



e-AESLA

REVISTA DIGITAL

E-AESLA n.º 6 (2020)

CONSEJO EDITORIAL DE E-AESLA, 6 (2020)

Esther Álvarez de la Fuente
Universidad de Valladolid

Susana Álvarez Álvarez
Universidad de Valladolid

Goretti Faya Ornia
Universidad de Valladolid

CONSEJO DE REDACCIÓN

Álvaro García Santa-Cecilia
Instituto Cervantes

Adquisición y aprendizaje:

Ana Llinares
Universidad Autónoma de Madrid

Irina Argüelles Álvarez
Universidad Politécnica de Madrid

Enseñanza de lenguas y diseño curricular:

María Belén Díez Bedmar
Universidad de Jaén

Isabel Alonso Belmonte
Universidad Autónoma de Madrid

Psicología del lenguaje, lenguaje infantil y psicolingüística:

María del Carmen Horno Chéliz
Universidad de Zaragoza

Sociolingüística:

Elena Fernández de Molina Ortés
Universidad de Granada

Análisis del discurso:

Begoña Bellés Fortuño
Universitat Jaume I

Lingüística de corpus, computacional e ingeniería lingüística:

Moisés Almela Sánchez
Universidad de Murcia

Lexicología y lexicografía:

Iraide Ibarretxe Antuñano
Universidad de Zaragoza

Traducción e interpretación:

Miguel Ángel Candel Mora
Universitat Politècnica de València

REVISORES ADICIONALES

Marta Aguilar, María Cecilia Ainciburu, Marta Albelda Marco, Egisvanda Isys De Almeida Sandes, Irene Alonso-Aparicio, María Esther Barros Díez, Antonio Benítez-Burraco, Ana Bocanegra-Valle, M^a Luisa Carrió, Ricardo Casañ Pitarch, María Luz Celaya, Germán Coloma, Carmen Conti Jiménez, Rocío Cruz, Rosa Currás Móstoles, Jéssica Elejalde Gómez, Arturo Escandón, Natalia Evnitskaya, Diana González Pastor, Almudena Hasan, Javier Herráez Pindado, Isabel Herrando Rodrigo, M. Dolores Jiménez López, Francisco Lorenzo, Bárbara Marqueta, Sofía Martín Laguna, Sergio Maruenda Bataller, Inés Olza, Julián Muñoz Pérez, Mauricio Narváez, Lourdes Pomposo, Diego Rascón, Alicia Ricart Vaya, Ana Rufat Sánchez, Davinia Sánchez García, Óscar Santos Sopena, Gema Sanz, Hanna Skorczynska, María del Carmen Vicente, Edyta Waluch de la Torre. Nuestro agradecimiento a todos ellos y a los revisores anónimos.

AGRADECIMIENTOS:

COMITÉ ORGANIZADOR DE AESLA 2019



Directora:

Raquel Fernández Fuertes
(Depto. de Filología inglesa)

Susana Álvarez Álvarez
(Depto. de Lengua española)

Esther Álvarez de la Fuente
(Depto. de Filología inglesa)

Carmen Cuellar Lázaro
(Depto. de francés y alemán)

Goretti Faya Ornia
(Depto. de Filología inglesa)

Laura Filardo Llamas
(Depto. de Filología inglesa)

Pedro Fuertes Olivera
(Depto. de Filología inglesa)

Tamara Gómez Carrero
(Depto. de Filología inglesa)

Eduardo Gómez Garzarán
(Depto. de Filología inglesa)

Lorena Hurtado Malillos
(Depto. de Lenguas modernas)

Elena Jiménez García
(Depto. de Didáctica de la lengua y su literatura)

Alejandro Martínez González
(Depto. de Filología inglesa)

Nieves Mendizábal de la Cruz
(Depto. de Lengua española)

Susana Merino Mañueco
(Depto. de Lengua española)

Leticia Moreno Pérez
(Depto. de Filología inglesa)

Sonja Mujcinovic
(Depto. de Filología inglesa)

Silvia Sánchez Calderón
(Depto. de Filología inglesa)

Francisco Javier Sanz Trigueros
(Depto. de Didáctica de la lengua y su literatura)

M^a Ángeles Sastre Ruano
(Depto. de Lengua española)

Radoslava Stankova Laykova
(Depto. de Filología inglesa)

Qianting Yuan
(Depto. de Filología inglesa)

COORDINADORES DE LOS PANELES CIENTÍFICOS DE AESLA
(2019)

Ana Llinares
Universidad Autónoma de Madrid

Elena Fernández de Molina Ortés
Universidad de Granada

Irina Argüelles Álvarez
*Universidad Politécnica de Madrid
Campus Sur*

Francisco Alonso Almeida
*Universidad de Las Palmas de Gran
Canaria*

María Belén Díez Bedmar
Universidad de Jaén

Begoña Bellés Fortuño
Universitat Jaume I, Castellón

Isabel Alonso Belmonte
Universidad Autónoma de Madrid

Moisés Almela Sánchez
Universidad de Murcia

Paloma López Zurita
Universidad de Cádiz

Iraide Ibarretxe Antuñano
Universidad de Zaragoza

María del Carmen Horno Chéliz
Universidad de Zaragoza

Miguel Ángel Candel Mora
Universitat Politècnica de València

PROYECTOS FINANCIADOS

La conferencia internacional de AESLA 2019 se organizó como parte de la actividad investigadora del grupo de investigación UVALAL (Universidad de Valladolid Laboratorio de Adquisición del Lenguaje) dentro del marco de dos proyectos de investigación financiados por la Junta de Castilla León y fondos FEDER (ref. VA009P17) y por el Ministerio de Ciencia, Innovación y Universidades y fondos FEDER (ref. PGC2018-097693-B-I00).



ÍNDICE

ADQUISICIÓN Y APRENDIZAJE

Corpus de aprendices de ELE y de nativos para el estudio de géneros conversacionales: las cartas profesionales en el nivel B2 del *MCER* (SOLTRAB-B2)

p. 1

Lourdes Díaz, Rosa María Lucha

Adquisición de las construcciones medias en español por hablantes de chino mandarín

p. 17

Bo Zhang

Cambios en la escritura de los estudiantes de un Grado de Maestro tras su participación en proyectos de innovación docente centrados en la didáctica de los géneros

p. 31

Isabel García Parejo, Juana Blanco Fernández

El proceso lector en español como lengua extranjera: fluidez lectora silenciosa y aptitud musical

p. 46

José Manuel Foncubierta

Asignación del género en español L3: perspectivas sobre la influencia de la adquisición y el uso en la organización del léxico multilingüe

p. 58

Hugues Lacroix, Anahí Alba de la Fuente

Cuestiones metodológicas en *gender studies*: ¿Qué ventajas conllevan métodos mixtos para la investigación del papel del género durante la pubertad en el aprendizaje de idiomas extranjeros?

p. 74

Patricia Uhl

ENSEÑANZA DE LENGUAS Y DISEÑO CURRICULAR

Contextos bilingües y sus implicaciones didácticas: fundamentos teóricos y herramientas básicas para medir la sensibilidad intercultural

p. 86

Óscar Santos Sopena

Español como lengua segunda en contextos de encierro: un taller de escritura académica

p. 104

Florencia Vecchione, Rocío Nahir Navarro, Camila Merlo

PSICOLOGÍA DEL LENGUAJE INFANTIL Y PSICOLINGÜÍSTICA

El lenguaje será ambiguo o no será: el porqué de la ambigüedad léxica y su estudio desde la evolución del lenguaje

p. 117

Natalia López Cortés

Estudio del proceso de adquisición del español en un niño bilingüe árabe-español: el sintagma nominal

p. 129

Khaled Ahmed

SOCIOLINGÜÍSTICA

Creencias y actitudes de estudiantes eslovacos hacia las variedades cultas del español

p. 149

Radka Svetozarovová

Naturaleza, funcionamiento, distribución y frecuencia de los estereotipos hacia España: el caso de las aulas indias

p. 165

Nabiel Ansari, Shweta Sahay

ANÁLISIS DEL DISCURSO

La posición del verbo en euskera: ¿van especialistas y profesorado en la misma dirección?

p. 180

Julian Maia-Larretxea

LINGÜÍSTICA DEL CORPUS, COMPUTACIONAL E INGENIERÍA LINGÜÍSTICA

Medidas para calcular la complejidad relativa a las lenguas naturales

p. 193

M^a Dolores Jiménez López

Codificación y etiquetado en los corpus de aprendices y su aplicación didáctica: la propuesta del corpus de interlengua española de aprendices sinohablantes (CINEAS)

p. 206

María de los Ángeles Calero Fernández, Maribel Serrano Zapata, M^a Begoña Gómez Devis

Diseño y creación de un corpus de aprendices de ELE en Japón (CELEN)

p. 223

Pilar Valverde

¿Es posible la detección automática de la violencia lingüística las redes sociales?

p. 241

Susana Campillo

LEXICOLOGÍA Y LEXICOGRAFÍA

Palabras cognadas en el vocabulario académico del inglés, español y estonio

p. 253

Mari Kruse

La información etimológica en los diccionarios generales de lengua del español y del italiano del siglo XX

p. 265

Maria Iannotti

TRADUCCIÓN E INTERPRETACIÓN

Entorno personal de aprendizaje y portafolios electrónicos en estudios de traducción

p. 278

Marcos Cánovas

Una aproximación a la traducción de la variación lingüística en *Gomorra – la serie*

p. 290

Matias Vedaschi, Laura Cruz García

PRÓLOGO

El presente volumen recoge una selección de las ponencias presentadas durante la celebración de la trigésimo séptima edición del Congreso Internacional de la Asociación Española de Lingüística Aplicada (Universidad de Valladolid, 27-29 de abril de 2019). El objetivo de AESLA 2019 es dar voz a las distintas perspectivas que el estudio del bilingüismo puede abordar a través de investigaciones de carácter teórico o metodológico, mostrando una gran variedad de estudios de investigación que pretenden explicar cómo funcionan determinadas cuestiones lingüísticas cuando dos o más lenguas entran en contacto.

Con el mismo lema que el mantenido durante AESLA 2019, *Bilingüismo y lenguas en contacto*, este volumen refleja a través de las distintas aportaciones el propósito fundamental de estudiar el lenguaje y cómo las lenguas interactúan en una gran variedad de contextos y desde ciertas perspectivas de la lingüística aplicada como son la adquisición y aprendizaje de lenguas, el diseño curricular, la lingüística del corpus y computacional, la ingeniería lingüística, la psicología del lenguaje, el lenguaje infantil, la traducción, así como aspectos sociolingüísticos, lexicográficos y discursivos asociados al estudio del lenguaje.

Nuestro más sincero agradecimiento al comité organizador y voluntarios del congreso que hicieron posible que la celebración de AESLA 2019 fuera un éxito, así como a los directores de los paneles científicos de AESLA, tanto por su gran trabajo y esfuerzo durante el desarrollo del congreso, como por realizar una labor encomiable en el seguimiento de la revisión de los manuscritos recibidos. Gracias también a los revisores externos que tanto han hecho por dar paso a las distintas comunicaciones presentadas en AESLA 2019 y por contribuir a una selección rigurosa de manuscritos para el presente volumen.

Por último, queremos agradecer a Raquel Fernández Fuertes, Sonja Mujcinovic y Eduardo Gómez Garzarán su gran ayuda personal y técnica para que la edición de este volumen haya sido posible.

Esther Álvarez de la Fuente, Susana Álvarez Álvarez y Goretti Faya Ornia

CORPUS DE APRENDICES DE ELE Y DE NATIVOS PARA EL ESTUDIO DE GÉNEROS CONVERSACIONALES: LAS CARTAS PROFESIONALES EN EL NIVEL B2 DEL MCER (SOLTRAB-B2)

Lourdes Díaz

Universidad Pompeu Fabra, Barcelona y CLiC-UB¹

Rosa María Lucha

Escuela Oficial de Idiomas de Barcelona-Drassanes, Barcelona

RESUMEN

En este trabajo presentamos los resultados de una investigación en curso sobre textos escritos de género conversacional (cartas de candidatura a una oferta de puesto de trabajo) producidos por un grupo de 34 aprendices de español como L2 de nivel B2 (MCER) y por 34 estudiantes de control nativos. Como es sabido, los géneros conversacionales son especialmente difíciles porque incorporan rasgos característicos de la interacción oral en la escritura, convenciones propias de los géneros escritos y convenciones de tratamiento profesional (cortesía), sujetas a variación entre lenguas. Presentes como género en el MCER o el PCIC desde los niveles B1 a C2, carecemos de descripciones sistemáticas empíricas de los patrones basadas en corpus tanto nativos como de aprendices de ELE/L2, a diferencia de lo que ocurre en otras lenguas, como el inglés (Swales 2004; Biber, Connor et al. 2007; Biber y Conrad 2009). Mediante la creación y análisis del corpus SOLTRAB-carta, ofrecemos i) un análisis funcional comparativo de textos completos de nativos y no nativos, y ii) mostramos que, si bien el léxico en sí no parece ser el problema de la diferencia de competencia entre nativos y aprendices en B2, sí lo son las convenciones pragmáticas del género, así como la extensión. Estas diferencias (i, ii) permiten distinguir entre producción nativa y no nativa y, por tanto, focalizar eficazmente la instrucción.

Palabras clave: corpus de aprendices, escritura ELE, adquisición de español L2, retórica contrastiva.

¹Esta investigación ha sido parcialmente financiada por el Ministerio de Economía y Competitividad, mediante el proyecto SOMEMBED-SLANG (TIN2015-71147-C2-2) de la Universidad de Barcelona.

ABSTRACT

In this paper we present the results of a current research on the writing of conversational genres, namely email and/or letter of application for a job. We compare the production by 68 subjects (34 native and 34 non-native B2 learners of L2 Spanish) from Soltrab corpus, and we provide: i) a functional analysis of the text structure of both groups (Swales 2004; Biber, Connor et al 2007); ii) evidence statistically supporting the claim that pragmatics and sociocultural differences are more relevant than lexis and fluidity in order to distinguish native from non-native production. Pedagogical implications teaching this genre are also driven.

Keywords: learner corpus, Spanish L2 writing, Spanish SLA, contrastive rhetoric

1. INTRODUCCIÓN: ANTECEDENTES Y OBJETIVOS

El interés por la lengua hablada y su abordaje desde los enfoques comunicativos, en la enseñanza y evaluación, ha ido de la mano de la distinción entre oralidad y escrituralidad, sus diferencias y sus ámbitos de interacción en la definición de géneros. En las lenguas segundas y extranjeras, además, se ha puesto el foco en los aspectos universales que poseen los escritos dialógicos o conversacionales (comunes a muchas culturas que ritualizan estas prácticas discursivas) y su plasmación en los géneros más habituales. De ahí el predominio de la carta o del correo electrónico (formal o informal) como formas dialogales en los materiales didácticos (manuales) y pruebas (exámenes) en la enseñanza de lenguas segundas o extranjeras (aunque también primeras) desde la década de los 90 del pasado siglo. La revisión más funcional de las clasificaciones textuales tradicionales (textos narrativos, descriptivos, argumentativos, expositivos, directivos), por lo general excesivamente atentas a las formas escritas (Werlich 1975; van Dijk 1983, 1984) hasta la década de los 90 (Adam 1991, 1992; Koch y Oesterreicher 2007, etc.), ha dado entrada a los géneros dialógicos y ha puesto en el foco las formas constituidas por secuencias combinadas. Asimismo, desde entonces, es el criterio funcional el que ha primado en la determinación de géneros (Martin 1985; Martin y Rose 2008; Paltridge 2006; Swales 2004, etc., para el inglés, en el ámbito funcionalista); como también la cooperación (dialogicidad), la referencialidad con respecto al momento de habla o situación, o aspectos como la simultaneidad o no con la producción, la inmediatez en la recepción, o la subjetividad y empatía (Koch y Oesterreicher 2007; Loureda 2003). En el panorama actual, los parámetros que se usan para identificar y establecer los géneros conversacionales se ajustan a una propuesta transversal donde el continuo oral/escrito se define por la presencia más o menos sustantiva de esos parámetros antes citados, que detallamos en la Figura 1. Particularmente, la toma de decisiones *top-down* o descendente que realiza el emisor acerca de cómo combinar e incluir estos aspectos mencionados resulta difícil de dominar en la L1 y es la verdadera piedra de toque de la competencia pragmático-discursiva en la L2, como muestran los estudios de corpus (ver Biber, Davis, Jones y Tracy-Ventura 2006, Lucha y Díaz 2017, Díaz y Lucha 2018, para el español). Estos, aunque todavía escasos en español, especialmente los que comparan corpus nativos y de aprendices para detectar parámetros, visibilizan los patrones en la L1 y ponen, también, de manifiesto la necesidad del foco didáctico en los géneros conversacionales como tales para la L2².

²Ver Adam (1991) y Lucha y Díaz (2017), Díaz y Lucha (2018) para una explicación más detallada sobre los parámetros y su incidencia en cartas y correos profesionales.

a- Grado de <i>publicidad</i> : esto es, el carácter más o menos público de la comunicación, para el que son relevantes el número de interlocutores y la existencia de público y sus dimensiones.
b- Grado de <i>familiaridad</i> entre los interlocutores: depende de la experiencia comunicativa previa conjunta, del grado de institucionalización de la comunicación, etc.
c- Grado de <i>implicación emocional</i> , que puede estar regida por el interlocutor (afectividad) y/o por el objeto de la comunicación (expresividad).
d- Grado de <i>anclaje</i> de los actos comunicativos en la situación o en la acción.
e- <i>Campo</i> referencial, para el que es decisiva la distancia de objetos y personas respecto al aquí y ahora del hablante.
f- <i>Inmediatez</i> física de los interlocutores (comunicación cara a cara frente a la distancia espacial y temporal).
g- Grado de <i>cooperación</i> , medido según las posibilidades de intervención de los receptores en la producción del discurso.
h- Grado de <i>dialogicidad</i> o de posibilidad de intercambiar papeles de emisor y receptor.
i- Grado de <i>espontaneidad</i> .
j- Grado de <i>fijación</i> temática.

Figura 1. Parámetros de caracterización de realizaciones lingüísticas en el continuo hablado/escrito según la condición comunicativa (Koch y Oesterreicher 2007:26)

Precisamente el carácter escaso de estos estudios basados en corpus para establecer a partir de los parámetros anteriores cuáles son los rasgos prototípicos para la carta *profesional*^B frente a la *privada*^A, y las dificultades que entraña su producción adecuada y sin deslizamiento entre ambos ámbitos y recursos (incluidos los de adecuación al registro y cortesía, como los más llamativos), nos han llevado a la recopilación del corpus *Soltrab*-carta/correo, de producciones nativas y no nativas de correos y cartas de solicitud de trabajo a partir de un anuncio. La función de *Soltrab* será permitirnos extraer el patrón característico del género a partir de parámetros como los indicados en la figura 1, establecer

³Distancia en la implicación emocional, escasa familiaridad entre interlocutores, distancia física, escasa cooperación en la producción y dialogicidad estrictamente regulada.

⁴Familiaridad, implicación emocional, espontaneidad, desarrollo temático libre, no muy planificado, con signos de cooperación o dialogicidad y anclaje en situaciones o emisiones previas.

y caracterizar sus realizaciones lingüísticas, esto es, describir las formas más características de cada secuencia funcional (*move*) y paso (*step*) en este género conversacional (Swales 2004) en español (peninsular). Siempre que sea posible, se podrá comparar la divergencia entre las opciones de la producción nativa y la no nativa con vistas a la didáctica de la lengua basada en el trabajo de géneros textuales en el aula.

2. PREGUNTAS DE INVESTIGACIÓN Y METODOLOGÍA

Las preguntas de investigación que nos formulamos son las siguientes:

- i) ¿Se pueden localizar los aspectos convencionalizados del género, en contexto, de forma sistematizada y significativa, desde el punto de vista pragmático-cultural, para llevar al aula? Dicho de otro modo, ¿hay un patrón común, sistematizable para el español (Swales 2004) que justifique la inclusión recurrente de este género en pruebas y modelos (p.e. PCIC) sobre competencia textual?
- ii) ¿Permitirá una muestra de 34 sujetos de cada tipo validar estadísticamente los descriptores previos (Díaz y Lucha 2018) hallados en la comparación entre hablantes nativos y no nativos con un grupo menor (N=17/17)?
- iii) ¿Serán cuantificables o estratificables esos aspectos (funcionales, sociopragmáticos) para definir o identificar la producción de un nivel o la de nativos y no nativos? (p.e. mayor adecuación en B2 que en B1, mayor fluidez en nativos que en B2 o C1) y ¿cómo?).

2.1. Metodología: contexto

Con el fin de investigar estas preguntas, decidimos llevar a cabo un estudio de la producción escrita de aprendices no nativos de nivel B2 (según el MCER) por dos razones. (i) La primera es porque se trata del nivel con el que se ejemplifica la tarea en el PCIC y en la mayoría de los exámenes consultados. (ii) La segunda, porque por las razones expuestas, encajaba fácilmente en la programación de diez semanas de un curso de B2.1 destinado a estudiantes de intercambio del programa Erasmus de la UPF. Diseñamos una recogida de producción escrita de los estudiantes no nativos (en adelante, NN) como actividad de aula. Paralelamente, dispusimos de un grupo de estudiantes

nativos que realizaron voluntariamente la tarea como grupo de control nativo (en adelante, N). Ambos grupos formaban parte del programa de Lenguas Aplicadas de la Universidad Pompeu Fabra de Barcelona.

2.2. Sujetos

Los 34 sujetos no nativos (NN) de este estudio participaban en un curso intensivo trimestral de español L2 de carácter general en el marco universitario. Todos habían realizado una prueba de nivel *online* estandarizada de la institución para la asignación de grupo, eran aprendices multilingües procedentes de facultades de ciencias humanas y sociales (50% hablantes nativos de lenguas anglogermánicas y 50% románicas, que tenían el inglés o el francés como L2 o L3, complementariamente) y eran de edad y formación similar.

Los 34 sujetos nativos pertenecían a la misma facultad y colaboraron de forma voluntaria en la tarea. Eran hablantes bilingües catalán-castellano y dominaban una lengua extranjera, predominantemente el inglés.

2.3. Tarea y técnica de elicitación de la tarea

Dado que en la programación del grupo de ELE aparecía el trabajo del área temática de la formación y salidas laborales, por un lado, y el debate oral y la escritura de currículos y cartas de solicitud de empleo, por otro, se programó con la profesora del grupo una secuencia de trabajo de *input* y de escenarios comunicativos que condujeran a una simulación en la que encajar la escritura de *una carta de candidatura para un puesto de trabajo* diseñada en el aula con los perfiles de los solicitantes y retomando el trabajo desarrollado previamente en la simulación. La escritura se realizó como parte de una tarea, en clase, de la que surge el corpus SOLTRAB-carta-NN (n=34). Esta parte del corpus, más la producción que, como tarea, realizaron los nativos (n=34), constituyen los datos que se analizan en este estudio.

2.4. Etiquetas: codificación funcional

Para la descripción y análisis de Soltrab hemos tomado una lista de etiquetas semántico-funcionales que proceden de la interpretación del contexto de

aparición, tanto a partir de la información léxica como funcional (ver Díaz y Lucha 2018). Dado que las cartas de solicitud de trabajo constituyen un género convencionalizado, donde la referencia explícita a la interacción tiene sus propias marcas rituales (fórmulas de saludo, despedida, tratamiento formal) y formato (fecha, firma, remitente, destinatario), se han previsto etiquetas específicas para todo ello, como muestra la figura 2. Las veintiuna etiquetas, introducidas en el texto, permiten recuentos y extracción de concordancias, etc. con la herramienta Antconc, de libre distribución. Para ejemplos de las etiquetas y una muestra de producción nativa y no nativa, ver Díaz y Lucha 2018.

1. E encabezamiento
2. DE dirección empresa
3. FE fecha
4. A asunto
5. FASP fórmulas de apertura saludo personal
6. PPA presentación personal apertura
7. ME motivación escrito
8. UBA ubicación anuncio
9. EI expresión de interés
10. VP valoración personal
11. ECV experiencia CV
12. EID experiencia idiomas
13. AP adecuación al perfil
14. UECV ubicación experiencia CV
15. ED expresar disponibilidad
16. CVA CV adjunto (o portfolio)
17. EE elogio a la empresa
18. PE petición entrevista (o respuesta)
19. FFPC fórmula final pre-cierre
20. FFC fórmula final Cierre
21. F firma

Figura 2. Etiquetas usadas en el análisis de Soltrab-carta

3. ANÁLISIS Y DISCUSIÓN DE DATOS

3.1 *Fluidez y diversidad léxica*

A continuación presentamos los resultados cuantitativos del grupo de sujetos nativos y no nativos que participaron en la investigación.

	Types	Tokens	Índice Uber
Nativos	5091	8177	74,39
No Nativos	5940	10152	68,96

Figura 3. Resultados de Types, Tokens e índice Uber U de los grupos Nativos y No Nativos

Por números absolutos, observamos (Figura 3) que el grupo de no nativos produce mayor número de palabras totales (*Tokens*). Esta diferencia sugiere, en un primer momento, que el grupo de sujetos no nativos posee una mayor fluidez léxica que el grupo nativo. Para comprobar si esa diferencia es significativa estadísticamente, recurrimos al análisis estadístico mediante *G-Stat 2.0*.

El primer paso fue verificar el ajuste a la bondad de las dos muestras mediante el test de Shapiro-Wilk para decidir qué test estadístico debíamos aplicar. Como vemos en la figura 4, el resultado del p-valor, inferior al índice de significación de 0.05, nos hizo rechazar la H_0 de ajuste a la bondad. Por tanto, el análisis estadístico se realizó mediante el test no paramétrico de Mann-Whitney (Figura 5) para la comparación de dos poblaciones, cuyo resultado muestra un p-valor de $0.0149 < 0.05$. Este resultado constata que la diferencia observada entre el grupo NN y N en la producción total de *tokens* es estadísticamente significativa. Por tanto, podemos afirmar que los sujetos no nativos son capaces de escribir textos más extensos que los sujetos nativos en igualdad de condiciones.

	p-valor
Nativos	$0.0215 < 0.05$
No Nativos	$0.0485 < 0.05$

Figura 4. Resultados del test Shapiro-Wilk para la comprobación de ajuste a la bondad

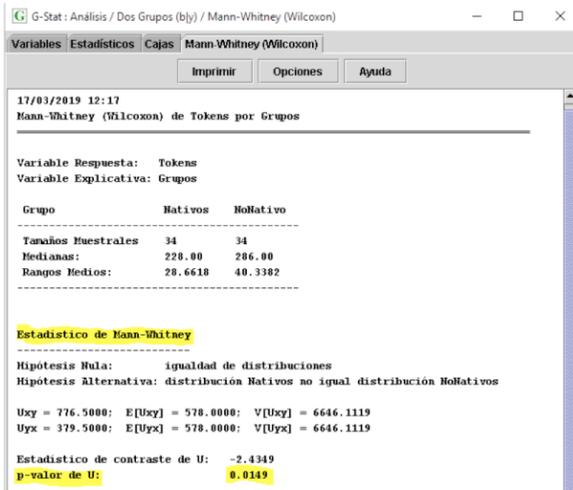


Figura 5. Resultado del test de Mann-Whitney para el análisis de la fluidez léxica

La figura 6 muestra gráficamente la distribución en ambos grupos.

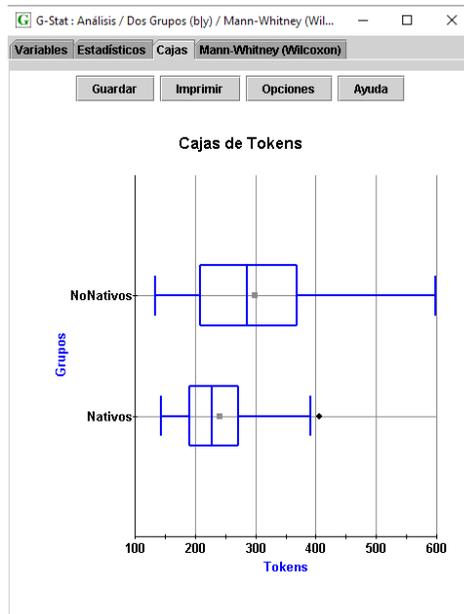


Figura 6. Diagrama de distribución de *tokens* por grupos

Van Esch *et al.* (2006) ya observaron esto en su investigación sobre expresión escrita en cuanto a fluidez. Los autores analizaron la evolución de los estudiantes no nativos a lo largo de un año académico en el que comprobaron que había un incremento considerable en el número de palabras en los textos del segundo año respecto del primero: “a dramatic increase in the number of words produced by the non-native writers.” (*op. cit.*: 113). Asimismo, al comparar los resultados del grupo de no nativos con los nativos, constataron que la mejora observada en los datos del segundo año de los no nativos estaba próxima o incluso superaba los resultados del grupo de sujetos nativos como, por ejemplo, en el número total de palabras. Como vemos, este resultado corrobora los resultados de nuestra investigación. Asimismo, en una investigación previa (Díaz y Lucha 2018), con una muestra inferior a la actual, ya pudimos comprobar que los sujetos no nativos mostraban una mayor fluidez léxica que los nativos con una diferencia significativa estadísticamente.

3.2 Análisis de la diversidad léxica

En nuestra investigación, las producciones escritas de nuestro corpus presentan, por un lado, una extensión mediana, pues no superan las 450 palabras, y, por otro lado, hay una gran diferencia de tamaño entre ellas, pues van desde las 130 hasta las 450 palabras. Por tanto, optamos por utilizar la medida de diversidad léxica *Uber U* (Dugast 1978, 1979, 1980; Vermeer 2000; Jarvis 2002) aconsejada en estos casos.

La fórmula para calcular el índice *Uber U* es:

$$U = \frac{(\log \text{tokens})^2}{\log \text{tokens} - \log \text{types}}$$

Tras comprobar el ajuste a la bondad de las dos muestras fue necesario utilizar un test paramétrico, en este caso el t-student. El resultado de este test (Figura 7) no permitió rechazar la H_0 de igualdad al mostrar un p-valor de $0.5108 > 0.05$, superior al valor de significación.

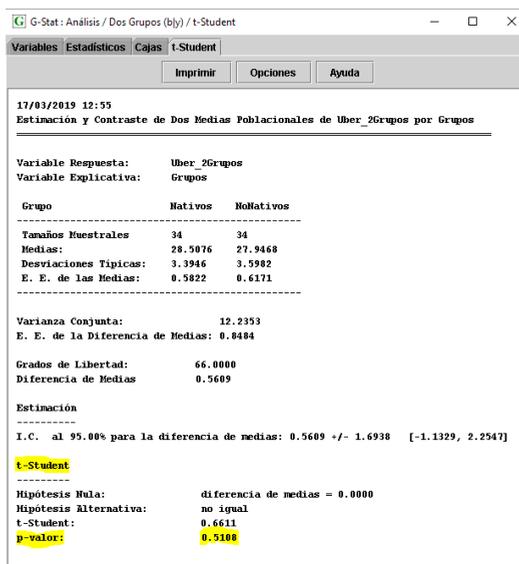


Figura 7. Resultado del test de t-Student para el análisis de la diversidad léxica

Este resultado corrobora el obtenido en nuestra investigación previa (Díaz y Lucha 2018). Asimismo, en esta misma línea, encontramos trabajos que también han comparado producciones escritas de nativos y no nativos, aunque hayan analizado otros aspectos de la competencia escrita. Entre ellos, resaltamos los de Teijeira *et al.* (2005), de Haan y van Esch (2005), Díaz y Lucha (2010, 2012). Todas esas investigaciones observaron cierta similitud entre ambos grupos, mostrando resultados que tampoco presentaban diferencia estadísticamente significativa. Teijeira *et al.* (*op. cit.*) resalta el hecho de que los sujetos no nativos, aun estando en un nivel no muy avanzado de sus estudios, son capaces de escribir textos coherentes. En este punto, podemos relacionar los resultados de nuestra investigación con los de Teijeira, pero con sujetos de nivel B2 según el MCER cuyos resultados son similares e incluso superiores a los sujetos nativos en cuanto a fluidez léxica.

3.3 Análisis de la presencia y distribución de etiquetas

El resultado de la comparación de las etiquetas que reflejan la estructura retórico-funcional en ambos corpus se recoge en las figuras 8 y 9 a continuación. En la figura 8 se han marcado las secuencias funcionales cuya coincidencia es

del 100% y en la figura 9 las coincidencias parciales, que reflejan diferencias pragmático- socioculturales que se perciben más allá del estilo. Esto ocurre, en especial, con el mayor predominio nativo de EID (experiencia en idiomas), ED (expresar disponibilidad). Por el contrario, es no nativo el predominio de EE (elogio a la empresa) o FFC (fórmula final cierre) además de la FFPC (pre-cierre), así como un mayor cuidado del anclaje de la información: A (indicación de asunto), UBA (referencia al anuncio).

FrecEtiquetasNNS_Antconc.txt: Blo				NativosEtiquetas_sub34-antconc.t			
Archivo	Edición	Formato	Ver	Archivo	Edición	Formato	Ver
#Word Types: 21				#Word Types: 21			
#Word Tokens: 955				#Word Tokens: 919			
#Search Hits: 0				#Search Hits: 0			
1	162	VP		1	179	VP	
2	148	ECV		2	123	EID	
3	95	AP		3	72	ECV	
4	80	EI		4	71	AP	
5	55	UECV		5	45	UECV	
6	40	E		6	40	F	
7	39	EID		7	39	EI	
8	35	F		8	38	ED	
9	35	FFPC		9	36	E	
10	34	FFC		10	36	FFPC	
11	33	FASP		11	34	FASP	
12	33	ME		12	34	ME	
13	31	UBA		13	32	PE	
14	25	A		14	31	FFC	
15	20	EE		15	27	CVA	
16	20	PE		16	22	PPA	
17	18	PPA		17	19	UBA	
18	17	CVA		18	18	A	
19	16	FE		19	12	FE	
20	13	DE		20	8	DE	
21	6	ED		21	3	EE	

Figura 8. Coincidencias al 100% en secuencias funcionales (*moves* obligatorios)

FrecEtiquetas:NNS_Antconc.txt: Blo				NativosEtiquetas_sub34-antconc.t			
Archivo	Edición	Formato	Ver	Archivo	Edición	Formato	Ver
#word Types: 21				#word Types: 21			
#word Tokens: 955				#word Tokens: 919			
#Search Hits: 0				#Search Hits: 0			
1	162	VP		1	179	VP	
2	148	ECV		2	123	EID	
3	95	AP		3	72	ECV	
4	80	EI		4	71	AP	
5	55	UECV		5	45	UECV	
6	40	E		6	40	F	
7	39	EID		7	39	EI	
8	35	F		8	38	ED	
9	35	FFPC		9	36	E	
10	34	FFC		10	36	FFPC	
11	33	FASP		11	34	FASP	
12	33	ME		12	34	ME	
13	31	UBA		13	32	PE	
14	25	AO		14	31	FFC	
15	20	EE		15	27	CVA	
16	20	PE		16	22	PPA	
17	18	PPA		17	19	UBA	
18	17	CVA		18	18	A	
19	16	FE		19	12	FE	
20	13	DE		20	8	DE	
21	6	ED		21	3	EE	

Figura 9. Coincidencias parciales que muestran distancia pragmático-cultural

4. CONCLUSIONES

Siguiendo a Biber y Conrad (2009:16-18), para establecer y caracterizar cómo son intralingüística e interlingüísticamente los géneros, los registros y los estilos —en nuestro caso para caracterizar las cartas de solicitud de trabajo en español nativo y no nativo—, no solo hay que comparar extractos de producción sino textos completos de nativos y no nativos en que basar comparaciones empíricas y responder preguntas de investigación como las nuestras. Esto es lo que hemos llevado a cabo mediante el análisis del corpus *Soltrab*, con el propósito de extraer implicaciones didácticas.

En cuanto a los aspectos más cuantitativos del léxico de *Soltrab*, hemos comprobado que, en relación con la fluidez, los sujetos del grupo no nativo escriben textos más extensos que los nativos, confirmando estadísticamente lo hallado en Díaz y Lucha (2018). Sin embargo, no hay diferencia significativa en cuanto a diversidad léxica entre ambos grupos. Por consiguiente, la implicación didáctica para un curso de ELE de carácter general es que lo que resultará

realmente relevante es el trabajo y evaluación del componente pragmático-funcional de los géneros y, sobre todo, el sociopragmático e intercultural (imagen, modalidad) caracterizadores de este género dialógico, dado que son las diferencias en porcentaje de etiquetas como VP, EE, EI y, sobre todo, ED las que han resultado relevantes para distinguir a los nativos de los aprendices de L2 en cuanto a adecuación a la situación y satisfacción de las funciones esperadas.

Respecto a las implicaciones para un curso de Español para fines específicos, de nuevo es más relevante el trabajo sociopragmático e intercultural que el de los aspectos considerados específicos del género carta profesional, puesto que hemos visto que etiquetas características del género como ECV, UECV, CVA, AP, FFC, FASP o ME, no permiten distinguir casi entre L1 y L2 en B2 y dejan escaso margen de aportación relevante al docente.

En ambos contextos de aprendizaje, por tanto, son los aspectos socioculturales y muy convencionales del género (retórica, colocaciones, fórmulas) los que merecen atención especial en el aula y son mejorables en B2, como han puesto de manifiesto etiquetas como FE, DE, FFPC y FFC.

Asimismo, el presente estudio ha permitido corroborar otro aspecto sociopragmático relevante para el género dialógico carta y correo de solicitud, como es la elección y peso de secuencias y funciones, esbozado en Díaz y Lucha (2018). Siguiendo a Koch y Oesterreicher (2007), propusimos un patrón que indicaba que si la elección de funciones se centraba en el receptor (patrón más propio de los nativos), daba pie a un texto interpretado como más cortés, más adecuado y más nativo; mientras que la selección de funciones más centradas en el emisor (presentación personal, valoración de su experiencia, adecuación al perfil, etc.), hacía el texto más expeditivo, más centrado en la autoimagen y menos nativo (incluso en B2).

REFERENCIAS BIBLIOGRÁFICAS

- ADAM, J. M. 1991. "Une typologie d'inspiration bajtinienne", ELA N° 83, pp. 7-17.
- ADAM, J. M. 1992. Les textes: types et prototypes. Paris: Nathan.
- ANTHONY, M. L. *Antconc* [<http://www.laurenceanthony.net/software/antconc/>].
- BIBER, D., CONNOR, U. y T. A. UPTON. (2007). Discourse on the move: using corpus analysis to describe discourse structure. Amsterdam: John Benjamins.
- BIBER, D. y CONRAD, S. (2009). *Register, Genre and Style*. Cambridge: Cambridge University Press.
- BIBER, D., M. DAVIES, J.K. JONES, N. TRACY y VENTURA. 2006. Spoken and written register variation in Spanish: a Multi-dimensional analysis. *Corpora* 1: 7-38.
- CONSEJO DE EUROPA. 2000. *Common European Reference Framework*. Estrasburgo: Consejo de Europa [versión española de 2002].
- DE HAAN, P. y K. VAN ESCH. 2005. "The development of writing in English and Spanish as foreign languages", *Assessing Writing*, 10: 100-116.
- DÍAZ, L. y LUCHA, R.M. 2010. "Teaching genres in an L2 writing course. The incidence of focus on form in written production or how to go a step further than natives", *Monografías MarcoEle*, 11: 20-36.
- DÍAZ, L. y LUCHA, R.M. 2012. "Escribo con acento en español (l2) ¿eso importa? ¿puedo obtener feedback automáticamente?", *Monografías MarcoEle*, (15), pp. 68-81.
- DÍAZ, L. y LUCHA, R.M. 2018. "Las cartas de solicitud de trabajo en ELE y la tradición discursiva en L1 y L2. A propósito de la didáctica de los textos conversacionales y sus géneros en el aula de ELE", *MarcoELE Revista de didáctica del español*, 26.
- DUGAST, D. 1978. "Sur quoi se fonde la notion d'étendue théorique du vocabulaire?", *Le français modern*, 1, pp. 25-32.
- DUGAST, D.1979. *Vocabulaire et stylistique*. 1. Théâtre et dialogue. Genève: Slatkine.

- DUGAST, D. 1980. *La statistique lexicale*. Geneva, Switzerland: Slatkine.
- JARVIS, S., GRANT, L., BIKOWSKI, D. y FERRIS, D. 2003. "Exploring multiple profiles of highly rated learner compositions", *Journal of Second Language Writing*, 12: 377-403.
- KOCH, P. y W. OESTERREICHER, 2007. *Lengua hablada en la Romania: español, francés, italiano*. Madrid: Gredos.
- LOUREDA, Ó., 2003. *Introducción a la tipología textual*. Madrid: Arco Libros, SL.
- LUCHA, R.M. y L. DÍAZ, 2017: "Tipos textuales expositivos, correos electrónicos y enseñanza de géneros textuales en ELE", *MarcoELE*, 25.
- MARTIN, J.R., 1985. *Factual writing: exploring and challenging social reality*. Geelong: Deakin University Press.
- MARTIN, J.R. y D. Rose, 2008. *Genre relations. Mapping Culture*. London: Equinox.
- PALTRIDGE, B., 2006. *Making sense of Discourse Analysis*. Brisbane. Merino Litographs.
- SWALES, J.M. 2004. *Research genres*. Cambridge: Cambridge University Press.
- TEJEIRA, M^a C.; ESCH, C. J. van; HAAN, P.J. M. de. 2005. "La coherencia y la cohesión en textos escritos por estudiantes de español como lengua extranjera". *Estudios de Lingüística Aplicada*, vol. 41, pp. 67-100.
- VAN DIJK, T., 1983. *La Ciencia del texto*. Barcelona: Paidós.
- VAN DIJK, T., 1984. *Texto y Contexto*. Madrid: Cátedra.
- VAN ESCH, K., DE HAAN, P., FRISSEN, L., GONZÁLEZ SANTERO, I. y DE LA TORRE MIRANDA, A. 2006. "Evolución en la competencia escrita de estudiantes de español como lengua extranjera", *Estudios de Lingüística Aplicada*, 43: 55-76.
- VERMEER, A. 2000. "Coming to grips with lexical richness in spontaneous speech data", *Language Testing*, 17, pp. 65-83.
- WERLICH, E., 1975. *Typologie der Texte*. Heidelberg: Quelle und Meyer.

