Automatic Modelling of Fundamental Frequency Using a Quadratic Spline Function. *Travaux de l'Institut de Phonétique d'Aix-En-Provence*, 75-85.
- Jun, S. A., Lee, S. H., Kim, K., & Lee, Y. J. (2000). Labeler agreement in transcribing Korean intonation with K-ToBI. En *Interspeech*, 211-214.
- Jun, S. A., Lee, S. H., Kim, K., & Lee, Y. J. (2010). Labeler agreement in transcribing korean intonation with K-ToBI. En *Interspeech* '10, 211-214.
- Lea, W. (1980). Prosodic aids to speech recognition. In W. Lea (Ed.), *Trends in Speech Recognition* (pp. 166–205). Englewood: Prentice-Hall.
- Lee, J., Kim, B. y Lee, G. (2002). Automatic corpus-based tone and break-index prediction using K-ToBI representation. ACM Transactions on Asian Language Information Processing (TALIP), 1(3), 207-224. doi:10.1145/772755.772757
- Martínez Celdrán, E., & Fernández Planas,
 A. M. (2003). Taxonomía de las
 estructuras entonativas de las
 modalidades declarativa e
 interrogativa del español estándar
 peninsular estándar según el modelo
 AM en habla de laboratorio. En E.
 Herrera y P. Martín (Eds.), La tonía:
 dimensiones fonéticas y fonológicas, 267294. México D. F.: El Colegio de
 México.
- Mertens, P. (2004). The Prosogram: Semi-Automatic Transcription of Prosody based on a Tonal Perception Model. En B. Bel y I. Marlien (Eds.), Proceedings of Speech Prosody 2004, 23-26. Nara (Japan). En http://bach.arts.kuleuven.be/pmert ens/papers/sp2004.pdf
- Noguchi, H. y Kiriyama, K. (1999).

 Automatic labeling of Japanese prosody using J-ToBI style description. En EUROSPEECH'99.

 Sixth European Conference on Speech Communication and Technology, 2259-2262. En http://20.210-193-52.unknown.qala.com.sg/archive/ar

- <u>chive_papers/eurospeech_1999/e99_</u> 2259.pdf
- Pamies, A., Fernández Planas, A. M., Martínez Celdrán, E., Ortega-Escandell, A., & Amorós Cespedes, M. C. (2002). Umbrales tonales en español peninsular. In *Actas del II Congreso de Fonética Experimental*, 272-278.
- Pierrehumbert, J. (1980). *The Phonology and Phonetics of English Intonation*. MIT, Cambridge, Massachusetts.
- Pierrehumbert, J. (1983). Automatic recognition of intonation patterns. En *Proceedings of the 21st annual meeting on Association for Computational Linguistics*, 85-90. Accesible en línea en http://dl.acm.org/citation.cfm?id=9 81328
- Prieto, P. (2002). Entonació: models, teoria, mètodes. Barcelona: Ariel.
- Prieto, P. (2014). The intonational phonology of Catalan. En S.-A. Jun (Ed.), *Prosodic typology*, vol. 2, 43-80. Oxford: Oxford University Press. En http://www.elebilab.com/documentos/archivos/publicaciones/3_GGT-08-04.pdf
- Prieto, P., & Roseano, P. (eds.). (2010). Transcription of Intonation of the Spanish Language. München: Lincom Europa.
- Rietveld, A. C. M. (1984). Syllaben, klemtonen en de automatische detectie van beklemtoonde syllaben in het Nederlands. Université de Nijmegen.
- Roseano, P. y Fernández Planas, A. M. (2013). Transcripció fonètica i fonològica de l'entonació: una proposta d'etiquetatge automàtic. Estudios de Fonética Experimental, XXII, 275-332. En http://www.raco.cat/index.php/EFE/article/view/275413
- Roseano, P., Fernández Planas, A. M., Elvira García, W., Cerdà Massó, R., & Martínez Celdrán, E. (aceptado). La entonación de las preguntas parciales en catalán. *Revista Española de Linguistica Aplicada*.
- Rosenberg, A. (2010). AutoBI a tool for automatic ToBI annotation. En INTERSPEECH 2010, 11th Annual Conference of the International Speech Communication Association, 146-149. Mihama, Japan. En

- http://eniac.cs.qc.cuny.edu/andrew/papers/autobi-is10.pdf
- Ross, K., & Ostendorf, M. (1996). Prediction of abstract prosodic labels for speech synthesis. Computer Speech and Language, 10(3), 155-185. En http://www.sciencedirect.com/science/article/pii/S0885230896900108
- Savino, M., Refice, M., & Daleno, D. (2002).