ADQUISICIÓN DE LAS CONSTRUCCIONES MEDIAS EN ESPAÑOL POR HABLANTES DE CHINO MANDARÍN

Bo Zhang
Universidad de Huelva

RESUMEN

Este artículo estudia la adquisición de construcciones medias en español como segunda lengua concentrándose en el uso del pronombre se y este combinado con el pronombre de objeto indirecto como estrategia para expresar el rechazo de responsabilidad, así como la afectación involuntaria. La investigación se lleva a cabo con ocho hablantes de español como segunda lengua y cuatro de ellos con título de máster en la facultad de Humanidades de la Universidad de Huelva mientras que el resto es del programa Erasmus para el alumnado de grado. Para efectuar el estudio, primero se les requiere a los participantes que respondan a la pregunta ¿Qué le pasó a...? de acuerdo con una serie de ilustraciones, con el fin de que determinen el uso de la estructura media. En segundo lugar, se analizan las respuestas obtenidas. Se comprueba que la construcción media que combina el uso del pronombre del objeto indirecto que representa el agente es de uso general en los hablantes nativos y de uso limitado en los hablantes de chino que tienen el español como L2 o LE. Por último, partiendo del análisis del resultado, unas reflexiones se realizan con objetivo de mejorar el diseño de los recursos didácticos de español en China.

Palabras clave: enseñanza-aprendizaje, adquisición, español como segunda lengua, chino, construcción media.

ABSTRACT

This article studies the acquisition of middle constructions in Spanish as a second language, concentrating on the use of the pronoun se and this combined with the indirect object pronoun as a strategy to express the rejection of responsibility as well as the involuntary affectation. The research is carried out with eight speakers of Spanish as a second language and four of them with a master's degree in the Faculty of Humanities at the University of Huelva, while the rest is from the Erasmus program for undergraduate students. To carry

out the study, participants are first required to answer the question, What happened to ...? according to a series of illustrations, in order to determine the use of the middle structure. Second, the answers obtained are analyzed. It is verified that the middle construction that combines the use of the pronoun of the indirect object that represents the agent is of general use in the native speakers and of limited use in the Chinese speakers who have Spanish as L2 or LE. Finally, based on the analysis of the result, some reflections are made with the aim of contributing to the refinement of teaching resources of Spanish in China.

Keywords: teaching-learning, acquisition, Spanish as a second language, Chinese, middle construction.

1. INTRODUCCIÓN

Según Schnell (2011) la presentación de la construcción media con *se* en el aula de español como segunda lengua, así como en los libros didácticos muestra carencias importantes y no detalla su uso en situaciones reales. En concreto, el autor investiga el tratamiento didáctico que se le da al pronombre *se* en cinco libros de enseñanza del español como lengua extranjera y concluye que todos ellos enfatizan el uso de *se* en oraciones reflexivas, enmarcadas en el tema de actividades diarias, pero solo tres hacen referencia también a oraciones impersonales y únicamente dos abordan el *se* para referirse a construcciones medias del tipo “La puerta se abrió” y “se me ha caído al suelo”, las cuales se describen como oraciones que expresan acciones accidentales o involuntarias. Teniendo en consideración dicha debilidad mejorable Guillén (2015) por medio de una investigación cuantitativa, indica que, desde el punto de vista de la adquisición, la construcción media que combina el uso del pronombre *se* con el pronombre de objeto indirecto que representa el agente es de uso general en los hablantes nativos y de uso muy limitado en los hablantes que tienen el español como L2 o LE. A partir de estas investigaciones previamente realizadas respecto a las construcciones medias en español, hemos revisado el material didáctico más importante de la enseñanza de español *Español Moderno I-IV* (Dong y Liu, 2014) para los hablantes chinos que tienen el español como L2 o LE y observamos que en general, todos ellos enfatizan el uso de *se* en oraciones pasivas reflexivas así como en oraciones impersonales tales como “se dice que” o “se habla chino” y por desgracia, no hemos descubierto una presentación precisa para referirse a las construcciones medias¹.

Teniendo en cuenta dicha laguna de enseñanza en los recursos didácticos de español en China, el presente estudio se centra en la adquisición de la construcción media por parte de ocho hablantes chinos de español como segunda lengua y cuatro de ellos con título de máster en la facultad de Humanidades de la Universidad de Hueva mientras que el resto es del programa *Erasmus* para el alumnado de grado. Empezaremos por la caracterización de las construcciones medias del español para luego situarlas en el marco de la adquisición del español como segunda lengua por los hablantes de chino mandarín (sección 2 y 3). Por consiguiente, se describirá la metodología de estudio y se precisarán los materiales utilizados para obtener el corpus de

¹La explicitación de dicho conocimiento gramático se introduce en la primera lección del Español Moderno III (Dong y Liu, 2014) pero no se encuentra precisa ni rigurosa por el hecho de que no diferencia bien las construcciones reflexivas y las pasivas reflejas de las medias.

respuestas (sección 4). Finalmente, se expondrán el análisis de los resultados y las respectivas conclusiones, así como las reflexiones respecto a mejorar el diseño de los recursos didácticos de español en China (sección 5).

2. LAS CONSTRUCCIONES MEDIAS EN ESPAÑOL

Molina (1974) establece dentro de las construcciones pronominales cuatro grandes tipos: impersonales, pasivas reflejas, medias y reflexivas. A base de esta clasificación indica con ejemplos concretos que las oraciones tales como “Se estropeó el coche al subir el puerto” y “Me alegro de que lo hayáis pasado bien” son de las construcciones medias.

Heredia (1998: 555) a partir del estudio de Molina (1974) se concentra en dos tipos de construcciones medias, s saber, uno con sujeto de persona² y otro con sujeto de cosa³ esquematizando dos implicaciones sintácticas que sirven como criterio identificador de la comprobación de la construcción media. Al referirse a la primera construcción formal indica el autor que, desde el punto de vista sintáctico, la construcción media presenta obligatoriamente un sujeto (que puede ser elíptico) y un verbo transitivo en forma pronominal, admite complementos tanto circunstanciales como preposicionales, pero rechaza la presencia del complemento directo:

(1) Sujeto + clítico + verbo transitivo – complemento directo (+ otros complementos)

Recurriendo a esta implicación, se excluye un grupo de construcciones formalmente parecidos de las medias⁴. No obstante, el autor enfatiza al mismo tiempo que se trata de un criterio general pero no suficiente para el reconocimiento de la construcción media a merced del hecho de que las pasivas reflejas, con una construcción bien semejante, tienden a confundir con las medias⁵. Exponiendo varias construcciones formalmente parecidas, recalca que la construcción media no existe ni puede existir referencia alguna a un

²Como se ha citado en el párrafo anterior “Me alegro de que lo hayáis pasado bien” (Molina 1974).

³“Se estropeó el coche al subir al puerto” (Molina 1974).

⁴Por ejemplo: La oración “estas cosas me preocupan” se queda excluida porque el *me* funciona como complemento directo.

⁵Por ejemplo: “La gripe se cura con aspirinas” y “En mi familia, se cura la gripe con remedios caseros” (Heredia 1998: 554) son distintas dado que la primera es de construcción media mientras que la segunda se considera la construcción de pasiva refleja.

agente humano.

(2) **se + pronombre personal átono + verbo transitivo + SN**

Para esta construcción, el autor hace hincapié en dos aspectos: primero, al clítico personal no puede atribuírsele en principio la función del complemento indirecto⁶; segundo, la resistencia del español al empleo del posesivo deja que este segundo patrón gramatical se haga extensivo a construcciones no propiamente medias⁷.

De hecho, las estrategias sintácticas nos facilitan la precisión de la construcción media. Es difícil, no obstante, dar cuenta cabal de todas las variaciones concernientes. En este sentido, Soler-Espiauba (1995: 402) por medio de su observación, complementa cuatro características generales para identificar las construcciones medias:

1. Sujeto gramatical inanimado
2. El objeto pasa a ser sujeto (Relación de identidad entre ambos).
3. La acción del verbo es independiente de la voluntad y tiene casi siempre resultados negativos no deseados⁸
4. El verdadero sujeto lógico se convierte en un pronombre dativo, receptor de la acción.

Coincidiendo en cierto modo con la postura de Heredia (1998)⁹, Maldonado (1999: 61) señala que, en las construcciones medias, la agentividad del verbo, así como el grado de elaboración de la fuerza inductiva se manifiesta de forma esquelética. Pues se trata de un elemento decisivo para que las medias se diferencien de las construcciones impersonales y de las pasivas. Así, a menor definición del agente inductor, mayor posibilidad de que el tema de las medias gane prominencia.

⁶Señala el autor Heredia (1974) que con el clítico personal se marca una especial relación —normalmente de pertenencia— entre el sujeto de la oración y la persona o ente implicados, si bien no puede atribuírsele en principio la función del complemento indirecto.

⁷Tomando como ejemplos dos oraciones: “se me pasó el dolor” y “se nos ha roto la lavadora” (1998), el autor propone que con un criterio identificador “a+pronombre reflexivo tónico+mismo” se obtiene la discriminación entre la media y la reflexiva debido al hecho de que las construcciones medias, por la intransitivación verbal que en ellas se produce, rechazan este refuerzo: “ * Se nos ha roto la lavadora a nosotros mismos”.

⁸Respecto a este punto, el autor menciona que la construcción también puede tener un sentido neutro en casos como “Se me ocurre una idea” o “Se le alegraron los ojos”.

⁹Heredia (1998: 555) indicó que la construcción media no existe ni puede existir referencia alguna a un agente humano.

Recorriendo los principales rasgos de las construcciones medias, Guillén (2015: 90) resume que sintácticamente, las estructuras medias se relacionan con las reflexivas por su uso de pronombres reflexivos de todas las personas gramaticales. En este sentido, recalca que además de *se*, las construcciones medias, al ser de paciente animado, pueden conllevar el pronombre reflexivo de todas las personas gramaticales¹⁰. Semánticamente, la autora señala que los verbos con carácter medio se relacionan con ciertos verbos transitivos a los que se les ha sustraído el rasgo de causatividad de la acción¹¹.

De hecho, la voz media, como indica Kemmer (1993: 2)¹², no es un fenómeno homogéneo, sino que se trata de un concepto semántico incluyendo a su vez una familia de construcciones medias que formalmente se muestran bien distintas. En nuestro estudio, tomaremos en cuenta las estructuras medias que se constituyen con los pronombres de distintas personas gramaticales (*se*, *os*, *nos*, etc.) y partiremos de la naturaleza de la voz media como operador de la afectación involuntaria, así como el rechazo de responsabilidad.

3. LAS CONSTRUCCIONES DEL *SE* EN LA ADQUISICIÓN DEL ESPAÑOL COMO SEGUNDA LENGUA

En lo que atañe al tema de adquisición, cabe adelantar que el presente texto, siguiendo el criterio de las investigaciones elaboradas por Ellis (1991), McLaughlin (1992) así como Larsen-Freeman y Long (1994), no aplica la distinción entre conceptos de aprendizaje y adquisición que estas, según Krashen (1992), se consideran compartimentos estancos. De hecho, actualmente, el término “adquisición” es utilizado para abarcar ambos contextos. Además, refiriéndose al impacto de la enseñanza en el aula, nos decantamos también por los planteamientos de Larsen-Freeman y Long (1994). En concreto, para ellos, la enseñanza/aprendizaje en clase puede ejercer en la velocidad con que los aprendices adquieren la lengua y, en los niveles de adquisición final.

¹⁰Ej: “Me alegro mucho de que lo hayáis pasado bien”; “Me enfermé”; “Ayer os empapasteis con el aguacero”.

¹¹Ej: “Marta perdió el libro”; “Se perdió el libro”.

¹²“The term “middle voice” has had a wide range of applications in the linguistic literature of this century. Sometimes it denotes a formal category (e.g. Valfells 1970), in keeping with its original use in referring to an inflectional category of the Classical Greek verb. In other cases it is purely semantic, as for example Lyons’ (1969: 373) characterization of middle voices as indicating that “the ‘action’ or ‘state’ affects the subject of the verb or his interests”. (Kemmer 1993: 1).

Tomando como referencia el modelo multidimensional de investigación desarrollado por el Grupo ZISA¹³, el estudio de Johnston (1995) expone de forma detallada las etapas de adquisición del español como segunda lengua. Acogiéndonos al enfoque de Johnston (1995), Guillén (2015: 92) concluye el planteamiento de este, en que el proceso de adquisición de español se divide en siete etapas¹⁴:

- Etapa1: Etapa de producción de palabras o fórmulas.
- Etapa2: Etapa de orden canónico: Sujeto-verbo-objeto.
- Etapa3: Etapa de sujeto al final, anteposición de adverbio. Cambio a posición final o a posición inicial
- Etapa4: Etapa de posición interna, tal es el caso de verbo-sujeto-objeto. Inversión, en el caso de la formulación de preguntas.
- Etapa5: Etapa del “orden libre” de palabras, uso de pronombres clíticos.
- Etapa6: Producción de clíticos y de su uso de secuencia.
- Etapa7: Uso de pronombres relativos y de cláusulas subordinadas.

Según esta indicación, conviene subrayar que el uso de la voz media pronominal correspondería a una estructura de adquisición tardía del español. Puede que cada etapa tenga elementos variables debido a las características del aprendiz (Johnston, 1995: 24), lo cierto es que los aprendices atraviesan todas las etapas del desarrollo hasta alcanzar la frontera meta. Por lo demás, la enseñanza no altera las secuencias del desarrollo (Larsen-Freeman y Long, 1994).

¹³ZISA: (Zweitsprachenwerb Italienischer and Spanischer Arbeiter) llevó a cabo un estudio de tiempo real con 45 adultos y otro longitudinal con 12, el cual se realizó durante dos años, con base en datos tomados de entrevistas de adquisición natural de alemán como segunda lengua por parte de hablantes nativos de español e italiano (Guillén 2015).

¹⁴Según Johnston (1995: 8) la teoría de “las etapas de la adquisición” derivada del “Procesamiento de la información” es aplicable a los aprendices de distintas comunidades de idioma. Pone como ejemplo el caso de los alumnos italianos, señala el autor que las etapas de adquisición de español como SL corresponden no solo al alumnado australiano sino también al italiano.

4. METODOLOGÍA

4.1 Descripción de la investigación

Se llevará a cabo un estudio de tiempo real a través del cual se estudia la actuación lingüística de los ocho hablantes de chino mandarín y los datos se recogerán en una sola sesión. Con motivo de obtener los datos para este estudio, nos dedicamos a un procedimiento en que se plantean preguntas de acuerdo con una serie de ilustraciones y así mismo, por medio del estímulo se esperan las respuestas correspondientes (Larsen-Freeman y Long 1994): los participantes que están expuestos a una serie de imágenes, deben responder a la pregunta “¿Qué le pasó a...?”, la cual pretende estimular la aparición de la estructura en estudio.

4.2 Participantes de la investigación

Los participantes están ubicados en el nivel de B2-C1 de acuerdo con el MCER (2002: 111). Partiendo de los descriptores de niveles comunes de referencia, cabe aclarar que en China continental se lleva otro sistema de certificado de idioma dedicado al alumnado universitario de carrera de español. Se trata de dos sesiones de exámenes: *EEE-4*¹⁵/*EEE-8*¹⁶ organizadas anualmente a nivel nacional, en que participan los estudiantes de la licenciatura de español una vez que han terminado el cuarto y octavo semestre, respectivamente, y que según fuentes pertinentes (Querol, 2014 y Zhou, 2017), equivaldrían a un nivel B2 y C1 de MCER. En nuestro caso, los participantes con título de máster que han aprobado ambos exámenes, tienen el nivel C1 de español mientras que el alumnado del grado, aún sin haber participado en el *EEE-8*, se ubica en el B2 por la prueba de *EEE-4*.

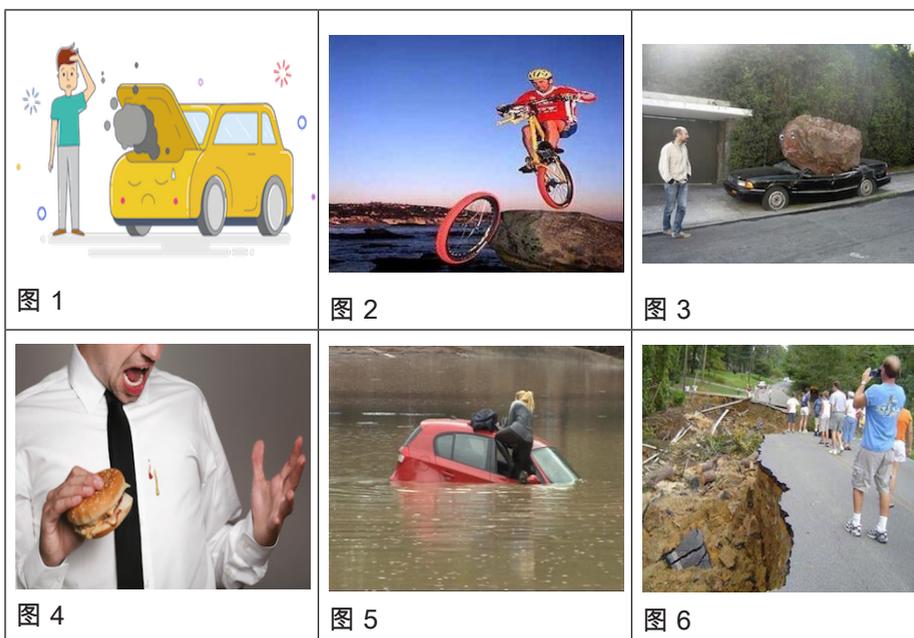
En el grupo de control, tenemos 8 participantes. Todos ellos son de España teniendo español como su lengua materna y estudian grado o máster en la Universidad de Huelva.

¹⁵*Examen de Español como Especialidad de Nivel 4*, una prueba administrada por el Comité Nacional de Español del Ministerio de Educación de China.

¹⁶*Examen de Español como Especialidad de Nivel 8*, una prueba de nivel superior administrada por el Comité Nacional de Español del Ministerio de Educación de China.

4.3 Procedimiento de la investigación

Al efectuar el test, a cada participante se le mostró la siguiente serie de ilustraciones. Se le requiere que la observe y describa concisamente qué le pasó a la persona involucrada. En lo que atañe a la selección de las ilustraciones, tenemos en consideración el planteamiento de Portilla (2006) en que se explicita una característica muy importante de los verbos implicados en las construcciones medias. Sostiene el autor que estas estructuras se tienden a construir con los verbos de proceso¹⁷. A su vez, propone una estrategia de prueba con el fin de la recuperación de este tipo de verbos, a saber, los verbos de proceso, por lo general, pueden responder a las preguntas tales como ¿qué pasó? o ¿qué le pasó a X?



Láminas de ilustraciones

¹⁷Según Vendler (1967), los verbos, de acuerdo con la información aspectual inherente que conllevan, se dividen en cuatro clases: estados, actividades, realizaciones y logros. A partir de dicha perspectiva, existe como guía gramatical otra clasificación de verbos en sentido más general que se trata de los verbos estáticos y dinámicos. Los estáticos son de la clase de estados que propone Vendler mientras que los dinámicos abarcan las clases de actividades, realizaciones y logros. Bajo esta concepción, los verbos procesos, ubicándose en el intervalo de las actividades y logros (realizaciones), se enfocan más en la transformación de estado.

5. ANÁLISIS DE RESULTADOS

A continuación, se organizan en tablas las respuestas de los hablantes¹⁸ y se dividen de primero estos en dos grupos según el nivel de español. La Tabla 1 muestra el número de aparición de la construcción media por parte de los hablantes de chino con nivel de español B2 que llevan al menos tres años aprendiendo este idioma como segunda lengua. Se evidencia que en las respuestas que incluyen la estructura objeto de estudio, los aprendices han tenido la noción del uso de pronombres clíticos.

Hablante 1	Hablante 2	Hablante 3	Hablante 4
4/6	1/6	2/6	5/6

Tabla 1. Número de aparición de la construcción media (Habla ntes de grado)

De la misma manera, la Tabla 2 muestra el número de aparición de la construcción media por parte de los hablantes de chino con nivel de español C1 que llevan al menos cinco años aprendiendo este idioma como segunda lengua.

Hablante 1	Hablante 2	Hablante 3	Hablante 4
2/6	2/6	3/6	6/6

Tabla 2. Número de aparición de la construcción media (Habla ntes de máster)

Al revisar las tablas, se evidencia que los números de apariciones de la construcción media de ambos grupos se encuentran bien similares: el de grado es 12 y el de máster es 13 (Tabla 3):

Número de hablante	Grado (B2)	Máster (C1)
Hablante 1	4	2
Hablante 2	1	2
Hablante 3	2	3
Hablante 4	5	6
Total	12	13

Tabla 3. Comparación del uso de la construcción media por parte de aprendices de grado y de máster

¹⁸A causa del límite requerido de la cantidad de letras, no expondremos las repuestas concretas de los participantes.

A base de esto, la Tabla 4 demuestra el número de aparición de la construcción media por parte de los nativos de español. Hay un total de 37 apariciones de la media.

Hablante 1	Hablante 2	Hablante 3	Hablante 4	Hablante 5	Hablante 6	Hablante 7	Hablante 8
5/6	4/6	4/6	5/6	5/6	3/6	5/6	6/6

Tabla 4. Número de aparición de la construcción media (Hablantes nativos)

Al sintetizar los resultados, conviene afirmar que la construcción media, sea la forma del **pronombre se + le + verbo** como la del **pronombre se + verbo**, es menos utilizada por los hablantes de español como segunda lengua que por los hablantes nativos (Tabla 5).

Hablantes	SE + le + verbo	SE + verbo	Total
Chinos	15	10	25
Nativos	23	14	37

Tabla 5. Comparación de aparición de la construcción media (Hablantes chinos y nativos)

6. CONCLUSIÓN

Una muestra reducida (16 hablantes incluyendo chinos y españoles) nos impide establecer generalizaciones, pero sí nos facilita observaciones. En primer lugar, se evidencia que la construcción media es menos utilizada por parte de los aprendices chinos de español como segunda lengua mientras que los hablantes nativos se decantan por esta estructura mostrando al mismo tiempo una resistencia al empleo del posesivo al referirse al objeto deíctico. En segundo lugar, recurriendo al planteamiento de Johnston (1995) en cuanto a las etapas de adquisición, se puede comprobar que la construcción media, objeto de estudio en este texto, pertenece al ámbito del uso de los pronombres (Etapa 6, producción de clíticos y de su uso en secuencia) que es de adquisición tardía. Por ello, se observa que los participantes chinos son propensos a adoptar alternativas para expresar el significado de afectación, así como el de accidentalidad, aplicando la construcción pasiva en vez de una estructura

media y, en varias ocasiones, tienen problemas con el uso del pronombre *se* cometiendo distintos tipos de errores. Por último, se observa que el alto nivel del dominio de español en nuestro estudio no diferencia significativamente a los aprendices de máster de los del grado respecto a la mejora de adquisición de las construcciones medias.

Al detenernos a reflexionar el porqué de las observaciones obtenidas anteriormente, por un lado, cabe recurrir a la indicación de Schnell (2011). El autor señala que la presencia de la construcción media con *se* muestra carencia importante en el aula de español como segunda lengua y no detalla su uso en situaciones reales. Revisando el principal recurso didáctico de español *Español Moderno* I-IV (Dong y Liu, 2014) que se circulan en el aula de ELE en China, se comprueba que dicha carencia también existe. Refiriéndose al uso de la construcción media en situaciones reales, se visibiliza que en dicho material se incluyen varias reglas simplificadas, así como unas explicitaciones imprecisas que equivalen la voz media a las construcciones reflexivas *y*, a las pasivas reflejas. Asumimos que este tipo de instrucción formal no apoya efectivamente la aceleración del proceso de adquisición de la construcción media en el aula de español. Por otro lado, para seguir la reflexión, conviene también tener en cuenta la situación práctica del ELE en China donde los profesores no nativos suelen encargarse de impartir la clase de gramática española. Así que como señala Cortés (2005), para los docentes no nativos el margen de improvisación no es tan amplio, y cuando más se alejan de los modelos y las normas que han aprendido, tanto más alto sienten que es el riesgo de cometer errores. Es decir, los docentes no nativos, impartiendo clase de español con más cautela, tienden a confiar más en la gramática normativa, así como en los recursos didácticos disponibles y, de esta manera, alejan sin darse cuenta muchas veces a sus alumnos del uso real del idioma aprendido. Considerando dicha situación concreta, defendemos que se fortalezca el contacto entre el profesorado nativo y el no nativo, y que para desarrollar el buen uso de las construcciones medias entre los aprendices chinos de habla español como LE o L2, sea indispensable la inclusión de más recursos pedagógicos elaborados por los nativos en el aula de ELE en China.

REFERENCIAS BIBLIOGRÁFICAS

- ALARCOS-LLORACH, E. (1970). *Estudios de gramática funcional del español*. Madrid: Editorial Gredos.
- CORTÉS MORENO, M. (2005). ¿Hay que enseñar gramática a los estudiantes de una lengua extranjera? *Revista Internacional de Filología y su Didáctica*. 28, 89-108.
- DONG YANSHENG, & LIU JIAN. 董燕生, 刘建. (2014). *Español Moderno I-IV. 现代西班牙语 1-4 册*. Beijing: Foreign Language Teaching and Research Press. 北京: 外语教学与研究出版社.
- ESCUTIA-LÓPEZ, M. (2010). El uso de *se* con verbos inacusativos por estudiantes avanzados de español como lengua extranjera: transferencia y reestructuración. *RESLA*. 23, 129-151.
- ELLIS, R. (1991). *Second language acquisition and language pedagogy*. Clevedon: Multilingual Matters.
- Guillén Solano, P. (2015) Adquisición de las construcciones medias en Español como segunda lengua. *Revista De Filología y Lingüística de la Universidad de Costa Rica*, 41(1), 87-101.
- HEREDIA, J.R. (1998). Sobre construcciones pronominales “medias” en español. *Español como lengua extranjera, enfoque comunicativo y gramática: Actas del IX congreso internacional de ASELE* (pp. 551-560). Santiago de Compostela: Universidad de Santiago de Compostela.
- INSTITUTO CERVANTES (2002). *Marco Común Europeo de Referencias para las Lenguas: Aprendizaje, enseñanza, evaluación*. Madrid: Grupo ANAYA.
- JOHNSTON, M. (1995). Stages of acquisition of Spanish as a second language. *ASLA*. 4, 1-28.
- KRASHEN, S. (1994). El modelo del monitor y la actuación de los adultos en L2. En J. Larsen-Freeman, D. y Long, M. (Eds.), *Introducción al estudio de la adquisición de segundas lenguas*. (I. Molina-Martos y P. Benítez-Pérez, Trans.). Madrid: Editorial Gredos.
- KEMMER, S. (1993). *The middle Voice*. Amsterdam: John Benjamins Publishing company.

- LARSEN-FREEMAN, D. y LONG, M. (1994). Introducción al estudio de la adquisición de segundas lenguas. (I. Molina-Martos y Benítez-Pérez, trans.). Madrid: Editorial Gredos.
- MALDONADO, R. (1999). *A media voz. Problemas conceptuales del clítico se*. México: Universidad Nacional Autónoma de México.
- MCLAUGHLIN, B. (1992). Algunas consideraciones metodológicas sobre el modelo del monitor. Por J. Muñoz-Liceras (Comp.). La adquisición de las lenguas extranjeras. (153-176). Madrid: Visor.
- MENDIKOETXEA, A. (1999). Construcciones con *se*: medias, pasivas e impersonales. Por I. Bosque y V. Demonte (Eds.), *Gramática descriptiva de la lengua española*. (1631- 1722). Madrid: Espasa.
- MOLINA REDONDO, J. A. (1974). *Uso de «se». Cuestiones sintácticas y léxicas*. Madrid: S.G.E.L. .
- PORTILLA, M. (2006). Los pronombres reflexivos clíticos como operadores de intransitivación en español. *Revista de Filología y Lingüística*. 32 (2), 185-201.
- QUEROL BATALLER, M. (2014). La especialidad de español como carrera universitaria en China. *Marco ELE*. 18, 1-12.
- QIANG Z. (2017). *El examen nacional de Español para los estudiantes de filología Española EEE-4: análisis y vinculación al MCER*. Tesis doctoral, Universidad de Barcelona. [Disponible en: <https://www.tdx.cat/handle/10803/405855>].
- SCHNELL, E. (2011). *El modelo interactivo de escucha y lectura para la enseñanza del se en español*. (Tesis de maestría). Universidad de Costa Rica.
- SOLÉ, Y. y SOLÉ, C. (1977). *Modern Spanish Syntax*. Massachusetts: D.C. Heath and Company.
- SOLER ESPIAUBA, D. (1995). El español: lengua posesiva, subjetiva e... irresponsable. (algunos aspectos de la función pronominal). *Tendencias Actuales en la Enseñanza del Español como Lengua Extranjera. Actas del VI Congreso de AESLE* (pp: 399-406). León: Universidad de León.
- VENDLER, Z. (1967). *Linguistic in Philosophy*. Ithaca, NY: Cornell University Press.

CAMBIOS EN LA ESCRITURA DE LOS ESTUDIANTES DE UN GRADO DE MAESTRO TRAS SU PARTICIPACIÓN EN PROYECTOS DE INNOVACIÓN DOCENTE CENTRADOS EN LA DIDÁCTICA DE LOS GÉNEROS¹

Isabel García Parejo

Universidad Complutense de Madrid

Juana María Blanco Fernández

Universidad de Castilla-La Mancha

RESUMEN

Este trabajo tiene como objetivo analizar el efecto del conocimiento e implementación del modelo de apoyo para la escritura Reading to Learn (Rose y Martin, 2012), de base sistémico funcional, en la calidad de las tareas escritas de unos estudiantes del Grado de Maestro. Participaron, a lo largo de dos cursos académicos, 50 estudiantes a los que se solicitó de manera voluntaria la escritura de 2 textos de la familia de los géneros narrativos y 2 textos de la familia de los informativos, antes y después de la instrucción en el modelo mencionado. El corpus se ha analizado desde la perspectiva sistémico-funcional (Halliday, 2004) i) buscando marcas relativas al campo, al tenor y al modo y ii) utilizando una escala de indicadores que han permitido puntuar hasta un máximo de 32 puntos cada texto. Tras aplicar la prueba de Wilcoxon se observa que existen diferencias en los géneros narrativos en todas las variables analizadas, aunque no así en los informativos.

Palabras clave: escritura, géneros, modelo *Reading to Learn*, formación del profesorado

ABSTRACT

The study analyzes the effect of knowledge and implementation of the scaffolding model for writing, Reading to Learn (Rose and Martin, 2012), influenced by systemic-functional linguistics and genre pedagogy on the written tasks of pre-service teachers in the Degree in Primary Education. In the study, implemented over two academic years, a writing

¹Este trabajo forma parte de los estudios realizados en el marco del proyecto Santander-UCM PR26/16-20348, titulado *Educación lingüística y formación del profesorado: incidencia del modelo Reading to learn de base sistémico-funcional en el Grado de Maestro de Educación Primaria.*

intervention was designed involving 50 students. Participants wrote 2 narrative texts and 2 factual texts, before and after the training in the model. The corpus was analyzed from a systemic-functional perspective (Halliday, 2004) i) looking for features of field, tenor and mode, and ii) using a scale of indicators with a maximum score of 32 points for each text. After applying the Wilcoxon test differences in the narrative genres in all the variables analyzed were observed, although this was not the case in the factual genres.

Keywords: writing, genres, *Reading to Learn* approach, teacher training

1. INTRODUCCIÓN

El estudio de la escritura académica en contexto hispanohablante se ha desarrollado principalmente a partir de las iniciativas formativas y de investigación de países como Argentina, Chile, Colombia o México (Navarro, 2017). En este contexto también destaca la labor de profesionales y grupos de investigación que desde distintas disciplinas, enfoques y centros universitarios han contribuido a atraer la atención al estudio de este campo en España (Montolío, 2008-2009, 2014; Castelló, 2007, 2015; Ballano y Muñoz, 2015; Núñez Cortés, 2015; Castelló *et al.*, 2016). Lo relevante, sin duda, es que su trayectoria viene siendo suficiente como para perfilar distintas orientaciones de investigación que contribuyen a esclarecer la situación del fenómeno en el contexto universitario de nuestro país, entre ellas, la que tiene que ver con la presentación y análisis del impacto de propuestas formativas para facilitar a los estudiantes la comprensión y producción de géneros académicos propios de su perfil profesional (en el contexto de formación de maestros véanse, entre otros, Jarpa, 2013; Guzmán y García, 2015, 2017; García Parejo *et al.*, 2017; Álvarez *et al.*, 2015; Romero y Jiménez, 2015; Heredia y Quirós, 2015; Vargas y Rivera, 2015).

Así, nuestro trabajo pretende explorar cuál es la incidencia que tienen sobre la escritura de los estudiantes del Grado de Maestro de Educación Primaria dos proyectos de innovación docente desarrollados en la Universidad Complutense de Madrid con la finalidad de conocer e implementar la pedagogía de género de base sistémico funcional o *Genre Pedagogy* (García Parejo *et al.*, 2017). Según Martin (1984: 25), un género consiste en una actividad orientada hacia un determinado propósito y formada por etapas, en la que los hablantes intervienen en tanto que miembros pertenecientes a una cultura. En línea con esta definición, Martin y Rose (2008) elaboran una clasificación de géneros discursivos escolares presentes en el currículo australiano y una pedagogía de escritura que permita a cualquier escolar afrontar con éxito las demandas de expresión escrita de los mismos. Su propuesta es conocida como modelo *Reading to Learn* (Rose y Acevedo 2017).

Con respecto a la taxonomía propuesta por los autores, su organización responde a tres propósitos fundamentales en torno a los cuales se distribuyen los distintos géneros: cautivar al lector, próximo a los géneros narrativos; informar al lector, que aglutina los géneros del tipo crónicas, explicaciones, informes y procedimientos; y, por último, evaluar, vinculado a las argumentaciones y a la

crítica.

Reading to Learn (en adelante *R2L*), por su parte, presenta los fundamentos propios de la pedagogía de género en que se sustenta: la lingüística sistémico-funcional, la sociología de la educación y el constructivismo. En relación con esta corriente, un concepto fundamental vinculado al modelo es el de *Scaffolding Learning Cycle* o Ciclo de Aprendizaje Andamiado, enfoque pedagógico que orienta la realización de cualquier tarea de aprendizaje en el aula a partir de su preparación previa y elaboración posterior y que está en la base de la secuencia de aplicación de las estrategias del modelo para las tareas de escritura. Estas estrategias se distribuyen en tres niveles según el grado de apoyo que ofrecen a los estudiantes y el nivel lingüístico desde el que operan: texto en contexto, oración en párrafo y palabra en cláusula (véase Rose y Acebedo, 2017).

2. CARACTERÍSTICAS DEL ESTUDIO

A partir del estudio de textos producidos por estudiantes de Grado que han cursado las asignaturas de Lengua española y Didáctica de la Lengua, los objetivos específicos de esta investigación parcial son dos, uno previo a la intervención docente: identificar las áreas donde ocurren la mayoría de los errores, y uno posterior: valorar si la innovación docente basada en el conocimiento explícito de los géneros y en su didáctica ha tenido influencia en la escritura de estos estudiantes, especialmente en las áreas conflictivas que pudieran detectarse en los primeros textos producidos.

2.1 Descripción de los participantes

Los participantes en este estudio (N=11) han sido seleccionados de una muestra total de 52 estudiantes de tercer curso del Grado de Maestro en Educación Primaria de las modalidades Bilingüe y General. Estos 11 participantes pertenecen al grupo General; de ellos, el 63,6% son mujeres y el 36,4% hombres, con una media de edad de 21,4 años. Estos 11 estudiantes han participado en todas las tareas propuestas a lo largo del curso 2016-2017 en las asignaturas de Lengua y Didáctica de la Lengua y, por ello, es posible analizar sus producciones antes y después de la intervención.

2.2 Metodología

Los datos de este estudio fueron recogidos en dos fases y se centraron en géneros escolares relevantes para el desarrollo profesional de futuros docentes, especialmente relatos e informes expositivos. Mientras que los relatos (por ejemplo, anécdotas y narraciones literarias) aparecen de forma sistemática en lecturas y actividades propuestas en manuales de Lengua en la educación primaria, los informes expositivos (especialmente la exposición descriptiva) son un género fundamental en la expresión del conocimiento en todos los niveles educativos y en todas las disciplinas académicas (Martin y Rose, 2008).

En una primera fase, antes de presentar el modelo *R2L*, se solicitó a los estudiantes la producción de un género narrativo, en concreto una anécdota relacionada con sus experiencias en las clases de Lengua (TNA) y de un género informativo, en concreto, una exposición descriptiva (TED1) para presentar las tres ciencias del lenguaje que han influido en la didáctica de la lengua. Al final del segundo semestre, tras la implementación de la secuencia didáctica *R2L*, se solicitó otro género narrativo, una narración literaria (TNN) y otra explicación descriptiva (TED2), originadas ambas a partir de las tareas de clase.

Para el análisis del corpus, se diseñó y validó con jueces expertos una escala analítica que incluye un total de 8 ítems organizados alrededor del estudio del *campo* (finalidad y etapas del género), del *tenor* (relación entre participantes, registro) y del *modo* (estructura y cohesión de las oraciones), (Halliday, 2004; Martin, 1984; Rose y Martin, 2012). Así, las categorías de la escala analítica son las mismas para todos los géneros, pero difieren en su nivel de concreción según el género solicitado, como se muestra a continuación:

Anécdota y Narración		Exposición descriptiva
CAMPO		1. Finalidad. ¿Interpretó adecuadamente la consigna que se le pedía?
	ETAPAS DEL GÉNERO	2. Orientación: ¿Incluye algún tipo de información espacial o temporal?
		3. Evento disruptivo: - ¿Aparece un evento insólito? (anécdota) - ¿Aparece una complicación que ha de resolverse? (narración)
		4. ¿Aparece relatada alguna reacción ante el evento disruptivo?
		2. Clasificación: ¿Introduce el tema con una oración tópico?
		3. Descripción: ¿Se anuncia cada nuevo tema requerido en la definición?
		4. ¿Describe el estudiante la entidad requerida explicando qué es?
TENOR	5. ¿Se puede identificar claramente en el texto quién es el escritor y cuál es la relación que tiene con la audiencia-destinatario?	
	6. ¿A lo largo del texto, el estudiante ha sido capaz de utilizar recursos léxicos para expresar y comunicar adecuadamente a la audiencia las actitudes y emociones de los personajes y del propio narrador?	6. ¿Utiliza el estudiante recursos léxicos para expresar valoraciones que sirven para explicar y clasificar, y para graduar los eventos o entidades requeridos?
MODO	7. ¿La estructura y complejidad de las oraciones es adecuada para la tarea requerida y el nivel académico del estudiante?	
	8. En lo que se refiere a la cohesión dentro de las oraciones y dentro del propio texto, ¿aparecen marcadores propios del género?	

Tabla 1. Categorías analizadas en los textos producidos por los estudiantes del Grado

Para cada ítem se propuso una escala de intervalo de 4 puntos (Kuiper *et al.*, 2017; Ferreira y Lantolf, 2008), y se redactaron los descriptores de rendimiento mínimo y máximo para que los evaluadores pudieran baremar las muestras sobre 32 puntos. Así, por ejemplo, para el ítem 6 de los géneros narrativos se debe calificar con la nota más baja (1) si no aparece expresada ninguna valoración ni reacción ante el episodio disruptivo, y con la nota más alta (4) si aparecen claramente expresadas diferentes actitudes y valoraciones por parte de los personajes y/o del narrador. Mientras que para el ítem 6 del género informativo, se debe calificar con la nota más baja (1) si aparece la valoración personal del estudiante sobre el tema, y con la nota más alta (4) si solo aparece valoración objetiva relacionada con la descripción-clasificación.

3. RESULTADOS SOBRE LA PRODUCCIÓN ESCRITA DE LOS ESTUDIANTES

Como se puede observar en la tabla que mostramos a continuación, la diferencia entre la media obtenida en el TNA (anécdota) y el TNN (narración literaria), así como entre el puntaje máximo y mínimo, es mayor que la que existe entre la obtenida en la exposición descriptiva 1 y la 2 (TED1, TED2).

	Media	Máximo	Mínimo	Desviación estándar
Anécdota TNA	19,18	27,00	12,00	4,73
Narración TNN	30,64	32,00	28,00	1,36
Exposición Descriptiva TED1	25,20	32,00	21,00	3,91
Exposición Descriptiva TED2	26,00	32,00	22,00	3,03

Tabla 2. Puntaje total de los textos producidos antes y después de instrucción didáctica

Esto pudiera estar relacionado con el hecho de que el tema de la exposición descriptiva fuera la definición de paradigmas lingüísticos que influyen en los cambios observados en la didáctica de la lengua a lo largo del s. XX. Este tema es conocido por los estudiantes, ya que les ha sido presentado en lecturas y solicitado en varios ejercicios, lo que convierte la tarea en la producción de un género secundario (Bajtín, 1999), organizado desde la experiencia en comunidades de prácticas académicas. Sin embargo, la anécdota, asociada a un género primario, y quizás, a la comunicación oral, no parece ofrecer a los estudiantes un referente cultural específico que puedan utilizar para organizar su texto.

A continuación, presentamos los datos relativos a las anécdotas y narraciones producidas por los estudiantes del Grado de Maestro.

3.1 Características de los relatos

En la clasificación de los géneros escolares propuesta por Martin y Rose (2008), la anécdota se incluye dentro de la familia de los 'relatos' cuya finalidad global es

cautivar al lector (relatos cotidianos, narraciones literarias, fábulas). Todos los géneros de esta familia se caracterizan por compartir una etapa de ‘orientación’, que incluye fases descriptivas del marco de la historia, del ambiente y de los personajes. Sin embargo, se diferencian en la etapa de registro de eventos y la propia naturaleza de los mismos, así como por presencia explícita de etapas de reacción y de interpretación.

En la medida en que los relatos presentan una serie de eventos necesarios para hacer progresar la acción, puede ocurrir que estos eventos sean esperados por el lector (u oyente) debido a los diferentes acontecimientos que han sucedido previamente (cf. relatos cotidianos). Pero pueden aparecer también como eventos inesperados, ya que toman la forma de problemas que se resuelven (cf. narraciones literarias) o de incidente que se interpreta y juzga (cf. fábulas). O quizás, aparecen eventos insólitos ante los que se produce una reacción, es el caso de las anécdotas.

3.1.1 Las anécdotas

Se solicitó a los estudiantes que relatasen una anécdota a partir de la siguiente consigna: “Qué recuerdas de tu clase de Lengua. Cuéntanos en 15 minutos una anécdota relacionada con tus experiencias en la clase de lengua cuando eras estudiante o, si lo prefieres, de tu experiencia durante tu periodo de Prácticas de Magisterio”. Esta tarea se solicitó en clase tras un debate y puesta en común alrededor de frases motivadoras de autor sobre ‘qué es la lengua’ y ‘qué es enseñar lengua’.

Teniendo en cuenta la caracterización del género ‘anécdota’, se esperaba un texto organizado alrededor de una etapa de orientación (situación) y de un evento disruptivo e insólito ante el que el autor reaccionase. En cada etapa, deberían poder encontrarse fases descriptivas y valorativas.

Sin embargo, como se puede observar en la tabla 3, es la presencia o no de las etapas características del género anécdota lo que condiciona el puntaje final, pues estas etapas permiten reconocer el género producido y, por tanto, su propósito. Es en estos ítems del campo donde se observan las puntuaciones más bajas:

	Media	Máximo	Mínimo	Desviación estándar
CAMPO TNA-1. finalidad	2,27	4,00	1,00	1,27
CAMPO TNA-2. etapa orientación	2,91	4,00	1,00	1,04
CAMPO TNA-3. evento disruptivo	2,00	3,00	1,00	0,89
CAMPO TNA-4. fases intermedias	1,91	3,00	1,00	0,83
TENOR TNA-5. identificación escritor, audiencia	3,09	4,00	2,00	0,70
TENOR TNA-6. relación y actitud escritor-audiencia	2,36	4,00	1,00	1,12
MODO TNA-7. estructura y complejidad oracional adecuada a modalidad y género	2,36	3,00	2,00	0,50
MODO TNA-8. cohesión clausulas-oraciones, párrafos	2,27	3,00	2,00	0,47

Tabla 3. Calificación media del campo, tenor y modo en las anécdotas producidas

Además, se observan medias bajas en lo que se refiere al modo, elemento que valora la complejidad oracional según nivel (estudiantes universitarios), así como la cohesión interna de todo el texto. Los dos siguientes ejemplos, transcritos respetando el original, son una muestra de extremos máximo y mínimo en las diferentes variables:

Ejemplo 1:

Este año en las prácticas, la profesora en uno de los ejercicios puso una serie de palabras y los alumnos tenían que leer en el diccionario su significado y con lo que ellos habían entendido tenían que formar una oración. A medida que los alumnos terminaban su examen yo iba leyendo y en uno de ellos veo que una de las frases que el alumno había formado era con la palabra *coz*. La frase era la siguiente: A Marta y Martín le dieron una *coz* los caballos en el campo.

Inmediatamente llamo al alumno, que era uno de los que más destacaba académicamente y le digo que porque me pone el verbo así y me dice: “El verbo dar, en pasado, *daron*”.

Ejemplo 2:

Los alumnos estaban leyendo Fray Perico y su borrico y tras cada capítulo la profesora les hacía preguntas sobre el tema, pero ví que no les preguntaba acerca del vocabulario, así que les hice yo varias preguntas y vi que no sabían lo que significaba y que tampoco preguntaban, así que al día siguiente decidí preguntarles las más difíciles y explicárselas.

En el primer ejemplo, el estudiante comienza presentando el marco donde se desarrolla el relato (año de prácticas, profesora y alumnos leyendo). En el segundo, también aparece un marco (profesora y alumnos leyendo), pero no se señala cuándo, por lo que la etapa de 'orientación' queda incompleta. En lo que se refiere a la presentación del evento insólito y a la organización de fases intermedias, el primer ejemplo incluye un episodio que llama la atención del futuro maestro y cierra con la respuesta del alumno del aula de primaria sin ningún comentario. Al cerrar de esta manera, implica al lector presuponiendo que este tendrá también una reacción ante el episodio. En el segundo texto, no aparece valoración explícita o implícita, el autor se limita a relatar unos hechos, por lo que no se ajusta a las características del género.

3.1.2 Las narraciones literarias

Si bien anécdotas y narraciones son dos géneros diferentes, ambos comparten etapas de orientación y de eventos disruptivos, así como estructuras oracionales asociadas a la modalidad escrita, a la descripción de diferentes ambientes y a la narración de eventos sucesivos. Como podemos observar en la tabla 4, tras la instrucción explícita los estudiantes mejoran:

	Media	Máximo	Mínimo	Desviación estándar
CAMPO TNN-1. finalidad	4,00	4,00	4,00	0,00
CAMPO TNN-2. etapa orientación	4,00	4,00	4,00	0,00
CAMPO TNN-3. evento disruptivo	3,91	4,00	3,00	0,30
CAMPO TNN-4. fases intermedias	4,00	4,00	4,00	0,00
TENOR TNN-5. identificación escritor, audiencia	4,00	4,00	4,00	0,00
TENOR TNN-6. relación y actitud escritor-audiencia	3,91	4,00	3,00	0,30
MODO TNN-7. estructura y complejidad oracional adecuada a modalidad y género	3,55	4,00	3,00	0,52
MODO TNN-8. cohesión cláusulas-oraciones, párrafos	3,27	4,00	2,00	0,65

Tabla 4. Calificación media del campo, tenor y modo en las narraciones producidas

Mostramos a continuación parte de la narración desarrollada por el autor de la anécdota 2 tras implementar estrategias de lectura y de escritura con el cuento de los hermanos Grimm, ‘El príncipe rana’:

Ejemplo 3:

Érase una vez un príncipe que volaba alrededor del castillo con su capa mágica. Era una capa de flores, su bien más preciado. Se divertía bastante hasta que una corriente de aire hizo que su capa cayese en una casa y él en un lago cercano. El príncipe trató de recuperarla, pero la casa

estaba rodeada de un muro tan alto que no se veía la parte de arriba [...]
De repente, salió de los arbustos un tejón horrible que dijo: No te sulfures. Yo te devolveré tu capa mágica si prometes quererme y dejarme vivir contigo. [...]

4. CONCLUSIONES

A pesar de las limitaciones del estudio (análisis de 44 textos y diferencia entre los géneros narrativos analizados), los resultados parecen apoyar la idea de que es necesario organizar el currículum de las asignaturas que forman parte del programa oficial del Grado de Maestro alrededor de programas docentes basados en una didáctica explícita de los géneros (Carlino, 2006). A través del análisis de la incidencia de nuestros proyectos de innovación sobre la escritura de los futuros docentes, hemos observado la dificultad que pueden encontrar estos estudiantes a la hora de organizar textos exigidos en su futura práctica docente (cf. relatos). Si se incluye el estudio explícito de los géneros, se favorecerá que los estudiantes mejoren en sus competencias discursivas académicas y profesionales. Además, el profesorado responsable de las diferentes materias del Grado podrá evaluar la calidad de los textos producidos por sus estudiantes teniendo en cuenta los diferentes ámbitos sociales y académicos en los que deban comunicarse.

En este sentido, consideramos que estudios de este tipo aportan información relevante para profundizar en los procesos de evaluación de la lengua escrita. Por un lado, el instrumento diseñado para analizar los textos (de base sistémico-funcional) y, por otro, el proceso seguido en las fases de diseño y aplicación del mismo (presencia de docentes y de expertos), permiten explorar las marcas características de cada género. Con todo, somos conscientes de que hay que seguir desarrollando este instrumento con vistas a incluir más variables dentro de cada una de las dimensiones propuestas (campo, tenor y modo).

De igual manera, para poder llegar a conclusiones generalizables, estudios futuros deberían integrar un mayor número de asignaturas del Grado, así como un mayor número de sujetos participantes. Estos resultados parciales pueden ayudar a generar entre la comunidad académica una reflexión acerca de los géneros discursivos, así como de las competencias asociadas a su conocimiento y uso, en la formación de los futuros maestros.

REFERENCIAS BIBLIOGRÁFICAS

- ÁLVAREZ, T., MATEO, T., SERRANO, M. P., GONZÁLEZ y M. A. 2015. “Diseño de la plataforma ‘RedacText 2.0’ para ayudar a escribir textos académicos e investigar sobre enseñanza y aprendizaje de la escritura”, *Revista Complutense de Educación*, 26(2): 425-445.
- BAJTÍN, M.M. 1999. *Estética de la creación verbal*. Madrid: Siglo XXI.
- BALLANO, I., MUÑOZ, I. 2015. *La escritura académica en las universidades españolas*. Bilbao: Servicio de Publicaciones de la Universidad de Deusto.
- CARLINO, P. 2006. “Ayudar a leer en los primeros años de universidad o de cómo convertir una asignatura en «materia de cabecera»”, *Educación en Ciencias*: 1-12.
- CASTELLÓ, M. 2015. “La investigación sobre escritura académica en los procesos de enseñanza-aprendizaje en la universidad Española”, *Cultura y Educación*, 27(3): 465-476.
- CASTELLÓ, M. (coord.) 2007. *Escribir y comunicarse en contextos científicos y académicos: conocimientos y estrategias*. Barcelona: Graó.
- CASTELLÓ, M., MATEOS, M., CASTELLS, N., IÑESTA, A., CUEVAS, I. y SOLÉ, I. 2016. “Academic Writing in Spain”. *Exploring European writing cultures. Country reports on genres, writing practices and languages used in european higher education*, O. Kruse, M. Chitez, B. Rodriguez, M. Castelló (coords.). Winterthur: ZHAW Zurich University of Applied Sciences. 201-209.
- FERREIRA, M. y LANTOLF, J. 2008. “A concept-based approach to teaching writing through genre analysis”. *Sociocultural theory and the teaching of second languages*, J. P. Lantolf, M. E. Poehner (eds.). Sheffield (UK) and Bristol (USA): Equinox. 285-320.
- GARCÍA PAREJO, I., AHERN, A. y GARCÍA BERMEJO, M.L. 2017. “La ‘pedagogía de los géneros discursivos’ en Grados de Formación del Profesorado: Presentación de dos proyectos de innovación docente”, *Lenguaje y Textos*, 46: 69-80.
- GUZMÁN, F. y GARCÍA, E. 2015. “La alfabetización académica en la Universidad. Un estudio predictivo”, *RELIEVE*, 21(1).
- GUZMÁN, F. y GARCÍA, E. 2017. “La alfabetización académica de los futuros maestros. Un estudio comparativo en varias universidades españolas”, *Revista de Investigación Educativa*, 35(2): 317-335.

- HALLIDAY, M.A.K. 2004. *An Introduction to Functional Grammar*, Third Edition. Revised by C. M.I.M. Matthiessen London: Hodder Arnold.
- HEREDIA, H. y QUIRÓS, F. 2015. “El taller de escritura académica como estrategia de alfabetización para la creación de textos específicos: memorias de prácticas y trabajos de fin de grado”. *Horizontes científicos y planificación académica en la didáctica de lenguas y literaturas*, X. Núñez, A. González, Pazos, C. y Donos, P. (coords.). V. N. Famalicão: Edições Húmus. 443-459.
- JARPA, M. 2013. “Una propuesta didáctica para el desarrollo de la escritura académica en universitarios”, *Revista Iberoamericana de Evaluación Educativa*, 6 (1): 29-48.
- KUIPER, C., SMIT, J., DE WACHTER, L. y ELEN, J. (2017). “Scaffolding tertiary students’ writing in a genre-based writing intervention”, *Journal of Writing Research*, 9(1): 27-59.
- MARTIN, J. (1984). “Language, register and genre”. *Children Writing: reader*, F. Christie (ed.). Geelong, Victoria, Australia: Deakin University Press. 21-30.
- MARTIN, J. R., ROSE, D. 2008. *Genre relations: mapping culture*. London: Equinox.
- MONTOLÍO, E. (coord.) 2008-2009. *Manual práctico de escritura académica*. Barcelona: Ariel. 3 vols.
- MONTOLÍO, E. (coord.) 2014. *Manual de escritura académica y profesional*. Barcelona: Ariel. 2 vols.
- NAVARRO, F. 2017. “Estudios latinoamericanos de la escritura en educación superior y contextos profesionales: hacia la configuración de un campo disciplinar propio”, *Lenguas Modernas*, 50: 9-14.
- NÚÑEZ CORTÉS, J. A. (coord.) 2015. *Escritura académica: de la teoría a la práctica*. Madrid: Pirámide.
- ROMERO, M., JIMÉNEZ, R. 2015. “La escritura académica como estrategia de centro en la Facultad de Ciencias de la Educación de Cádiz”. *La escritura académica en las universidades españolas*, I. Ballano, I. Muñoz (coords.). Bilbao: Servicio de Publicaciones de la Universidad de Deusto. 63-95.
- ROSE, D. y ACEVEDO, C. 2017. “Aprender a escribir, Leer para aprender: origen y desarrollo de proyectos para la mejora de la lectura y la escritura en Australia”, *Lenguaje y Textos*, 46: 7-18.

- ROSE, D., MARTIN, J. 2012. Learning to Write, Reading to Learn: Genre, Knowledge and Pedagogy in the Sydney School. Sheffield (UK) and Bristol (USA): Equinox.
- VARGAS, J. M. y RIVERA, P. 2015. "Writing Programs o la planificación de la escritura académica en el ámbito universitario: una práctica para la mejora de la expresión escrita en estudiantes de nuevo ingreso", *Horizontes científicos y planificación académica en la didáctica de lenguas y literaturas*, X. Núñez, A. González, Pazos, C., Donos, P. (coords.). V. N. Famalicão: Edições Húmus. 525-535.

EL PROCESO LECTOR EN ESPAÑOL COMO LENGUA EXTRANJERA: FLUIDEZ LECTORA SILENCIOSA Y APTITUD MUSICAL

José Manuel Foncubierta

Universidad de Huelva / KU Leuven

RESUMEN

Este estudio muestra el modo en que en la fluidez lectora silenciosa en L2 interactúa con las habilidades lectoras adquiridas en L1, con la conciencia fonológica en L2 y con variables no lingüísticas como la aptitud musical, de acuerdo con su naturaleza acústica.

Participaron 117 estudiantes universitarios italianos de español como lengua extranjera (ELE). Los estudiantes tuvieron que completar una batería de pruebas cuyos resultados permitieron medir su fluidez lectora silenciosa en L1 y en L2 (Test of silent contextual reading fluency), su conciencia fonológica en L2 (Spoonerism Test) y su sensibilidad musical, de acuerdo con el reconocimiento del ritmo, el tono, el tempo y la melodía (Mini-Proms Test). Los análisis de correlaciones y el análisis de regresión lineal muestran la existencia de una relación significativa entre variables lingüísticas y no lingüísticas, lo cual contribuye a explicar los perfiles lectores fluyentes en L2.

Palabras-clave: Fluidez lectora silenciosa, Conciencia fonológica, Español lengua extranjera, Aptitud musical, Diferencias individuales

ABSTRACT

This paper shows the way in which silent reading fluency in L2 interacts with reading skills acquired in L1, with phonological awareness in L2 and with non-linguistic variables such as musical aptitude, according to its acoustic nature.

The study involved 117 Italian university students of Spanish as a foreign language (ELE). The students completed a battery of tests whose results allowed to measure their L1 and L2 silent reading fluency (Test of silent contextual reading fluency), their phonological awareness (Spoonerism Test) and their musical sensibility, consisting in the recognition of rhythm, tone, tempo and melody (Mini-Proms Test Battery). The correlation and covariance analyzes employed show the existence of a significant relationship between linguistic and non-linguistic variables, which helps to explain the profiles of fluent readers in L2.

Keywords: Silent reading fluency, Phonological awareness, Spanish as a foreign language, Musical aptitude, Individual differences.

1. INTRODUCCIÓN

La fluidez es un componente crucial para el desarrollo de la competencia lectora y de la implicación motivacional del lector con la lectura, además de ser fundamental para la detección de perfiles lectores con problemas (Gagliano *et al.*, 2015; Ciuffo *et al.*, 2017). La lectura competente requiere la coordinación e integración automática de los mecanismos de procesamiento necesarios para devolver a la oralidad los elementos lingüísticos que decodificamos de manera visual a través de la lectura del texto escrito. Tradicionalmente, este “aspecto de la automaticidad” (Groen, Veenendaal y Verhoeven, 2018) se ha medido con tareas en las que el lector lee un texto con velocidad, y sin errores, en un espacio de tiempo determinado. Sin embargo, desde la aparición del informe de *National Reading Panel* (2000), la investigación ha ido centrándose también en otros aspectos de la fluidez como la expresividad o prosodia lectora (Allington, 2006). Como señala Nation (2009: 135), “la fluidez no significa simplemente llegar a leer más rápido, sino cambiar la talla y la naturaleza de la unidad básica con la que el lector trabaja”. Para el lector no fluyente en L2, la unidad básica es subléxica “la/c-a/s-a”, mientras que para el lector fluyente es probable que sea la palabra prosódica “lacasa”. Cuando se lee, las palabras impresas están separadas por espacios en blanco, mientras que en la oralidad son un continuo (son un continuo) y no solo el conocimiento gramatical en L2 ayuda al lector a comprender. El lector en L2 puede beneficiarse de habilidades lectoras adquiridas en L1, tal y como proponen la Hipótesis de la interdependencia lingüística de Cummins (1979) o la *Hipótesis de la codificación lingüística* de Sparks (1995). Pero el lector en L2 también necesita devolver al texto su estructura melódica y, para eso, ha de realizar el adecuado procesamiento sintáctico, las pausas correctas para el reconocimiento de los grupos fónicos, las variaciones de volumen para las inferencias o la detección de las actitudes emocionales y la comprensión lectora (Groen, Veenendaal y Verhoeven, 2018). Todo ello precisa de cierta *sensibilidad musical* o aptitud hacia las propiedades acústicas y melódicas del lenguaje.

La sensibilidad auditiva o aptitud musical, entendida como capacidad natural de las personas para procesar y discriminar el input musical (*non-musician*), frente a la cualidad de las personas formadas musicalmente (*musician*), ha sido tema de interés por cuanto el potencial musical de personas aculturadas musicalmente (*musical sleepers*) puede ofrecer datos más objetivos para evaluar las diferencias individuales en la percepción de patrones musicales (Law &

Zentner, 2012; Zentner & Strauss, 2017).

La música y la prosodia del habla comparten similitudes estructurales fundamentales para el desarrollo de la fluidez lectora (Patel, 2008). Tanto en la entonación de los enunciados como en los contornos melódicos de la música se producen cambios temporales en la frecuencia fundamental, basados en una estructuración jerárquica por agrupamientos de unidades menores que determinan el ritmo (Besson, Barbaraoux & Dittinger, 2017). Por ello, si leer con fluidez, aún en silencio, significa sonorizar internamente los símbolos escritos siguiendo los contornos melódicos específicos de la lengua y los patrones rítmicos necesarios para integrar la información y lograr la comprensión lectora, es muy probable que la aptitud del lector de L2 hacia la discriminación de estructuras sonoras no lingüísticas esté relacionada con la fluidez lectora silenciosa. Slevc & Miyake (2006) estudiaron la relación de la aptitud musical con la competencia lingüística en una L2 (fonología, sintaxis y vocabulario) con estudiantes adultos sin formación musical. Sus resultados mostraron que la aptitud musical (o tener buen oído) facilitaba la adquisición de la L2, especialmente en el nivel de la discriminación y producción fonológica, no tanto en el léxico ni en la sintaxis. El estudio de Milovanov, Pietilä, Tervaniemi & Esquef (2010) muestra resultados similares con estudiantes adultos finlandeses aprendiendo inglés. Otros estudios como los de Zhang & McBride-Chang (2010) planean la hipótesis de que la sensibilidad auditiva dé forma a la capacidad de discriminar y categorizar los sonidos del habla.

2. DISEÑO DE LA INVESTIGACIÓN

El objetivo de este estudio se dirige a expandir nuestro conocimiento sobre el desarrollo de la fluidez lectora silenciosa, al estudiar el reconocimiento visual de palabras (reconocimiento de palabras en un texto) y su relación con aspectos auditivos como la conciencia fonológica y la aptitud musical. El estudio se estructura en torno a cuatro preguntas de investigación:

1. ¿Existe una relación entre el nivel de fluidez lectora silenciosa en L1 y en L2?
2. ¿La conciencia fonológica en L2 se relaciona con la aptitud musical?
3. ¿La aptitud musical se relaciona con la fluidez lectora silenciosa en L1 y en L2?

4. ¿La conciencia fonológica en L2 y la fluidez lectora silenciosa en L1 pueden predecir la fluidez lectora silenciosa en L2?

En este estudio reportamos los resultados de una muestra de lectores adultos italianos con diferente nivel de fluidez lectora silenciosa en español como lengua extranjera.

2.1 Participantes

Formaron parte del estudio un total de 124 sujetos, de los cuales solo 117 respondieron todas las pruebas que se suministraron. De estos, el 34.7% (n = 43) eran hombres, y el 65.3% (n = 81) eran mujeres. El mínimo de edad observado fue de 21 y el máximo de 25 años, con una media de 21.49 años (Sd = .768). Todos los sujetos eran estudiantes de primer año de Licenciatura en Mediación Lingüística y Cultural en la Universidad de Macerata. Como hablantes nativos de italiano, nunca habían estudiado español en España, y reconocieron no sufrir ningún tipo de discapacidad lectora. Solo un 4.7% había recibido algún tipo de formación musical, como aprender a tocar algún instrumento musical.

2.2 Instrumentos

Test de Fluidez Lectora Silenciosa (FLS)

En primer lugar, se suministró el Test de Fluidez Lectora Silenciosa. Se trata de una adaptación de Test Of Silent Contextual Reading Fluency (Hammill, Wiederholt, & Allen, 2006). Esta prueba mide la rapidez con la que los estudiantes pueden identificar palabras individuales dentro de un texto. Las palabras impresas aparecen sin espacios ni signos de puntuación. A diferencia del TOSCRF, las letras aparecen en minúscula siguiendo el precepto de Hiebert & Reutzel (2014: 37) de que “las palabras minúsculas ofrecen al ojo lector un Skyline de las palabras”. Los estudiantes tuvieron tres minutos para reconocer tantas palabras como les fuera posible utilizando barras de separación con el bolígrafo. El resultado total fue derivado del número de palabras correctamente identificadas en el espacio de tiempo señalado. Según Hammill *et al.* (2006), la validez estimada de esta prueba varía de 0.67 a 0.85 en relación con otras

medidas de lectura validadas.

Se trataba de dos fragmentos diferentes de la *Declaración de los Derechos Humanos*, uno en español y otro en italiano. La prueba la realizaron primero con el texto en L2 y luego en L1.

Conciencia Fonológica (Spoonerism task, Perin, 1983)

A continuación, se suministró Spoonerism Task. Los alumnos escuchaban nombre y apellido de personas famosas en la cultura española (por ejemplo, **Penélope Cruz**) y se les pedía que cambiaran la consonante inicial del nombre por la consonante inicial del apellido (produciendo **Cenélope Pruz**). Una vez que se les pronunciaba el nombre, solo disponían de 4 segundos para responder. Había 18 pares de nombres. El test fue administrado individualmente. Se usó una grabadora Olympus Ws-650S para la recogida de datos. El tiempo de recogida de datos fue de 2 minutos por cada sujeto.

Aptitud musical (Mini-PROMS Test Battery, Zenter & Strauss, 2017)

Por último, se les suministró la versión reducida de Proms, un test sobre aptitud musical desarrollado por la Universidad de Innsbruck. Compuesto por una batería de subtests, Mini-Proms permite medir la aptitud musical mediante la discriminación auditiva de diferentes estructuras acústicas musicales: Tuning, Melody, Accent and Tempo. El subtest de Tuning se basa en determinar si dos tonos consecutivos son similares o no respecto al primer tono. Melody es un subtest consistente en discriminar secuencias y patrones en función de su armonía. Accent evalúa la capacidad de detección y retención de patrones rítmicos en una secuencia de 5 a 12 beats y Tempo compara la capacidad de distinguir dos secuencias con ritmo monótono, sin ningún acento y con distinta duración. El test se suministró en grupo en una sala de ordenadores, compartiendo el enlace que nos había proporcionado la Universidad de Innsbruck. El tiempo de recolección de datos fue de 20 o 25 minutos.

2.3 Análisis de datos

Primero se procedió a estudiar la normalidad de las variables mediante la prueba de Kolmogorov-Smirnov, con el fin de determinar el uso de pruebas paramétricas o no paramétricas. Se optó por el uso de pruebas no paramétricas, dada la no normalidad de las variables estudiadas ($p < .05$). Con el fin de responder las hipótesis relacionales, se llevó a cabo un análisis correlacional bilateral mediante la prueba Rho de Spearman.

Para responder a las preguntas acerca de la influencia de la conciencia fonológica y de la FLS en L1 sobre la FLS en L2, se implementó un estudio de regresión lineal.

El software utilizado para llevar a cabo todos los análisis de datos mencionados ha sido el SPSS v23.

3. RESULTADOS

En el estudio correlacional (Tabla 1) se observa que todas las variables correlacionan de forma directa significativamente y con una elevada magnitud. Señalar que la correlación más elevada es entre Aptitud Musical y Fluidez Lectora Silenciosa en L1, mientras que la más baja es entre Conciencia Fonológica y Fluidez Lectora Silenciosa en L1.

Tabla 1: Correlaciones Rho (ρ) de Spearman					
		FLS L1	FLS L2	AM	CF
L1 Fluidez Lectora Silenciosa		1	.750*	.761*	.645*
L2 Fluidez Lectora Silenciosa			1	.758*	.668*
Aptitud Musical				1	.671*
Conciencia Fonológica					1
* la correlación es significativas $p < .01$					

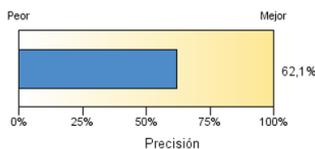
Con el modelo de regresión lineal, se observó que el modelo explicaba el 62.1% de la varianza (R^2 ajustado) de la variable dependiente (FLS L2), resultando el modelo de ajuste significativo ($F = 95.447$, $p < .01$), al igual que los coeficientes del modelo, tal y como se muestran en la tabla 2.

Tabla 2: Coeficientes del modelo de regresión lineal.

	Coeficientes no estandarizados		Coeficientes estandarizados	p	95,0% IC	
	B	SE	Beta		L. Inf	L. Sup
(Cte)	-74.851	23.014		.002	-120.442	-29.261
FLS L1	.700	.072	.693	.000	0.556	0.843
CF	4.787	2.311	.149	.041	0.21	9.364

a Variable dependiente: Fluidez Lectora Silenciosa en L2

Con lo que quedaría como resultado la siguiente ecuación de regresión, tanto en coeficientes no estandarizados como estandarizados:



Importancia del predictor
Destino: Silent Reading Fluency L2

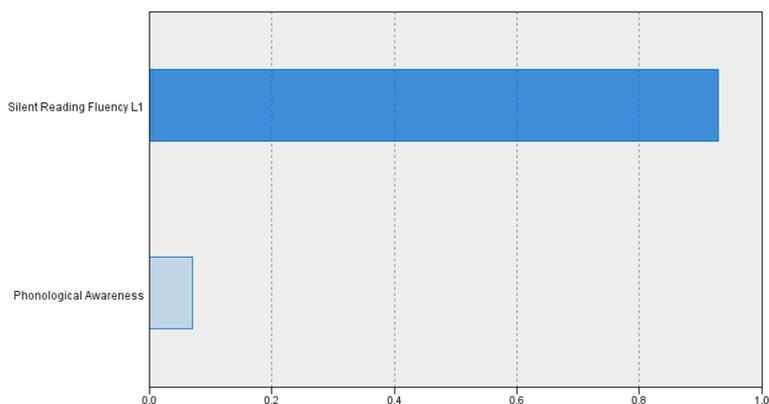


Figura 1. Porcentaje de varianza explicada e importancia de los predictores del análisis de regresión lineal

4. DISCUSIÓN

En nuestra primera pregunta de investigación nos planteamos si podría observarse una transferencia de habilidades en FLS de L1 a L2. El estudio correlacional ($r = .750$) indica que hay una correlación elevada entre la destreza lectora adquirida en lengua materna y el desarrollo de la destreza lectora en L2. Tal y como propone la *Hipótesis de la interdependencia lingüística* de Cummins (1979) o la *Hipótesis de la codificación lingüística* de Sparks (1995), lo normal es que un buen lector en L1 también lo sea en L2. Algo que puede observarse especialmente cuando se comparan dos lenguas afines y transparentes como es el caso del español y del italiano.

La segunda pregunta planteaba si la sensibilidad auditiva, medida como aptitud musical, se relacionaba con la destreza de procesamiento fonológico en la L2. Los resultados confirman una correlación de rango medio/alto ($r = .671$). Estudios previos, como los de Slevc & Miyake (2006) o Milovanov, Pietilä, Tervaniemi & Esquef (2010), ya habían demostrado las relaciones entre la aptitud musical y el aprendizaje de segundas lenguas en adultos, especialmente en lo relacionado con el nivel fonológico y la adquisición de habilidades orales en la L2. Nuestro estudio confirma la importancia de la aptitud musical en el desarrollo de la conciencia fonológica.

En cuanto a la tercera pregunta, los resultados del análisis correlacional indican que la sensibilidad auditiva o aptitud musical del lector mantiene una relación positiva no sólo con la conciencia fonológica en L2, sino también con fluidez lectora silenciosa en L1 ($r = .761$) y con la fluidez lectora silenciosa en L2 ($r = .758$). Estudios dedicados a la relación entre el sistema auditivo general y el procesamiento del habla, como el de Zhang & McBride-Chang (2010), llegaron a plantear la posibilidad de que fenómenos de percepción del habla de dominio lingüístico y el mecanismo de procesamiento auditivo general se relacionaran, siendo así que la sensibilidad auditiva general sería subyacente a habilidades como la percepción de los rasgos acústicos del habla, sea en el nivel segmental como en el suprasegmental. Las correlaciones entre la aptitud musical y la fluidez lectora silenciosa (reconocimiento visual de palabras), muestran que la sensibilidad auditiva puede ser un factor determinante en el desarrollo de la fluidez lectora silenciosa tanto en L1 como en L2, siendo esta la mayor contribución de este estudio.

Finalmente, la última pregunta de investigación trataba de determinar el peso de la variable visual (fluidez lectora silenciosa en L1) y de la variable oral

(conciencia fonológica en L2) a la hora de explicar la fluidez lectora silenciosa en L2. Los resultados del análisis de regresión lineal indican que es la pista visual la que orienta al lector italiano cuando lee en español, siendo así que el factor de la oralidad predice en menor grado los resultados obtenidos en el test de fluidez lectora silenciosa. Sin embargo, el resultado del análisis correlacional entre conciencia fonológica en L2 y la fluidez lectora silenciosa en L2 ($r=.668$) nos indica que la correlación entre ambas variables es fuerte, cuando la fluidez lectora silenciosa en L1 no es considerada. Así, a pesar de una mayor contribución de la destreza visual en el reconocimiento de palabras, la conciencia fonológica del lector parece contribuir a mejorar la competencia lectora del estudiante de español italiano.

4. CONCLUSIONES

El objetivo principal de este estudio ha sido explorar los diferentes factores que afectan al desarrollo de la fluidez lectora silenciosa, con especial énfasis en la aptitud musical y la conciencia fonológica de estudiantes italianos aprendiendo español en la universidad de Macerata.

Hasta donde llega nuestra revisión de la literatura existente, la fluidez lectora silenciosa ha sido estudiada en lectores italianos adultos con dislexia (Gagliano *et al.*, 2015) en estudiantes jóvenes de secundaria y en estudiantes universitarios, en contextos de monolingüismo (Ciuffo *et al.*, 2017). Asimismo, se han abordado estudios sobre aptitud musical y aprendizaje en segundas lenguas en estudiantes italianos adolescentes (Picciotti *et al.*, 2018) y sobre problemas de discriminación de patrones rítmicos en lectores adultos con dislexia (Cancer & Antonietti, 2018), pero no se había abordado la relación de la aptitud musical con la fluidez lectora silenciosa en estudiantes italianos adultos aprendiendo español en la universidad.

Los resultados obtenidos nos permiten concluir que el desarrollo de la fluidez lectora silenciosa en el ámbito del aprendizaje del español como lengua extranjera es engañosamente visual. La similitud de los sistemas ortográficos de lenguas estudiadas (italiano y español), guía al lector en su proceso de lectura. Sin embargo, la sensibilidad auditiva del lector no nativo, su aptitud musical en el caso del presente estudio, demuestra ser un factor determinante en el desarrollo de la fluidez lectora. En futuros estudios, sería interesante conocer el modo en que se comportan estas variables en lenguas con menor grado de afinidad.

AGRADECIMIENTOS

El autor agradece a la Universidad de Innsbruck por la ayuda en la recolección de datos para la aptitud musical, al concederle permiso para suministrar el test Mini-Proms. Y a la Dra. Marisa Martínez Pésico (Universidad de Macerata, Italia) por su colaboración, al facilitarle la recogida de datos en su universidad.

FINANCIACIÓN

Este estudio ha sido financiado por el Ministerio de Economía y Competitividad de España a través del proyecto I+D+i «Aptitud musical, fluidez lectora y percepción intercultural de estudiantes universitarios europeos» Referencia: FFI2016-75452-R 1.

REFERENCIAS BIBLIOGRÁFICAS

- ALLINGTON, R. L. (2006). Fluency: Still waiting after all these years. In S. J. Samuels & A. E. Farstrup (Eds.), *What research has to say about fluency instruction*. Newark, DE: International Reading Association.
- BESSON, M., BARBAROUX, M., & DITTINGER, E. (2017). Music in the brain: Music and language processing. In R. Ashley & R. Timmers (Eds.), *The Routledge Companion to Music Cognition*. New York, NY, NY: Routledge / Taylor & Francis.
- CANCER, A. & ANTONIETTI, A. (2018). Rapid Automatized Naming, Verbal Working Memory, and Rhythm Discrimination as Predictors of Reading in Italian Undergraduate Students with and without Dyslexia. *Brain Sciences*. 8, 87
- CIUFFO, M., MYERS, J., INGRASSIA, M., MILANESE, A., VENUTI, M., ALQUINO, A., ... ANTONELLA GAGLIANO, B. (2017). How fast can we read in the mind? Developmental trajectories of silent reading fluency. *Reading and Writing*, 30(8), 1667–1686.
- CUMMINS, J. (1979). *Linguistic Interdependence and the Educational Development of Bilingual Children* Author (s): James Cummins Source : *Review of Educational Research* , Vol . 49 , No . 2 (Spring , 1979) , pp . 222-251 Published by : American Educational Research Associati. 49(2), 222–251.
- GAGLIANO, A., CIUFFO, M., INGRASSIA, M., GHIDONI, E., ANGELINI, D., BENEDETTO, L., ... STELLA, G. (2015). Silent reading fluency: Implications for the assessment of adults with developmental dyslexia. *Journal of Clinical and Experimental Neuropsychology*, 37(9), 972–980. <https://doi.org/10.1080/13803395.2015.1072498>
- GROEN, M. A., VEENENDAAL, N. J., & VERHOEVEN, L. (2019). The role of prosody in reading comprehension: evidence from poor comprehenders. *Journal of Research in Reading*, 42(1), 37–57.
- HAMMILL, D. D., WIEDERHOLT, J. L., & ALLEN, E. A. (2006). *Test of silent contextual reading fluency*. Austin, TX: PRO-ED.
- HIEBERT, E. & REUTZEL, D.R. (2014). *Revisiting silent reading: New directions for teachers and researchers*: Santa Cruz, CA : I. TextProject, Ed.

- LAW, L. N. C., & ZENTNER, M. (2012). *Assessing Musical Abilities Objectively: Construction and Validation of the Profile of Music Perception Skills*. 7(12).
- MILOVANOV, R., PIETILÄ, P., TERVANIEMI, M., & ESQUEF, P. A. A. (2010). Foreign language pronunciation skills and musical aptitude: A study of Finnish adults with higher education. *Learning and Individual Differences*, 20(1), 56–60.
- NICHD. (2000). *Report of the National Reading Panel*. Washington, DC: U.S. Government Printing Office.
- NATION, P. (2009). Reading Faster. *International Journal of English Studies*, 9(2), 131–144.
- PATEL, A. D. (2008). *Music, Language, and the Brain*. New York, NY: Oxford University Press.
- PERIN, D. (1983). Phonemic segmentation and spelling. *British Journal of Psychology*, 74(1), 129–144.
- PICCIOTTI *et al.* (2018). Correlation between musical aptitude and learning foreign languages: an epidemiological study in secondary school Italian students. *Acta Otorhinolaryngol Ital.* 38(1), 51–55.
- SLEVC, L. R., & Miyake, A. (2006). *Individual Differences in Second-Language Proficiency Does Musical Ability Matter? Psychological Science*, 17, 675–681.
- SPARKS, R. L. (1995). Examining the linguistic coding differences hypothesis to explain individual differences in foreign language learning. *Annals of Dyslexia*, 45, 187–214.
- ZENTNER, M., & STRAUSS, H. (2017). Assessing musical ability quickly and objectively: development and validation of the Short-PROMS and the Mini-PROMS. *Annals of the New York Academy of Sciences*.
- ZHANG, J., & McBRIDE-CHANG, C. (2010). Auditory sensitivity, speech perception, and reading development and impairment. *Educational Psychology Review*, 22(3), 323–338.

ASIGNACIÓN DEL GÉNERO EN ESPAÑOL L3: PERSPECTIVAS SOBRE LA INFLUENCIA DE LA ADQUISICIÓN Y EL USO EN LA ORGANIZACIÓN DEL LÉXICO MULTILINGÜE¹

Hugues Lacroix
Université de Montréal (Canadá)
Université Bordeaux-Montaigne
IKER UMR-5478

Anahí Alba de la Fuente
Université de Montréal (Canadá)

RESUMEN

Este estudio se centra en la concordancia de género dentro del sintagma determinante en las etapas iniciales de adquisición del español como lengua tercera (L3) por parte de niños bilingües francés/leuskera. Nuestros objetivos se articulan en torno a dos ejes: (1) investigar los efectos de variables generalmente relacionadas con la concordancia de género, como la morfofonología, el criterio analógico y el género por defecto, y (2) explorar la influencia interlingüística a través de factores como la tipología, el uso y el orden de adquisición. Presentamos los resultados de una prueba escrita de descripción de imágenes desarrollada en el marco de un estudio longitudinal.

Palabras clave: L3 español, género gramatical, influencia interlingüística, estrategias metalingüísticas.

ABSTRACT

The current study focuses on gender agreement within the determiner phrase in the first stages of acquisition of Spanish as a third language by Basque/French bilingual children. Our objective is twofold: (1) to investigate the effects of variables usually related to gender, with special attention to morphophonology, the use of a default gender and the analogical criterion, and (2) to explore crosslinguistic influence through the most commonly studied

¹Este trabajo se enmarca en la investigación de los proyectos de investigación siguientes: VA009P17 financiado la Junta de Castilla y León y FEDER; PGC2018-097693-B-I00 financiado por el Ministerio de Ciencia, Innovación y Universidades y FEDER; y Proyecto de doctorado 201873 financiado por el FRQSC (Fonds de recherche du Québec – Société et culture).

factors in the field of TLA, namely, typology, use and order of acquisition. We present the results of a written picture description task that is part of a larger, longitudinal study.

Keywords: L3 Spanish, Gramatical Gender, crosslinguistic influence, metalinguistic strategies.

1. INTRODUCCIÓN

En las dos últimas décadas, y partiendo de trabajos pioneros en los que se centraban en los aprendientes de L3 como tales y no como aprendientes de (otra) L2 (Cenoz, Hufeisen, & Jessner, 2001; De Angelis, 2007; entre otros), el estudio de la Adquisición de lenguas terceras (ALT) se ha ido estableciendo como un campo propio, diferenciado de la Adquisición de lenguas segundas (ALS). Una de estas diferencias es la multiplicidad de posibilidades de transferencia en ALT vs ALS, que constituye el foco del presente trabajo². Más concretamente, analizamos la concordancia de género en las etapas iniciales de adquisición del español como L3 por parte de niños bilingües francés/euskera. En la sección 2, presentamos los fundamentos teóricos de nuestro trabajo sobre las propuestas de ALT y el fenómeno del estudio. La sección 3 incluye las preguntas de investigación y la metodología empleada en nuestro estudio; en la sección 4, presentamos los resultados principales, seguidos de una discusión sobre esos mismos en la sección 5.

2. FUNDAMENTOS TEÓRICOS Y ESTUDIOS ANTERIORES

2.1 Adquisición de L3

Con una mejor delimitación entre los campos de ALS y de ALT, se han propuesto una serie de modelos e hipótesis propios para determinar las diferentes posibilidades en cuanto a la influencia de las lenguas previamente adquiridas en el proceso de adquisición de la L3. Según la hipótesis del *L2 Status Factor* (*L2SF*; Bardel & Falk, 2012), el orden y el contexto de adquisición

²Por restricciones de espacio, no podemos detenernos en la presentación de las diferencias entre ALS y ALT, por lo que remitimos al lector a De Angelis (2007) para una discusión más exhaustiva sobre el tema.

de las lenguas previas juega un papel importante, por lo que la L2 —la lengua adquirida en un contexto más formal y más similar al contexto de adquisición de la L3— sería la única fuente de influencia posible en el proceso de adquisición de la L3. En cambio, otros modelos, como el *Typological Primacy Model* (TPM; Rothman, 2015), el *Cumulative-Enhancement Model* (CEM; Flynn, Foley, & Vinnitskaya, 2004) y el *Linguistic Proximity Model* (LPM; Westergaard, Mitrofanova, Mykhaylyk, & Rodina, 2017) otorgan un papel predominante a la proximidad entre las lenguas, ya sea a nivel global (TPM) o a nivel estructural (CEM, LPM). En el caso del CEM, se considera que las lenguas previas solo pueden tener un impacto positivo o neutro, mientras que el LPM y el TPM no excluyen la posibilidad de influencia negativa. Volveremos sobre las predicciones de cada modelo a la hora de formular nuestras hipótesis con el trio de lenguas francés-euskera-español.