 Methods and Tools for Prosodic
 Analysis of a Spoken Italian Corpus.
 En Proceedings of the I International
 Conference on Language Resources and
 Evaluation, 307-312. En http://lrecconf.org/proceedings/lrec2002/pdf/101.pdf
- Sridhar, V. (2008). Exploiting acoustic and syntactic features for automatic prosody labeling in a maximum entropy framework. En *IEEE Transactions on Audio, Speech, and Language Processing, 16*(4), 797-811. En http://ieeexplore.ieee.org/xpls/abs-all.jsp?arnumber=4453862
- Syrdal, A. K., & McGory, J. T. (2000). Intertranscriber reliability of ToBI prosodic labeling. En *INTERSPEECH*, 235-238.
- Syrdal, A. K., Hirschberg, J., McGory, J., & Beckman, M. (2001). Automatic ToBI prediction and alignment to speed manual labeling of prosody. *Speech Communication*, 33(1), 135-151. En http://www.sciencedirect.com/science/article/pii/S016763930000073X
- Wagner, A. (2008). Automatic labeling of prosody. En *Proceedings of the 2nd ISCA Workshop on Experimental Linguistics, ExLing 2008, 25-27.* Athens, Greece. En http://isca-speech.org/archive_open/archive_papers/exling2008/exl8_221.pdf
- Wightman, C., & Ostendorf, M. (1994).

 Automatic labeling of prosodic patterns. En IEEE Transactions on Speech and Audio Processing, vol. 2, 469-481.

 En http://ieeexplore.ieee.org/xpls/abs-all.jsp?arnumber=326607
- Yoon, T., Chavarría, S., Cole, J., & Hasegawa-Johnson, M. (2004). Intertranscriber Reliability of Prosodic Labeling on Telephone Conversation Using ToBI. Proceedings of ICSA International Conference on Spoken Language Processing, Interspeech 2004, Jeju, Korea, 2729-2732.

Agradecimientos.

Este trabajo ha sido financiado mediante los fondos concedidos por el Ministerio de Economía y Competitividad para el proyecto AmperCat con referencia FFI2012-35998 y la beca predoctoral APIF-2012 de la Universidad de Barcelona.

Estudio comparativo de la información que aportan los rasgos articulatorios y acústicos de los fonemas al reconocimiento automático de habla

SIMÓN, JAVIER; LLOMBART GIL, JORGE; MIGUEL ARTIAGA, ANTONIO; LLEIDA SOLANO, EDUARDO

1 Introducción

En este experimento, hemos querido comprobar si nuestro modelo acústico de reconocimiento de voz, basado en redes neuronales diseñadas con rasgos de naturaleza articulatoria (Llombart, Miguel & Lleida, 2014), funciona mejor con el uso de rasgos acústicos o con la combinación de ambos.

1.1 Sobre la historia del concepto de rasgo fonológico

En este apartado, realizaremos un breve repaso histórico. El concepto de rasgo en la teoría fonológica se desarrolla plenamente en Jakobson, Fant, & Halle (1967 [1952]) donde se realiza el establecimiento de un conjunto de rasgos distintivos de naturaleza universal. No obstante, ya se venía fraguando en

Roman Jakobson, al menos desde la década de los años 30 del siglo XX¹ y alguna intuición similar puede observarse tambien en Jan Baudoin de Courtenay (Jakobson, 1971: 419), en los *Grundzüge der Phonologie* de Trubetzkoy (Anderson, 1985: 117 y ss.).

Al otro lado del Átlántico, no existió un gran entusiasmo sobre unidades menores al fonema en la lingüística estructuralista norteamericana. Explícitamente, en lo que sabemos, solo Hockett (1955: 132, 134 y 158) habla de rasgo distintivo y lo utiliza en su análisis fonológico. El final del ocaso del estructuralismo norteamericano coincidió con el nacimiento de la lingüística generativa. El ámbito de esta también se extendió al dominio fonológico con la coincidencia de dos estudiosos: Noam Chomsky y Morris Halle. Este último ya había trabajado en fonología junto a Jakobson(Jakobson, 1952). A partir de 1956, Chomsky y Halle, entre otros, generan un agrio debate en contra de las concepciones que sobre gramática, y fonología, postuló el estructuralismo norteamericano (Newmeyer, 1986: 28 y ss.). La colaboración entre ambos autores dio como resultado su obra *The Sound Pattern of English*, cuya primera edición se publicó en 1968.