2.2 El género gramatical

El género gramatical es una propiedad léxica intrínseca de los sustantivos (Corbett, 1991). Mientras el euskera no expresa morfosintácticamente el género gramatical, el francés y el español comparten un sistema de género dual (masculino-femenino). Ahora bien, aunque se trata de una propiedad intrínseca del sustantivo, el rasgo de género es observable gracias a la concordancia con las palabras que lo acompañan (determinante, adjetivo), más que en el sustantivo propiamente dicho. Los ejemplos a continuación muestran, por una parte, la concordancia de género en español (1) y en francés (2) y, por otra parte, la ausencia de marcas de género en euskera (3).

- | | | |
|-----|-----------------------|----------------|
| (1) | a. una casa blanca | <i>Español</i> |
| | b. un guante blanco | |
| (2) | a. une maison blanche | <i>Francés</i> |
| | b. un gant blanc | |
| (3) | a. etxe zuri bat | <i>Euskera</i> |
| | b. eskularru zuri bat | |

Tanto en francés como en español, dentro de un sintagma determinante (SD), el determinante y el adjetivo deben concordar en género con el sustantivo que modifican, una regla de concordancia afianzada en torno a los tres años en los casos de adquisición típica de la L1. En cambio, en adquisición de L2, este fenómeno representa un verdadero desafío (Dewaele & Véronique, 2001; Franceschina, 2005; entre otros). A continuación, resumiremos algunos de los principales trabajos sobre género en L2.

2.2.1 La concordancia de género en L2

Los estudios anteriores sobre concordancia en L2, muestran que los aprendientes tienen más aciertos en la concordancia con los determinantes que con los adjetivos (Dewaele & Véronique, 2001; Montrul, Foote, & Perpiñán, 2008; entre otros). Estos resultados sobre L2 son compatibles con los datos de adquisición de L1 que muestran que la concordancia con los determinantes precede la concordancia con los adjetivos (Hernández Pina, 1984).

A su vez, el papel de la terminación de los sustantivos también ha sido muy estudiado en el campo de ALS, y los resultados apuntan hacia una mayor facilidad con las terminaciones consideradas como transparentes (-a para el femenino, -o para el masculino) que lleva a un mayor acierto en la concordancia (Franceschina, 2005; Montrul et al., 2008) y a tiempos de reacción más rápidos en el reconocimiento de sustantivos con terminación transparente (Foote, 2014).

2.2.2 El masculino por defecto y la congruencia interlingüística

Además de los patrones de concordancia y los efectos de la morfofonología, encontramos dos nociones recurrentes en los estudios de género en L2: el uso de un género por defecto y la aplicación del criterio analógico. Desde un punto de vista teórico, se considera que el género no marcado del español es el masculino (Harris, 1991). Desde el punto de vista de la adquisición (L1/L2/L3), varios estudios muestran una tendencia a sobreutilizar el masculino, sobre todo al inicio del proceso de adquisición (Foote, 2014; Montrul *et al.*, 2008; White, Valenzuela, Kozłowska-Macgregor, & Leung, 2004; entre otros).

Por su parte, el concepto de *criterio analógico* procede de los estudios sobre alternancia de código en los cuales se estudiaban SD mezclados (p. ej. *ella*

book) para ver qué tipo de mezcla era favorecida por los hablantes (Liceras, Fernández Fuertes, Perales, Pérez-Tattam, & Spradlin, 2008; entre otros). En el ejemplo anterior, la preferencia por *el book* se tomaba como muestra de aplicación del criterio analógico, ya que el determinante concuerda con el equivalente de *book* en español (“libro”). Los resultados de esos estudios son mixtos y varían según el perfil de los hablantes. Mientras que los hablantes bilingües dominantes de una lengua con género (por ejemplo, L1 español/L2 inglés) tienden a preferir los SD que respetan el criterio analógico, los dominantes de una lengua sin género (L1 inglés/L2 español) tienden a preferir los SD con un determinante masculino (Klassen & M. Liceras, 2017). Esta diferencia se podría atribuir al hecho de que las representaciones del rasgo de género se suelen considerar como más estables en la L1 de los hablantes (Franceschina, 2005).

En los estudios donde se mide el tiempo de reacción en bilingües que hablan dos lenguas con género (p. ej. italiano-español), se encontraron reacciones más rápidas cuando el género de un sustantivo era congruente de una lengua a otra (mismo género) que cuando era incongruente (Klassen, 2016; Paolieri, Padilla, Koreneva, Morales, & Macizo, 2019; entre otros). En bilingües que hablan una lengua sin género y otra con género (p. ej. euskera-español), se han observado diferencias en el procesamiento del género en base a la lengua dominante (Caffarra, Barber, Molinaro, & Carreiras, 2017). Los autores atribuyen la diferencia observada a la mayor prominencia de la morfología en una lengua aglutinante como el euskera. En todos los casos, la presencia o ausencia de género en la L1 o lengua dominante de los hablantes parece tener un efecto sobre la asignación y el procesamiento del género en la L2.

3. DISEÑO EXPERIMENTAL

Nuestras preguntas de investigación eran las siguientes:

(P1) ¿Cuáles son los factores que provocan más errores de género en el SD en español L3?

(P2) ¿Cuáles son los factores que condicionan la influencia interlingüística (a partir de la L1 o L2) en español L3?

(P3) ¿Cómo evolucionan los efectos iniciales de la influencia interlingüística a lo largo del proceso de adquisición de español L3?

Por extensión, también nos interesa explorar cómo los factores de la pregunta (P1) interactúan con la influencia de la pregunta (P2). En cuanto a la pregunta (P1), consideramos que los participantes deberían reproducir de manera global lo observado en los estudios de L2 presentados en las secciones 2.2.1 y 2.2.2, es decir: (1a) más errores en los adjetivos, (1b) con el género femenino, (1c) las terminaciones opacas y (1d) el género incongruente. En lo relativo a la influencia interlingüística, (2a) si la proximidad de las lenguas (sea estructural o tipológica de manera más global) es el factor más importante, todos los participantes deberían mostrar la misma influencia del francés, sea esa su L1 o L2. En cambio, (2b) si el estatus de la L2 es el factor más relevante, los participantes deberían mostrar una influencia del francés solo si se trata de su L2. En cuanto a la pregunta (P3), si bien esperamos que los resultados mejoren a lo largo del proceso de instrucción, la evolución específica de cada factor resulta más difícil de predecir, dada la escasez de estudios anteriores de L3 sobre nuestro objeto de estudio.

3.1 Participantes

Nuestro estudio se llevó a cabo en el País Vasco norte (suroeste de Francia) con alumnos de sexto grado del colegio, cuya lengua de instrucción es el euskera, con 4-5 horas de cursos de francés a la semana. Al inicio de nuestro estudio, los participantes estaban en las etapas iniciales de su proceso de adquisición de español como L3 (dos meses de enseñanza, 4 horas a la semana). Nuestro estudio incluye también un grupo de control de alumnos de Valladolid (España) de edad comparable, cuya lengua materna es el español (grupo ES).

Los padres de los alumnos completaron un formulario de consentimiento y un cuestionario de antecedentes lingüísticos junto con sus hijos y a partir de la información de los cuestionarios, los participantes eran asignados a uno de los siguientes grupos, según la(s) lengua(s) hablada(s) en casa: francés y euskera (FR/EU), euskera (EU) o francés exclusivamente (FR), como muestra la Tabla 1. En los tres grupos experimentales, la edad de adquisición de la L2 o de la 2L1 variaba de un grupo a otro de manera significativa ($M=2.1$, $p < .001$), pero era particularmente alta en el grupo EU, lo que nos indica que esos participantes habían adquirido su L2 en un contexto más formal, en la escuela (o en el jardín de infancia, o al inicio de la escuela primaria).

	ES CONTROL (n=21)	EU (n=14)	FR (n=15)	FR/EU (n=23)
Edad	11.4 (0.2)	11.3 (0.3)	11.4 (0.3)	11.3 (0.3)
Edad de adquisición de la L2/2L1	—	4.1 (1.8)	2.5 (1.8)	0.6 (1.3)
Tiempo de instrucción en español (h)	—	28 (0)	28 (0)	28 (0)
Resultado en la prueba de nivel en español (%)	98 (2.2)	10.6 (9)	10.4 (6.6)	8.1 (4.5)

Tabla 1. Perfil de los distintos grupos de participantes según la(s) lengua(s) hablada(s) en casa (valores promedios, acompañados de la desviación estándar entre paréntesis).

3.2 Materiales

Todos los participantes completaron una tarea escrita de descripción de imágenes, en la cual tenían que escribir una respuesta formada por un determinante, un sustantivo y un adjetivo de color para cada imagen que veían (por ejemplo, «una» + «casa» + «roja»). Todas las imágenes representaban un sustantivo concreto inanimado y el sustantivo asociado a la imagen aparecía siempre en la diapositiva, para evitar respuestas en blanco. La tarea incluía 6 ítems de práctica, 48 distractores y 24 ítems experimentales distribuidos según los siguientes criterios: el género (masculino / femenino), la transparencia de la terminación (transparente / opaca) y la congruencia de género con el francés (congruente/incongruente). Todos los sustantivos seleccionados procedían de trabajos anteriores sobre el género gramatical (Foote, 2014; Montrul *et al.*, 2008), de manuales de ELE de nivel elemental o del inventario específico de los niveles A1-A2 del *Plan Curricular* (Instituto Cervantes, 2007). Además de los factores antes mencionados, consideramos los criterios de ilustrabilidad, de familiaridad y de concreción extraídos del corpus ESPal (Duchon, Perea, Sebastián-Gallés, Martí, & Carreiras, 2013) para establecer la lista final de ítems.

3.3 Protocolo

Los participantes completaron la tarea en un aula de clase, individualmente y en silencio. Los estímulos fueron proyectados en un PowerPoint para grupos de 15 a 28 participantes a la vez y cada participante tenía que escribir su respuesta en una hoja de respuestas individual. Después de la práctica con los seis primeros ítems, el administrador de la tarea (el primer autor del presente trabajo) dejaba los ítems de la tarea sucederse automáticamente a un ritmo de 25 segundos cada uno.

Repetimos el mismo protocolo con los mismos participantes después de 140 horas (6 meses más tarde, al final del primer año) de instrucción en español y después de 200 horas (en marzo del año siguiente).

4. RESULTADOS

Cada respuesta contenía potencialmente dos marcas de concordancia, una en el determinante, otra en el adjetivo. Por ejemplo, para la respuesta “una casa roja”, “una” y “roja” fueron codificadas como dos marcas independientes (0 en caso de error, o 1 en caso de acierto). Para comprobar nuestras hipótesis, llevamos a cabo una serie de regresiones logísticas a efectos mixtos usando el paquete *lme4* (Bates, Mächler, Bolker, & Walker, 2015) del programa *R* (R Core Team, 2018). Presentaremos primero los resultados de la primera recogida de datos, para después comparar la evolución de los factores a través de las tres recogidas.

4.1 Los primeros pasos en español L3, la primera recogida

Una primera regresión contaba con la variable grupo (ES, EU, FR y FR/EU) como único efecto fijo para comprobar que había una diferencia significativa entre el grupo de control y los demás grupos y para averiguar si existía una diferencia a nivel global entre los grupos experimentales. El grupo de control obtiene una tasa de concordancia de 99%, muy encima de los resultados de los grupos experimentales que obtienen un promedio de 55.9% ($p < .001$). El grupo EU obtiene los mejores resultados y la diferencia entre los grupos experimentales es significativa en las comparaciones FR – EU ($z = -2.3, p = .02$) et FR/EU – EU ($z = -4.3, p < .001$), pero solo marginalmente

significativa en el contraste FR/EU – FR ($z = -1.8, p = 0.07$). Esta primera aproximación a los datos nos invita a indagar más en las posibles diferencias de manera más detallada entre los grupos.

En cuanto a las hipótesis de la pregunta 1, la figura 1 muestra las tasas de concordancia en las diferentes condiciones para cada grupo experimental. Para comprobar la significatividad de los resultados que aparecen en esta Figura 1, aplicamos otro modelo de regresión logística a efectos mixtos usando los efectos fijos siguientes: grupo, género del sustantivo (masculino/femenino), dominio de concordancia (determinante/adjetivo), transparencia de la terminación (transparente/opaca), congruencia de género con el francés (congruente/incongruente) y las interacciones grupo-transparencia y grupo-congruencia. Comprobamos también las demás interacciones posibles (grupo-género y grupo-dominio), pero no eran significativas y las descartamos de nuestro modelo final.

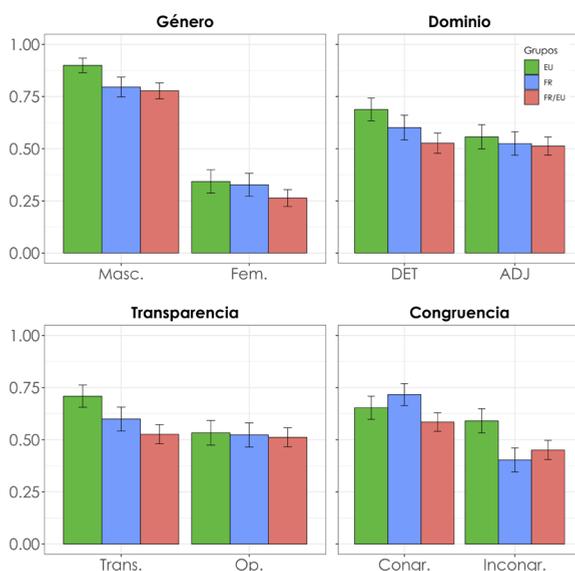


Figura 1. Tasa de concordancia según de los diferentes criterios de nuestras hipótesis.

En la primera gráfica de la Figura 1, se puede observar que el género afecta a los tres grupos de manera similar, ya que todos obtienen valores bastante altos para el masculino y relativamente bajos para el femenino, dando lugar a una diferencia de gran amplitud y muy significativa ($z = -8,749$, $p < .001$). En la gráfica sobre el dominio, los valores son más altos en la concordancia con el determinante que con el adjetivo. Esta diferencia es muy significativa a nivel global, pero de amplitud menor ($z = -3.319$, $p < .001$). La gráfica de la transparencia muestra resultados en torno al 50% en la condición opaca para todos los grupos y una mayor variabilidad en la condición transparente. Esta diferencia es significativa a nivel global ($z = -5.696$, $p < .001$) y también en la interacción con los grupos. En comparación con el nivel de referencia EU:Transparente, las diferencias con las condiciones FR:Opaca y FR/EU:Opaca son significativas ($z = 2.408$, $p = .016$; $z = 4.132$, $p < .001$; respectivamente). Finalmente, en cuanto a la congruencia, se puede observar en la última gráfica de la figura 1 que todos los grupos obtienen mejores resultados en la condición congruente, pero dada la fuerte interacción entre ese factor y el factor grupo, la diferencia no es significativa a nivel global ($z = -1.159$, $p = .25$). Dicha interacción entre grupo y congruencia solo es significativa para el grupo FR ($z = -2.833$, $p = .005$), que muestra la mayor variación de una condición a otra, pasando del mejor promedio en la condición congruente (71.6%) al peor en la condición incongruente (40.3%).

De estos resultados destacan sobre todo las interacciones Grupo-Transparencia y Grupo-Congruencia.

4.2 La evolución

De manera global, todos los grupos mejoran de una recogida a otra, pasando de un promedio de 56 % de aciertos en sus etapas iniciales, a 70 % (tras 140 horas de instrucción en español) y 76 % (tras 200 horas). La figura 2 muestra la evolución de los diferentes factores por separado³. Para analizar más en detalle la evolución de los diferentes factores, aplicamos el mismo modelo de regresión expuesto en la sección 4.1 a los datos de cada nueva recogida.

³La gráfica sobre el dominio de concordancia no aparece en la Figura 2 por cuestiones de espacio y porque mostraba una evolución bastante lineal. A partir de la segunda recolección el efecto ya no es significativo ($z = -1.131$, $p = 0.26$).

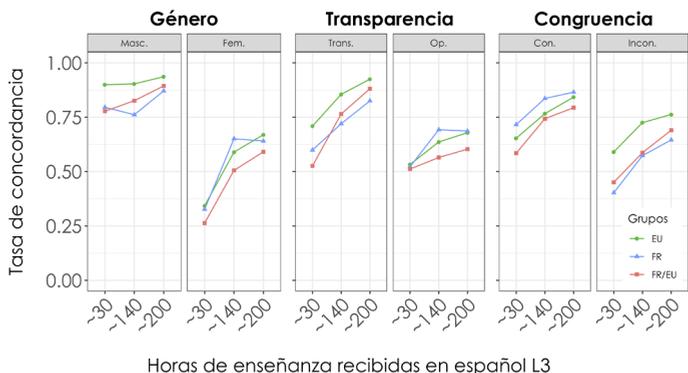


Figura 2. Evolución de la tasa de concordancia según de los diferentes criterios de nuestras hipótesis.

De manera general, lo observado en las etapas iniciales se mantiene en la segunda recogida de datos, pero con amplitud menor.

La primera gráfica nos revela una evolución no lineal del uso del masculino en la condición femenina. Si bien en la primera recolección los participantes sobreutilizaban el masculino con los sustantivos femeninos (69,8 % de masculino de media en los grupos experimentales), a partir de la segunda recogida todos los grupos empiezan a usar el femenino en los contextos adecuados con más frecuencia (por encima del 50 %).

En cuanto a la transparencia, el grupo EU se mantiene por encima de los demás en la condición transparente y esa interacción grupo-transparencia sigue siendo significativa en la segunda recogida ($z = 4.778$, $p < 0.001$). En la tercera recogida, la amplitud y la significatividad del efecto disminuyen ($z = 3.016$, $p = 0.003$). Con respecto a la congruencia, el grupo FR sigue teniendo mayor dificultad en la condición incongruente y esa interacción entre congruencia y grupo se mantiene de manera más duradera, tanto en cuanto al tamaño del efecto como a su significatividad ($z = -2.979$, $p = 0.003$ en la segunda recogida; $z = -2.128$ $p = 0.03$ en la tercera).

5. DISCUSIÓN

Como hemos visto, de manera general, nuestras cuatro sub-hipótesis (1a-d) se confirman en las etapas iniciales de adquisición del español en nuestros participantes. De manera global, se observan tasas de errores más altas con los sustantivos femeninos, opacos e incongruentes (con el género del francés) y los errores son más frecuentes con los adjetivos que con los determinantes. Estos resultados confirman lo que se ha observado en estudios anteriores de ALS (ver secciones 2.2.1 y 2.2.2). Además, los datos longitudinales nos permitieron ver que, si bien los participantes tendían a sobreutilizar el género masculino en sus etapas iniciales, esta estrategia perdía prominencia a medida que iban progresando en su proceso de adquisición. Esta progresión nos lleva a afirmar que, si bien nuestros datos reflejan un uso del masculino por defecto, se trata de una estrategia inicial, en ausencia de otro tipo de estrategias (como el uso de indicios morfofonológicos).

Lo más novedoso de nuestros resultados se encuentra en la compleja influencia interlingüística que nos permite observar un estudio de ALT. Hemos visto que según el orden/contexto de adquisición y el uso de las dos primeras lenguas, nuestros participantes mostraban efectos distintos en lo que se refiere a la transparencia de la terminación y la congruencia de género con el francés. Concretamente, los participantes del grupo EU, que adquirieron el francés más tarde y en un contexto más formal, eran más sensibles a la terminación de los sustantivos que los otros grupos, algo que se observa, además, desde la primera etapa. Proponemos dos explicaciones posibles a esa mayor sensibilidad a los indicios morfofonológicos del español por parte del grupo EU. Por una parte, podría representar una estrategia que han desarrollado en un contexto de adquisición más formal del francés como L2 (en línea con el *L2SF*; Bardel & Falk, 2012). Recordemos que el francés comparte los mismos géneros que el español y en la que la forma (terminación) también tiene cierto peso a la hora de asignar un género a un sustantivo. Por otra parte, esta estrategia podría proceder del uso predominante del euskera en su vida cotidiana, una lengua sin género gramatical, pero en la que la morfología juega un papel central. Por lo tanto, esos participantes del grupo EU podrían tener una mayor sensibilidad a la morfología de manera general y una representación menos estable del rasgo de género (Caffarra et al., 2017). Nuestro diseño experimental no nos permite desambiguar entre las dos posibles explicaciones, pero en todos los casos podemos afirmar que no se trata de una influencia interlingüística del

francés o del euskera per se, sino más bien de una estrategia metalingüística fruto de la experiencia lingüística de los hablantes.

En cuanto al grupo FR y su mayor sensibilidad a la congruencia de género entre el francés y el español, se podría atribuir a la proximidad tipológica entre esas dos lenguas, como sugiere el *TPM* (Rothman, 2015). Sin embargo, este modelo no predice ninguna diferencia en función del orden de adquisición, pero en nuestros resultados sí se observan diferencias en ese sentido. Así pues, sugerimos que la mayor sensibilidad a la congruencia del grupo FR ocurre porque sus representaciones de género son más estables y más activas al contar ya con este rasgo en su L1 (Franceschina, 2005). Por lo tanto, esas representaciones son más difíciles de inhibir cuando se usa un sustantivo en español cuyo género contrasta con el de su equivalente en francés.

6. CONCLUSIONES

En línea con los trabajos anteriores sobre ALS, nuestros participantes trilingües también se mostraron sensibles a factores como la opacidad, la congruencia, el dominio de concordancia y el género por defecto a la hora de establecer relaciones de concordancia dentro del DP (P1). Nuestros datos muestran también que la lengua hablada en casa tuvo un impacto en los resultados de nuestros participantes (P2) y que los efectos iniciales observados tras 28 horas de instrucción se iban difuminando a medida que los participantes avanzaban en su proceso de aprendizaje (P3).

En cuanto a las futuras líneas de investigación, sería interesante aplicar el mismo protocolo con participantes que hablen, en lugar del euskera, otra lengua sin género y sin una morfología tan rica (por ejemplo, el inglés). Eso nos permitiría discernir mejor entre los conocimientos metalingüísticos procedentes de la adquisición de la L2 y el impacto de tener como L1 una lengua aglutinante.

REFERENCIAS BIBLIOGRÁFICAS

- BARDEL, C. y FALK, Y. (2012). The L2 status factor and the declarative/procedural distinction. En J. Cabrelli Amaro, S. Flynn, & J. Rothman (Eds.), *Third language acquisition in adulthood* (Vol. 46, pp. 61-78). <https://doi.org/10.1075/sibil.46.06bar>.
- BATES, D., MÄCHLER, M., BOLKER, B. y WALKER, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>.
- CAFFARRA, S., BARBER, H., MOLINARO, N. y CARREIRAS, M. (2017). When the end matters: Influence of gender cues during agreement computation in bilinguals. *Language, Cognition and Neuroscience*, 32(9), 1069-1085. <https://doi.org/10.1080/23273798.2017.1283426>.
- CENOS, J., HUFSEIN, B. y JESSNER, U. (2001). *Cross-linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives*. Clevedon, UK: Multilingual Matters.
- CORBETT, G. (1991). *Gender*. <https://doi.org/10.1017/CBO9781139166119>.
- DE ANGELIS, G. (2007). *Third or additional language acquisition*. Clevedon, UK: Multilingual Matters.
- DEWAELE, J.-M. y VÉRONIQUE, D. (2001). Gender assignment and gender agreement in advanced French interlanguage: A cross-sectional study. *Bilingualism: Language and Cognition*, 4(03), 275–297.
- DUCHON, A., PEREA, M., SEBASTIÁN-GALLÉS, N., MARTÍ, A. y CARREIRAS, M. (2013). EsPal: One-stop shopping for Spanish word properties. *Behavior research methods*, 45(4), 1246–1258. <https://doi.org/10.3758/s13428-013-0326-1>.
- FLYNN, S., FOLEY, C. y VINNITSKAYA, I. (2004). The Cumulative-Enhancement Model for Language Acquisition: Comparing Adults' and Children's Patterns of Development in First, Second and Third Language Acquisition of Relative Clauses. *International Journal of Multilingualism*, 1(1), 3-16. <https://doi.org/10.1080/14790710408668175>.
- FOOTE, R. (2014). Age of Acquisition and Sensitivity to Gender in Spanish Word Recognition. *Language Acquisition*, 21(4), 365-385. <https://doi.org/10.1080/10489223.2014.892948>.
- FRANCESCHINA, F. (2005). *Fossilized Second Language Grammars: The acquisition of grammatical gender*. <https://doi.org/10.1075/lald.38>.

- HARRIS, J. W. (1991). The Exponence of Gender in Spanish. *Linguistic Inquiry*, 22(1), 27-62.
- HERNÁNDEZ PINA, F. (1984). *Teorías psicosociolingüísticas y su aplicación a la adquisición del español como lengua materna*. Madrid: Siglo XXI.
- INSTITUTO CERVANTES (2007). *Plan curricular del Instituto Cervantes: Niveles de referencia para el español*. Madrid: Edelsa.
- KLASSEN, R. (2016). The representation of asymmetric grammatical gender systems in the bilingual mental lexicon. *Probus*, 28(1). <https://doi.org/10.1515/probus-2016-0002>.
- KLASSEN, R. y M. LICERAS, J. (2017). The representation of gender in the mind of Spanish-English bilinguals: Insights from code-switched Adjectival Predicates. *Borealis – An International Journal of Hispanic Linguistics*, 6(1), 77. <https://doi.org/10.7557/1.6.1.4100>.
- LICERAS, J. M., FERNÁNDEZ FUERTES, R., PERALES, S., PÉREZ-TATTAM, R. y SPRADLIN, K. T. (2008). Gender and gender agreement in bilingual native and non-native grammars: A view from child and adult functional-lexical mixings. *Lingua*, 118(6), 827-851. <https://doi.org/10.1016/j.lingua.2007.05.006>.
- MONTRUL, S., FOOTE, R. y PERPIÑÁN, S. (2008). Gender agreement in adult second language learners and Spanish heritage speakers: The effects of age and context of acquisition. *Language Learning*, 58(3), 503–553. <https://doi.org/10.1111/j.1467-9922.2008.00449>.
- PAOLIERI, D., PADILLA, F., KORENEVA, O., MORALES, L. y MACIZO, P. (2019). Gender congruency effects in Russian–Spanish and Italian–Spanish bilinguals: The role of language proximity and concreteness of words. *Bilingualism: Language and Cognition*, 22(1), 112-129. <https://doi.org/10.1017/S1366728917000591>.
- R CORE TEAM (2018). *R: A Language and Environment for Statistical Computing*. URL: <https://www.R-project.org/>.
- ROTHMAN, J. (2015). Linguistic and cognitive motivations for the Typological Primacy Model (TPM) of third language (L3) transfer: Timing of acquisition and proficiency considered. *Bilingualism: Language and Cognition*, 18(2), 179–190. <https://doi.org/10.1017/S136672891300059X>.

- WESTERGAARD, M., MITROFANOVA, N., MYKHAYLYK, R. y RODINA, Y. (2017). Crosslinguistic influence in the acquisition of a third language: The Linguistic Proximity Model. *International Journal of Bilingualism*, 21(6), 666-682. <https://doi.org/10.1177/1367006916648859>.
- WHITE, L., VALENZUELA, E., KOZLOWSKA-MACGREGOR, M. y LEUNG, Y.-K. I. (2004). Gender and number agreement in nonnative Spanish. *Applied Psycholinguistics*, 25(1), 105-133. <https://doi.org/10.1017/S0142716404001067>.

CUESTIONES METODOLÓGICAS EN *GENDER STUDIES*: ¿QUÉ VENTAJAS CONLLEVAN LOS MÉTODOS MIXTOS PARA LA INVESTIGACIÓN DEL PAPEL DEL GÉNERO DURANTE LA PUBERTAD EN EL APRENDIZAJE DE IDIOMAS EXTRANJEROS?

Patricia Uhl

Universidad de Erlangen-Núremberg

RESUMEN

Por un lado, estudios sistemáticos con base cuantitativa amplia realizados en Alemania sobre el rendimiento lingüístico en lenguas extranjeras (p. ej. Klieme, 2008; Köller et al., 2010) han mostrado que, en general, las chicas superan académicamente a los chicos. Por otro lado, Schmenk (2002, 2009) demuestra que ni chicas ni chicos pueden ser considerados como “mejores aprendientes”. Ante esta situación, una investigación que examine por qué estas diferencias según el género se han encontrado en algunos pero no en todos los estudios sobre el aprendizaje de lenguas extranjeras, es, a todas luces, necesaria. En el presente artículo, la interseccionalidad del género y sus consecuencias para estudios en estos campos serán discutidas a fin de desarrollar un diseño de investigación innovador como base para un estudio que se está llevando a cabo en Alemania en el grado 9 de las clases de francés y español.

Palabras clave: género, aprendizaje de idiomas extranjeros, interseccionalidad, métodos mixtos

ABSTRACT

On the one hand, systematic performance measurements conducted in the context of extensive studies of foreign language learning carried out in Germany (e.g. Klieme, 2008; Köller et al., 2010) have shown that boys are generally outperformed by girls. On the other hand, Schmenk (2002, 2009) provides evidence showing that neither boys nor girls can be considered as ‘better learners’. Therefore research examining why gender differences have been found in some, but not in all studies of foreign language learning is of great importance. In the present article, the intersectionality of gender and its consequences for studies in these fields will be discussed in order to develop an innovative research design which is the basis for a study which is carried out in Germany in grade 9 in the context of French and Spanish classes.

Keywords: gender, language learning, intersectionality, mixed methods

1. PREJUICIOS Y MITOS ACTUALES SOBRE EL PAPEL DEL GÉNERO EN EL APRENDIZAJE DE IDIOMAS EXTRANJEROS

El prejuicio de que las asignaturas de lenguas extranjeras modernas, y particularmente las lenguas románicas, son “asignaturas para chicas” parece estar bastante extendido en la sociedad alemana. Como razones de ese fenómeno se presentan muy a menudo datos estadísticos: el número de alumnas de francés y de español y el de las profesoras de estas dos lenguas en Alemania es significativamente más alto que las respectivas cifras del alumnado y el profesorado masculinos (cf. Grein, 2012: 173). Además, los resultados de los rendimientos lingüísticos de las alumnas son, por lo general, mejores en los grandes estudios educativos. Precisamente, en el caso del francés se puede citar por ejemplo el estudio *Ländervergleich* (Köller *et al.*, 2010); por su parte, en el caso del inglés se vio una importante ventaja de las alumnas del noveno grado en el estudio DESI (cf. Hartig/Jude, 2008). Sin embargo, en cuanto al español como lengua extranjera, aún escasean los estudios sobre posibles diferencias según el género en Alemania.

Tomando como referencia el francés, se ha de añadir a lo anterior que muchos alumnos consideran que esta lengua “suena de manera femenina” (cf. Leupold, 2007: 15). Por último y mirando más allá de las instituciones educativas, el prejuicio de que el uso de lenguas extranjeras en el mundo del trabajo sea más bien un campo femenino parece estar bastante extendido; de hecho, no solo en Alemania, sino también en España o Inglaterra (cf. Braun/Schwemer, 2013: 2; López Rúa, 2006: 100; Williams *et al.*, 2002).

En resumen, se puede observar que todavía existen no pocos estereotipos en cuanto al género, su papel en el aprendizaje y el uso de lenguas extranjeras; de igual modo, se pone de manifiesto, tanto para el francés como para el español, la falta de estudios específicos sobre la relación entre el factor del género y el aprendizaje de idiomas. Así las cosas, los futuros estudios deberían cuestionar estos estereotipos y dar informaciones actuales sobre posibles diferencias según el género en los diferentes campos del aprendizaje y en diferentes competencias y destrezas lingüísticas. Para aclarar más el objeto y las necesidades de investigación, en lo que sigue se tratarán las características específicas de la categoría género y la metodología de algunos estudios ejemplares ya existentes (cf. Uhl, 2019).

2. GÉNERO COMO CATEGORÍA INTERDEPENDIENTE E INTERSECCIONAL

La discusión sobre la categoría de género cuenta con una larga trayectoria en la investigación sociológica y educativa y, por tanto, hay que escoger una definición entre las muchas posibles que existen del término. Por un lado, se habla de “sexo” (ingl. *sex*) para denominar la categoría biológica y la diferenciación entre masculino, femenino y diverso. Por otro lado, se ha establecido la categoría social, cultural y psicológica de “género” (ingl. *gender*) para designar el resultado de su construcción sociocultural y psicológica. Para referirse al mismo proceso de construcción, se emplea también el término *Doing gender*, el cual demuestra adecuadamente el carácter dinámico, activo e inestable de tal constructo, lo que supone como consecuencia lógica que los procesos de construcción de género difieran de una época a otra, de un espacio cultural a otro y, también, de una edad del hombre a otra. De esta manera, la diferenciación entre “masculino”, “femenino” y “diverso”, la cual antes parecía tan fácil, se convierte en un constructo de gran complejidad que requiere una metodología empírica muy elaborada para ser identificada y descrita. En cuanto al ámbito de la educación, la siguiente definición del término de género parece ser adecuada porque muestra la complejidad y la interseccionalidad con otros factores y, a su vez, la falta de delimitaciones claras entre *sex* y *gender*:

“The term ‘gender issues’ refers ‘to the social and educational aspects of the pupils’ being male or female; that is to say, to the real and perceived biological, genetic, cultural, educational, and indeed lifelong implications of the gender” (Portwood, 2000: 34, cit. según García Gómez, 2005: 80).

El término “interseccionalidad”, que se creó en el contexto de la investigación sobre desigualdades sociales (cf. Crenshaw, 1989: 139), es definido por Brah y Phoenix (2004: 76) como los efectos complejos, irreducibles, variados y variables que surgen cuando múltiples ejes de diferenciación –por ejemplo, económicos, políticos, culturales, psíquicos, subjetivos y experimentales– se intersecan en contextos históricamente específicos. La interseccionalidad hace hincapié en el hecho de que las diferentes dimensiones de la vida social no pueden ser divididas en categorías discretas.

En resumen, hay que analizar los estudios ya existentes teniendo presente la pregunta de si investigan diferencias según el sexo o según el género, si se conforman con modelos unilaterales en los cuales tratan la categoría del género

como un factor aislado o si aplican un enfoque integrativo e interseccional a su diseño de investigación –tal como propone por ejemplo López Rúa (2006). Ella concibe sobre la base de un estudio-meta el *Gender network of interactive variables in FLL* en el que establece todos los factores que interactúan con el género en el caso del aprendizaje de lenguas extranjeras. Así, determina factores neurológicos y cognitivos como, entre otros, la inteligencia verbal y la aptitud para aprender lenguas extranjeras, junto a factores afectivos como, por ejemplo, actitudes, motivaciones y factores de personalidad. Además, a estos factores internos añade otros externos, tales como las influencias del entorno (familia y *peers*) y las de tipo social y educativo. De manera específica, se plantea que el profesor puede fomentar o no las capacidades lingüísticas y comunicar por varios canales las expectativas a nivel de comportamiento y rendimiento lingüístico. Lo último conduce, en combinación con los demás factores, al comportamiento lingüístico esperado y, por ende, al correspondiente éxito en el aprendizaje y a la correspondiente competencia en la lengua extranjera, la cual influye de nuevo –entre otras cosas– sobre la selección de la profesión (cf. López Rúa, 2006). Este modelo se debería completar, en nuestra opinión, poniendo atención aún mayor a los ejes de interseccionalidad, es decir, al trasfondo lingüístico, muy a menudo medido al usar las entidades “monolingüismo” vs “plurilingüismo”, (cf. Steinlen, 2018) y al trasfondo socioeconómico, el cual se puede medir por ejemplo con ayuda de la categorización del ISCO-08.¹

3. ESTADO ACTUAL DE LA INVESTIGACIÓN

Un análisis del estado actual de la investigación muestra que existen numerosos estudios sobre el campo temático “género y aprendizaje de idiomas extranjeros” que destacan por su gran heterogeneidad: investigan sobre diferentes idiomas extranjeros en diferentes países, tienen en cuenta diferentes destrezas lingüísticas y diversas variables individuales del aprendiente (p.ej. la relación entre género y motivación para aprender una lengua extranjera), al tiempo que escogen y aplican diferentes métodos de investigación, lo que, como no podía ser de otra manera, lleva a diferentes resultados de investigación. Por ello, cierto número de estudios empíricos atestiguan una superioridad de las aprendientes femeninas en el aprendizaje de idiomas (cf. p.ej. Klieme, 2008;

¹En cuánto al género como triada de categorías en el ámbito de la investigación sobre interseccionalidad, véase también Bührmann (2009: 35-36).

Klieme *et al.*, 2010; Köller *et al.*, 2010), mientras que otros estudios presentan como resultado la superioridad de los aprendientes masculinos (cf. p. ej. Bügel/Buunk, 1996; Edelenbos/Vinje, 2000; Lynn *et al.*, 2005). Asimismo, no faltan estudios que no detectan ninguna diferencia según el género en las diversas competencias lingüísticas (cf. p.ej. Hyde/McKinley, 1997; Jiménez/Terrazas, 2005-2008). Por último, hay estudios que demuestran mayores diferencias dentro del grupo femenino y dentro del grupo masculino que entre estos dos grupos (cf. p.ej. Moss, 2007; Schmenk, 2002).

En suma, se puede decir que no hay un consenso actual sobre la influencia del género en el aprendizaje de lenguas extranjeras ni en el rendimiento, lo que se puede explicar por el hecho de tratarse de un factor indirecto e interseccional. Por consiguiente, un mero análisis de los resultados aporta poco a los intereses existentes de comprensión de la relación entre el factor género y el aprendizaje de idiomas. Esta situación obliga a analizar la metodología de los estudios anteriores para aproximarse más al objeto de investigación; por ello, examinamos a continuación tres estudios concretos. Como el enfoque del análisis de estos tres estudios está en su índole metodológico, se permite abarcar también estudios que investigan en las competencias del inglés y que provienen de diferentes países, aunque el estudio descrito en este artículo sea concebido para investigar en un contexto muy concreto: el efecto según el género en francés y español en el grado nueve en el sur de Alemania (cf. Uhl, 2019).

El primer estudio de interés metodológico, el estudio cuantitativo DESI (2003/2004) abarca el rendimiento lingüístico en inglés de aproximadamente 11.000 alumnos y alumnas del grado 9 en Alemania, sin concentrarse exclusivamente en el aspecto del género (cf. Nold/Rossa, 2007: 181). Por eso sólo compara los rendimientos lingüísticos de los aprendientes masculinos y femeninos sin ir más allá, es decir, sin comparar las interacciones entre el género y los demás factores que detecta López Rúa (2006). Aun así, es un estudio importante para el tema del género porque distingue muchas destrezas lingüísticas específicas: la comprensión lectora, la comprensión auditiva, la sociopragmática, la gramática, la reconstrucción de textos, la destreza de escribir y la destreza de hablar (y dentro de ésta la sintaxis, el vocabulario, la fluidez y la pronunciación). En la lengua extranjera inglesa, los resultados en función del género difieren bastante según la competencia parcial: en el estudio DESI, las chicas llevan, en general, una gran ventaja en su rendimiento en inglés; si bien, primero, su grado difiere según la competencia parcial (cf. Hartig/Jude, 2008:

204) y, segundo, en la destreza de hablar y especialmente en las competencias parciales de pronunciación y fluidez son los aprendientes masculinos los que superan a las aprendientes femeninas. Como se trata de un estudio longitudinal, se revela que los chicos desarrollan sus competencias lingüísticas sobre todo a través del progreso en la producción oral y precisamente a partir de esto elaboran Meißner *et al.* (2008: 158) la hipótesis de que los chicos van remontando a través de su progreso en la producción oral y llegan, así, a alcanzar un poco más tarde el nivel de las chicas en las demás competencias. Por eso suponen que el desarrollo de los chicos muchas veces se lleva a cabo con algún desfase de tiempo, pero que, al final de la pubertad, chicas y chicos llegan al mismo nivel de competencia lingüística.

Un estudio más orientado al tema del género es el estudio transversal de Fuchs (2013). Tiene un enfoque completamente diferente e investiga sobre diferencias según el género en cuanto a factores afectivos, especialmente el autoconcepto, el interés y la motivación en los grados 8 y 9 en la asignatura de inglés en Alemania. Como abarca los dos grados, se pueden hacer algunas conjeturas sobre los progresos afectivos de los dos géneros partiendo de las indicaciones en los cuestionarios. Sin embargo, como no hace constar el rendimiento en inglés, falta el establecimiento de una relación entre el lado afectivo y el rendimiento lingüístico. Además, en su conclusión, Fuchs (2013) propone la aplicación de métodos mixtos que combinan enfoques cuantitativos y cualitativos para conseguir un resultado más completo, sobre todo en lo que se refiere a la búsqueda de explicaciones de puntos en común y diferencias. Además, se puede añadir que para superar la mera comparación entre aprendientes masculinos y femeninos, una investigación cualitativa suplementaria podría servir para acercarse más a la construcción, tanto individual como colectiva y estereotipizada a nivel psicológico, cultural y social del género en el ámbito de las clases de lenguas extranjeras institucionalizadas a la edad de la pubertad. Se podría realizar de tal manera que después del estudio cuantitativo principal, se llevarían a cabo entrevistas individuales cualitativas que podrían referirse especialmente a características situacionales concretas de las clases o a características motivacionales de los aprendientes (cf. Fuchs, 2014: 196).

El último estudio que destaca notablemente por su metodología es el de Kissau (2006). Examina a aproximadamente 500 aprendientes de francés del grado 9 en Ontario, Canadá. Se trata de un estudio que justamente aplica *mixed methods*, a la vez que investiga los factores internos como la motivación y

los intereses, así como los factores externos, como por ejemplo la percepción de las expectativas por parte de los padres. Dado que Kissau (2006) se concentra sobre todo en factores afectivos y que su objetivo es la investigación en diferencias según el género, tampoco alcanza a reflejar el modelo de López Rúa (2006), debido a que no hace constar el rendimiento lingüístico. En nuestra opinión, registrar a la vez los rendimientos sería muy importante, puesto que, según Meißner *et al.* (2008) las diferentes variables afectivas pueden tener a menudo diferentes efectos en aprendientes masculinos y femeninos; incluso, es probablemente que tengan efectos completamente independientes del género.