En la idea primigenia de Jakobson et al. (1976) el listado de rasgos es de naturaleza acústica, aunque señalan que cada uno de estos rasgos tiene correlatos articulatorios y perceptivos. No obstante, advierten del peligro de considerar una relación biunívoca entre la naturaleza de los rasgos. Un mismo acústico puede conseguirse con diversas configuraciones articulatorias, pero una configuración articulatoria determinada solo puede producir un resultado acústico. Lo mismo sucede en la relación acústicoperceptivo. Varias ondas diferentes pueden producir el mismo resultado perceptivo, pero no a la inversa. Su opción por la naturaleza acústica de los rasgos viene determinada por la importancia que estos autores dan al receptor. Tal y como estaba el desarrollo de las ciencias fonéticas, fonológicas y relacionadas, lo más próximo al receptor que podía estudiarse eran las características acústicas del sonido. Dejaban las propiedades perceptuales y auditivas como una tarea urgente para el futuro (Jakobson et al., 1967: 12-13).

Por el contrario, Halle, en su obra posterior junto a Chomsky (Chomsky & Halle, 1991)², defiende una aproximación articulatoria para la definición de los rasgos. Ahora ya no interesa el oyente, como ocurría en el trabajo conjunto con Jakobson y Fant. En este trabajo posterior señalan que "the individual features [...] represent the phonetic capabilities of man" (p. 299). Por tanto, hacen referencia a una teoría articulatoria de los rasgos distintivos y así es como los describen. No obstante, no niegan tampoco la importancia de los correlatos acústico y perceptivo de cada rasgo. Simplemente, consideran que un tratamiento más a fondo de estas relaciones agrandaría demasiado ese capítulo, que ya de por sí lo consideran "a digression from the main theme of our book" (p. 299).

Así pues, tenemos dos tipos de enfoques en cuanto a la naturaleza de los rasgos distintivos: acústico y articulatorio. Hasta ahora nuestros modelos de reconocimiento de voz con redes neuronales habían trabajado solo con rasgos articulatorios (Llombart *et al.*, 2014). Ahora bien, esta aproximación parece una vía de doble camino. A partir de un material con datos acústicos, se extraen mediante algoritmos los datos articulatorios. Después, a partir de estos, se clasifican los fonemas. Parece lógico pensar que si tratamos directamente el

¹ Para ello, puede verse el índice de Jakobson (1962) donde se puede rastrear el uso de este término en su obra.

² Fue publicada por primera vez en 1968.

material acústico, con sus rasgos acústicos, se deberían obtener mejores resultados en la tasa de reconocimiento.

Esta hipótesis es la que hemos comprobado en este experimento. Para ello, hemos utilizado los rasgos acústicos del español postulados por Alarcos Llorach (1965) a partir de los rasgos universales de Jakobson *et al.* (1967), como se puede ver en la Tabla 1.

	o	а	е	u	i	٨	ı	١	r	g	x	k	ŋ	j	s	tʃ	m	b	f	р	n	d	θ	t
Vocal /No vocal	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Consonante/No consonante	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Denso/Difuso	+	+	-	-	-	+	-			+	+	+	+	+	+	+	-	-	-	-	-	-	-	-
Grave/Agudo	+	±	-	+	-					+	+	+	-	-	-	-	+	+	+	+	-	-	-	-
Nasal /Oral										(-)	(-)	(-)	+	-	-	-	+	-	-	-	+	-	-	-
Continuo/Interrupto						+	+	-	-		+	-			+	-			+	-			+	-
Sonoro (flojo)/Sordo (tenso)								+	-	+	(-)	-		+	(-)	-		+	(-)	-		+	(-)	-

Tabla 1. Rasgos acústicos del español. (Según Alarcos LLorach)

A título comparativo, mostramos a continuación la Tabla 2 con los rasgos articulatorios utilizados habitualmente por nosotros. Estos están basados en una estrategia similar a la usada en otros trabajos (Kirchhoff, Fink, & Sagerer, 2002).

	IPA	a	β	ð	е	f	γ	i	'n	k	1	λ	m	n	0	р	r	٢	S	t	θ	tſ	u	Х
Sonoridad	Sonoro	+	+	+	+		+	+	+		+	+	+	+	+		+	+					+	
	Sordo					+				+						+			+	+	+	+		+
Posición	Alta							+															+	
	Media				+																			
	Ваја	+													+									
	Labial		+			+							+			+								
	Dental			+																	+			
	Alveolar										+			+			+	+	+	+				
	Palatal								+			+										+		
	Velar						+			+														+
	Glotal																							
Modo	Vocal	+			+			+							+								+	
	Nasal								+				+	+										
	Fricativa					+													+		+	+		+
	Aprox-Lateral		+	+			+				+	+						+						
	Plosivo									+						+	+			+				
Labialización	Labializada														+								+	
Vocálica	No Labializada	+			+			+																
Localización	Anterior	+			+			+																
Vocálica	Posterior														+								+	