4. REFLEXIÓN SOBRE LA APLICACIÓN DE MÉTODOS CUANTITATIVOS Y CUALITATIVOS

La pregunta de con qué métodos se puede investigar de manera interseccional todavía no ha encontrado una respuesta clara (cf. Knapp, 2008: 44). Esta carencia proviene sobre todo de la complejidad que conlleva ese concepto: por un lado, no se trata de una mera adición de categorías, sino de la búsqueda de conexiones y correlaciones entre ellas; por otro lado, hay que tener en cuenta que estas categorías se han ido desarrollando en función de su historia: piénsese en categorías como roles o estereotipos de género vinculados, a su vez, a sucesos actuales, y piénsese en categorías como la motivación. Por ello, no cabe sino suponer que las categorías interseccionales tienen diferentes efectos en diferentes contextos.

La triangulación de métodos en el sentido de una consecución de la aplicación de métodos cuantitativos y cualitativos en la cual dos series de datos se hacen constar con dos métodos diferentes dentro del mismo grupo (*Across-Method*) puede abarcar esta complejidad, dado que diferentes aspectos del mismo fenómeno pueden ser observados e interpretados desde dos perspectivas diferentes (cf. Elsner/Viebrock, 2015).

Los datos cuantitativos en forma de resultados de tests lingüísticos estandarizados, de tests cognitivos estandarizados y en forma de cuestionarios rellenos que incluyen la biografía lingüística, el trasfondo socioeconómico y varios constructos afectivos pueden ser evaluados por varios procedimientos estadísticos. Los tests que comprueban hipótesis de diferencia son apropiados para averiguar posibles disparidades entre aprendientes femeninas y aprendientes masculinos (p. ej. Test-T, ANOVA etc. cf. Larson-Hall, 2016). Los análisis de correlación, los análisis de regresión y los análisis discriminantes

lineales promueven más detalles sobre relaciones causa-efecto.

En cuanto a lo cualitativo, Flick (2016: 44) recomienda escoger el cinco por ciento de la muestra que lleva a cabo los tests cuantitativos para una entrevista cualitativa. La clave está en escoger a sujetos de estudio que son representativos para ciertos aspectos, por ejemplo su rendimiento lingüístico o sus estereotipos según el género (cf. Morse/Niehaus, 2009: 63-65). Una integración más profunda de las dos series de datos puede ser lograda al provocar una reflexión más profunda sobre partes específicas del cuestionario y al centrarse en los ámbitos en los cuales aparecen diferencias según el género (cf. Settineri, 2015: 17). Con la ayuda de los métodos cualitativos, se puede lograr discernir fenómenos sociales y psicológicos y considerar estructuras de sentido que no se revelan al analizar los datos estadísticos (cf. Uhl, 2019).

5. CONCLUSIÓN

El presente artículo ha tenido como objetivo presentar una serie de reflexiones metodológicas –especialmente pertinentes ante algunas aparentes contradicciones en la bibliografía existente sobre el tema– que podrían llevar a resolver el problema ante el cual nos expone el concepto de la interseccionalidad. La metodología propuesta puede ser una contribución a la investigación en cuanto a la categoría del género y a su papel en los procesos de aprendizaje de lenguas en un contexto institucionalizado: se ha visto que un diseño de metodología mixta que combina enfoques cuantitativos y cualitativos puede ser una solución para acercarse a la complejidad que supone una entidad interseccional. Como retos en este procedimiento cabe apuntar la exigencia de compaginar los dos enfoques de tal manera que lo cualitativo sea una profundización de lo cuantitativo. Además, a la hora de interpretar los resultados, el reto está en la búsqueda de una relación lógica entre los datos cuantitativos y los cualitativos. Aún así, el enfoque triangular tiene la ventaja de llegar a conclusiones que van más allá de los datos estadísticos y que, por tanto, permiten una investigación interseccional de la categoría del género.

REFERENCIAS BIBLIOGRÁFICAS

- BRAH, A. y PHOENIX, A. 2004. "Ain't I a woman? Revisiting intersectionality". *Journal of International Woman's Studies*, 5: 74-86.
- BRAUN, C. y SCHWEMER, K. 2013. "Jungen im Französischunterricht fördern". *Der fremdsprachliche Unterricht Französisch*, 122: 2-10.
- BÜHRMANN, A. D. 2009. "Intersectionality – ein Forschungsfeld auf dem Weg zum Paradigma? Tendenzen, Herausforderungen und Perspektiven der Forschung über Intersektionalität". *GENDER*, 2: 28-44.
- BÜGEL, K. y BUUNK, B. P. 1996. "Sex Differences in Foreign Language Text Comprehension: The Role of Interest and Prior Knowledge". *Modern Language Journal*, 80: 15-31.
- CRENSHAW, K. 1989. *Demarginalizing the intersection of race and sex: A black feminist critique of antidesecrimination doctrine, feminist theory and antiracist politics*. University of Chicago Legal Forum 1989, 139-167.
- EDELENBOS, P. y VINJÉ, M. P. 2000. "The assessment of a foreign language at the end of primary (elementary) education". *Language Testing*, 17/2: 144-162.
- ELSNER, D. y VIEBROCK, B. 2015. "Triangulation in der Fremdsprachenforschung". En: Elsner, D. y Viebrock, B. (eds.), *Triangulation in der Fremdsprachenforschung*. Frankfurt a. Main: Lang, 7-16.
- FLICK, U. 2016. *Qualitative Sozialforschung*. Reinbek: Rowohlt.
- FUCHS, S. 2013. *Geschlechterunterschiede bei motivationalen Faktoren im Kontext des Englischunterrichts: Eine empirische Studie zu Motivation, Selbstkonzept und Interesse im Fach Englisch in der Sekundarstufe I*. Frankfurt a. Main: Lang.
- FUCHS, S. 2014. "Geschlechterunterschiede bei motivationalen Faktoren im Kontext des Englischunterrichts der Sekundarstufe I: Ausgewählte Ergebnisse einer empirischen Studie zu Motivation und Interesse." *Zeitschrift für Fremdsprachenforschung (ZFF)*, 25/2: 175-205.
- GARCÍA GÓMEZ, A. 2006. "Sneaking gender into the EFL class". *Porta Linguarum*, 5: 77-89.

- GREIN, M. 2012. "Geschlechterforschung und Fachdidaktik Französisch (mit Hinweisen auf Fachdidaktik Spanisch)". En: M. Kampshoff y C. Wiepcke, (eds.), *Handbuch Geschlechterforschung und Fachdidaktik*. Wiesbaden: Springer, 169-183.
- HARTIG, J. y JUDE, N. 2008. "Sprachkompetenzen von Mädchen und Jungen." En: DESI-Konsortium (ed.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch: Ergebnisse der DESI-Studie*. Weinheim: Beltz, 202-207.
- HYDE, J. S. y MCKINLEY, N. M. 1997. "Gender Differences in Cognition: Results from Meta-Analysis." En: P. J. Caplan, M. Crawford, J. S. Hyde y J. T. E. Richardson, (eds.), *Gender Differences in Human Cognition*. New York: Oxford University Press, 30-51.
- JIMÉNEZ CATALÁN, R. M. y TERRAZAS GALLEGO, M. 2005-2008. "The Receptive Vocabulary of English Foreign Language Young Learners". *Journal of English Studies*, 5-6: 173-191.
- KISSAU, S. 2006. "Gender differences in second language motivation: An investigation of micro- and macro-level influences". *Canadian Journal of Applied Linguistics (CJAL) / Revue canadienne de linguistique appliquée (RCLA)*, 9/1: 73-96.
- KLIEME, E. 2008. "Systemmonitoring für den Sprachunterricht". En: E. Klieme (ed.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch: Ergebnisse der DESI-Studie*. Weinheim: Beltz, 1-10.
- KLIEME, E., ARTELT, C., HARTIG, J., JUDE, N., KÖLLER, O., PRENZEL, M., SCHNEIDER, W. y STANAT, P. (eds.) 2010. *PISA 2009: Bilanz nach einem Jahrzehnt*. Münster: Waxmann.
- KÖLLER, O., KNIGGE, M. y TESCH, B. (eds.) 2010. *Sprachliche Kompetenzen im Ländervergleich: Befunde des ersten Ländervergleichs zur Überprüfung der Bildungsstandards für den Mittleren Schulabschluss in den Fächern Deutsch, Englisch und Französisch*. Münster: Waxmann.
- KNAPP, G.-A. 2008. "Intersectionality – ein neues Paradigma der Geschlechterforschung?" En: R. Casale, B. Rendtorff (eds.), *Was kommt nach der Genderforschung? Zur Zukunft feministischer Theoriebildung*. Bielefeld: Transcript, 33-53.

- LARSON-HALL, J. 2016. *A Guide to Doing Statistics in Second Language Research Using SPSS and R*. London: Routledge.
- LEUPOLD, E. 2007. *Französischunterricht als Lernort für Sprache und Kultur: Prinzipien und Praxisbeispiele*. Seelze-Velber: Klett-Kallmeyer.
- LÓPEZ RÚA, P. 2006. "The Sex Variable in Foreign Language Learning: An Integrative Approach". *Porta Linguarum*, 6: 99-114.
- LYNN, R., FERGUSON, D. M. y HORWOOD, J. 2005. "Sex differences on the WISC-R in New Zealand". *Personality and Individual Differences*, 39: 103-114.
- MEISSNER, F.-J., BECKMANN, CH. y SCHRÖDER-SURA, A. 2008. *Mehrsprachigkeit fördern, Vielfalt und Reichtum Europas in der Schule nutzen (MES): Zwei deutsche Stichproben einer internationalen Studie in den Klassen 5 und 9 zu Sprachen und Fremdsprachenunterricht*. Tübingen: Narr.
- MORSE, J. M., NIEHAUS, L. 2009. *Mixed Method Design: Principles and Procedures*. Walnut Creek: Left Coast Press.
- MOSS, G. 2007. *Literacy and Gender. Researching Texts, Contexts, and Readers*. London/New York: Routledge.
- NOLD, G., ROSSA y H. 2007. "Hörverstehen". En: B. Beck y E. Klieme (eds.), *Sprachliche Kompetenzen: Konzepte und Messung: DESI-Studie*. Weinheim: Beltz, 178-196.
- SCHMENK, B. 2002. *Geschlechtsspezifisches Fremdsprachenlernen? Zur Konstruktion geschlechtstypischer Lerner- und Lernbilder in der Fremdsprachenforschung*. Tübingen: Stauffenburg.
- SCHMENK, B. 2009. "Vorsicht Stereotype! Gender und Fremdsprachenlernende". *Praxis Fremdsprachenunterricht*, 6: 4-7.
- SETTINIERI, J. 2015. "Forschst Du noch, oder triangulierst Du schon?" En: D. Elsner y B. Viebrock (eds.), *Triangulation in der Fremdsprachenforschung*. Frankfurt a. Main: Lang, 17-35.
- STEINLEN, A. 2018. "Grundschulen mit verschiedenen Englischprogrammen: Geschlecht und Mehrsprachigkeit auf dem Prüfstand". *Zeitschrift für Fremdsprachenforschung*, 29/1: 3-26.

- UHL, P. 2019. "Fremdsprachenforschung *extended*: Die Erforschung der intersektionalen Kategorien Geschlecht und *Gender* mit Hilfe interdisziplinärer Ansätze und Triangulation". En: E. Wilden y H. Rossa, 2019. *Fremdsprachenforschung als interdisziplinäres Projekt*. Frankfurt a. M.: Peter Lang. 207-224.
- WILLIAMS, M., BURDEN, R. y ANVERS, U. 2002. "French is the Language of Love and Stuff: student perceptions of issues related to motivation in learning a foreign language". *British Educational Research Journal*, 28: 503-328.

CONTEXTOS BILINGÜES Y SUS IMPLICACIONES DIDÁCTICAS: FUNDAMENTOS TEÓRICOS Y HERRAMIENTAS BÁSICAS PARA MEDIR LA SENSIBILIDAD INTERCULTURAL

Oscar O. Santos-Sopena

Universidad Politécnica de Madrid

RESUMEN

En contexto bilingüe, la competencia intercultural es primordial en el aprendizaje del inglés como lengua extranjera (EFL). Para ello, el avance de la adquisición de competencias comunicativas (L2/L3) en relación con el componente sociocultural, y partiendo de un enfoque multidisciplinar, se convierten en fundamentos teóricos a tener presente. En este estudio, enmarcado dentro de dos proyectos de innovación docente, se revisa el concepto de sensibilidad intercultural; además, de seleccionar y rediseñar dos herramientas para su cálculo, para posteriormente, describir las implicaciones didácticas más relevantes. Como reflexión final, vemos la importancia en la medición de la sensibilidad intercultural como ayuda y mejora de la comunicación oral/intercultural y la adquisición de la L2/L3 en contextos globales y plurilingües.

Palabras clave: competencia intercultural, competencia comunicativa, sensibilidad intercultural, aula multicultural, ESP/EFL

ABSTRACT

This project focuses on the magnitude of intercultural competence when teaching English as a Foreign Language (EFL) in bilingual settings. The elaboration of an interdisciplinary approach to English Second/Third Language Acquisition is crucial for understanding communicative and intercultural competences. The aim of this paper is to analyze, redesign, and develop the different two methods of measuring intercultural sensitivity in bilingual contexts and the didactic implications of these measurements. This educational innovation project aims to improve communication and intercultural perceptions and improve the verbal interaction of program participants in global and plurilingual contexts.

Keywords: Intercultural competence, communicative approach, intercultural sensitivity, multicultural classroom, ESP/EFL

1. INTRODUCCIÓN

El desarrollo de competencias profesionales y de transformación social en relación con el aprendizaje de lenguas, como las competencias interculturales y comunicativas en contextos bilingües, son esenciales en un aula global y de diversidad cultural como las que nos encontramos en la actualidad. En este estudio, enmarcado dentro de los proyectos ERASMUS+ 2016-1-ES01-KA201-025491 y PIE UPM 2017-18/2018-19, se pretende analizar las diferentes formas de medición de la sensibilidad intercultural en contextos bilingües en el marco de ambos programas y sus implicaciones didácticas. El objetivo general del estudio fue revisar el concepto de sensibilidad intercultural a partir de estos tres ejes de acción:

- (1) Establecer las herramientas para medir la sensibilidad intercultural en contextos bilingües tan amplios que van desde la educación primaria a la universitaria;
- (2) Desarrollar formas de medición de la evolución de la sensibilidad intercultural a partir de las herramientas propuestas y en contexto bilingüe; y
- (3) Proporcionar diversas implicaciones didácticas para una lengua global como es el inglés.

2. MARCO TEÓRICO: COMPETENCIA INTERCULTURAL Y COMUNICATIVA EN LA ENSEÑANZA DE LENGUAS

El *MCER* (2002) examina la competencia intercultural dentro de su relación con el conocimiento propio del componente sociocultural a partir de la conexión existente entre la sociedad/comunidad, la lengua y su(s) cultura(s) ya que “there is no language without culture” (Springer, 2005, p. 2). De ahí que la competencia intercultural esté íntimamente ligada a los contextos plurilingües propios de nuestro contexto de aprendizaje en el que los procesos comunicativos son imprescindibles dentro del proceso de adquisición:

El alumno no adquiere dos formas de actuar y de comunicarse distintas y que no se relacionan, sino que se convierte en plurilingüe y desarrolla una interculturalidad. Las competencias lingüística y cultural respecto a cada lengua se modifican mediante el conocimiento de la otra lengua y contribuyen a crear una conciencia, unas destrezas y unas capacidades interculturales. Permiten que el individuo desarrolle una personalidad más rica y compleja, y mejoran la

capacidad de aprendizaje posterior de lenguas y de apertura a nuevas experiencias culturales (MCER, 2002, p. 47).

Por ello, hay que tener presente de qué manera se genera la comunicación intercultural en la enseñanza de lenguas ya que esta parte de la negociación de significados socioculturales a partir de la comprensión de los sistemas sociales. Por lo tanto, “el conocimiento de los valores y las creencias compartidas por grupos sociales de otros países y regiones, como, por ejemplo, las creencias religiosas, los tabúes, la historia común asumida, etc., resulta esencial para la comunicación intercultural” (MCER, 2002, p. 11) el cual suscita una aproximación a la diversidad cultural propia de una lengua global como es el inglés:

The current state of the literature to be murky in terms of the clarity of the Intercultural Communicative Competence construct. Definitions of the construct vary considerably as to whether it is a trait, skill, or performance outcome. In addition, current measurements of ICC overly rely on self report methods, which have a number of flaws that result in less than optimal assessment (Griffith *et al.*, 2016, p. 1).

Siguiendo esta línea de pensamiento dentro de la competencia intercultural, y sus formas de medición, aparece la sensibilidad intercultural definida en términos generales por Chen y Starosta como la habilidad de carácter individual que ayuda a potenciar la apreciación y *por ende* la comprensión de los contrastes culturales, propiciando la comunicación intercultural (1997, p. 5). Si profundizamos brevemente en el concepto encontramos dos acercamientos teóricos que ayudan a ver la amplitud del término *per se*:

Por un lado, se entiende por sensibilidad intercultural “the way people construe cultural difference and [...] the varying kinds of experience that accompany these constructions” (Bennett, 1993, p. 24). Precisamente dicho encuentro, construcción y entendimiento cultural se conecta con la definición de la competencia intercultural que ofrece Deardorff (2006): “the ability to interact effectively and appropriately in intercultural situations, based on specific attitudes, intercultural knowledge, skills and reflection” (p. 247).

Por otro lado, Chen y Starosta (1997) señalaron que la competencia intercultural parte de un acercamiento a las emociones y actitudes frente al encuentro cultural (componente cognitivo, afectivo y conductual/pedagógico). Estas tres dimensiones deben desarrollarse de manera paralela para alcanzar

un grado de sensibilidad intercultural adecuado dentro de adquisición de la competencia intercultural (Gómez Barreto, 2011; Roldán Mejías, 2016).

Por lo tanto, si tenemos presente una definición amplia y transversal de la propia competencia intercultural es necesario examinar la competencia comunicativa. El avance de la adquisición de estas competencias en L1/L2/L3 desde un enfoque multidisciplinar y cotidiano es fundamental a nivel teórico (Schulz, 2001; Ellis, 2003; Lee & Van Patten, 2003; Yamada, 2009; Stonebanks, 2010). Como señala Littlewood:

El comunicador más eficaz en una lengua extranjera no es siempre la persona que mejor maneja las estructuras. Con frecuencia es la persona que tiene mayor habilidad en una situación comunicativa en la que se incluye tanto a sí misma como a su oyente [...] Igual que en el caso de la comprensión, el estudiante necesita entonces adquirir no sólo un repertorio de elementos lingüísticos sino también un repertorio de estrategias para usarlas en situaciones concretas (Littlewood, 1996, p. 4).

Es decir, a competencia comunicativa parte de la comunicación funcional y de interacción social (Richards y Rodgers, 1998). Por ejemplo, una conversación en la oficina, una discusión sobre un tema de actualidad, un diálogo improvisado en una tienda, un debate político, etc. son el germen de una comunicación real y práctica (Martín Peris, 2008). Por ello, el enfoque comunicativo se convierte en un ejercicio interdisciplinar que debe incluir una diversidad de temas y procesos de aprendizaje en el que se fomenta las destrezas de comunicación oral y escrita de la L2/L3. Por lo tanto, el objetivo es potenciar el carácter intrínseco del “hablante plurilingüe e intercultural” (Kramsch y Hua, 2016; Byram, Gribkova y Starkey, 2002). Así pues, en dicha experiencia se debe fomentar la comunicación intercultural: “an encounter with another person (or group of people) who is perceived to have different cultural affiliations from oneself” (Hubert *et al.*, 2014, p. 16).

En este sentido es necesario mencionar las 5C's de *The American Council on the Teaching of Foreign Languages* (ACTFL). Es decir, categorías como la comunicación, la cultura, la conexión, la comparación y la comunidad que se convierten en categorías esenciales por reforzar el vínculo entre ambas competencias, la intercultural y la comunicativa, y por desarrollar “the ability of the users of a language to understand and use language accurately and appropriately in context” (Buján Sánchez, 2016).



Figura 1. 5C's ACTFL [www.actfl.org]

3. REDISEÑO Y ADAPTACIÓN CUESTIONARIOS QUE MIDEN LA SENSIBILIDAD INTERCULTURAL

Los cuestionarios que se rediseñaron y adaptaron en ambos proyectos de innovación docente midieron la sensibilidad intercultural de sus participantes; siempre dependiendo de las particularidades de cada uno de los dos programas que vamos a describir brevemente en este apartado.

A la hora de rediseñar y adaptar estos dos cuestionarios se consideraron los dos acercamientos teóricos revisados en la sección anterior. Recordemos que, en primer lugar, el de Bennett (1993) parte de un acercamiento cognitivo de los procesos de adaptación cultural que ayuda a identificar y apreciar las diferencias culturales a partir de parámetros evolutivos y métricos. De ahí que aparezcan seis: 3 etnocéntricos (negación, defensa y minimización) y 3 etnorelativos (aceptación, adaptación e integración).

En segundo lugar, desde la perspectiva de Chen y Starosta (1997) nos ofrece un acercamiento a una dimensión afectiva mucho más amplia y que se acerca a la definición de la competencia intercultural propuesta en la sección anterior. Esta basada en el reconocimiento, tolerancia, respeto e integración de las diferencias culturales. Es decir, esta medición refleja una multiculturalidad interactiva dentro de la experiencia intercultural. Por ello, cuando medimos la sensibilidad intercultural partimos de dos enfoques: en primer lugar, el transcultural que fomenta la comunicación entre diferentes culturas dentro

del proceso de adquisición. Y, en segundo lugar, el multicultural que examina y considera las variables que afloran en las diferentes culturas en la región de estudio (cultura/s meta) y, por lo tanto, se resalta la diversidad sociocultural y lingüística presente (Spychała, 2014).

Por ello, estos dos instrumentos en fase de pilotaje están adaptados y compuestos de nueve ítems de escala tipo Likert (1=No es cierto, 2=A veces es cierto, 3=Bastantes veces es cierto, 4=Siempre es cierto) que miden la competencias intercultural y comunicativa; además de la sección con preguntas sociodemográficas que, en ambos casos, y por la naturaleza de los proyectos, se plantean de forma distinta como veremos a continuación (Figuras 2, 5, 6 y 7).

Por lo tanto, es primordial describir de forma breve los dos proyectos de innovación docente para entender la naturaleza del rediseño de los cuestionarios propuestos.

3.1 CLIL Teaching and Inclusive Education: “Erasmus+ Project KA2 2016-1-ES01-KA201-025491 (2016-2018)”

Este primer proyecto innovación docente dentro del marco de *Erasmus+ Project KA2* tiene como objetivos:

1. Medir la sensibilidad intercultural inicial de los participantes;
2. Establecer parámetros de medición adicional y aplicabilidad didáctica en contextos monolingües (UK), bilingües (Italia y España) y plurilingües (Letonia); y
3. Potenciar la competencia intercultural en este tipo de programas.

De ahí lo vital que fue rediseñar y adaptar el cuestionario original de Chen y Starosta a los participantes del proyecto. Estos fueron estudiantes de 9-11 años de los siguientes centros en España y Europa: CEIP Andrés García-Soler (Lorca, España), Riga Classical Gymnasium, (Riga, Letonia), Scuola Padre Pío (Altamura, Italia), Redhills Community Primary School (Exeter, UK) y Offwell Primary School and Nursery, (Honiton, UK)¹.

¹La Universidad Nebrija (Madrid, España) fue la encargada de establecer los parámetros y diseño de la investigación propuesta (medir la evolución de la sensibilidad intercultural de los participantes). Para más información sobre el proyecto consultar el siguiente enlace: <http://clilnebrija.blogspot.com/>.

About myself...

I am a ...  

I am _____ years old

 I speak _____ languages

I was _____ years old when I started learning a second language



I have lived in _____ countries

Figura 2. Cuestionario demográfico

QUESTIONNAIRE

Read each sentence carefully and mark with an X the degree to which you agree or disagree. There are no right or wrong answers. If you have any questions or if you don't understand the meaning of any words, raise your hand and the teacher or the person in charge will help you.

	 Strongly Disagree 1	 Disagree 2	 Agree 4	 Strongly Agree 5
1. I enjoy interacting with people from other countries.				
2. I think people from other countries and cultures are not willing to listen or to tolerate other people's views.				
3. I do not like to work with people from other cultures or customs.				
4. I like to talk in front of people from other countries/cultures.				
5. I enjoy interacting with people from other cultures (for instance, playing games, etc.)				
6. I feel confident when interacting with people from other countries or cultures.				
7. I am open to people from other cultures.				
8. I would not accept the opinions of people from other cultures.				
9. I have a positive attitude towards people from other cultures.				

Figura 3. Cuestionario sensibilidad intercultural 1

El cuestionario (Figura 3) se rediseñó para ayudar a potenciar las siguientes categorías presentes en la definición de la competencia y sensibilidad intercultural propuesta:

- Énfasis en la interacción intercultural siguiendo los estándares europeos del aula multicultural europea;
- Respeto por las diferencias culturales de los países participantes/ implicados en el proyecto; y
- Confianza en la experiencia intercultural dentro de la competencia comunicativa.

3.2 Aula multicultural como espacio de lengua y cultura aplicada: “Ayudas a la innovación educativa y a la mejora de la calidad de la enseñanza”

Este segundo programa pedagógico para las asignaturas de inglés para fines específicos parte de dos proyectos subvencionados el Área de Innovación Educativa de la Universidad Politécnica de Madrid². Estos pretendieron fomentar la comunicación en inglés para fines profesionales y la comprensión de los numerosos contextos plurilingües e interculturales que aparecen en el marco universitario propio de los grados de ingeniería. Para ello, se estructuró el proyecto en tres ejes fundamentales.

En primer lugar, *Aulas Puente* como programa que conecta a los estudiantes universitarios de herencia y/o experiencia internacional con estudiantes de infantil/primaria de Madrid a través de un intercambio y un aprendizaje colaborativo e intercultural.

En segundo lugar, para aquellos estudiantes que carecen de dicha herencia/ experiencia internacional participaron en el programa *Conversifi* que conecta a estudiantes de lenguas extranjeras con hablantes nativos (inglés-español) en sesiones de 15 minutos de práctica en la lengua seleccionada vía videollamadas (usando el móvil y/o el ordenador). La participación en este programa se

²Los dos proyectos de innovación educativa son los siguientes:

- (1) 2017-18: Aprendizaje basado en retos para la mejora de las destrezas comunicativas en la asignatura *English for Professional and Academic Communication* (EPAC): Creación del Aula Multicultural.
- (2) 2019: Aprendizaje en entornos colaborativos como mejora de las destrezas comunicativas e interculturales en las asignaturas de español: Desarrollo e impulso del Aula Multicultural (Espacio de Lengua y Cultura Aplicada).

forja a través del intercambio lingüístico gracias a una aplicación móvil o el ordenador durante los fines de semana. Participan estudiantes de universidades de EE.UU., México y España como Tufts University, Yale University, Harvard University, MIT, Boston College, UNAM, UPM, Universidad de Cantabria, entre otras.

Finalmente, se han diseñado tres seminarios en torno al componente multicultural que ayudan a adquirir las competencias interculturales propuestas³.

Como instrumento de recogida de datos se realizó una encuesta/cuestionario a los participantes en la que se midió la sensibilidad intercultural. A continuación, se muestra los diferentes ítems que miden la sensibilidad intercultural desde una perspectiva amplia dentro de las dos competencias objeto de estudio y reflejadas en el marco teórico descrito anteriormente (Figura 4).

En las Figuras 5, 6 y 7 se incluyó las variables sociodemográficas haciendo un especial hincapié en la variable de identidad de género al inicio del cuestionario (Figura 5) y país de nacimiento tanto el propio como el familiar al final de cuestionario (Figuras 6 y 7). Se escogió ese orden para no condicionar al participante en las respuestas anteriores en torno a la competencia intercultural y comunicativa.

Intercultural Sensitivity Initial Questionnaire

We invite our students to fill out this questionnaire concerning intercultural communication. There are no right or wrong answers. Please, record your first impression by indicating the degree to which you agree or disagree with the statement.

1 = strongly disagree

2 = disagree

3 = uncertain

4 = agree

5 = strongly agree

Thank you for your cooperation!

Figura 4. Cuestionario sensibilidad intercultural 2

³Para más información sobre el proyecto se puede consultar el siguiente enlace: www.aulamulticultural.es.

Demographic questions

Which category below includes your age?

- 18 or younger
- 19-24
- 25-30
- 31-39
- 40 or older

To which gender do you most identify?

- Female
- Male
- Transgender Female/Trans Woman/ Male-to-Female (MTF)
- Transgender Male/Trans Man/ Female-to-Male (FTM)
- Gender Variant/Non-Conforming
- Not listed
- Prefer not to answer/Choose not to disclose

Are you White, Black, Asian, Latino, or some other race?

- White
- Black
- Asian
- Latino
- Other:

Figura 5. Cuestionario demográfico 2.1

Question 10

In what country were you born?

- Spain
- Other: _____

Figura 6. Cuestionario demográfico 2.2

Question 11

In what country were your family born?

Spain

Other: _____

Figura 7. Cuestionario demográfico 2.3

Los participantes han de señalar el grado de acuerdo o desacuerdo con cada uno de los 9 ítems propuestos. Para ello, se maneja una escala sobre 5 puntos (1=totalmente en desacuerdo, 2=en desacuerdo, 3=no estoy seguro, 4=de acuerdo, 5=totalmente de acuerdo). La puntuación mínima es de 9 puntos, esta revela una *actitud en contra*, y la máxima de 45 puntos, cercana a una *actitud muy favorable*. Asimismo, la puntuación de 27 (el 60 %) muestra una *actitud indiferente o indecisa*. Además, otra clasificación interesante es la que plantea De Santos (2003) dividida en 5 categorías/factores: 1) *Interaction Engagement*, 2) *Tolerance*, 3) *Interaction Confidence*, 4) *Respect of Cultural Differences* y 5) *Attention*.

En las Figuras 8, 9, 10 y 11 se incluyó las variables de interacción y comunicación intercultural vinculadas a la identidad cultural en las primeras tres preguntas (Figuras 8, 9 y 10) y en el énfasis de la experiencia intercultural en relación con la comunicación verbal y no verbal en la última (Figura 11). Las Figura 8 y 11 ayudan a calcular una actitud favorable o muy favorable en el Factor 1 (*Interaction Engagement*) y las Figuras 9 y 10 en el Factor 3 (*Interaction Confidence*).

Question 1

I enjoy interacting with people from different cultures.

1 2 3 4 5

Strongly disagree Strongly agree

Figura 8. Cuestionario sensibilidad intercultural 1

Question 2

I feel confident when interacting with people from different cultures.

1 2 3 4 5

Strongly disagree Strongly agree

Figura 9. Cuestionario sensibilidad intercultural 2

Question 3

I always know what to say when interacting with people from different cultures.

1 2 3 4 5

Strongly disagree Strongly agree

Figura 10. Cuestionario sensibilidad intercultural 3

Question 4

I often show my understanding through verbal or nonverbal cues when interacting with people from different cultures.

1 2 3 4 5

Strongly disagree Strongly agree

Figura 11. Cuestionario sensibilidad intercultural 4

En las Figuras 12, 13, 14 y 16 se incluyó las variables respeto, impresión, percepción y comprensión intercultural. Las Figuras 12 y 16 ayudan a calcular una actitud favorable o muy favorable en el Factor 4 (*Respect of Cultural Differences*), la Figura 13 en el Factor 1 (*Interaction Engagement*) y las Figuras 14 y 15 en el Factor 5 (*Attention*).

Question 5

I respect the values of people from different cultures.

1 2 3 4 5

Strongly disagree Strongly agree

Figura 12. Cuestionario sensibilidad intercultural 5

Question 6

I tend to wait before forming an impression of people from different cultures.

1 2 3 4 5

Strongly disagree Strongly agree

Figura 13. Cuestionario sensibilidad intercultural 6

Question 7

I am very observant and perceptive when interacting with people from different cultures.

1 2 3 4 5

Strongly disagree Strongly agree

Figura 14. Cuestionario sensibilidad intercultural 7

Question 8

I try to learn as much as I can when interacting with people from different cultures.

1 2 3 4 5

Strongly disagree Strongly agree

Figura 15. Cuestionario sensibilidad intercultural 8

Question 9

I have an understanding of the role culture and racism play in the development of identity and world views among minority groups.

1 2 3 4 5

Strongly disagree Strongly agree

Figura 16. Cuestionario sensibilidad intercultural 9

El cuestionario planteado ayuda a perfilar la posible evolución de la sensibilidad intercultural en contextos globales y plurilingües. A su vez, propicia la mejora de la comunicación intercultural, la gestión de las emociones y de confianza durante la adquisición de la L2/L3, así como la interacción entre los participantes ya que gracias a este tipo de medición la adaptación pedagógica es más precisa y transversal.

4. IMPLICACIONES DIDÁCTICAS Y CONCLUSIONES

Para esta investigación, y siguiendo los objetivos propuestos, se revisó el concepto de sensibilidad intercultural en relación con la competencias intercultural y comunicativa, es decir, la comunicación intercultural en programas de adquisición del inglés como L2/L3. Asimismo, se rediseñó dos herramientas para su cálculo y para cada proyecto. A continuación, se van a describir las implicaciones didácticas más relevantes de los dos proyectos de forma conjunta ya que las recomendaciones didácticas y sus conclusiones han seguido la misma línea⁴.

⁴Estos son los estudios que se han elaborado y publicado hasta el momento en el marco de los dos proyectos y de los que hemos partido para la escritura de este artículo:

(1) MAVROU, I., LÓPEZ MEDINA, B. 2019. "Attentional Capacity and Written Production in English (L1 and L2) and Spanish (L1) within the Framework of Erasmus+ Project", *Empirical Studies in Multilingualism: Analysing Contexts and Outcomes*, A. Jiménez-Muñoz, A. C. Lahuerta Martínez (eds.). Frankfurt: Peter Lang, 103-129.

(2) SANTOS-SOPENA, Ò. 2018. "Aprendizaje basado en retos para la mejora de las destrezas comunicativas en la asignatura *English for Professional and Academic Communication* (EPAC)", Portal de Innovación Educativa (UPM). <https://innovacioneducativa.upm.es/jornadas-2018>.

Por lo tanto, las implicaciones didácticas que se plantean giran en torno al desarrollo de la metodología activa. Son ejemplos, el aprendizaje basado en retos y el enfoque basado en tareas / proyectos. Ambos escenarios ayudan a adquirir ciertas competencias profesionales a partir de la práctica; el papel de las políticas de intercentros / instituciones; el enfoque multidisciplinar; y el uso de las WEB 2.0. que en todo momento se fomentan en ambos programas.

Otra propuesta didáctica por subrayar es el papel de la comunicación intercultural como acercamiento pedagógico en la construcción de actividades L2 / L3 (dentro y fuera del aula; diseñadas e incluidas en cada proyecto). Por ejemplo, la resolución de un problema / situación cotidiana de forma colaborativa, el debate como herramienta de transformación social, el teatro aplicado a la enseñanza de lenguas como instrumento primario dentro del arte+educación y/o la acción de medición de la sensibilidad intercultural, entre otras. Esta última cuestión es muy importante ya que la actividad que se genera en torno a completar los cuestionarios propuestos *per se* es de gran interés pedagógico y ayuda a mejorar la competencia comunicativa tanto la oral como la escrita, y sin olvidar una destreza como la comprensión lectora por estar el cuestionario en L2 / L3.

La innovación educativa en este tipo de programas parte de la internacionalización de los programas, equipos de trabajo e instituciones implicadas. Por lo tanto, otra implicación didáctica esencial es el carácter interdisciplinar y práctico de los procesos colaborativos propuestos y que enfatizan las siguientes cuestiones: la creatividad, la motivación, los retos docentes y el desafío pedagógico de diversidad y acción sociocultural. Asimismo, este tipo de acciones propician el acercamiento del aula a la sociedad dentro de la interacción y compromiso cultural; un énfasis y respeto por las diferencias culturales; y la confianza y disfrute en la interacción.

Como conclusiones principales tenemos:

- Se acentúa la internacionalización los participantes/estudiantes;
- Se potencia el aspecto social y de mediación intercultural, desarrollando

(3) SANTOS-SOPENA, Ò. 2019. "Implicaciones didácticas en la complejidad sintáctica y léxica: una aproximación desde la competencia intercultural de la población migrante", *La competencia textual y la complejidad de los textos: entre la lingüística y la didáctica*. A. Schrott, B. Tesch (eds.). Hannover: Ibidem (en prensa).

(4) SANTOS-SOPENA, Ò. 2019. "Aprendizaje interdisciplinario como mejora de las destrezas comunicativas e interculturales en las asignaturas de inglés para fines específicos", *Nuevas visiones en educación y empresa para una nueva sociedad*. UES. Madrid: Editorial Fragua (en prensa).

la competencia intercultural y de diversidad cultural como herramienta de transformación social;

- Se desarrolla la pragmática intercultural y comunicativa a partir de la interacción social en entornos de enseñanza-aprendizaje bilingües/plurilingües;

Por lo tanto, dichas conclusiones manifiestan la trascendencia de este tipo de proyectos para la mejora de la adquisición de la lengua inglesa en contextos globales y plurilingües dentro y fuera del aula que ayudan a desarrollar diversas prácticas pedagógicas de inmersión lingüística. Por ello, el aumento en la visión pluricultural de los participantes, dentro de un mundo globalizado y en constante cambio, fomenta la comunicación y el entendimiento de los contextos multiculturales a los que se enfrenta.

REFERENCIAS BIBLIOGRÁFICAS

- BENNETT, M. J. 1993. "Towards Ethnorelativism: A Developmental Model of Intercultural Sensitivity", *Education for the Intercultural Experience*, R. M. Paige (ed.). Yarmouth, ME: Intercultural. 21–71.
- BUJÁN SÁNCHEZ, L. DE LAS M. 2016. *Pragmatic competence in Spanish students of EFL: making requests and apologies* (Trabajo Fin de Master). Facultad de Filología. Universidad de La Coruña.
- BYRAM, M., GRIBKOVA, B., STARKEY y H. 2002. *Développer la dimension interculturelle dans l'enseignement des langues. Une introduction à l'usage des enseignants*. Strasbourg: Conseil de l'Europe.
- CONSEJO DE EUROPA, 2002. *El Marco Común Europeo de Referencia para las lenguas (MCER)*. Madrid: Ministerio de Cultura y Deportes y Editorial Anaya.
- CHEN, G. M. y STAROSTA, W. J. 1997. "A Review of the Concept of Intercultural Sensitivity", *Human Communication*, 1: 1-16.
- DEARDOFF, D. K., 2006. "Identification and Assessment of Intercultural Competence as a Student Outcome of Internationalization", *Journal of Studies of International Education*, (10)3: 241-266.
- ELLIS, R. 2003. *Task-based Language Learning and Teaching*. Oxford: Oxford University.
- GÓMEZ BARRETO, I. 2011. *La competencia Intercultural en la formación inicial de los maestros de educación infantil en la Universidad de Castilla la Mancha* (Tesis doctoral). Departamento de didáctica y organización escolar y didácticas especiales. Universidad Nacional de Educación a Distancia.
- GRIFFITH, R.L., WOLFELD, L., ARMON, B.K., RIOS, J., LIU y O.L. 2016. "Assessing Intercultural Competence in Higher Education: Existing Research and Future Directions", *ETS Research Report*.
- HUBERT *et al.* 2014. *Developing Intercultural Competence through Education*. Strasbourg: Council of Europe Publishing.
- KLIPPEL, F. 1994. "Cultural Aspects in Foreign Language Teaching", *Journal for the Study of British Cultures*, 1(1): 49-61.

- KRAMSCH, C. y ZHU H. 2016. "Language, Culture and Language Teaching", *Routledge Handbook of English Language Teaching*, G. Hall (ed.). London: Routledge. 38-50.
- LEE, J. F., VAN PATTEN y B. 2003. *Making Communicative Language Teaching Happen*. New York: McGraw-H.
- LITTLEWOOD, W. 1996. *La enseñanza comunicativa de idiomas. Introducción al enfoque comunicativo*. Cambridge: Cambridge University Press.
- MARTÍN PERIS, E. 2008. *Diccionario de términos clave de ELE*. Madrid: SGEL.
- RICHARDS, J. S. y RODGERS T. S. 1998. *Enfoques y métodos en la enseñanza de idiomas*. Cambridge: Cambridge University Press.
- ROLDÁN MEJÍAS, C. 2016. "Competencia intercultural docente en los planes de formación inicial del profesorado de ELE de la Comunidad de Madrid", *Revista Nebrija de Lingüística Aplicada a la Enseñanza de Lenguas*, 10(21): 18-26.
- STONEBANKS, C. 2010. "Consequences of Perceived Ethnic Identities", *Miseducation of the West*. London: PRAEGE.
- YAMADA, M. 2009. "The Role of Social Presence in Learner-Centered Communicative Language Learning Using Synchronous Computer-mediated Communication: Experimental Study", *Computers & Education*, 52(2009): 820-833.
- SCHULZ, R. A. 2001. "Cultural Differences in Student and Teacher Perceptions Concerning the Role of Grammar Instruction and Corrective Feedback: USA-Colombia", *The Modern Language Journal*, 85(2): 244-258.
- SPRINGER, B. 2005. "La comunicación intercultural y la enseñanza de lenguas", *2º Encuentro Práctico de Profesores de Español como Lengua Extranjera*. 1-21.
- SPYCHAŁA, M. 2014. "La problemática de la comunicación intercultural: consideraciones didácticas para la enseñanza de las lenguas", *Studia Romanica Posnaniensia*, Adam Mickiewicz University Press, Poznań, XLI/1: 87-100.