Tabla 2. Rasgos articulatorios del español

1.2 Fundamentos de Redes Neuronales Artificiales

El aprendizaje automático se basa en un conjunto de disciplinas computacionales que ha adoptado un buen grado de uso en los últimos

tiempos. Este tipo de disciplinas consiste, a grandes rasgos, en extraer a partir de un conjunto de datos una representación útil de cierta información presente en dichos datos. Dicha representación nos permite programar sistemas automáticos para ejercer diferentes acciones. La ventaja de estos sistemas es que, una vez que son capaces de obtener cierta representación a partir de datos conocidos, podrán obtener dicha representación a partir de datos que nunca antes han visto, de la misma forma que hace con los que ya ha aprendido. En otras palabras, el sistema aprende por sí mismo un conocimiento sobre unos datos que luego puede aplicar. El ejemplo más claro es el modo en que aprendemos nosotros mismos. En la infancia aprendemos las formas geométricas despreciando los detalles particulares y adquiriendo un conocimiento categórico más abstracto. El objetivo de los sistemas automáticos de reconocimiento es imitar este proceso de aprendizaje donde se desprecien datos irrelevantes y se extraigan las características básicas diferenciadoras (Bishop, 2006; Duda, Hart & Stork, 2012).

Estas técnicas no solo pueden ser usadas para realizar determinadas acciones, sino que nos pueden ayudar a comprender mejor qué tipo de información tienen en común un conjunto de datos; por ejemplo, si esa información es representativa, si es discriminativa para clasificar en diferentes subconjuntos, etc. Siguiendo con el ejemplo anterior, si un sistema con aprendizaje automático es capaz de aprender a clasificar formas geométricas y se equivoca muy poco, confunde una forma de cada mil o diez mil, quiere decir que hay algún tipo de información que ha aprendido que hace que las formas geométricas se puedan distinguir muy bien, por ejemplo el número de lados. De este experimento podríamos deducir que el número de lados es una información muy discriminante para la clasificación de formas geométricas. Y por supuesto este tipo de estudios se pueden llevar a cualquier área, como en el caso que nos ocupa, la clasificación de fonemas a partir de una base de datos de frases grabadas. Esta etapa es crucial en un sistema de aprendizaje automático; en ella determinamos qué características o tipo de información ayudan a resolver mejor el problema. En nuestro caso será elegir qué rasgos, acústicos o articulatorios, son más relevantes en la clasificación de fonemas.

De las diversas técnicas en el ámbito del aprendizaje automático este trabajo se centra en las redes neuronales artificiales. Las redes neuronales artificiales son modelos computacionales basados en el sistema nervioso. Desde los primeros modelos descritos por McCulloch y Pitts en 1943 (McCulloch & Pitts, 1990), pasando por el perceptrón simple desarrollado por Rosenblat (1958), el algoritmo de entrenamiento *backpropagation* (Werbos, 1974) hasta las últimas técnicas de redes profundas (Hinton *et al.*, 2012), las redes han evolucionado mucho, obteniendo unos resultados en la actualidad muy satisfactorios tanto en el ámbito de procesado de imagen como en el de voz. Esto se produce gracias a que cada vez se entiende mejor su comportamiento y se han conseguido estructuras que funcionan muy bien, imitando o aproximando estructuras naturales del sistema nervioso con estos modelos matemáticos, aunque en este ámbito falta mucho camino por recorrer.

El funcionamiento de una neurona, de una forma muy simplificada, se basa en que recibe estímulos o impulsos nerviosos a través de las dendritas. La neurona acumula las respuestas a dichos impulsos nerviosos hasta que, cuando ha recibido suficiente estimulación, genera un impulso nervioso que transmite a otras neuronas a través del axón.

Del mismo modo, funciona el modelo matemático que componen las redes neuronales artificiales. Se van sumando las entradas, x_1 a x_M a la neurona de

forma ponderada, multiplicándolas por unos pesos, $w_1 a w_M$. Posteriormente le sumamos un valor β y el resultado será lo que se va acumulando en la neurona debido a las respuestas de la entrada. Para obtener la activación final, o la respuesta de la neurona, se pasa el resultado de la suma ponderada por una función no lineal, de esta forma si la suma supera cierto valor a la salida de la neurona tendremos la activación, o lo que en las neuronas biológicas sería el impulso nervioso. En la Figura 1 se puede ver una representación del modelo matemático de una neurona artificial y, a la derecha, vemos algunas funciones de activación que se usan. En estas funciones de activación vemos cómo para ciertos valores de entrada (en el eje de ordenadas) apenas se obtienen valores de salida (en el eje de abscisas). Para controlar como las entradas generan la salida se modifican los valores de los pesos w_1 a w_M y del valor β . Es la búsqueda de estos valores para obtener la salida que queremos lo que se llama proceso de entrenamiento o aprendizaje.