ESPAÑOL COMO LENGUA SEGUNDA EN CONTEXTOS DE ENCIERRO: UN TALLER DE ESCRITURA ACADÉMICA

Florencia Vecchione

Instituto de Lingüística de la Universidad de Buenos Aires, Argentina

Rocío Nahir Navarro

Universidad de Buenos Aires, Argentina

Camila Merlo

Instituto de Lingüística de la Universidad de Buenos Aires, Argentina

RESUMEN

El programa de la Universidad de Buenos Aires UBAXXII de educación en cárceles fue creado con el fin de que las personas privadas de libertad ambulatoria tuvieran acceso a la educación superior. Se identificaba la falta de manejo de la variedad estándar del español —particularmente la variedad académica— por parte de los estudiantes como una de las principales causas de deserción en el Centro Universitario Devoto (CUD). Por este motivo, se resolvió llevar a cabo el Taller de comprensión y producción de textos académicos, cuyo objetivo era el de enseñar el manejo de la lengua académica a los alumnos que estaban comenzando sus estudios universitarios.

Desde el enfoque de enseñanza del español como lengua segunda y extranjera (ELE/EL2), consideramos que el manejo oral y escrito de la lengua académica supone el aprendizaje de una nueva lengua, distinta a la que el individuo usa en situaciones cotidianas. Y es por este motivo que su aprendizaje debe darse a través de una instrucción explícita y sistemática de las competencias lingüísticas y elementos gramaticales que la componen (Acuña et al., 2003). En el contexto de la cárcel, entendemos que este enfoque cobra especial relevancia, ya que conviven distintas variedades de lengua. En este trabajo se presentará el marco de intervención del Taller de Comprensión y Producción de textos académicos llevado adelante en el Centro Universitario Devoto en el año 2017 y los resultados obtenidos por los estudiantes¹.

Palabras clave: lengua académica, ELE, CUD

¹La presente comunicación es parte una investigación que se realiza en el marco del proyecto UBACYT “Diagnóstico, planificación, evaluación y formación para la enseñanza de segundas lenguas: estudiantes brasileños, profesores qom y estudiantes en contexto de encierro”, en la Universidad de Buenos Aires.

ABSTRACT

The programme UBA XXII of Inmate Education from the University of Buenos Aires was created to grant prison inmates access to higher education. A lack of competence in the use of a Spanish standard variety —particularly, an academic variety— was identified as one of the main causes for high dropping out rates in Centro Universitario Devoto (CUD). For this reason, it was decided that a workshop on reading and writing academic texts was to be held, and the workshop's aim was to teach the use of academic language to students that were beginning their university studies.

From the perspective of teaching Spanish as a Second and Foreign Language, we consider that both the oral and written use of academic language entails the learning of a new language/variety, different from the one used in everyday life. And it is for this reason that its learning should be through the explicit and systematic teaching of linguistic competences and the grammatical elements it is made of (Acuña et. al., 2003). In the context of prison, we understand that this perspective is especially relevant as there are very different varieties. In this paper we will present the intervention framework of the Workshop on Reading and Writing Academic Texts that took place in Centro Universitario Devoto during 2017, and the results yielded by students.

Key words: Academic Spanish, Spanish as a Second Language, Inmate Education

1. INTRODUCCIÓN

El Centro Universitario Devoto (CUD) funciona dentro del Programa de la Universidad de Buenos Aires UBA XXII de Educación en Cárceles, y reúne a aproximadamente trescientos estudiantes privados de su libertad que tienen la posibilidad de realizar sus estudios universitarios en contextos de encierro. El programa, dependiente de la Secretaría de Asuntos Académicos del Rectorado, nace en el marco de un convenio celebrado el 17 de diciembre de 1985 entre el Rector de la Universidad de Buenos Aires y la Dirección Nacional del Servicio Penitenciario Federal (SPF).

A partir de las dificultades académicas que muchos estudiantes del programa presentaban, y de acuerdo con el pedido de docentes, alumnos y funcionarios, se elaboró un diagnóstico que nos permitió, desde la Facultad de Filosofía y Letras y el Instituto Nacional de Antropología y Pensamiento Latinoamericano, diseñar un taller de comprensión y producción de textos académicos. El taller estuvo dirigido por Leonor Acuña, coordinado por Lidia Chang y destinado a estudiantes del Centro Universitario Devoto que estaban ingresando al Ciclo Básico Común² de la Universidad de Buenos Aires. Para poder definir los objetivos del taller, nos preguntamos cuáles eran las trayectorias educativas de los alumnos, cómo había sido su experiencia con respecto a la educación formal y cuál era su relación con la lengua académica.

El objetivo de nuestro trabajo es presentar el enfoque de Español como Lengua Segunda y Extranjera para el diagnóstico, la planificación y la evaluación de la comprensión y producción de textos académicos por parte de los estudiantes del Centro Universitario Devoto. Presentaremos el marco de intervención del equipo docente y un balance de los resultados obtenidos por los estudiantes.

2. LA ESCRITURA ACADÉMICA EN EL NIVEL SUPERIOR

En el sistema educativo argentino existe cierta preocupación, desde los años 80 hasta la fecha, en torno a las competencias en materia de lectura y escritura con las que los estudiantes ingresan en la universidad. Hay un acuerdo general en la academia (Arnoux *et al.*, 2002; Adelstein & Kuguel, 2004; Moyano,

²El Ciclo Básico Común de la UBA comprende seis materias cuya aprobación es obligatoria para cursar las carreras de grado de la universidad.

2007) de que la universidad debe brindar a quienes ingresan herramientas para la comprensión y producción de los textos que circulan en este ámbito.

Desde el campo de estudios de la enseñanza del español como lengua extranjera con fines académicos (Vázquez, 2004; Pastor Cesteros, 2006) se desarrollan propuestas que pretenden entrenar a los estudiantes en el manejo del discurso académico de la lengua meta, y que surgen principalmente para atender a las necesidades de estudiantes extranjeros en universidades de habla hispana, en el contexto de programas europeos de movilidad.

2.1 El enfoque de enseñanza del español como lengua segunda y extranjera

Partiendo de un enfoque de enseñanza del español como lengua segunda y extranjera (ELE/EL2), consideramos que el manejo oral y escrito de la lengua académica supone el aprendizaje de una nueva variedad de lengua, distinta a la que el individuo usa en situaciones cotidianas. Partimos de la premisa de diferenciar la lengua estándar, entendida como el resultado de una intervención directa y deliberada sobre la lengua, de la lengua materna. Esta última es aquella adquirida de modo natural en la infancia mediante la interacción con la familia y la comunidad de habla, y sobre todo, se diferencia de la estándar por *no* ser la lengua aprendida mediante la instrucción escolar (Trudgill y Hernández Campoy, 2007). Desde esta perspectiva, entonces, cualquier persona se enfrenta al aprendizaje de una segunda lengua en el transcurso de su escolarización, que consiste en definitiva en un proceso cuyo objetivo es que el destinatario logre alcanzar un cierto nivel de proficiencia en la lengua estándar o académica.

En nuestra experiencia de trabajo con estudiantes en contextos de encierro, nos encontramos con hablantes de variedades que, si bien no son objeto de estudio del presente artículo, presentan una considerable distancia lingüística respecto de la variedad estándar o académica. Nos centramos en que de esta desconexión o alternancia (Appel y Muysken, 1987) entre la lengua materna y la estándar se desprenden resultados académicos no satisfactorios como hipótesis para la elaboración de una propuesta de enseñanza.

En la Argentina, el 83 % de la población carcelaria ingresa en el servicio penitenciario sin haber completado sus estudios secundarios³. Los destinatarios

³De acuerdo al informe del año 2016 del Sistema Nacional de Estadísticas sobre Ejecución de la Pena, Dirección Nacional de Política Criminal en Materia de Justicia y Legislación Penal, Ministerio de Justicia y Derechos Humanos.

del taller, en muchos casos, no habían completado la escolarización obligatoria antes de ingresar en el sistema penitenciario. Son adultos, en su gran mayoría provenientes de clases bajas, marginados social y económicamente. Ellos mismos, como se expondrá en el diagnóstico, identifican dificultades para desempeñarse exitosamente en las situaciones comunicativas propias de un ámbito universitario.

Adoptar el enfoque de ELE/EL2 tiene como finalidad instruir en el nivel pragmático (qué se espera de nosotros en ciertas situaciones), en el nivel gramatical (qué aspectos lingüísticos y gramaticales son propios de los géneros académicos) y en el nivel cultural (qué géneros componen el discurso académico y qué sentido tienen, cuáles son los temas que aborda la universidad, etc.).

3. LA PROPUESTA DEL TALLER

Para cumplir con la meta de nivel de lengua necesaria para ingresar y permanecer en las carreras de la Universidad de Buenos Aires, y partiendo de la situación sociolingüística de los estudiantes del CUD, consideramos necesario trabajar en la escritura a través de los géneros discursivos propios de la academia, partiendo de una noción de géneros discursivos como respuestas habituales a situaciones recurrentes delimitadas socialmente (Bazerman, 2015).

La trayectoria de vida de los estudiantes estaba marcada por la exclusión de aquellos espacios en los que circulan los géneros académicos, ya que estos géneros forman parte de la praxis social en ámbitos determinados y son reforzados por la estructura de poder institucional (Hart-Davidson, 2015). El hecho de no manejar los géneros en los que se basa la enseñanza del nivel superior supone una barrera imposible de sortear sin una exposición sistemática y un trabajo explícito sobre el género, y por ende, el reconocimiento de los tipos de mensajes que puede contener determinados textos, los tipos de situaciones en las que puede utilizarse, etc. (Bazerman op. cit.).

En la edición del año 2017, hubo ciento veintidós inscriptos al taller que realizaron la evaluación diagnóstica y se sumaron a alguno de los tres niveles. Las clases se dictaron durante todo el año lectivo, dividido en dos cuatrimestres, es decir que en total fueron unas 120 horas de clases; cuatro hora semanales durante aproximadamente treinta semanas. De los alumnos que lo cursaron, sesenta y siete finalizaron el curso y obtuvieron la certificación.

3.1 El diagnóstico y la planificación

Para la elaboración de un diagnóstico sobre el cual planificar el trabajo del año en el taller, se llevaron adelante reuniones con las autoridades del programa y docentes del ciclo básico, se entrevistó a los estudiantes y se diseñó una evaluación de diagnóstico de la que participaron.

A través de la realización de entrevistas, pudimos recoger las experiencias y trayectorias socioeducativas y dar cuenta de las principales preocupaciones de los estudiantes con respecto a su capacidad para manejar textos académicos. El análisis de las entrevistas realizadas a los estudiantes arroja los siguientes datos:

- a. Llevaban mucho tiempo alejados de espacios de enseñanza formal, o habían completado sus estudios dentro del sistema penitenciario.
- b. No contaban con las herramientas necesarias a la hora de enfrentarse a las tareas de comprensión y producción de textos en el marco de las materias del Ciclo Básico Común.
- c. Esa dificultad para manejar textos académicos se traducía en un impedimento para regularizar y aprobar las materias del Ciclo Básico Común y, así, acceder a las carreras universitarias.

Esta diversidad de trayectorias debía traducirse en una propuesta que las abarque y que propicie una enseñanza centrada en las diferentes necesidades lingüísticas. Es por esto que, como parte del diagnóstico, el equipo diseñó una evaluación para determinar las capacidades de uso de la lengua académica por parte de los estudiantes. Para esto, se tomó el modelo del Certificado de Español: Lengua y Uso (CELU)⁴ que mide el nivel de lengua —aunque no presenta preguntas específicas sobre lengua y gramática—, y la utilización que se hace de la misma. La evaluación distingue los niveles básico, intermedio y avanzado en el mismo examen. Esta división por niveles permite formular distintos objetivos de acuerdo con el manejo de la lengua académica que los estudiantes poseen, sin presuponer sus conocimientos previos ni su nivel

⁴El Certificado de Español: Lengua y Uso (CELU) es un examen administrado desde noviembre de 2004 por un consorcio de cincuenta y nueve universidades nacionales argentinas. Cuenta con el aval del Ministerio de Educación, Cultura, Ciencia y Tecnología, y el Ministerio de Relaciones Exteriores. Se trata de un examen que acredita la competencia que tiene el hablante de español como lengua segunda y extranjera para comunicarse oralmente y por escrito, de manera efectiva en diversas situaciones. Se trata de un examen que consta de dos partes, oral y escrita, y evalúa las cuatro destrezas (leer, escribir, hablar y escuchar) de forma integrada.

de lengua (McNamara, 2008). En estas evaluaciones, que consistían en dos ejercicios —uno de lectura de un texto de divulgación científica y otro de producción escrita—, se solicitaba que los estudiantes identificaran cierta información del texto de divulgación y la utilizaran para crear un texto nuevo en un género no académico. A partir de estas consignas se puso en juego, entonces, la comprensión del texto y la competencia en la escritura así como también el nivel de manejo de los géneros académicos.

Una vez realizado el diagnóstico, se formularon los siguientes objetivos para el diseño del curso (Acuña *et al.* 1991): el objetivo general fue lograr que el trabajo fuera transversal, es decir, que el aprendizaje de los estudiantes en el taller se pudiera trasladar a su rendimiento en las distintas materias del CBC; en cuanto a los objetivos específicos, la planificación del trabajo con los estudiantes estaba orientada a que alcanzar un nivel de usuario independiente, de acuerdo con las formulaciones del Marco común europeo de referencia para las lenguas⁵.

B1	Es capaz de entender las ideas principales de textos complejos que traten de temas tanto concretos como abstractos, incluso si son de carácter técnico siempre que estén dentro de su campo de especialización. Puede relacionarse con hablantes nativos con un grado suficiente de fluidez y naturalidad de modo que la comunicación se realice sin esfuerzo por parte de ninguno de los interlocutores. Puede producir textos claros y detallados sobre temas diversos así como defender un punto de vista sobre temas generales indicando los pros y los contras de las distintas opciones.
B2	Es capaz de comprender los puntos principales de textos claros y en lengua estándar si tratan sobre cuestiones que le son conocidas, ya sea en situaciones de trabajo, de estudio o de ocio. Sabe desenvolverse en la mayor parte de las situaciones que pueden surgir durante un viaje por zonas donde se utiliza la lengua. Es capaz de producir textos sencillos y coherentes sobre temas que le son familiares o en los que tiene un interés personal. Puede describir experiencias, acontecimientos, deseos y aspiraciones, así como justificar brevemente sus opiniones o explicar sus planes.

⁵INSTITUTO CERVANTES. 2002. Traducción y adaptación del *Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación*.

4. LA IMPLEMENTACIÓN

A partir de la evaluación de diagnóstico de los estudiantes, los dividimos en tres niveles (elemental, intermedio y avanzado), y se propuso un trabajo diferenciado para cada uno de ellos. El planteo de la implementación se basó en la enseñanza a partir de géneros académicos y se incluyó también trabajo de contenido gramatical. A lo largo del semestre los estudiantes fueron confeccionando su portafolio que incluía las producciones escritas de todo el curso y serviría como muestra del proceso de aprendizaje. El taller se dividió en tres partes cuyos encuentros abordaron los géneros seleccionados para cada nivel. En la primera clase de cada segmento se presentaron las características de cada género. Tanto la lectura como la escritura de los textos tuvieron lugar dentro del espacio del taller, aunque también se fomentó el trabajo individual por fuera de la clase. Las correcciones por parte de las docentes fueron semanales y se focalizaron en la adecuación del texto al género elegido, el uso de las herramientas gramaticales y de coherencia y cohesión trabajadas en la clase. A continuación detallamos los criterios para la división en niveles y los aspectos abordados en cada uno.

4.1 El nivel elemental

En este nivel agrupamos a los estudiantes que presentaban dificultades para distinguir entre géneros académicos y no académicos, y para adecuar el registro de lengua entre coloquial y formal. Generalmente, empleaban estrategias como copiar textualmente la información solicitada y no eran capaces de elaborar un nuevo texto a partir de ella. Sus producciones no presentaban coherencia y cohesión textual, y en algunos casos presentaban faltas de concordancia en número.

En cuanto a los géneros académicos, se trabajó con el resumen, el informe y la respuesta a consignas de examen. Se presentaron las características de la tipología textual, y a partir de la lectura de artículos y capítulos de manuales de las materias del Ciclo Básico, los estudiantes debían producir textos pertenecientes a estos géneros. Se incorporó al taller la dimensión del trabajo con la gramática: ortografía, puntuación y uso de conectores lógicos.

4.2 El nivel intermedio

En el nivel intermedio se agrupó a los estudiantes que podían reconocer los géneros académicos y las diferencias con otros géneros lingüísticos pero no tenían un manejo o comprensión cabal de los géneros académicos. Si bien en sus producciones se observaba la identificación de las ideas principales de los textos, tenían dificultades para expresarlas de manera cohesiva en un texto nuevo.

En este grupo se trabajó en la comprensión de textos académicos complejos a través del reconocimiento de hipótesis y argumentos y la producción de textos nuevos, respetando la especificidad de los géneros resumen, respuesta a consignas de examen e informe. En la producción, se trabajó con la estructuración de dichos textos a través del uso de conectores y marcadores discursivos y se enfatizó en los recursos para organizar la información. Se trabajaron cuestiones de estilo como la puntuación y el uso de comas.

4.3 El nivel avanzado

Aquellos estudiantes que fueron agrupados en este nivel eran capaces de distinguir entre diferentes géneros y adecuar el registro de sus escritos. Sin embargo, en la producción no manejaban las tipologías textuales de géneros académicos diferenciados.

Se presentaron los géneros de informe, respuesta de examen y monografía. El trabajo se enfocó en la coordinación de la precisión comunicativa, la gramaticalidad y la adaptación cultural. Por ejemplo, en la monografía estas tres dimensiones se pusieron activamente en juego porque suponen: a) la investigación y el desarrollo de una temática relevante para el ámbito académico, b) una expresión clara y elaborada que de cuenta de un manejo avanzado de los elementos textuales de cohesión y gramaticales en general y c) una completa adecuación al registro académico y adaptación al género en cuestión.

5. SOBRE LOS RESULTADOS Y LA EVALUACIÓN

Los resultados de la implementación del taller fueron medidos en función de los avances evidenciados en cada portafolio de los alumnos. Como el taller se dividió en tres bloques en los que se abordó un género en particular, al finalizar cada segmento los estudiantes realizaron una actividad de evaluación en la que debían producir un texto que se ajustara al género estudiado.

A partir de la observación de estos trabajos podemos sostener que, en el nivel elemental, por una parte, los estudiantes realizaron producciones escritas en un registro adecuado y siguiendo las pautas de los géneros académicos conocidos. Por otra parte, corrigieron significativamente la ortografía y puntuación. Sin embargo, al finalizar el taller seguían presentando dificultades para lograr una mayor coherencia y cohesión de los textos más complejos.

En el nivel intermedio, las producciones eran cohesivas y estructuradas producto de la utilización de conectores lógicos variados y organizadores de la información. Resta profundizar el trabajo con la precisión comunicativa, ya que en ocasiones los estudiantes recurrían a formulaciones demasiado generalizadas.

El grupo avanzado logró el manejo fluido de los géneros académicos de informe y consigna de examen, y en sus producciones observamos una mayor precisión comunicativa. En cuanto al trabajo con el género de la monografía, persistían algunas dificultades en el aspecto gramatical ya que contaban con un repertorio reducido de elementos textuales de cohesión.

Además, incorporamos la experiencia de los alumnos, y por ello incluimos una sistematización de las entrevistas de fin de curso. Con el fin de presentar una evaluación de la cursada en base a la experiencia previa con respecto a la educación formal y los objetivos de los estudiantes con respecto al curso, realizamos las siguientes preguntas:

1. Antes de empezar el CBC en el Programa UBA XXII ¿Cuál fue el último año que estudiaste en una institución de educación formal?
2. ¿Por qué empezaste el taller de lectura y escritura?
3. ¿Para qué te sirvió el taller?

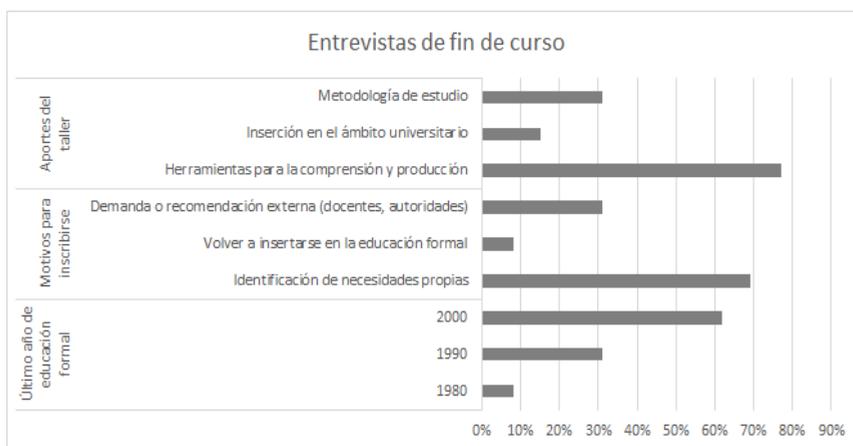


Figura 1. Autoevaluación de los estudiantes sobre la experiencia del taller

De las respuestas se desprende que la motivación principal a la hora de tomar el taller fue el reconocimiento de las propias falencias que tenían los estudiantes al momento de abordar textos de géneros específicos que circulan en el ámbito académico. Resulta relevante destacar que, frente a las falencias que ellos mismos reconocían tener, encontraron en su mayoría que el taller les brindó las herramientas necesarias para la comprensión y producción de textos. La diversidad de experiencias y trayectorias que se mencionó en un principio se ve reflejada en la cantidad de estudiantes cuyo último año de educación formal había sido en las décadas de los ochenta y noventa.

6. CONCLUSIONES

La problemática en torno a las competencias en lectura y escritura de quienes ingresan a la universidad requiere especial atención en los casos en los que los estudiantes no tienen un nivel de manejo de la lengua estándar que les permita continuar sus trayectorias educativas en el nivel superior. En el caso de la educación en contextos de encierro, la atención de esta dificultad se vuelve más urgente. Se trata de una población en situación de marginalidad, sujeta a un alto nivel de estigmatización social y cuya posibilidad de acceso a la educación superior se ve muy limitada fuera de este programa. Por este motivo, resulta esencial pensar la enseñanza explícita de la lengua académica como una política lingüística y educativa de inclusión y permanencia en la universidad.

REFERENCIAS BIBLIOGRÁFICAS

- ACUÑA L., FERNÁNDEZ C., MAMARIAN D., MENEGOTTO A., PALACIOS M. 1991. “Niveles de enseñanza del español como segunda lengua”, Primeras Jornadas Nacionales *De la teoría lingüística a la enseñanza de la lengua*. Rosario, Santa Fe.
- ACUÑA L. 2003. *Módulo n° 2 El diagnóstico sociolingüístico*. Buenos Aires: Programa DIRLI.
- ACUÑA L. 2012. “Destinatarios y modelos de lengua: sobre el CELUy el desarrollo de ELSE Argentina”, en E. R Sturza., C. Ivani & I Valesca Brasil (org.). *Português e Espanhol: Esboços, Percepções e Entremeios*. Santa María: UFSM, pp. 223-236.
- ADELSTEIN A., KUGUEL, I. 2004. *Los textos académicos en el nivel universitario*. Los Polvorines: Universidad de General Sarmiento.
- APPEL, R. & MUYSKEN, P. 1987. *Bilingüismo y contacto de lenguas*. Barcelona: Ariel.
- BAZERMAN, C. 2015. “Writing speaks to situations through recognizable forms”. En Adler-Kassner, Linda & Wardle, Elizabeth (Eds.) *Naming what we know*, Utah University Press, pp. 62-69.
- CIAPUSCIO G., ADELSTEIN A., GALLARDO S. 2010. “El texto especializado: propuesta teórica y prácticas de capacitación académica y profesional en Argentina”. En G. Parodi (Ed.), *Alfabetización académica y profesional en el siglo XXI. Leer y escribir desde las disciplinas*. Santiago de Chile: Ariel.
- HART-DAVIDSON B. 2015. “Genres are enacted by writers and readers” en Adler-Kassner, Linda & Wardle, Elizabeth (Eds.) *Naming what we know*, Utah University Press.
- INSTITUTO CERVANTES. 2002. Traducción y adaptación del *Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación*. URL: <http://cvs.cervantes.es/obref/marco>
- MCNAMARA T. 2008. “Language assessment as a social practice”, conferencia IV Coloquio CELU. San Martín: UNSAM.

- MOYANO E. 2007. "Enseñanza de habilidades discursivas en español en contexto preuniversitario: una aproximación desde la LSF". *Signos*, 40 (65), 573-608.
- PASTOR CESTEROS S. 2006. "La enseñanza del español como lengua vehicular en contextos académicos", en *MarcoELE* n° 2.
- TRUDGILL P., HERNÁNDEZ CAMPOY J. M. 2007. *Diccionario de sociolingüística*. Madrid: Gredos.
- VÁZQUEZ G. 2004. "La enseñanza del español con fines académicos" en J. Sánchez Lobato e I. Santos Gargallo (Eds.) *Vademécum para la formación de profesores. Enseñar español como segunda lengua y lengua extranjera*. Madrid: SGEL, 1129-1147.

EL LENGUAJE SERÁ AMBIGUO O NO SERÁ: EL PORQUÉ DE LA AMBIGÜEDAD LÉXICA Y SU ESTUDIO DESDE LA EVOLUCIÓN DEL LENGUAJE

Natalia López Cortés
Universidad de Zaragoza

RESUMEN

Aunque no hay acuerdo entre aquellos que piensan que es un elemento positivo y los que opinan que es un elemento negativo, no se puede negar que la ambigüedad léxica es un fenómeno común en el lenguaje. Por ello, un estudio que pretenda entender la ambigüedad en profundidad no puede estar completo sin responder a dos preguntas clave: por qué existe este fenómeno y cómo pudo haber surgido. En este trabajo se ofrece, por un lado, una aproximación a las diferentes teorías que consideran la ambigüedad como un elemento necesario para el correcto funcionamiento del sistema lingüístico; y, por otro, una propuesta preliminar para el estudio de la ambigüedad desde un punto de vista evolutivo, que plantea la idea de dos fases en su surgimiento, correspondientes a los dos tipos de polisemia.

Palabras clave: ambigüedad, significado, evolución del lenguaje, lexicón mental, redes semánticas

ABSTRACT

Although there is no agreement between those who claim that it is a positive element and those who think that it is a negative one, it cannot be denied that lexical ambiguity is a common phenomenon in language. Therefore, a study which aims to fully understand ambiguity cannot be complete without answering two key questions: why this phenomenon exists and how it may have arisen. This paper shows, on the one hand, an approach to different theories that consider ambiguity as an essential element for the proper functioning of the linguistic system; and, on the other hand, a preliminary proposal to study ambiguity from an evolutionary point of view, which contemplate the idea of two phases in its emergence, related to both types of polysemy.

Keywords: ambiguity, meaning, language evolution, mental lexicon, semantic networks

1. INTRODUCCIÓN

La ambigüedad, ya sea como resultado de un proceso fonético, gramatical o léxico (Ullmann, 1976), está presente de manera constante en las lenguas. Nuestro objeto de estudio, la ambigüedad léxica, derivada de las palabras homónimas y polisémicas, es un fenómeno producido por la convergencia en una única forma lingüística de más de un significado, tengan o no tengan relación entre sí y sea o no su origen etimológico el mismo. Por tanto, se puede afirmar que la ambigüedad es un hecho lingüístico cotidiano y constante, con el que los hablantes lidiamos regularmente. Por ello, un estudio sobre la ambigüedad léxica no puede estar completo sin plantearse el porqué de su existencia ni cómo surgió en un primer momento. Presentar un estudio preliminar sobre estos dos asuntos es el objetivo de este trabajo.

2. ¿POR QUÉ EXISTE LA AMBIGÜEDAD?

En un sistema de comunicación ideal, ¿existiría la ambigüedad? Esta pregunta ha sido planteada varias veces en los últimos años. El hecho de que una única forma lingüística transmita varios significados diferentes parece ser, en un primer momento, un obstáculo para una comunicación totalmente eficaz.

La ambigüedad no parece ser una propiedad necesaria del lenguaje; de hecho, Ullmann (1976: 204) afirma que un fenómeno como la homonimia no tiene ventajas y que, de hecho, si se eliminara totalmente, la lengua sería “un medio más eficaz”. Otros autores, como Cohen (2006), consideran la ambigüedad como un obstáculo, un accidente o un defecto del sistema; para Vivanco (2003: 47), “se trata de un fenómeno constitutivo de la lengua que supone un inconveniente para todos los idiomas”.

Resulta interesante el uso del término ‘constitutivo’, pues se deja entrever la idea de que la ambigüedad en la lengua, sea positiva o negativa, se produce: la ambigüedad existe y es, como dicen Solé et al. (2010: 22) “an important, universal and yet apparently undesirable property of language”. Para Wasow et al. (2005) es sorprendente el hecho de que, a lo largo de la evolución del lenguaje, en ningún momento se haya eliminado la ambigüedad: su teoría es que si fuera un elemento que pudiera distorsionar la comunicación, debería haber sido al menos reducida; pero si esto no ha ocurrido se debe deducir que la ambigüedad tiene alguna función.

Desde diferentes perspectivas teóricas se ha demostrado que la ambigüedad no es solo un elemento deseable sino también necesario. El tema ha sido estudiado desde diferentes disciplinas, desde las matemáticas (Shannon, 1948; Zipf, 1949) hasta la teoría de redes (Solé, 2009), pasando por la lingüística (Piantadosi et al., 2012; Srinivasan y Rabagliati, 2015). Todos estos datos apuntan a que la ambigüedad es un elemento básico en el lenguaje, puesto que supone un factor de economía, reduce el coste de procesamiento y cohesionan el lexicón mental, haciendo que se produzca un salto cualitativo en la organización del vocabulario.

Quizá esta última sea la cuestión más interesante. La ambigüedad léxica tiene una función clara en las redes semánticas puesto que es la responsable de que se dé el umbral de percolación, esto es, el número crítico de conexiones necesario para que la red posea la propiedad del mundo pequeño. Esta propiedad es común a todas las redes complejas y es uno de los conceptos básicos dentro de estos estudios sobre la complejidad. Hace referencia a que los nodos de una red están tan interconectados que el sistema es un mundo pequeño, en el que se puede saltar de un elemento a otro con rapidez, como si hubiera atajos entre los diferentes puntos.

Así, las conexiones que se producen entre palabras ambiguas permiten establecer diferentes atajos entre elementos que, de otra manera, estarían totalmente separados entre sí. A partir de diferentes simulaciones realizadas con la base de datos de WordNet (Fellbaum, 1998) se ha comprobado que, si se eliminan las palabras ambiguas de una red semántica, el número crítico de conexiones se reduce, el umbral de percolación no se alcanza y se pierde la propiedad del mundo pequeño. La red deja de ser, por tanto, una red compleja (v. Solé, 2009). La ambigüedad actúa como una suerte de puente, haciendo que la navegación por la red del lenguaje sea más rápida, más eficaz y mucho menos costosa.

Además, la ambigüedad juega un papel importante para la economía del sistema (Piantadosi *et al.*, 2012; López-Cortés, 2018): el vincular una unidad con más de un significado hace que se produzca una reutilización de elementos y se reduzca así el coste de procesamiento de las unidades léxicas y de la selección de sus significados. Esto se traduce en que un sistema no ambiguo puede ser mejorado al introducir formas ambiguas que ahorren esfuerzo, uniendo significados posibles a formas únicas más sencillas. El sistema se convierte en más eficiente, una vez más, gracias a la ambigüedad.

Esta idea está relacionada con la propuesta de Srinivasan y Rabagliati

(2015), quienes consideran la polisemia como una herramienta esencial para la construcción y desarrollo del lexicon: la polisemia reduce la arbitrariedad producida al mapear un significado a una forma lingüística y supone un mecanismo de ahorro que hace más fácil a los niños el aprendizaje.

Desde un punto de vista comunicativo, la ambigüedad supone un equilibrio que minimiza el esfuerzo tanto del hablante como del oyente. En la comunicación existen dos fuerzas opuestas: la unificación y la diversificación (Zipf, 1949). El emisor quiere unificar el código (cuántos más significados transmita una expresión, mejor) mientras que el receptor quiere diversificarlo (cuántos menos significados transmita una expresión, mejor); y ambos buscan lograr el mínimo esfuerzo posible: si hay ambigüedad, para el emisor es muy fácil producir una expresión puesto que no tiene que esforzarse en escoger un significado; si no hay ambigüedad, para el receptor es muy fácil interpretar cualquier expresión puesto que no tiene que esforzarse en desambiguarla. De esta manera, surge una tensión entre los deseos del emisor y del receptor. La ambigüedad se da como resultado de un compromiso entre las necesidades de los participantes en la comunicación, quienes minimizan su esfuerzo (Corominas-Murtra *et al.*, 2011).

La ambigüedad parece ser, por tanto, un elemento estructural con ciertas funciones de cohesión del sistema y una de las bases que garantiza, o al menos posibilita, su eficacia. Este repaso de algunas de las respuestas a la existencia de la ambigüedad tiene una continuación lógica clara: cómo pudo surgir en un primer momento este fenómeno. Por ello, en el siguiente apartado se va a proponer un acercamiento al fenómeno desde el punto de vista de la evolución del lenguaje.

3. LA AMBIGÜEDAD DESDE UN PUNTO DE VISTA EVOLUTIVO

3.1 Marco teórico: las Ciencias de la Complejidad

El marco teórico en el que se basa esta propuesta sobre el surgimiento de la ambigüedad es el de las Ciencias de la Complejidad, en especial el de autores como Longa (2001, 2011), que aplica las bases de la disciplina para el estudio del lenguaje y su evolución, estableciendo como punto de partida la idea de que el lenguaje es un sistema complejo. Lo que hace que un sistema sea complejo

es que tenga un orden jerárquico: en un sistema complejo los elementos se engarzan y se integran, convirtiéndose en elementos de orden superior y haciendo que el sistema sea más eficaz. Simon (1962), a través del estudio de diferentes estructuras y sistemas, descubre que trabajar jerárquicamente con elementos estructurales es más eficaz¹. La jerarquía es, por tanto, la base para lograr la eficacia de cualquier sistema.

La idea clave de esta teoría es que este orden jerárquico se alcanza a través de una auto-organización del sistema: el propio sistema tiene una motivación intrínseca para organizarse jerárquicamente, y esto no se produce por acción de una causa o agente externo. Este agente ha pasado a la historia con nombre propio: la selección natural.

Así, el modelo neodarwinista se basa en la selección natural como factor plenamente responsable del surgimiento de cualquier cambio: por ello, la evolución del lenguaje, así como el origen de cualquier cambio a nivel evolutivo, habría de ser necesariamente gradual, continuista y adaptacionista. Por otro lado, desde la óptica chomskiana se opta por un cambio brusco que facilita la aparición de rasgos exclusivamente humanos, sobre los que opera la selección natural hasta un punto determinado. Así, existen dos visiones diferentes sobre el alcance de la selección natural: o bien como la única explicación posible de la existencia del diseño complejo, donde se situaría el lenguaje (apoyada por autores como Pinker y Bloom, 1990), o bien como un filtro más que opera sobre un sistema complejo, derivado de un cambio brusco².

La idea de que el lenguaje no surge como una adaptación al medio es lo que tienen en común la visión chomskiana con la visión de las Ciencias de la Complejidad. Desde la perspectiva de esta disciplina, el diseño complejo y, en consecuencia, el lenguaje, no surgen a través de la selección natural, sino como una necesidad interna del propio sistema. En concreto, el sistema necesita estructurarse debido a un aumento en el número de los elementos que lo componen³. Para ello se produce una auto-organización del sistema, que

¹Sobre este tema resulta realmente ilustrativa la parábola de los relojeros Hora y Tempus, propuesta por Simon (1962) en su clásico trabajo sobre complejidad.

²Según Chomsky, el lenguaje no surge como una adaptación al medio, sino como un epifenómeno: el cerebro se reorganiza a consecuencia de un aumento en su tamaño.

³Se debe recordar que esta propiedad no es exclusiva del lenguaje, sino que sucede también con diferentes estructuras. Longa (2001: 5) cita el ejemplo de Goodwin (1994) basado en las colonias de hormigas: cuando hay pocos individuos, el comportamiento es impredecible y caótico, pero cuando se alcanza un número crítico de componentes de la colonia, se produce un cambio brusco que lleva del caos al orden.

hace que se genere una estructura o jerarquía. De esta manera, la complejidad (y la jerarquía) emerge⁴ gracias a la auto-organización.

Longa (2001, 2011) aplica todos estos conceptos al lenguaje y afirma que, en el caso del lenguaje, es el aumento de elementos léxicos el responsable de que se produzca la auto-organización. Antes de esto, existía un ‘lenguaje primitivo’, caracterizado por su linealidad y por estar vinculado a lo referencial, es decir, por estar directamente ligado con el medio externo inmediato. La propuesta de Longa (2001, 2011) es que este protolenguaje se supera al producirse un aumento de los elementos léxicos que lo conforman. Así, los elementos léxicos alcanzan un umbral crítico y esto hace que sea necesario una auto-organización del sistema, en forma de emergencia de la estructura jerárquica⁵.

El por qué se produce este aumento se explica, por un lado, teniendo en cuenta que la jerarquía dota al sistema de mayor eficiencia a la hora de lidiar con los elementos, tal y como se ha mencionado antes, y por tanto el sistema lingüístico será más eficiente lidiando con los elementos léxicos. Y esta mayor eficiencia es necesaria para mejorar la capacidad representacional del lenguaje (Longa, 2001: 11). Es esta capacidad representacional, esencial para el lenguaje, lo que motiva la auto-organización del sistema y permite pasar de un protolenguaje basado en lo referencial, lo vinculado directamente al nicho ecológico (Longa, 2001: 9) a un lenguaje pleno que, como consecuencia, va más allá de lo referencial. Tal y como dice Chomsky (1988: 189), esto permitió a nuestros antepasados superar la simple reacción a los estímulos. De esta manera, con el paso al lenguaje pleno se posibilita el representar (y, en consecuencia, comunicar) conceptos no basados en el medio⁶.

En el siguiente esquema se presenta un resumen de las ideas expuestas hasta el momento. En una primera fase, existe un protolenguaje lineal, sin estructura y, debido a un aumento de los elementos léxicos, que alcanzan un nivel crítico, se produce una auto-organización del sistema. Este proceso conduce a la emergencia del lenguaje pleno, caracterizado por la jerarquía y la existencia de una estructura. Esta autoorganización se produce, entonces,

⁴El concepto de emergencia es clave en el estudio de los sistemas complejos y se basa en el hecho de que la suma de los elementos que componen estos sistemas hace que surjan propiedades que a nivel de los elementos no pueden predecirse.

⁵La estructura jerárquica se vincula con la sintaxis y con la recursividad.

⁶En relación con esto, se ha de comentar que parece haber acuerdo en el hecho de que la semántica referencial ha de ser la más antigua desde un punto de vista evolutivo, puesto que tal y como dicen Benítez Burraco y Barceló-Coblijn (2015: 121-122) “términos abstractos para conceptos como ‘felicidad’ o ‘infinitud’ habrían aparecido mucho más tarde”.

como una dinámica interna del sistema, no como una respuesta o adaptación al medio. Además, se pasa de un protolenguaje ligado a la referencialidad a un lenguaje pleno no exclusivamente ligado a la referencialidad, sino centrado en lo representacional, lo que abarca desde la realidad hasta la irrealidad.

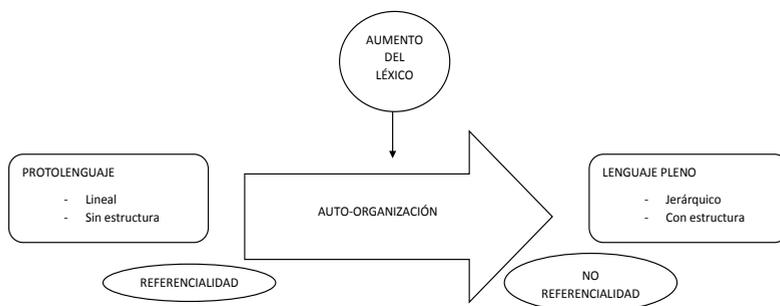


Figura 1. Paso del protolenguaje al lenguaje pleno a través de la auto-organización

3.2 El surgimiento de la ambigüedad

Es en este marco en el que se inserta nuestro estudio de la ambigüedad desde un punto de vista evolutivo. Nuestra propuesta consiste en la diferenciación de dos fases en el surgimiento de la ambigüedad.

3.2.1 Primera fase de la ambigüedad

La ambigüedad, es decir, la transmisión de varios significados diferentes a través de una expresión única es, ante todo, como se ha señalado en el apartado 2, un mecanismo de economía. Así, mediante este proceso, se reconcilian dos necesidades: por un lado, la capacidad semiótica del ser humano y la necesidad de nombrar lo que nos rodea y, por otro lado, el espacio limitado de almacenamiento para el léxico. Esto explica una primera fase en el surgimiento de la ambigüedad: este fenómeno supone una solución al aumento del léxico, pues lo reduce y limita. Por tanto, esta primera fase podría haberse producido antes de la auto-organización, como un mecanismo para lidiar con los elementos léxicos.

Otro asunto distinto es el modo en el que se genera la ambigüedad o, dicho

con otras palabras, cómo se pasa de un léxico monosémico a uno polisémico. Una posible respuesta al respecto es que, en esta primera fase de ambigüedad, los elementos léxicos llegan a la ambigüedad por polisemia metonímica: una sola forma reconcilia varios significados que se relacionan entre sí de manera “parte-todo”. Esta idea entronca con la propuesta de la existencia de la polisemia regular, llevada a cabo por Pustejovsky (1995), y según la cual existen unos patrones regulares de creación polisémica: por ejemplo, el de ‘animal’ y ‘carne de animal’. Esta relación es una relación metonímica esperable en cualquier lengua y que supone un mecanismo de economía⁷.