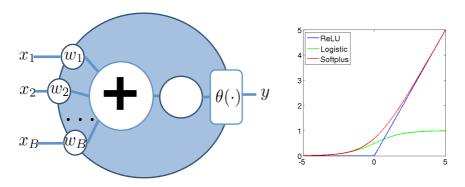


Figura 1. Izquierda: Representación del modelo matemático de una neurona. Derecha: Algunas funciones de activación

$$y = \Theta\left(\sum_{i=1}^{M} w_i \cdot x_i + \beta\right) \tag{1}$$

En la ecuación 1 observamos el modelo matemático de una neurona donde θ se refiere a la función de activación de salida. Varios ejemplos de estas funciones se pueden ver en la Figura 1, donde la más usual es la función Logistic aunque la función Unidad Lineal Rectificada ($Rectified\ Linear\ Unit,\ ReLU$) ha adquirido gran relevancia en los últimos años.

Una neurona aislada es un modelo demasiado simple para obtener una buena representación de la información, por ello se usan estructuras que combinan muchas neuronas. Una de estas estructuras es el perceptrón multicapa. Este consiste en colocar varias neuronas en una capa, donde todas ellas comparten las mismas entradas. Estas capas se concatenan unas a continuación de las otras conectando las salidas de una capa a las entradas de la capa siguiente. Mediante este proceso se consigue aumentar en gran medida la capacidad de representación. En este tipo de redes normalmente se diferencia entre tres tipos de capas. La primera es la capa de entrada, donde las entradas a las neuronas son los datos. Después se pueden encontrar una o varias capas ocultas, cuyas entradas son las salidas de la capa anterior. Se llaman ocultas porque no podemos interpretar su representación de los datos. Por último, está la capa de salida, cuyas salidas suponen la información que buscamos. En la Figura 2 vemos un ejemplo de la estructura típica con tres capas.

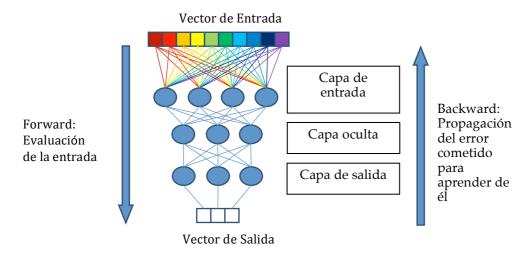


Figura 2. Estructura de perceptrón multicapa

Cuando se usa una red para la evaluación de una entrada, cada neurona produce una respuesta. Este proceso se repite consecutivamente en cada capa hasta alcanzar la capa de salida; esta etapa se denomina *forward*. Una de las aplicaciones típicas de las redes neuronales es usarlas para clasificar las entradas observadas en grupos homogéneos. Estos tienen que ser definidos por el investigador antes de realizar cada uno de los experimentos. En la salida codificamos cada uno de los grupos definidos para ese experimento de una forma binaria, en este caso cada uno de los rasgos acústicos y articulatorios.

Para enseñar a la red se usa el algoritmo *backpropagation* que, resumido de una forma muy sencilla, se basa en enseñar por prueba y error. Para ello partimos de un importante conjunto de datos anotados en función de las definiciones establecidas por el investigador, como se ha mencionado arriba. Después se hace la evaluación de los datos en la red neuronal y cuantificamos si son aciertos o fallos. Los aciertos se refuerzan. En el caso de los fallos se corrigen los parámetros de la red para mejorar el comportamiento de la red en las siguientes evaluaciones. Este proceso se repite hasta que el número de errores detectados sea suficientemente bajo para la aplicación o el sistema deje de mejorar.

2 Descripción del experimento

Para este estudio nos basamos en la base de datos *Albayzin* que está fonéticamente balanceada (Moreno *et al.*, 1993). De todo el corpus se ha seleccionado el corpus fonético que consiste en dos partes: una parte de entrenamiento y otra de evaluación. El entrenamiento está formado por 200 frases. Cuatro locutores pronuncian las 200 frases y 160 locutores pronuncian únicamente 25 frases de las 200, de forma que se mantenga balanceada la cantidad de fonemas. Para la evaluación 40 locutores diferentes a los anteriores pronuncian cada uno 50 frases diferentes a las del entrenamiento. Cada frase se ha etiquetado usando 25 fonemas del castellano.