Además, el hecho de que la polisemia metonímica sea lo esperable en una primera fase del surgimiento de la ambigüedad puede tener que ver con el nicho ecológico, es decir, la posición que ocupamos como especie en un ecosistema. En un primer momento, el protolenguaje es, como se ha comentado anteriormente, totalmente referencial, ligado plenamente al contexto externo. El protolenguaje sirve para satisfacer las necesidades derivadas del nicho ecológico; de esta manera, solo se usan representaciones relacionadas de manera directa con el medio, esto es, con lo necesario para sobrevivir (Longa, 2001: 9).

3.2.2 Segunda fase de la ambigüedad

La segunda fase de la ambigüedad se produciría una vez ya se ha dado la auto-organización del sistema y ha surgido, en consecuencia, el lenguaje pleno. Este se caracteriza, como ya se ha mencionado, por su estructura jerárquica. Es en este momento en el que se pierde la referencialidad y el lenguaje se convierte en una herramienta para construir no solo lo directamente vinculado con el medio inmediato, sino todo lo que va más allá: lo abstracto o la ‘irrealidad’, como dice Longa (2001).

⁷Un acercamiento experimental al tema de la polisemia como fenómeno universal, elaborado por Youn *et al.* (2016), arroja resultados interesantes para el tema que nos ocupa: así, a través de la técnica de las redes semánticas y partiendo de una base de 81 idiomas diferentes, descubrieron que existen conceptos que son más polisémicos que otros (2016: 1766), apuntando a la existencia de unos patrones universales de polisemia. Los vínculos más frecuentes de las redes señaladas en su trabajo, elaboradas a partir de léxico celeste y del paisaje, son la metonimia, la hiperonimia y la co-hiponimia. Así, la polisemia metonímica podría considerarse un fenómeno regular y universal, ligado al protolenguaje y a la referencialidad.

Nuestra propuesta es que es en este contexto en el que surge la capacidad metafórica que deriva en un nuevo tipo de ambigüedad: la polisemia metafórica. Este recurso permite establecer relaciones entre conceptos más alejados entre sí y más alejados del nicho ecológico, puesto que ya se ha superado la referencialidad. La polisemia metafórica sigue teniendo una función positiva para la economía del sistema, pero ya no está tan constreñida por ello: al haberse auto-organizado el sistema, ha surgido una jerarquía que nos permite lidiar con el alto número de elementos léxicos.

En definitiva, la metáfora cognitiva surgiría como consecuencia de la pérdida de referencialidad del sistema, lograda a través de su auto-organización. La metáfora sería un proceso general posibilitado por el lenguaje.

3.2.3 Resumen

A continuación, y a modo de recapitulación de las ideas hasta aquí expuestas, se retoma el esquema anteriormente presentado, añadiendo la información sobre el surgimiento de la ambigüedad:

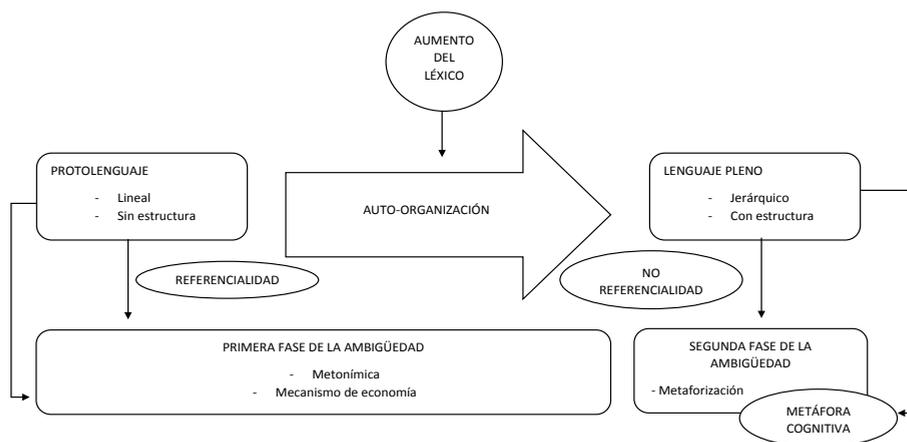


Figura 2. Surgimiento de la ambigüedad

Como vemos en la figura anterior, la propuesta que presentamos considera que la ambigüedad surge en una primera fase (en el protolenguaje) como un mecanismo de economía. De esta manera, se daba una solución al aumento del léxico, que acabaría produciendo la auto-organización. Esta primera

ambigüedad es exclusivamente metonímica, puesto que estamos en momento en el que el protolenguaje es totalmente referencial. La segunda fase de la ambigüedad, consistente en la polisemia metafórica, se posibilita mediante la auto-organización del sistema y el surgimiento del lenguaje pleno. En este punto, al perderse la necesidad de la referencialidad, surge la capacidad metafórica, como un recurso supeditado al lenguaje pleno y su estructura.

4. CONCLUSIONES

En este trabajo se ha presentado un estudio preliminar a cerca de dos preguntas claves sobre la ambigüedad léxica: por qué existe y cómo pudo surgir⁸.

Para responder a la primera pregunta, se han presentado diferentes aproximaciones, que consideran la ambigüedad como un elemento esencial a la hora de estructurar y cohesionar el lexicón mental, como un elemento comunicativo que equilibra los deseos de los participantes de la comunicación y como un mecanismo que facilita la navegación por las redes semánticas.

En cuanto a la segunda pregunta, se ha planteado un primer acercamiento al tema de la ambigüedad desde un punto de vista evolutivo, enmarcado en la rama de las Ciencias de la Complejidad. Así, se han distinguido dos fases en su surgimiento que podrían corresponder, además, con los dos tipos de polisemia.

Tras todas estas reflexiones podemos afirmar que, como dice el título del artículo, el lenguaje será ambiguo o no será.

⁸Tal y como ha señalado un revisor anónimo, somos conscientes de las limitaciones de nuestro estudio. Sin embargo, esta aproximación a la ambigüedad desde un punto de vista evolutivo es simplemente un primer paso de nuestro estudio, que necesariamente deberá ser ampliado con bibliografía más allá de las Ciencias de la Complejidad (por ejemplo, sobre el modelo evo-devo).

REFERENCIAS BIBLIOGRÁFICAS

- BENÍTEZ BURRACO A. y BARCELÓ-COBLIJN LL. 2015. *El origen del lenguaje*. Madrid: Editorial Síntesis.
- CHOMSKY N. 1988. *Language and problems of knowledge*. Cambridge, MA: MIT Press.
- COHEN A. 2006. "Why ambiguity?", *Between 40 and 60 Puzzles for Manfred Krifka*, H. Gaertner, S. Beck, R. Eckardt, R. Musan, B. Stiebels (eds.). Berlin: Zentrum für Allgemeine Sprachwissenschaft, Typologie und Universalienforschung.
- COROMINAS-MURTRA B., FORTUNY ANDREU J. y SOLÉ R. 2011. "Emergence of Zipf's Law in the Evolution of Communication", *Physical review. Statistical, nonlinear, and soft matter physics*, 83(3): 1-7.
- FELLBAUM C (ed). 1998. *WordNet: An Electronic Lexical Database*. Cambridge, MA: MIT Press.
- LONGA V.M. 2001. "Sciences of complexity and language origins: an alternative to natural selection", *Journal of Literary Semantics*, 30(1): 1-17.
- LONGA V.M. 2011. "La evolución del lenguaje desde la perspectiva chomskiana", *Revista UIS-Humanidades*, 39(2): 11-37.
- LÓPEZ-CORTÉS N. 2018. "Reducir, reutilizar y reciclar las palabras, o de por qué existe la ambigüedad", *Ciencia Cognitiva*, 12(3): 67-69.
- PIANTADOSI S., TILY H. y GIBSON E. 2012. "The communicative function of ambiguity in language", *Cognition*, 122(3): 280-291.
- PINKER S. y BLOOM P. 1990. "Natural language and natural selection", *Behavioral and Brain Sciences*, 13(4): 707-784.
- PUSTEJOVSKY J. 1995. *The generative lexicon*. Cambridge, MA: The MIT Press.
- SHANNON C. 1948. "A mathematical theory of communication", *Bell System Technical Journal*, 27: 623-656.
- SIMON H.A. 1962. "The architecture of Complexity", *Proceedings of the American Philosophical Society*, 106(6): 467-482.
- SOLÉ R. 2009. *Redes complejas: del genoma a Internet*. Barcelona: Tusquets Editores.

- SOLÉ R., COROMINAS-MURTRA B., VALVERDE S. y STEELS L. 2010. "Language Net works: their structure, function and evolution", *Complexity*, 15(6): 20-26.
- SRINIVASAN M., RABAGLIATI, H. 2015. "How concepts and conventions structure the lexicon: cross-linguistic evidence from polysemy", *Lingua*, 157: 124-152.
- ULLMANN S. 1976. *Semántica. Introducción a la ciencia del significado*. Madrid: Aguilar.
- VIVANCO CERVERO V. 2003. *Homonimia y polisemia: teoría semántica y aplicación lexicográfica*. Buenos Aires: Ediciones del Sur.
- WASOW T., PERFORS A., BEAVER D. 2005. "The puzzle of ambiguity", *Morphology and the Web of Grammar: Essays in Memory of Steven G. Lapointe*, C.O. Orgun, P. Sell (eds.). Chicago: CSLI Publications. 265-282.
- YOUN H., SUTTON L., SMITH E., MOORE C., WILKINS J. F., MADDIESON I., CROFT W., BHATTACHARYA T. 2016. "On the universal structure of human lexical semantics", *Proceedings of the National Academy of Sciences*, 113(7): 1766-1771.
- ZIPF G. 1949. *Human behavior and the principle of least effort*. New York: Addison-Wesley.

ESTUDIO DEL PROCESO DE ADQUISICIÓN DEL ESPAÑOL EN UN NIÑO BILINGÜE ÁRABE – ESPAÑOL: EL SINTAGMA NOMINAL

Khaled Ahmed Azzouz
Universidad de Valladolid

ABSTRACT

The importance of studies on language acquisition. According to Miguel Serra (2000) linguistic acquisition is investigated by psycholinguists, linguists and experts from other areas; since with the study of these aspects can on the one hand, offer solutions for children who have difficulties in acquiring the language—regardless of what the cause is—and facilitate their learning; and, on the other hand, to reveal very significant data that serve for the study of the learning of the second languages.

Keywords: Acquisition – bilingualism – language – grammar – noun phrase.

RESUMEN

La importancia de los estudios sobre la adquisición del lenguaje. Según Miguel Serra (2000) la adquisición lingüística es investigada por psicolingüistas, lingüistas, sociolingüistas y expertos de otras áreas; ya que con el estudio de dichos aspectos pueden por una parte, ofrecer soluciones para los niños que tengan dificultades en adquirir el lenguaje—independientemente de cuál sea la causa— y facilitar su aprendizaje; y, por otra parte, revelar datos muy significativos que sirven para el estudio del aprendizaje de las segundas lenguas.

Palabras clave: Adquisición – bilingüismo – lenguaje – gramática – sintagma nominal.

1. OBJETIVOS DEL ESTUDIO

Los objetivos de nuestro estudio son; recoger los datos del proceso de adquisición lingüística del español desde los 3 años hasta los 5 años y 7 meses de un niño nacido en España en el seno de una familia árabe, así como describir la adquisición de la morfología flexiva del sintagma nominal y el cuadro cronológico de la adquisición de la gramática producido por el sujeto y comparar todo ello con los de otros estudios anteriores.

2. ESTADO DE CUESTIÓN

Los estudios longitudinales sobre la adquisición del español más importantes que hemos tenido en cuenta a la hora de realizar nuestro estudio son: López Ornat (1977) (1994), Hernández Pina (1984), Pérez Pereira (1991), Mariscal (1977) y Gerardo Aguado (2010) (2014), entre otros, para abordar la adquisición en general, dado que los estudios sobre la adquisición del sintagma nominal en niños bilingües árabe / español son casi nulos. Excepcionalmente hemos encontrado un estudio sobre el sintagma determinante en dos sujetos marroquíes, es el trabajo longitudinal de Estrella Nicolás de Benito (2014).

3. METODOLOGÍA

La metodología longitudinal utilizada se ha pretendido adecuar al tema del estudio. Las grabaciones empezaron el mes de noviembre del 2010 y finalizaron en marzo 2013. Los datos de nuestro estudio se componen de 57 filmaciones, realizadas por el padre quien en la mayoría de los casos se ha encargado de realizar las grabaciones audiovisuales, aunque la madre ha registrado algunas filmaciones de 2012 (por ejemplo, cuando la familia estaba de vacaciones en Libia).

El corpus contiene todos los datos de las conversaciones del sujeto con las personas implicadas, las producciones del sujeto basadas en escenas casi siempre espontáneas, (como cuando juega en casa y fuera en casa con sus amigos). Los datos del corpus están ordenados cronológicamente con la fecha de la realización de la filmación.

En cada vídeo, tenemos las transcripciones del sujeto y de las personas implicadas, el análisis morfológico y el análisis sintáctico.

Exponemos una muestra del corpus según el formato y la versión de la página de CHILDES que estamos trabajando para su publicación.

4. MUESTRA DEL CORPUS

Hemos empezado a trabajar los datos del corpus según el modelo de CHILDES. Estamos en ello para que los datos del corpus estén disponibles para los investigadores y la comunidad científica.

@Begin:
@Participants: Chi Target_ Child, INV Fahter Investgator
@Age of CHI: 3;1:0
@Birth of Participant: 26/10/2007
@Birth of Participant: 10/03/1974
@Coder: Khaled Azzouz
@Education: Pre-Kindergarten
@Language: Spanish and Arabic
@SES of Chi:middle
@Filename: SAM1
Sex of Participant: male
@Date: 2/11/2010
@Location: Valladild. Spain
@Situation: En casa juega con su padre.

- *CHI: ره تراه بابا [= x *A ver papá, A ver papá*]. (Su padre puso la cámara y él quiere ver lo que hace su padre).
- *INV: كعمر العب ما في شي [= x *Siéntate y juegas no hay nada*].
- *CHI: Mira papá mira mira.
- *INV: تعال اهنيه العب اهنيه [= x *Ven aquí a jugar aquí*].
- *CHI: Mira papa mira.
- *INV: اه تبي الخيط [= x *Ah quieres la cuerda*].
- *CHI: خيط [= x *cuerda*].
- *INV: من عنده الخيط؟ [= x *¿Quién tiene la cuerda?*].
- *CHI: Tú.
- *INV: اني معنديش [= x *Yo no lo tengo*].
- *INV: غدوة نشريلك [= x *Mañana te compro una cuerda*].
- *CHI: برا خيط بره [= x *La cuerda fuera*]. (Se refiere a comprarlo de fuera).
- *CHI: Mila شوف [= x *mira*] # mila [mira]. (Dice las dos palabras simultáneamente en árabe y en español y enseñando a su padre su juguete).
- *INV: هات نوا اندير هولي [= x *Dame te lo hago*].
- *CHI: Mira # Mira.
- *INV: غدوة نشريلك وحدة جديدة [= x *Mañana te compro otro nuevo*].
- *INV: ¿Quién te compró este juguete?
- *CHI: Sara
- *INV: Sara
- *CHI: ايه [= x *sí*]
- *INV: شن قلنا احني على الشبشب؟ [= x *¿Qué hemos dicho sobre llevar la zapatilla?*].
- *CHI: Mira
- *INV: راجي بعدين اندير هولك [= x *Luego te lo hago*].
- *INV: ¿Quién te lo traído esta juguete?
- *CHI: Sara.
- *INV: هذي ما تخذ مش [= x *Este no funciona papa*].
- *CHI: no xxxxxx بيبي [= x *bebe*].
- *INV: امتع البيبي # ما فيهاش [= x *es del bebé*].
- *CHI: امطعي بيبي اعيط [= x *bebé llora*] [= x *mía*].
- *INV: حطها فوق الطاولة. [= x *ponla sobre la mes.*] [= x *¿Es tuyo este?*]. امطعك انت هذي [= x *así luego la lleva*].
- *CHI: ريموت امطعي [= x *Mando a distancia mía*].
- *INV: لكن اني [= x *Quiere un coche con mando de distancia?*] [= x *Pero te dije antes si recoges el mando del suelo te daba el coche con el mando*] [= x *Te lo doy para jugar?*]. نعطيك نعطها لك تلعب بها
- *CHI: باهي [= x *vale*].

- *INV: هاك العيب [= x *Toma juegas*].
- *CHI: بابا بابا [= x *papá papá*]. (Enseñando a su padre el juguete).
- *INV: هذي تبي بتاريه غدوة نشريلك باتريات [= x *Este necesita batería mañana te compra baterías*].
- *INV: ¿Cómo te llamas?
- *CHI: Ahmed.
- *INV: و خوك [= x *Y Tu hermano*] ¿cómo se llama? (La palabra hermano se lo dijo en árabe).
- *CHI: Mohamed.
- *INV: ¿Y mama como se llama?
- *CHI: Mama.
- *INV: لا عندها اسم ثاني [= x *No, tiene un nombre*].
- *CHI: Amira.
- *INV: ايوه اميرة [= x *Exactamente Amira*] ¿Y papá como se llama?
- *CHI: Jaled.
- *INV: ¿Y tú cómo te llamas?
- *CHI: Ahmed.
- *CHI: هذه بابا معاش تخدم شوف غدوة نشرو بتريات [= x *Este papá ya no funciona, mira mañana compramos baterías*].
- *CHI: NO.
- *INV: ¿Como no!
- *INV: اندسها بيش غدوة نلعبو بيها [= x *La guardamos para que mañana juguemos con ella*] غدوة نشريلك بتريات [= x *Mañana te compro baterías*].
- *CHI: ببي ايعيط [= x *Bebé llora*] (Va corriendo para ver el bebé).
- *INV: بشويه [= x *Despacito*]
- *CHI: الضئى [= x *La luz*] (Viene corriendo diciendo la luz).
- *INV: ¿خيره الضئى شن دخله [= x *¿Qué pasa a la luz?*].
- *INV: تعال خوذ الرموت من الارض خير ما يتكسر [= x *Ven a coger el mando del suelo para que no se rompa*, كان هذا معاش انخليك تلعب بيه [= x *Sino no te dejo jugar con el mando*] اطفيه و حطه ايوه [= x *pon lo arriba en la mesa*] حطه فوق الطاولة [= x *apagalo y pon lo, muy bien*] فوق اكثر يتكسر [= x *arriba más se rompe*].
- *CHI: Hola, Dora Dora.
- P: ركب حتى التانيات زيد غادي و ديرهم [= x *Monta también las otras, vete más allí y lo haz*].
- *CHI: Dola.
- *CHI: aaaam mmmm.
- @End:

5. ADQUISICIÓN DE LA MORFOLOGÍA LIGADA Y LIBRE DEL ESPAÑOL

5.1 Adquisición del género

El género se refiere en español a las dos categorías de masculino y femenino. Se puede marcar con la terminación de las vocales *a*, *e* u *o* y las consonantes *d*, *l*, *n*, *r* o *z*. La norma es particularmente complicada, puesto que hay excepciones y existen palabras que terminan con la misma vocal *o* con la misma consonante y pueden ser masculinas o femeninas. Por ejemplo, en el caso de la vocal *e* y de la consonante *z* pueden marcar tanto una palabra masculina como una femenina (*el coche y la calle; la nariz y el lápiz*). La adquisición del género en los niños, según investigadores como Aguado (2010) (2014), Hernández Pina (1984), o Sonia Mariscal (1996) (2014), no plantea problemas. Hernández Pina (1984) afirma que el sujeto no distinguía el género de los primeros nombres que emitió hasta la etapa de dos palabras.

5.1.1 En el caso de nuestro estudio

Hemos notado que el sujeto emite los primeros nombres sin darse cuenta de que existe una norma para femenino-masculino, hasta la etapa de dos palabras, donde ya distinguía el género de las palabras, pero cometió errores de todo tipo en la concordancia de género.

Ha podido llegar a reconocer el género de las palabras terminadas en consonantes como *l*, *r*, *n*, y *z*. En los siguientes ejemplos, se puede asegurar que el niño distingue las palabras masculinas terminadas en *o* de las palabras femeninas *a* que van antepuestas de un artículo determinado o indeterminado. Del mismo modo, distingue el género de las palabras terminadas en *e*, unas veces masculinas y otras, femeninas:

5.1.2 Uso correcto de determinantes con nombres

ON15: Jugando xxxxxxxx la pupa. (3:02)

ON73: Quiero un coche. (3:06)

- ON60: Papá papá quiero el coche. (3:07)
 ON51: Cuando la noche xxxxxxxxxx una y dos xxxxxx mayores. (3:07)
 ON52: La oreja. (3:09)
 ON7: Es una mariquita. (3:09)
 ON44: El dormir, voy mirar el fútbol. (3:10)
 ON51: En la calle. (4:00)
 ON35: Un pantalón. (4:01)

5.1.3 Errores de comisión

Hemos observado que el niño en aquellos casos en que las palabras no terminaban ni en *a* ni en *o* mostraba dificultades en la utilización del género. Este hecho, lo podemos observar en los siguientes ejemplos: con *e* (*la pie*) 3:07, en los casos que debería seguir la regla general del uso del género: *a* implica emplear un artículo femenino y *o* implica el uso del artículo masculino. El sujeto todavía presenta ejemplos de inestabilidad en el uso de la regla. Esto genera errores de comisión como los siguientes (*una amarillo*) 4:06 (*Un caja*) 3:07 (*este pala*) 3:07 (*El pelota pequeña*) 3:0 (*unas peces*). Aunque también, se han registrado casos en los que no se aplica la correlación habitual en la lengua española entre vocal de tema *o/a* y género masculino y femenino, respectivamente, y que el sujeto produce correctamente (*el agua*) 3:06 (*En la mano*) 3:10 (*la foto*) 4:04. Asimismo, se ha registrado un caso de uso de artículo con verbo infinitivo (*El dormir*).

5.2 La adquisición del número

La adquisición del número con las palabras terminadas en consonantes se ha producido en una etapa muy tardía y ha salido a la luz con algunas consonantes —como *r*, *l* y *n*— aplicando la marca *es* para el plural. Hernández Pina (1984) como otros autores —Aguado Alonso (2010) y (2014); Mariscal Altares (1996)— confirman, como lo hacemos en nuestro estudio, que la adquisición del número ha sido normal cuando el sustantivo termina en vocal y, en menor medida, cuando acaba en consonante.

La producción de palabras terminadas en *s* y *es* ha sido escasa. Las palabras terminadas en las consonantes *d* y *z*, no han aparecido.

En el caso de la adquisición del número de nuestro sujeto, la primera palabra producida en plural termina con *es* en una edad muy temprana a los 3:02 años y parece que la ha aprendido sin aplicar la regla. Se produce antes de que aparezcan palabras terminadas en *o / a*. Por ejemplo, *coches* la ha producido cuando estaba jugando a carreras de coches.

5.2.1 Error en concordancia del número

En cuanto a los errores de concordancia de número, el sujeto comete errores en: el uso de un determinante femenino con un nombre masculino en plural (*una mocos*) (3:06); un caso de uso de determinante masculino de plural con un nombre masculino en singular (*los niño*) (3:07); también un determinante masculino con un nombre masculino en plural (*El bebes*) (4:05) y, otro caso, pero al contrario (*mucho esqueletos*) (5:01). A continuación, exponemos los enunciados con errores:

ON70: Mira que tengo una mocos aquí. (3:06)

ON17: bailando como los niño. (3:07)

ON4: El bebes. (4:05)

ON90: Y he visto mucho esqueletos. (5:01)

5.3. La adquisición de los pronombres

Nuestro sujeto de estudio usa los pronombres personales para dar referencia a la persona que habla con él y del tiempo. Como ejemplos encontramos: 3:01 (ON5: *tú*); 3:02 (ON7: *mira yo he pintado Nupa*) (ON66: *yo mío*) a la edad 3:04 (ON3: *pues, yo la quiero así*); (ON31: *ahora tú*) a la edad 3:05 (ON26: *yo aquí*) 3:07 (ON34: *yo partido*) 3:07. Los pronombres demostrativos y los posesivos son los primeros pronombres en aparecer en una etapa muy temprana, a los 3:02 meses. Los pronombres de primera y segunda persona del singular *yo tú* de forma exhaustiva, en comparación con el resto de los pronombres personales.

El uso de *yo tú* del singular en la mayoría de los casos se han producido con adverbios de tiempo (*aquí, ahora*) y a veces con verbos conjugados en la 1ª P, la 2ª P, y la 3ª P.

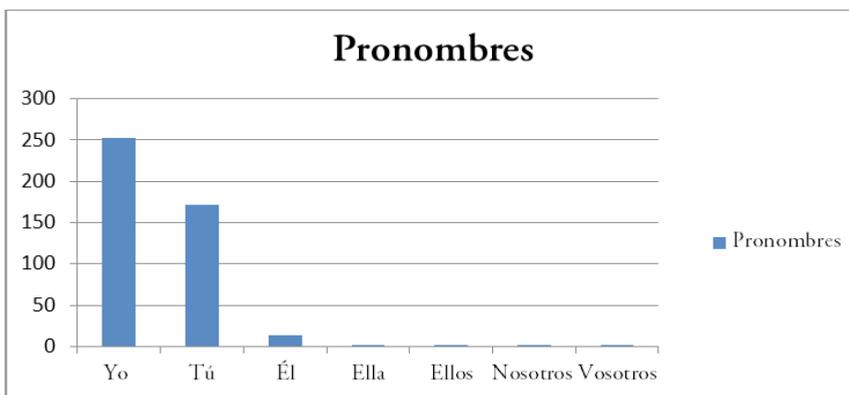


Gráfico 1. Pronombres adquiridos por el sujeto

En los estudios de Hernández Pina (1984); Larrañaga (1994); Shum, Conde y Díaz (1992); y Aguado Alonso (2010) (2014) todos concidieron en el uso frecuente de *yo* y *tú* en comparación con los otros pronombres personales, y afirmaron la aparición de los pronombres personales con los verbos en la etapa de dos palabras.

Hemos notado en nuestro estudio que en la mayoría de los casos con verbos el sujeto omite el pronombre personal. Creemos que la versión adulta del padre (que emplea el español con el niño) y la madre (que emplea el árabe) han podido influir. Estas dos lenguas, el español y el árabe, son lenguas *pro drop*, por lo tanto, el sujeto casi nunca emplea los pronombres personales en función de sujeto con los verbos.

También es reseñable la aparición de los pronombres demostrativos y los pronombres posesivos (esto, este, ese, esta, esa, mío, mía) en comparación con los determinantes demostrativos y posesivos, y en comparación con otros estudios de otros sujetos. En cuanto a los pronombres interrogativos aparecidos son: *qué*, *por qué*, *de quién*, *para quién*, *cuál* y el pronombre *dónde* y son usados con frecuencia.

Hernández Pina afirmó la aparición en caso de su estudio de *de qué* y *por qué* en la etapa temprana. Este hecho coincide con nuestro estudio. El pronombre interrogativo *cuál* fue usado por el sujeto mientras que en el estudio de Hernández Pina no aparece.

Los pronombres personales de objeto directo e indirecto *la*, *lo*, *le* han aparecido con un uso regular en la etapa temprana, antepuestos en la mayoría

de los casos a verbos en modo indicativo y, en algunas ocasiones, pospuesto a verbos en imperativo.

Los pronombres *me*, *te* y *se* han aparecido muy temprano; mientras que los pronombres *os* y *nos* lo hicieron en una etapa muy tardía.

Hernández Pina (1984) afirmó la aparición de los pronombres de formas objetivas *me*, *te* y *se* en la etapa temprana y la ausencia de los pronombres *os* y *nos* en el repertorio del sujeto. Asimismo, en el estudio de Shum, Conde y Díaz (1992) los pronombres *os* y *nos* tampoco han aparecido.

Los pronombres tónicos y átonos en los datos de la adquisición de estas categorías en nuestro sujeto no se han diferenciado de los datos comparados de otros sujetos presentados en nuestro trabajo. Lo mismo podemos afirmar durante la adquisición de los pronombres donde el sujeto cometió errores de tipología diversa.

edad	personales	numerales	posesivos	demostrativos	indefinidos	interrogativos	exclama- tivos
3:01	tú						
3:2	yo	dos tres cuatro cien	mío mía	esto este	otro todos	qué	
3:4	la te			ese esta		por qué	
3:4							
3:5	lo me				otra una	de quién para quién quién	
3:6	le	una			alguien	cuál	
3:7	él						
3:8	ella a mi se						
3:9	ellos					dónde	
3:10				esa			
3:11							
4:00	nos						

4:10	conmigo						
4:11	vosotros						
5:00	los						
5:01	os a ti				algunos		
5:02	nosotros						

Tabla 1. Edad de aparición de los pronombres

5.4. La adquisición de las preposiciones

Las preposiciones son palabras invariables, es decir, no tiene ni género ni número, su función principal es relacional. Según Serra (2000:371), las preposiciones se clasifican en dos tipos: preposiciones léxicas o plenas (*sobre*) y funcionales o vacías (*de*).

Edad	3:02	3:04	3:06	3:07	3:08	3:09	4:03
preposición	de	en	a cd	de	hasta	con	a ci
	a + infinitivo		por				
	para+ infinitivo						

Tabla 2. Preposiciones adquiridas por el sujeto

Exponemos las preposiciones adquiridas por nuestro sujeto y por otros sujetos estudiados por otros investigadores. De este modo, podemos obtener una idea clara sobre la adquisición de las preposiciones. Casi todos los sujetos pasan por el mismo proceso de adquisición en cuanto a esta categoría.

Preposición Sujeto	Hernández Pina	Aguado (2014) ¹	Gili Gaya ²	Aguado (2010) ³	Pilar Larrañaga	Melina Aparici
de	a	en	a	A	para	con
a + infinitivo para + infinitivo	en	con	con	En	en	a
en	de	a	de	De	de	para
por	para	a cd	en	Para	a	de
a cd	por	de	para	Con	por	en
de tiempo	hasta	para	por	Por		
por		sin	sin	hacia		
hasta		hasta	hasta			
con						
a ci						

Tabla 3. Preposiciones adquiridas por nuestro sujeto y por otros sujetos

En cuanto a las múltiples divergencias entre los investigadores en la aparición de las preposiciones en mente, encontramos que en el caso de nuestro sujeto la primera preposición en aparecer ha sido *de*, mientras que dos investigadores señalan que fue la preposición *en* y, otros dos, que ha sido la preposición *a*. Asimismo, encontramos divergencias en la aparición de las preposiciones *a*, *en*, *de* y *con*. También exponemos el siguiente gráfico, donde podemos observar la frecuencia de uso de la como prototipo de preposición o prototipo de determinante, y la frecuencia de uso del resto de las preposiciones.

¹El estudio de Gerardo Aguado (2014) mencionado por Mariscal Altaras se trata de un estudio sobre 24 sujetos de la frecuencia de uso de las preposiciones.

²El estudio de las preposiciones de Gili Gaya está extraído del libro Miguel Serra.

³El estudio de Gerardo Aguado (2010) es un estudio sobre 35 niños.

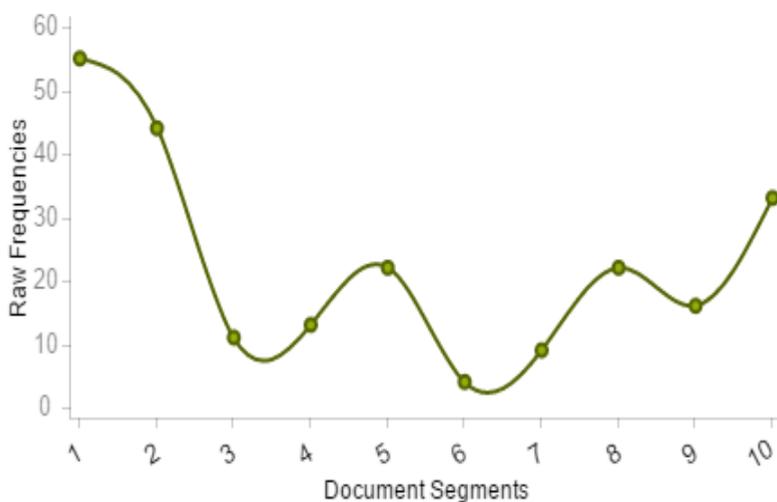


Gráfico 2. Partícula fónica [ə]

La partícula fónica [ə] prototipo de preposición o prototipo de determinante ha sido mencionada por muchos investigadores. Hernández Pina (1984: 256) la partícula *a* antepuesta de nombre y antepuesta de verbos (*a dormir*) expresiones fijas. Solias Arís (2013:3512) mencionó apoyos fónicos precediendo al nombre que han recibido el nombre de *protodeterminantes*, indicando dirección, lugar, procedencia, objetivo, etc. Delante de verbos, en estos casos se les denomina protopreposiciones. La forma fónica es la de /o/, /a/, /e/ o /ə/, con alternancia entre ellos en el mismo sujeto. Según Mariscal Altares (2014: 267), la omisión del artículo con el nombre incluía una pieza fónica antepuesta a los nombres, forma de las vocales /a/e/o/ —rara vez /u/ y nunca /i/— y posteriormente, en la forma adulta de /el/la/un/una/.

Teniendo en cuenta las referencias anteriores, se podría confirmar que existe una producción temprana en forma de preposición generalizada a través del uso de la partícula fónica [ə] que sirve como prototipo de preposición.

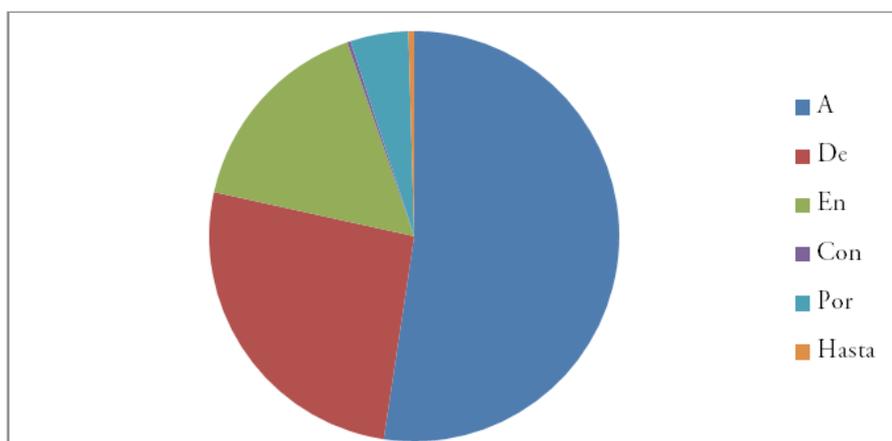


Gráfico 3. Preposiciones adquiridas por el sujeto

5.5. Adquisición de los determinantes:

Los determinantes determinados *la* y *el* han aparecido con más frecuencia en el habla del sujeto en comparación con los indeterminados *un* y *una*.

Nuestros datos coinciden en parte con los datos de Hernández Pina (1984) sobre su afirmación en que los primeros artículos adquiridos son *el* y *la*; en cambio, la autora afirma que los indeterminados más usados son *un* y *una*. Díaz y Cortés (1992) afirman que los determinantes determinados e indeterminados son las primeras categorías adquiridas después del género y el número.

Han aparecido las formas 1ª P y 2ª P del plural *mis*, *tus*; mientras que en el estudio de Gerardo Aguado Alonso solo aparecen las singulares *mi*, *tu* y *su*. Hernández Pina (1984) también afirmó la aparición de los determinantes posesivos *mi*, *tu* y *su* en singular y *mis* y *tus* en plural. Hemos notado el escaso uso de los demostrativos en comparación con los determinantes determinados e indeterminados.

No han aparecido las formas de lejanía *aquel*, *aquella* en singular y en plural; mientras que en el caso de Rafael de Hernández Pina, casi ha producido todos los demostrativos menos *estos* y *aquellos*, los plurales no figuraban en el repertorio del sujeto.

Los determinantes numerales aparecidos fueron: *dos*, *cinco*, *cuatro* y *primero*, este último lo usó de forma incorrecta antepuesto a un sustantivo (*primero gol*).

Los determinantes indefinidos aparecidos en nuestro caso son: *otra, otro, muchos y muchas*.

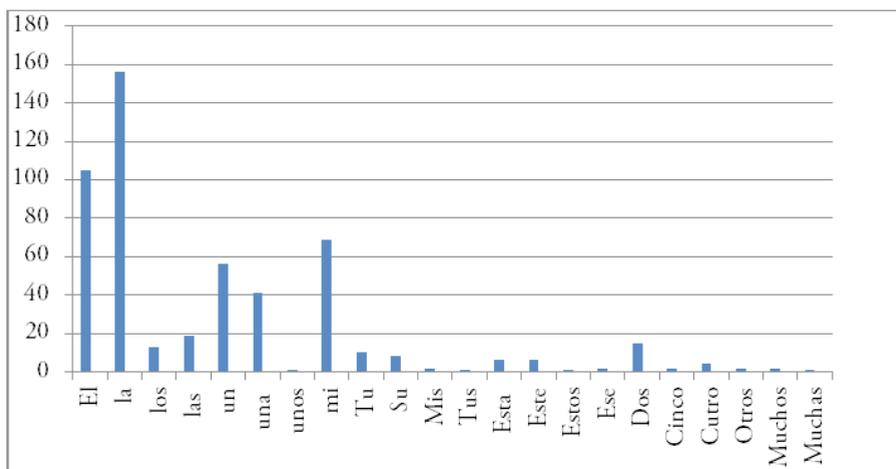


Gráfico 4. Determinantes adquiridos por el sujeto

Artículos				Posesivo		Demostrativo		Numeral		Indefinido	
indefinido		definido									
Una	3:09	La	3:02	Mi	3:05	Esta	3:02	Dos	3:02	Otra	3:05
Un	3:06	El	3:02	Tu	3:08	Este	3:02	Cinco	3:11	Otro	4:45
Una	4:09	Los	3:06	Su	3:06	Ese	4:09	Cuatro	3:11	Muchas	4:02
Unos	4:11	Las	4:01	Mis	4:05	Estos	4:03	Primer	3:08	Muchos	4:04
				Tus	4:05						

Tabla 4. Determinantes adquiridos por el sujeto

6. ASPECTO SOCIOLINGÜÍSTICO

Hemos notado varias cuestiones de gran importancia en el estudio. Una de ellas es el uso de una lengua en lugar de la otra. Nuestro sujeto antes de estar expuesto a la lengua española ha producido sus primeras palabras en árabe.

Según nuestros datos, podemos afirmar el uso de deícticos por parte del niño —como adverbios de lugar y tiempo, pronombres demostrativos y verbos— para comunicarse. Nosotros creemos que los deícticos son una estrategia que emplea el niño para comunicarse como fase previa. Las primeras palabras que usaba el sujeto son los deícticos en árabe como (ليه *sí*, شوف *mira*, لا *no*, امتعي *mío*, باهي *vale*) y otras palabras más, todas estas palabras deícticas mencionadas las ha remplazado por palabras deícticas en el español después de un periodo de tiempo muy corto. Creemos que los deícticos son una estrategia que emplea el niño para comunicarse como fase previa. Muchos autores, como Idiazabal (1994), Hernández Pina (1984), Pérez Perera (1996), han afirmado también en sus estudios en sujetos monolingües y bilingües el uso de deícticos.

Durante el desarrollo de la adquisición de la lengua española el sujeto rechaza el uso de su lengua materna el árabe, aunque sus padres siempre le hablaban en la lengua árabe.

7. ALTERNANCIA DE CÓDIGOS

El uso de la alternancia de códigos es un proceso muy frecuente en la adquisición de dos lenguas. Hemos podido mostrar en nuestro estudio la tipología de alternancia de código con ejemplos producidos por el sujeto en la primera fase de la adquisición del español y también hemos mencionado la fase en la que han desaparecido las alternancias de códigos. En principio, el sujeto utiliza algunas palabras deícticas de la lengua materna el árabe. A la edad 3:2 hace una mezcla entre dos elementos gramaticales (ON12: هكي *así vamos*) se trata de un verbo conjugado en la 3 P^o en plural en español y un adverbio de modo producido en la lengua árabe. En el segundo caso, (ON2: *Estrella* هني *est*), usa un pronombre demostrativo producido en la lengua árabe con un sustantivo en la lengua española. Otros casos como los siguientes: (ON60: *Manzana no pez* حوته *n*) (ON61: *Este no* هذه *este* هذه *no*).

8. CONCLUSIONES

Muchos investigadores argumentan sobre la adquisición del lenguaje en los niños que en el proceso de los estudios del lenguaje casi siempre existen diferencias entre los sujetos. Las diferencias se deben a varias razones como al nivel socio-cultural, al nivel socio-económico, y la edad.

En nuestro estudio, hemos incidido en el análisis de la adquisición del sintagma nominal y de sus estructuras gramaticales a través de tres elementos relacionados con el proceso de aprendizaje: uso correcto-comisión-omisión. También, hemos empleado la comparación con otros sujetos.

En este estudio de adquisición de la lengua española según nuestros datos, sobre la aparición del género y el número estamos de acuerdo con Hernández Pina (1984), Mariscal (1996), Pérez Pereira (1996), Aguado (2010; 2014), López Ornat (1994), y Solís Arís (2013) en que el número se adquiere una vez adquiere el género. Determinados elementos de categorías gramaticales como pronombres, preposiciones y determinantes aparecen en caso de nuestro sujeto y no aparecen en los otros sujetos estudiados.

Hemos notado diferencia en algunas categorías gramaticales, que todavía no ha adquirido el sujeto, pero la expectativa es que aunque no aparezcan en este periodo de estudio, las aprenderá en el futuro.

En el caso del cambio de una lengua a otra, creemos que esta situación de rechazo, por una parte, es una repuesta por parte del sujeto en la que ve la lengua de sus padres es una lengua minoritaria, y por otra parte, por la fuerte exposición lingüística que ha tenido el sujeto en la sociedad donde vive.