Los audios de estas frases están grabados en formato wav. Para usar estos audios como entrada a las redes neuronales se calculan los coeficientes Mel Frequency Cepstral Coefficients (MFCC) (Huang, Acero & Hon, 2001), con una ventana de análisis de 25 ms, que se desplaza 10 ms para obtener cada trama. No obstante, para obtener una representación más completa del fonema, a la red neuronal se le presenta la trama etiquetada para clasificar más las cinco

tramas anteriores y las cinco posteriores. Así obtenemos una representación con un contexto total de 110 ms. Hay estudios que demuestran que existe un contexto de 100 ms con información relevante para el reconocimiento de habla alrededor de cada una de estas tramas (Yang, 1999). Se puede visualizar la relevancia de este contexto mediante la estimación de la información mutua (Segura, Benítez, Torre & Rubio, 2001; Yang, van Vuuren & Hermansky, 1999).

En este estudio se van a tratar tres tipos de información diferentes; por ello, la base de datos se encuentra etiquetada tres veces de acuerdo con estos tipos de información. La primera es mediante los fonemas; la segunda es de acuerdo con los rasgos articulatorios; y, por último, la tercera teniendo en cuenta los rasgos acústicos; ambos tipos de rasgos han sido explicados arriba. La transcripción de las frases es manual y el etiquetado fonético se realiza mediante un diccionario canónico. El etiquetado de los rasgos se ha realizado de forma automática; para los rasgos acústicos se ha usado la Tabla 1 y para los rasgos articulatorios la Tabla 2.

En un reconocedor automático del habla basado en redes neuronales, el primer paso consiste en usar la red neuronal como clasificador de fonemas. En este trabajo queremos mejorar esta importante etapa. Para ello proponemos la utilización de algunos rasgos que no se habían utilizado en esta etapa en trabajos previos. Para mejorar el sistema planteamos distintas alternativas que validamos experimentalmente. En los experimentos comparamos la información que aportan estas distintas opciones de uso de rasgos acústicos y articulatorios.

Una vez entrenada la red que clasifica entre los diferentes fonemas, se entrena una red por cada rasgo articulatorio y acústico por separado. Según las clasificaciones explicadas anteriormente serán cinco redes neuronales para los rasgos articulatorios y siete para los rasgos acústicos. Una vez clasificados todos los rasgos se irá combinando la información que aportan con los fonemas de forma que se podrían corregir algunos errores con los datos adicionales. El proceso de combinación de toda la información usa de nuevo una red neuronal que clasifica fonemas, pero cuya entrada no son las tramas de audio, sino la salida de las redes que queremos combinar. Por ejemplo, supongamos que la entrada a esta red de fusión es el fonema /a/, que es sonoro, bajo, vocal, no labializado, etc. De acuerdo con esto la red de fusión aprenderá a relacionar los fonemas con sus descripciones. En el caso de que, debido a un error, el fonema no coincida perfectamente con la descripción, la red podrá decidir qué fonema es el que más se ajusta. Gracias a estos mecanismos se puede reducir la tasa de error.

En la experimentación vamos a ensayar distintas combinaciones de rasgos acústicos y articulatorios junto a información fonética. Si el clasificador que combina información fonética con la de los rasgos se equivoca menos que el clasificador de fonemas inicial, significa que los rasgos aportan información útil al reconocimiento.

3 Resultados experimentales

Para la evaluación de las prestaciones en este estudio, valoraremos la tasa de error por trama, es decir, el número de tramas que el clasificador etiqueta de forma diferente a la etiqueta correcta frente al número total de tramas. Este error lo denominaremos Tasa de Error de Trama (*Frame Error Rate, FER*), donde el 0 % *FER* se considera el resultado ideal.

En la Tabla 3 vemos los resultados de los primeros experimentos realizados. La red neuronal que clasifica fonemas y que usamos como referencia obtiene un 26,28 % de *FER*. Un primer experimento consiste en la fusión de los fonemas con toda la información de los rasgos acústicos y lo mismo para los rasgos articulatorios. En el caso de los rasgos articulatorios se obtiene un 23,72 % de *FER* y en el caso de los acústicos se obtiene un 23,70 %. Esto supone un 9,8 % de mejora relativa frente al caso de referencia, y nos dice que el uso de esta información adicional supone una mejora en el sistema. Sin embargo, no hay una diferencia significativa entre el experimento acústico y articulatorio que nos diga cuál de los dos tipos de rasgos aporta más información.

A continuación se ha fusionado con la información de los fonemas la información articulatoria y la acústica de forma conjunta. En este caso se ha obtenido un 20,34 % de FER que supone un 22,6 % de mejora relativa. Este resultado además nos dice que las informaciones articulatoria y acústica pueden complementarse mutuamente y cada una aporta una información diferente a la otra. Si aportaran el mismo tipo de información el resultado no variaría frente a usar solo una u otra.

Información Usada	Frame Error Rate FER [%]
Fonemas	26,28
Fonemas + Rasgos Articulatorios	23,72
Fonemas + Rasgos Acústicos	23,70
Fonemas + Rasgos Articulatorios + Rasgos Acústicos	20,34

Tabla 3. Experimentos de combinación de rasgos acústicos y articulatorios

En la Tabla 4 vemos cómo mejora la combinación de la información fonética con cada uno de los rasgos individualmente. Vemos que el rasgo que más ayuda a la clasificación es el rasgo posición de tipo articulatorio. Cabe destacar que este resultado es mejor que la fusión con todos los rasgos articulatorios puesto que la combinación de demasiadas características puede añadir incertidumbre (Bishop, 2006). Por ello en los experimentos siguientes vamos a explorar distintas combinaciones de rasgos intentando encontrar la que mejor resultado obtiene.

Fonemas + Rasgo Articulatorio FER [%]												
Sonoridad	oridad Posición Modo Labialización vocálica Localización vocálica											
22,70	21,94	22,15	22,	.01	22,22							
	Fonemas + Rasgo Acústico FER [%]											
Vocal	Vocal Consonante		Grave/	Nasal/	Continuo/	Flojo/						
Vocai Consonante		Difuso	Agudo	Oral	Interrupto	Tenso						
23,98	23,62	23,64	23,28	23,59	23,72	23,91						

Tabla 4. Ratio de error por trama del clasificador de fonemas final, fusionando la salida del clasificador de fonemas incicial con un rasgo

El siguiente experimento consiste en añadir a la mejor fusión anterior otro rasgo para comprobar si sigue mejorando el resultado. En la Tabla 5 vemos que de nuevo en todos los casos se mejora. En este caso es un rasgo acústico el que aporta una mayor mejora lo que puede apoyar la idea de que los rasgos acústicos complementan a los rasgos articulatorios en cierta medida.

Por último se ha añadido al mejor caso anterior otro rasgo. En la Tabla 6 se puede observar que con otro rasgo se puede seguir mejorando. No obstante, ya se encuentran varios casos en los que no se mejora, incluso se empeora. Esto da la intuición de que la combinación o el orden de las fusiones pueden influir en el resultado final, quedando patentes relaciones entre rasgos que no son aparentes en un primer momento. No obstante, realizar todas las

Fonemas + Posición + Rasgo Articulatorio FER [%] Sonoridad Modo Labialización vocálica Localización vocálica 21,68 21,42 21,4721,61 Fonemas + Posición + Rasgo Acústico FER [%] Denso/ Grave/ Nasal/ Continuo/ Flojo/ Vocal Consonante Difuso Agudo Oral Interrupto Tenso 21,42 21,56 21,48 21,06 21,54 21,36 21,39

combinaciones posibles con los doce rasgos estudiados resulta intratable por el elevado número de combinaciones diferentes.

Tabla 5. Ratio de error por trama del clasificador de fonemas final, fusionando la salida del clasificador de fonemas incicial con la posición articulatoria y un rasgo

Fonemas + Posición + Grave/Agudo + Rasgo Articulatorio FER [%]												
Sonoridad	Modo	Labializa	Lo	calización vocálica								
21,28*	21,05	2		21,37*								
	Fonemas + Posición + Grave/Agudo + Rasgo Acústico FER [%]											
Vocal	Vocal Consonante		Nasal/	Cor	ntinuo/	Flojo/						
Vocai		Difuso	Oral	Inte	errupto	Tenso						
21,17*	21,17* 21,07*		20,95	2	1,09*	21,04						

Tabla 6.Ratio de error por trama del clasificador de fonemas final, fusionando la salida del clasificador de fonemas incicial con la posición articulatoria, si es Grave o Agudo y un rasgo. Con asterisco los que no mejoran frente al caso anterior

4 Conclusiones

Existen varios estudios en los cuales se demuestra que el uso de rasgos articulatorios mejora el reconocimento automático de habla (Kirchhoff, Fink & Sagerer, 2002; Leung, Mak & Sun-Yuan Kung, 2004). Junto a estos trabajos, nuestro estudio muestra que los rasgos acústicos también pueden mejoran el reconocimiento automático del habla y ofrecen una complementaria a la articulatoria. Además se puede profundizar en las relaciones entre ambos tipos de rasgos para obtener un mejor resultado, aunque exhaustiva requiere una gran cantidad búsqueda computacionales. Mediante la fusión de todos los rasgos, tanto acústicos como articulatorios, se obtiene el mejor resultadoen este trabajo. No obstante, con el uso de solo tres rasgos se obtiene un resultado muy próximo, lo que obliga a valorar el coste de la evaluación frente a la mejora de las prestaciones.

5 Agradecimientos

Este estudio ha sido financiado por el Gobierno Español y Unión Europea (FEDER) por los proyectos TIN2011-28169-C05-02 e INNPACTO IPT-2011-1696-390000

6 Bibliografía

Alarcos Llorach, E. (1965). *Fonología Española* (4.ª ed.). Madrid: Gredos.

Anderson, S. R. (1985). Phonology in the Twentieth Century: Theories of Rules and Theories of Representations. Chicago: The University of Chicago Press.

Bishop, C. M. (2006). Pattern Recognition and Machine Learning (Information Science and Statistics). Springer. Retrieved from

http://www.amazon.com/Pattern-Recognition-Learning-Information-Statistics/dp/0387310738

Chomsky, N., & Halle, M. (1991). *The sound pattern of English* (Vol. 1st MIT Pr). Cambridge, Mass: MIT Press.

Duda, R. O., Hart, P. E., & Stork, D. G. (2012). Pattern Classification, 680. En

- http://books.google.com/books?hl= es&lr=&id=Br33IRC3PkQC&pgis=1
- Hinton, G., Deng, L., Yu, D., Dahl, G., Mohamed, A., Jaitly, N., & Kingsbury, B. (2012). Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups. *IEEE Signal Processing Magazine*, 29(6), 82-97. doi:10.1109/MSP.2012.2205597
- Hockett, C. F. (1955). *A Manual of Phonology*. Baltimore: Waverly Press.
- Huang, X., Acero, A., & Hon, H.-W. (2001). Spoken Language Processing: A Guide to Theory, Algorithm and System Development. Prentice Hall, 1008. En http://www.amazon.com/Spoken-Language-Processing-Algorithm-Development/dp/0130226165
- Jakobson, R. (1952). Preliminaries to Speech Analysis: The Distinctive Features and Their Correlates. En http://books.google.es/books/about/Preliminaries_to_Speech_Analysis.html?id=9zHVAAAAMAAJ&pgis=1
- Jakobson, R. (1962). *Selected writings*. The Hague: Mouton.
- Jakobson, R. (1971). The Kazan' School of Polish Linguistics and its Place in the International Development of Phonology. En Selected Writings II: Word and Language. The Hague: Mouton, 394-428.
- Kirchhoff, K., Fink, G. A., & Sagerer, G. (2002). Combining acoustic and articulatory feature information for robust speech recognition. *Speech Communication*, 37(3-4), 303.319. doi:10.1016/S0167-6393(01)00020-6
- Leung, K.-Y., Mak, M.-W., & Sun-Yuan Kung. (2004). Applying articulatory features to telephone-based speaker verification. En 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 1, I-85-8. IEEE. doi:10.1109/ICASSP.2004.1325928
- Llombart, J., Miguel, A., & Lleida, E. (2014). Articulatory Feature Extraction from Voice and Their Impact on Hybrid Acoustic Models. *IberSPEECH 2014*. *LNAI 8854*, 138–147.
- McCulloch, W., & Pitts, W. (1990). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biology*, 52(1-2), 99-115. doi:10.1016/S0092-8240(05)80006-0
- Moreno, A., Poch, D., Bonafonte, A., LLeida, E., Llisterri, J., Mariño, J., &

- Nadeu, C. (1993). Albayzín speech database: Design of the phonetic corpus. Proceedings of the 3rd European Conference on Speech Communication and Technology, 1, 175-178.
- Newmeyer, F. J. (1986). *Linguistic Theory in America* (2.ª ed.). San Diego: Academic Press.
- Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65, 386-408. doi:10.1037/h0042519
- Segura, J., Benitez, M., Torre, A. de la, & Rubio, A. (2001). Feature extraction from time-frequency matrices for robust speech recognition. *INTERSPEECH*. En http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.20.3888&rep=rep1&type=pdf
- Werbos, P. (1974). Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences.
- Yang, H. (1999). Relevancy of time-frequency features for phonetic classification measured by mutual information. *Proc. IEEE ICASSP*, 8921(2), 225-228. En http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=758103
- Yang, H., van Vuuren, S., & Hermansky, H. (1999). Relevancy of time-frequency features for phonetic classification measured by mutual information. In 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. ICASSP99 (Cat. No.99CH36258) (Vol. 1, pp. 225-228 vol.1). IEEE. doi:10.1109/ICASSP.1999.758103