REFERENCIAS BIBLIOGRÁFICAS

- AGUADO OREA, Javier. (2000). *Adquisición de los complementos pronominales personales en español*: Universidad Autónoma de Madrid Serv.
- AGUADO ALONSO, Gerardo. (2014). *El desarrollo de la morfosintaxis en el niño*. Madrid: General Paradiñas.
- AGUDO ALONSO, Gerardo. (2010). El desarrollo del lenguaje de 0 a 3 años, bases para un diseño curricular en la educación infantil. Madrid: Matería.
- AGUIRRE MARTÍNEZ, María del Carmen y MARISCAL ALTARES, Sonia. (2001). *Cómo adquieren los niños la gramática de su lengua: perspectivas teóricas*. Madrid: Paracuellos de Jarama.
- AGUIRRE MARTÍNEZ, Carmen y MARISCAL ALTARES, Sonia. (2001). *Como adquieren los niños la gramática de su lengua: perspectivas teóricas*. Madrid: Uned.
- ÁLVAREZ DE LA FUENTE, Esther. (2007). *Análisis lingüístico de la traducción natural: datos de producción de niños gemelos bilingües inglés / español*. Tesis doctoral.
- ÁLVAREZ MARTÍNEZ, María Angeles. (1989). *Pronombre, (El). Personales, artículos, demostrativo, posesivo*. Madrid. Arco libros.
- ALVINO FANTINI, E. (1982). *La adquisición de lenguaje en un niño bilingüe. Barcelona*. Herder.
- AMUSATEGI ROTAETXE, Karmele. (1988). *Sociolingüística*. Madrid: Síntesis.
- APARICI AZNAR, Melina, Gloria DÍAZ y Montserrat CORTÉS. (1996). *El orden de la adquisición de morfemas en catalán y castellano*. En M. Pérez Pereira (ed.), *Estudios sobre la adquisición del castellano, catalán, euskera y gallego*, Santiago de Compostela, Universidad de Santiago de Compostela.
- APPEL, René y Pieter MUYSKEN. (1988). *Bilingüismo y contacto de lenguas*. Editorial: Ariel.
- APPEL, René y Pieter MUYSKEN. (1987). *Language Contact and Bilingualism, El Bilingüismo y contacto de lenguas*, Traducción y versión española de Anexo M. Lorenzo Suárez y Clara I. Bouzada Fernández. 1º Edición: septiembre 1996.

- ARADOS PERAITA, Herminia. (1989). *Adquisición del lenguaje*. Madrid: Universidad Nacional de Educación a Distancia.
- AROSTEGUI ETXEBARRIA, Maitena. (1995). *El bilingüismo en el estado español*. Bilbao: Ediciones FBV, S.I.
- BAKER, Colin. (1997). *Fundamentos de educación bilingüe y bilingüismo*. Traducción de Ángel Alonso –Cortés. Madrid: Cátedra.
- BARALO OTTONELLO, Marta. (1999). *La adquisición del español como lengua extranjera*. Madrid: Arco Libros.
- BARREÑAS, H. (1999). *Ausencia de reglas gramaticales en las primeras producciones lingüísticas de niños vascos, Infancia y aprendizaje*. (núm 86, página 91-117).
- BARRERA, L y FRACA, L. (1991). *Psicolingüística y desarrollo del español*. Caracas: Monte Ávila.
- BERKO GLEASON, Nan Berstein Ratner. (1999). *Psicolingüística*. Traducción Ana Esquinas Rychen. 2 edición.
- BERNÁRDEZ, Enrique. (1999). *¿Qué son las lenguas?* Madrid: Alianza.
- CAPDEVILA, M. y SERRA, M. (1996). *Estudio de las categorías funcionales en niños catalanes*. Madrid: Unid.
- CHOMSKY, N. (1957). *Syntactic structures*, La Haya, Mouton & co (segunda impresión 1962).
- CONTRERAS GONZÁLEZ, María. (2003). *Dificultades lingüísticas en los trastornos del lenguaje frente a la desventaja medioambiental*. Universidad de Granada. Tesis doctoral.
- CORTÉS, Montserrat. (1993). *Algunas reflexiones en torno a la adquisición e la estructura argumental*. Anuario de Psicología. 1993, no 57, 13-27 O 1993, Facultat de Psicologia Universitat de Barcelona.
- CREGO GARCÍA, M^a Victorina. (2003). *Reflexiones sobre la clase de palabra pronombre*. Universidad de Santiago de Compostela. Campus de Lugo.
- DI PIETRO, Robert. (1986). *Estructuras Lingüísticas en contraste. Versión original Language Structures In Contrast*. Versión en español de Felisa M, Puszkin de Siegel. Madrid: Gredos.
- DIEZ-ITZA, Eliseo. (1992). *Adquisición del lenguaje Principado de Asturias*. Oviedo: Pentalfa.

- DIEZ-ITZA, Eliseo. (1993). *Adquisición del lenguaje*. Oviedo: Pentalfa.
- DIEZ-ITZA, Eliseo, Verónica MARTÍNEZ, Manuela MIRANDA. (2005). *Análisis y codificación fonológicos de corpus de habla infantil con el proyecto CHILDES: Errores de omisión en el desarrollo normal y en Síndrome de Dowan*, en José Manuel Oro, Jesús Varela Zapata y Jodee Anderson (eds.) *La Enseñanza de Lenguas en una Europa Multicultural. Actas del XXI Congreso de la Asociación Española de Lingüística Aplicada*, Santiago de Compostela, Universidad de Santiago de Compostela, págs. 272-285.
- DORCASBERRO SIGNORET, Aline. (2003). *Bilingüismo y cognición: ¿cuándo iniciar el bilingüismo en el aula?*, *Perfiles Educativos*, vol. XXV, núm. 102, 2003, pp. 6-21
- ESCANDELL VIDAL, M. Victoria. (2011). *Invitación a la lingüística*. Madrid. Centro de estudios: Ramón Areces.

CREENCIAS Y ACTITUDES DE ESTUDIANTES ESLOVACOS HACIA LAS VARIEDADES CULTAS DEL ESPAÑOL: CASTELLANA, ANDALUZA Y RIOPLATENSE

Radka Svetozarovová
Universidad de Granada

RESUMEN

El presente trabajo estudia las actitudes y creencias de alumnos eslovacos procedentes de secciones bilingües eslovaco-españolas hacia las distintas variedades geográficas cultas del español aplicando la metodología del proyecto PRECAVES-XXI (encuesta en línea, basada en la técnica de pares falsos). Nuestros objetivos fueron establecer si los alumnos eslovacos son capaces de reconocer las variedades del español a partir de una grabación, cuál de ellas consideran la más prestigiosa y cuál es la que mejor valoran de forma directa, mediante la valoración de la grabación. Tras el análisis de los resultados pudimos concluir que son las variedades castellana, andaluza y rioplatense las que más fácilmente se reconocen y que es la castellana la que más prestigiosa se considera y que mejores resultados obtiene en la valoración directa, seguida de la variedad rioplatense y andaluza. Por último, se discuten las posibles causas de estos resultados y su relevancia en el ámbito de la enseñanza del español.

Palabras clave: PRECAVES-XXI, alumnos eslovacos, actitudes lingüísticas, creencias lingüísticas, variedades cultas de español

ABSTRACT

This paper examines the attitudes and beliefs of Slovak students from Slovak-Spanish bilingual high school sections towards the different geographical varieties of Spanish, by applying the methodology of the PRECAVES-XXI project (online survey, based on the matched-guise technique). Our main goals were to establish whether Slovakian students were able to recognize Spanish varieties from a recording, which of them they considered to be the most prestigious one and which one they valued the best in the direct evaluation (through the evaluation of the recording). After analyzing the results, we were able to conclude that the Castilian, Andalusian and Rioplatense varieties were the easiest to recognize and that it was the Castilian variety that was considered to be the most prestigious one and which obtained the best results in the direct evaluation, followed by the Rioplatense and Andalusian varieties. Finally, the possible causes of these results and their relevance in the field of teaching Spanish as a foreign language are discussed.

Keywords: PRECAVES-XXI, Slovak students, linguistic attitudes, linguistic beliefs, educated varieties of Spanish

1. INTRODUCCIÓN

El estudio de actitudes y creencias hacia las variedades del español es relativamente reciente, no obstante, ya cuenta con una base bastante sólida (podemos mencionar el trabajo de Chiquito y Quesada Pacheco (2014), Santana Marrero (2018a) o el volumen monográfico del Boletín de Filología de la Universidad de Chile (Varios, 2018). Este tipo de estudios se ha basado hasta ahora casi exclusivamente en hablantes nativos. Consideramos, sin embargo, que conocer el punto de vista de hablantes de español como lengua extranjera (ELE) podría resultar de gran interés, no solo porque constituyen un grupo numeroso e importante dentro de la comunidad de habla hispana, sino también por la posible aplicación de los resultados del presente estudio en la enseñanza de ELE y en la política lingüística española en general. Por ello, basándonos en el proyecto PRECAVES-XXI, analizamos las actitudes y creencias de hablantes de ELE procedentes de Secciones Bilingües (SSBB) de Eslovaquia hacia las variedades normativas del español, en concreto, la variedad castellana, andaluza y rioplatense.

1. 1. Actitudes hacia las variedades geográficas del español

En la actualidad, existe variedad de opiniones acerca de la definición del término “actitud”; no obstante, podemos destacar dos modelos mayoritarios: el conductista y el mentalista. Mientras que el primero entiende la actitud como comportamiento, el segundo la entiende como un conjunto de creencias o valoraciones previas que predisponen al individuo a actuar de una determinada manera (Almeida, 2003: 181). Por tanto, desde esta última perspectiva, no es posible observarla directamente y, como consecuencia, su estudio es más complejo; sin embargo, este modelo permite, en cierto modo, predecir el comportamiento del individuo (López Morales, 2004: 287-288). Por ello, nos basamos en el modelo mentalista de la actitud.

Por lo general, se reconocen tres componentes de esta: el afectivo (relacionado con los sentimientos), el cognoscitivo (relacionado con el conocimiento, las creencias, los estereotipos...) y el conductual (relacionado con el comportamiento del hablante) (López Morales, 2004: 288). Así, un hablante puede considerar que una variedad lingüística es fea o bonita (componente afectivo), que es propia de una zona geográfica (componente cognoscitivo), y puede rechazarla o aceptarla (componente conductual). No obstante, el

componente cognoscitivo no tiene por qué reflejar la realidad (López Morales, 2004: 291), ya que un hablante puede considerar que una variedad es, por ejemplo, de México, cuando no es así.

Por último, cabe mencionar que la percepción de las distintas variedades guarda una estrecha relación con su prestigio, ya que las lenguas con mejor estatus y más prestigio tienden a despertar actitudes más positivas (Campos Bandrés, 2018: 103). En este sentido, la escolarización juega un papel muy importante, ya que permite establecer la diferencia entre usos entendidos como correctos e incorrectos (Moreno Fernández, 2012: 92-93), lo cual, en nuestra opinión, resulta aún más importante en el caso de hablantes extranjeros.

1. 2. Educación secundaria bilingüe eslovaco-española: Secciones Bilingües

La muestra con la que trabajamos en este estudio se compone únicamente de estudiantes procedentes de SSBB eslovaco-españolas. Si bien es cierto que se trata de una muestra muy específica, nos interesa por tres razones:

- Los estudiantes de las SSBB tienen un alto nivel del español (al terminar sus estudios obtienen el certificado de nivel C1) y entran en contacto casi diario con este idioma.
- Tienen un profundo conocimiento de la cultura hispana gracias a la variedad de actividades culturales y viajes a países hispanohablantes que organizan los centros.
- Por último, como menciona Moreno Fernández (2012: 216), en la formación de la conciencia lingüística hacia variedades distintas de la propia es importante el contacto con sus hablantes nativos, probablemente aún más en el caso de hablantes extranjeros, ya que el español no es su lengua materna. En nuestro caso, esta condición se ve satisfecha por el hecho de que el Ministerio de Educación, Cultura y Deporte en España (MECD) convoca anualmente plazas para profesores de nacionalidad española para impartir clases en las SSBB de Eslovaquia; aunque es cierto que el alumno de las SSBB suele entrar en contacto también con hablantes nativos hispanoamericanos, ya sea mediante una actividad cultural o en clase.

1.3. Variación geográfica en el contexto de ELE en Secciones Bilingües de Eslovaquia

En el ámbito de ELE, el *Marco Común Europeo de Referencia para las Lenguas* (MCER) es de mención obligatoria; sin embargo, al tratarse de un documento base, únicamente indica que el alumno debería tener conocimiento acerca de la variación geográfica de la lengua meta y que debería ser capaz de entender un discurso en cualquier acento, incluso cuando este no le resulte familiar (2002). El *Plan Curricular del Instituto Cervantes* (PCIC) (2008), sobre todo en lo que respecta al apartado *Norma lingüística y variedades del español*, reconoce que parte de la variedad centro-norte peninsular para elaborar sus inventarios, pero también señala que recoge rasgos pertenecientes a otras variedades, es más, hace hincapié en que la variedad castellana no es la única variedad normativa y, dependiendo del contexto y las necesidades de los alumnos, el docente puede elegir qué importancia les otorgará a los distintos rasgos geolingüísticos del español.

No obstante, el documento más importante en nuestro caso es el *Currículo de lengua y literatura españolas. Cursos III, IV y V. Secciones Bilingües de Eslovaquia* (2016), donde se especifica que el alumno de las SSBB eslovacas debería conocer y saber ubicar las distintas variedades geográficas del español, conocer sus rasgos característicos y ser capaz de reconocerlas y entenderlas. Por tanto, podemos concluir que un alumno procedente de una SSBB de Eslovaquia debería, en teoría, poseer conocimiento suficiente sobre la variación geográfica española y ser capaz de reconocer las distintas variedades del español.

1.4. Variedades geográficas del español

Siguiendo las bases del proyecto PRECAVES-XXI (Cestero y Paredes, 2018b: 19), reconocemos las ocho variedades principales del español, propuestas por Moreno Fernández (2019):

a) Variedades españolas:

- a. castellana (centro y norte de la Península)
- b. andaluza (Andalucía; al tener en cuenta sobre todo el criterio fonético, incluimos aquí también las hablas de transición de Murcia y Extremadura)
- c. canaria (las Islas Canarias)

b) Variedades hispanoamericanas:

- a. mexicana y centroamericana (México, Guatemala, Honduras, El Salvador y Nicaragua)
- b. caribeña (Antillas, Costa Rica, Panamá, la mayor parte de Venezuela y la costa de Colombia)
- c. andina (la mayor parte de Colombia, Ecuador, Perú, Bolivia y la zona andina de Venezuela)
- d. chilena (Chile)
- e. rioplatense (Uruguay, Paraguay, Argentina).

2. OBJETIVOS

Si bien es cierto que el proyecto PRECAVES-XXI permite recoger mucha más información de la que aquí se presenta, a efectos de este trabajo, y por no extendernos demasiado, nos planteamos únicamente tres objetivos:

- ¿Cuál es la variedad que los alumnos consideran más prestigiosa?
- ¿Son capaces de reconocer las distintas variedades a partir de una grabación?
- ¿Cómo valoran las variedades de forma directa?¹

3. METODOLOGÍA

Como ya hemos mencionado, nuestro estudio está basado en el proyecto PRECAVES-XXI, cuyo objetivo es estudiar las actitudes y creencias de los hablantes hacia las variedades geográficas cultas del español. Dicho estudio se lleva a cabo mediante una encuesta anónima y en línea, basada en la técnica de pares falsos, que consiste en dos partes: en la primera se recogen datos socio-geográficos del encuestado, y en la segunda, el encuestado escuchará 16 grabaciones de las 8 variedades cultas del español. Cada variedad aparece dos veces: una en discurso leído y una en discurso espontáneo, y dependiendo del código que se le asigne al encuestado, este escuchará a lo largo de toda la encuesta ¹Este tipo de valoración consiste en evaluar directamente la grabación que el encuestado escucha, frente a la valoración indirecta que, en el caso del proyecto PRECAVES-XXI, consiste en la evaluación de la persona que habla, el país del que el encuestado considera que este hablante procede y la cultura de la región correspondiente.

o voz de hombre o voz de mujer. En este trabajo describimos únicamente aquellas partes de la encuesta que están relacionadas con el reconocimiento y la valoración directa de las variedades; la descripción metodológica del proyecto, así como las capturas de pantalla de la encuesta completa se pueden consultar en el artículo de Cestero y Paredes, publicado en el *Boletín de Filología* (2018b).

Antes de iniciar el análisis de las grabaciones, el informante debe responder a la pregunta “En su opinión, ¿en qué región o zona se habla mejor el español?”. En la segunda parte, tras escuchar cada grabación, el encuestado debe valorar de manera directa e indirecta la variedad. En el caso de la valoración directa, en una escala de 1 a 6 de diferencial semántico se valoran las siguientes características de la variedad: áspera/suave, monótona/variada, rural/urbana, lenta/rápida y confusa/clara para las características cognitivas, y desagradable/agradable, complicada/sencilla, distante/cercana, dura/blanda, aburrida/divertida y bonita/fea para las características afectivas. Los polos negativos y positivos se van intercambiando, con el fin de evitar respuestas automáticas. A continuación, se le pide al encuestado que destaque de manera libre el rasgo que más y que menos le haya gustado. Por último, el alumno debe responder, de manera libre, de qué región considera que es el hablante de la grabación.

4. COMPOSICIÓN DE LA MUESTRA Y VARIABLES

La composición de la muestra con la que trabajamos se puede observar en la tabla 1. Cabe mencionar que todos los alumnos encuestados, o bien estaban realizando sus estudios en una de las SSBB de Eslovaquia, o bien ya habían terminado sus estudios secundarios en una de estas secciones y en el momento de realizar la encuesta estaban matriculados en una carrera universitaria, filológica o no filológica. El muestreo se realizó mediante la técnica “bola de nieve”, es decir, de forma accidental no probabilística.

SECCIONES BILINGÜES		N
SEXO	Hombres	10
	Mujeres	19
ESTUDIOS	1: Preuniversitarios (no finalizados)	15
	2: Universitarios (no finalizados)	10
	3: Universitarios con formación filológica (no finalizados)	4
TOTAL		29

Tabla 1. Composición de la muestra

En cuanto a las variables sociales, el sexo y la edad no se tienen en cuenta, puesto que, en el caso de la primera, el número de mujeres es mucho mayor que el de los hombres, y en el caso de la segunda, todos los participantes pertenecen a la misma generación, ya que su edad está comprendida entre los 17 y 24 años. La variable estudios, al igual que las variables propias de la grabación (voz y discurso) finalmente no resultaron significativas, por lo que no se incluyen.

5. RESULTADOS

5. 1. *¿Cuál es la variedad más prestigiosa?*

En la primera pregunta, al encuestado se le pidió que, mediante una respuesta libre, indicara cuál es la región en la que considera que se habla mejor el español. Como se puede observar en la tabla 2, los alumnos extranjeros siguen un patrón muy parecido al de los hablantes nativos europeos, quienes también tienden a considerar la variedad castellana como el modelo de prestigio, lo que se puede comprobar en trabajos ya realizados en el marco del proyecto PRECAVES-XXI (Cestero y Paredes, 2018a: 55; Hernández Cabrera y Samper Hernández, 2018: 186; Manjón-Cabeza, 2018: 152; Méndez Guerrero, 2018: 95; Santana Marrero, 2018b: 121).

Hay que destacar que únicamente 3 de los encuestados consideran que no existe jerarquía entre las variedades, siendo todos ellos estudiantes de carreras filológicas. Por tanto, el análisis cualitativo nos parece indicar que la variable estudios podría ser significativa; no obstante, para confirmarlo, será imprescindible ampliar la muestra en futuros estudios.

	n	%
Castellana	15	57,69 %
Andaluza	3	11,54 %
Centroamericana	1	3,85 %
General España	2	7,69 %
No hay/matices	3	11,54 %
Otros	1	3,85 %
NS/NC	1	3,85 %
Total	26	100,00 %

Tabla 2. Variedades consideradas más prestigiosas

5.2. Reconocimiento de las variedades

Antes de analizar cómo valoran los informantes las distintas variedades de manera directa, es imprescindible saber si son capaces de reconocerlas. Esto se debe a que, para la valoración directa, solo tendremos en cuenta las variedades reconocidas, para evitar una posible interferencia entre lo que verdaderamente se escucha y lo que se cree escuchar.

Dividimos el acierto en cuatro niveles: acierto exacto (cuando se contesta “Islas Canarias”, “Tenerife”, etc. para la variedad canaria), acierto genérico (ej.: “España” para una variedad española), error (ej.: “Madrid” para la variedad rioplatense) y error, pero manteniéndose dentro del mismo grupo (innovador o conservador) al que pertenece la variedad. Para considerar este segundo tipo de error aparte, partimos de la idea de que existen variedades conservadoras (castellana, mexicana y andina) e innovadoras (andaluza, canaria, caribeña, chilena y rioplatense) que se caracterizan, en el caso de las primeras, por el mantenimiento de ciertos rasgos lingüísticos (principalmente fonéticos) y en el caso de las segundas, por su pérdida, aspiración o debilitamiento. Este tipo de error se daría cuando el encuestado contesta, por ejemplo, “Castilla” para la variedad mexicana o “Andalucía” para la variedad rioplatense, ya que se está manteniendo dentro del mismo grupo, (conservador en el caso de las variedades castellana y mexicana e innovador en el caso de la andaluza y la rioplatense). Creemos que es interesante considerar este tipo de error aparte, ya

que el hecho de que el encuestado no acierte, pero sí se mantenga dentro del mismo grupo al que pertenece la variedad escuchada, puede significar que está empleando sus conocimientos lingüísticos y fonéticos para etiquetarla.

	n	%
Acierto exacto	72	19,46 %
Acierto genérico	79	21,35 %
Error con mantenimiento conservación/innovación	79	21,35 %
Error (sin mantenimiento conservación/innovación)	140	37,84 %
Error (total)	219	59,19 %
Todas	370	100,00 %

Tabla 3. Reconocimiento de las variedades

Como puede observarse en la tabla 3, el reconocimiento general es muy bajo, lo que es de esperar en el caso del hablante extranjero; sin embargo, hay que destacar que el porcentaje de error con mantenimiento del grupo innovador/conservador constituye un 21,35 %, lo que quiere decir que, en cierto modo, los alumnos sí utilizan sus conocimientos lingüísticos.

Ahora bien, lo que resulta aún más interesante es saber cuáles son las variedades más reconocidas. En la figura 1 puede observarse que las únicas tres variedades que llegan a un porcentaje considerable para el estudio son la castellana, andaluza y rioplatense, mientras que las demás apenas se reconocen. Una de las posibles explicaciones es que las variedades castellana y rioplatense poseen rasgos característicos que las diferencian de las demás: la distinción entre /s/ y /θ/ (Moreno Fernández, 2019: 119) y el yeísmo con pronunciación tensa (Moreno Fernández, 2019: 340-341), respectivamente. En el caso de la variedad andaluza, podría ser la cercanía geográfica que guardan los encuestados con la variedad (ámbito europeo peninsular) y su carácter innovador los que les permiten reconocerla, aunque haría falta un estudio más pormenorizado para establecer si verdaderamente son estos los factores condicionantes.

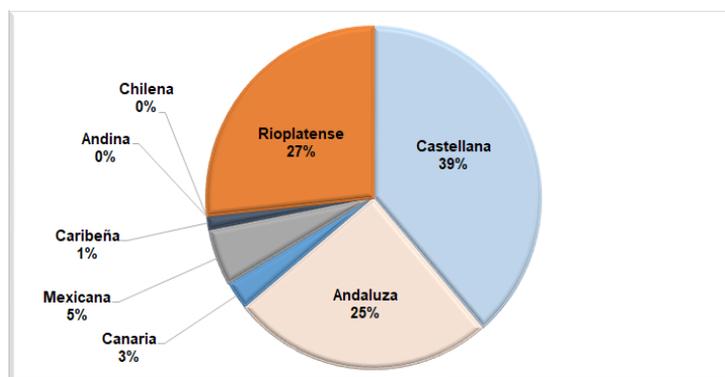


Figura 1. Las variedades más reconocidas

5.3. Valoración directa de las variedades

En este apartado analizamos la valoración directa de las tres variedades más reconocidas: castellana, andaluza y rioplatense. Los cálculos se realizaron mediante el paquete estadístico SPSS 20, utilizando la prueba ANOVA para la comparación de medias y el test de Bonferroni para establecer entre qué grupos se encuentran diferencias significativas.

Nos interesa saber, en primer lugar, cuál es la media general, afectiva y cognitiva de las valoraciones de las tres variedades. En la tabla 4 se muestran las medias de una escala de 1 a 6:

Variedad grabación	Media cognitiva		Media afectiva		Media general	
	Media	Desviación	Media	Desviación	Media	Desviación
Castellana	4,16	,61	4,05	,81	4,10	,61
Andaluza	3,69	,59	3,45	,87	3,56	,63
Rioplatense	4,04	,48	3,69	,86	3,85	,65

Tabla 4. Valoración directa de las variedades

Tras realizar la prueba ANOVA del factor “variedad”, las diferencias sí resultan significativas (con $p=0,049$ para la media afectiva, $p=0,021$ para la media cognitiva y $p=0,015$ para la media general). No obstante, mediante el test de Bonferroni pudimos comprobar que únicamente existe diferencia significativa entre la variedad castellana y andaluza ($p=0,05$ para la media afectiva, $p=0,019$ para la media cognitiva y $p=0,012$ para la media general). De forma gráfica, las medias quedan reflejadas en la figura 2:

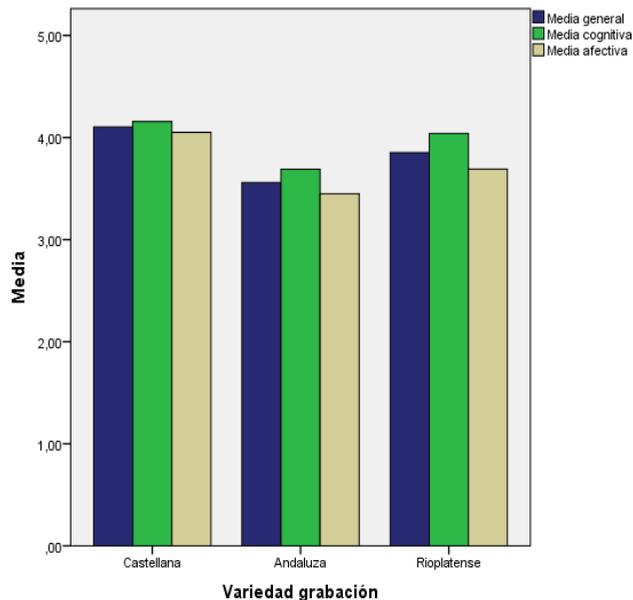


Figura 2. Medias de la valoración directa de las variedades

Ahora bien, nos interesa saber cómo valora el informante las características concretas de cada variedad, mencionadas en el apartado de la metodología. En la figura 3 se reflejan las medias de la escala de 1 a 6 de diferencial semántico.

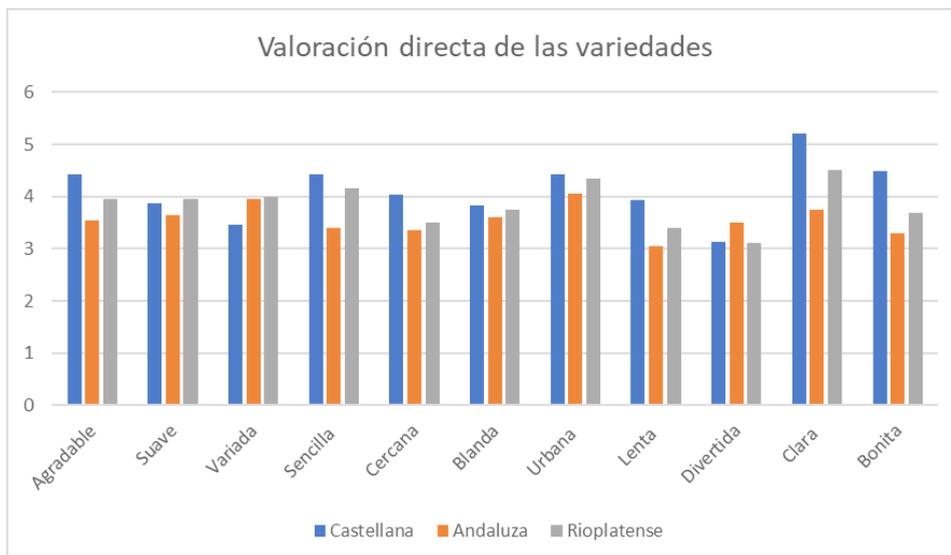


Figura 3: Valoración directa de las variedades

Como se puede observar en las figuras 2 y 3, la valoración general de las variedades es positiva, ya que se mantiene en todo momento por encima del valor medio, 3,5. Sin embargo, no en todas las características se ha mantenido la misma proporción. Este es el caso, por ejemplo, de la característica variada/monótona, donde la variedad castellana queda valorada como “más monótona” que las restantes dos; así ocurre también con “divertida/aburrada”, donde la variedad andaluza resulta como la más divertida de las tres, y en cuyo caso habría que preguntarse si lo adecuado sería interpretar que el alumno simplemente considera esta variedad más variada que las otras dos, o si se podría tratar de cierto reflejo del estereotipo del “andaluz divertido”. No obstante, en ninguna de las dos características la diferencia entre las medias resulta significativa. No es así en el caso de sencilla/complicada, clara/confusa y bonita/fea, donde se encuentra diferencia significativa entre la media de la variedad castellana y andaluza (con $p=0,022$, $p=0,0002$ y $p=0,002$ respectivamente, obtenido mediante el test de Bonferroni). En los tres casos,

los datos apuntan a que la variedad castellana es considerada más sencilla, clara y bonita que la andaluza, quedando la rioplatense entre ambas. Estos datos podrían apuntar a que los alumnos extranjeros tienden a valorar mejor precisamente aquellas variedades que les resultan más claras y sencillas, lo que probablemente esté unido también al carácter conservador de estas. Es más, si nos fijamos en las respuestas a preguntas abiertas, donde al encuestado se le pide que destaque los rasgos que más y que menos le hayan gustado de la variedad, nos encontramos, efectivamente, con cierta tendencia a valorar mejor lo conservador. En el caso de la variedad castellana, los rasgos destacados como positivos son la pronunciación fricativa de la /x/ y africada de la /tʃ/, la no omisión de las consonantes, en especial la /s/, y la claridad de la variedad. Muchos de los encuestados consideran que el hablante “pronuncia todo muy bien”. En cambio, los aspectos peor valorados fueron su carácter lento, duro y distante, aunque la mayoría de los encuestados consideró que no había ningún rasgo negativo que destacar. En el caso de la variedad andaluza, se destacó como positiva su variedad y suavidad, y como aspecto negativo la omisión de la /s/ implosiva y la desafricación de la /tʃ/. En cuanto a su velocidad, esta fue considerada positivamente como adecuada, no obstante, en varios casos se destacó como rasgo negativo un exceso de pausas y falta de una velocidad constante. Si bien es cierto que esta evaluación resulta contradictoria, se puede deber a que los encuestados consideraron adecuada la velocidad de la grabación para su comprensión, a pesar de que esta no les pareció constante. Por último, en el caso de la variedad rioplatense no podemos destacar una línea clara, ya que los dos rasgos más repetidos, la entonación y el yeísmo rehilado, fueron considerados como positivos para unos y negativos para otros.

6. CONCLUSIONES

En el presente estudio hemos analizado el reconocimiento de las distintas variedades geográficas del español a partir de una grabación y la valoración directa de las tres variedades más reconocidas, castellana, andaluza y rioplatense, por parte de estudiantes procedentes de secciones bilingües eslovaco-españolas.

De manera general, podemos concluir que los alumnos parecen tener muy poca familiaridad con las variedades canaria, mexicana, caribeña, andina y chilena, dado su reconocimiento muy bajo.

En cuanto a la valoración directa de las variedades, pudimos comprobar que los alumnos eslovacos siguen unos patrones de prestigio muy parecidos a los

hablantes nativos, puesto que consideran la variedad castellana como la más prestigiosa en casi todos los aspectos. No obstante, el análisis pormenorizado de las características desglosadas desvela que son el carácter sencillo/complicado, confuso/claro y bonito/feo donde mayores diferencias pueden encontrarse, lo que nos hace pensar en que el alumno eslovaco valora de manera positiva aquellas variedades de español que mejor entiende, en este caso, la castellana. Además, tras el análisis de las respuestas libres, parece ser que es el carácter conservador de una variedad el que mejor se valora, frente a las aspiraciones y pérdidas de elementos fonéticos. En cambio, la monotonía de las variedades se valora de forma negativa, frente a su carácter variado, que concretamente es el único caso en el que la variedad andaluza (considerada más variada) parece superar ligeramente a la castellana (evaluada como monótona). La falta de una tendencia clara en las valoraciones libres de la variedad rioplatense y su posición central entre las dos variedades peninsulares nos lleva a pensar que los alumnos realmente no tienen una opinión concreta acerca de la misma, posiblemente por falta de familiaridad; sin embargo, sí la reconocen, ya que se caracteriza por un rasgo que no comparte con ninguna otra variedad.

Por último, nos parece importante destacar que, aunque en los últimos años se está promoviendo de manera muy activa el carácter plurinormativo del español, los alumnos eslovacos de ELE de SSBB aún parecen tener lagunas en este aspecto. Por tanto, consideramos importante que el docente de ele, uno de los principales transmisores del prestigio lingüístico en la comunidad no nativa, aborde de manera sistemática la enseñanza de las distintas variedades de español y no se limite a destacar aquellos rasgos que presentan desviación con respecto a la variedad que se enseña, haciendo hincapié en la validez de cada modelo culto. Consideramos que este conocimiento le ayudaría al alumno no solo a cumplir con los objetivos que se especifican en el currículo de las SSBB, sino también entender mejor el ya mencionado carácter pluricéntrico del español.

REFERENCIAS BIBLIOGRÁFICAS

- ALMEIDA, M. 2003. *Sociolingüística*. La Laguna: Universidad de La Laguna.
- CAMPOS BANDRÉS, I. O. 2018. *Lengua minorizada y enseñanza: actitudes, metodologías y resultados de aprendizaje en el caso del aragonés*. Zaragoza: Prensas de la Universidad de Zaragoza.
- CESTERO, A. M. y PAREDES, F. 2018a. “Creencias y actitudes de los jóvenes universitarios del centro-norte de España hacia las variedades cultas del español”, *Boletín de Filología de la Universidad de Chile*, 53(2): 45-86.
- CESTERO, A. M. y PAREDES, F. 2018b. “Creencias y actitudes hacia las variedades cultas del español actual: el proyecto PRECAVES XXI”, *Boletín de Filología de la Universidad de Chile*, 53(2): 11-43.
- CHIQUITO, A. B. y QUESADA PACHECO, M. Á (eds.). 2014. *Actitudes lingüísticas de los hispanohablantes hacia el idioma español y sus variantes*. Bergen: University of Bergen.
- HERNÁNDEZ CABRERA y C. E., SAMPER HERNÁNDEZ, M. 2018. “Creencias y actitudes de los jóvenes universitarios canarios hacia las variedades cultas del español”, *Boletín de Filología de la Universidad de Chile*, 53(2): 179-208.
- INSTITUTO CERVANTES, 2008. *Plan curricular del Instituto Cervantes: niveles de referencia para el español.*, Instituto Cervantes. Web. 14 de mayo de 2019.
https://cvc.cervantes.es/ENSEÑANZA/biblioteca_ele/plan_curricular/default.htm.
- LÓPEZ MORALES, H. 2014. *Sociolingüística*. Madrid: Gredos.
- MANJÓN-CABEZA CRUZ, A. 2018. “Creencias y actitudes de los jóvenes universitarios granadinos hacia las variedades cultas del español”, *Boletín de Filología de la Universidad de Chile*, 53(2): 145-177.
- MÉNDEZ GUERRERO, B. 2018. “Creencias y actitudes de los jóvenes universitarios mallorquines hacia las variedades cultas del español”, *Boletín de Filología de la Universidad de Chile*, 53(2): 87-114.
- MINISTERIO DE EDUCACIÓN, CULTURA Y DEPORTE, 2016. *Currículo de lengua y literatura españolas. Cursos III, IV y V. Secciones Bilingües de Eslovaquia*. Secretaría General Técnica del MECD, Subdirección General de Información y Publicaciones.

- MORENO FERNÁNDEZ, F. 2019. *La lengua española en su geografía: Manual de dialectología hispánica*. Madrid: Arco Libros.
- MORENO FERNÁNDEZ, F. 2012. *Sociolingüística cognitiva: Propositiones, escolios y debates*. Madrid: Editorial Iberoamericana/ Vervuert.
- SANTANA MARRERO, J. 2018a. “Creencias y actitudes de jóvenes universitarios sevillanos hacia las variedades normativas del español de España: andaluza, canaria y castellana”, *Pragmática sociocultural: Revista Internacional sobre Lingüística del Español*, 6(1): 71-97.
- SANTANA MARRERO, J. 2018b. “Creencias y actitudes de los jóvenes universitarios sevillanos hacia las variedades cultas del español”, *Boletín de Filología de la Universidad de Chile*, 53(2): 115-144.
- SECRETARÍA GENERAL TÉCNICA DEL MECD, SUBDIRECCIÓN GENERAL DE INFORMACIÓN Y PUBLICACIONES. 2002. *Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación*. Secretaría General Técnica del MECD, Subdirección General de Información y Publicaciones. Madrid.
- VARIOS. 2018. *Percepción de las variedades cultas del español: creencias y actitudes de jóvenes universitarios hispanohablantes*, 53 (2). (Monográfico). Santiago de Chile: Universidad de Chile.

NATURALEZA, FUNCIONAMIENTO, DISTRIBUCIÓN Y FRECUENCIA DE LOS ESTEREOTIPOS HACIA ESPAÑA: EL CASO DE LAS AULAS INDIAS

Nabiel Ansari

Shweta Sahay

Jawaharlal Nehru University (India)

RESUMEN

El presente artículo pretende indagar sobre cuáles son los estereotipos culturales sobre España y su gente que se han ido consolidado en la mente de los estudiantes indios que aprenden español como lengua extranjera.

Con este fin, se ha llevado a cabo un estudio empírico en el que han participado más de 170 estudiantes indios de Grado y 15 profesores indios procedentes de las cuatro universidades públicas más importantes de la India en cuanto a la enseñanza del español.

En los resultados encontramos, por un lado, una generalización excesiva en los alumnos en su percepción de las costumbres españolas, y por otro, la pervivencia de estereotipos españoles anticuados que todavía continúan presentes en la mente de los estudiantes de ELE de la India.

Palabras clave: cultura, estudiantes universitarios, estereotipos españoles, aulas indias

ABSTRACT

This article aims to investigate what are the cultural stereotypes about Spain and its people that have been consolidated in the minds of Indian students who learn Spanish as a foreign language.

Keeping this objective in mind, an empirical study was carried out in which more than 170 Indian undergraduate students and 15 professors from four most prestigious public universities in India for Spanish learning and teaching participated.

In the results, we find out that on the one hand there is an overgeneralization among students in their perception of Spanish customs, and on the other hand, there exists outdated Spanish stereotypes that still persist in the mind of Indian students learning Spanish in India.

Keywords: culture, university students, stereotypes, Spanish stereotypes, Indian classroom

1. INTRODUCCIÓN

Todos los aprendices de una lengua extranjera entran en el aula con sus propias predisposiciones, suposiciones y creencias hacia la cultura meta. Sus procesos de aprendizaje parten de nociones y actitudes preconcebidas generadas en sus respectivas culturas de origen y son dichas preconcepciones las que actúan como base para entender la nueva cultura de la lengua extranjera.

En India, tanto los estudiantes como los profesores traen diferentes antecedentes culturales y diferentes actitudes hacia “el otro”. Parece pertinente señalar que, en el ámbito de la enseñanza de español en la India, existen muchos estereotipos en los alumnos que aprenden español como lengua extranjera que no se encuentran en consonancia con la realidad del español, los españoles y la España del siglo XXI.

Los estereotipos se han convertido en uno de los problemas principales de la actual sociedad globalizada ya que ahora el contacto entre culturas es más frecuente. En la India, la información se difunde con mucha más velocidad y se tiende a reflexionar menos sobre el contenido que se recibe a través de las redes sociales. Así que, en esta época, se corre más riesgo de que los estereotipos se fosilicen, si no se presta atención a los procesos cognitivos de la percepción de la información de los alumnos.

El objetivo de nuestra investigación es descubrir el grado de conocimiento de la cultura española en los alumnos indios de español como lengua extranjera, además de encontrar maneras de desarrollar la competencia comunicativa intercultural en los alumnos mediante los estereotipos que tienen en cuanto a la lengua y a la cultura española; y, asimismo, analizar el rol de los profesores en cuanto a la enseñanza de los aspectos culturales.

De este modo, uno de los objetivos principales de este estudio consiste en realizar un análisis sobre los estereotipos y prejuicios presentes en los estudiantes indios sobre España, el idioma español y los españoles. Esto es importante porque nos ayudará a determinar el porqué, el cómo surgieron y cuáles son las razones por las que todavía están presentes. Dicho análisis también determinará hasta qué punto dichas ideas representan la cultura española.

2. LOS ESTEREOTIPOS EN LA ENSEÑANZA DE IDIOMAS

La palabra *estereotipo* fue utilizada por primera vez en 1922 por Walter Lippmann en su libro *Public Opinion*. La usó para describir “juicios hechos de otros a base de su pertenencia a un grupo étnico” (Brunsch: 2005).

La palabra estereotipo proviene del griego: *stereos*, que significa ‘sólido’ o ‘rígido’ y *tipos*, que significa ‘carácter’, ‘tipo’ o ‘modelo’. Por su parte, el *DLE* define el estereotipo como una “imagen o idea aceptada comúnmente por un grupo o sociedad con carácter inmutable”.

Schneider (2004: 17) afirma que “no hay un consenso sobre qué son los estereotipos”. Su definición generalizada es que son “las creencias categorizadas que tenemos de las personas” (p. 29). Una definición más detallada de Hilton y Von Hippel (1996: 240) los define como “creencias sobre las características, atributos y comportamientos de los miembros de ciertos grupos”.

Por otro lado, los estereotipos también se pueden tratar en términos de ventajas y desventajas. Brislin (1986: 44) los define como una espada de doble filo: por un lado, un “aspecto útil e importante de pensamiento inteligente y eficiente” y, por el otro, “cualquier categorización de elementos individuales preocupados con las personas que enmascaran las diferencias entre esos elementos”. Aunque la misma idea de los estereotipos a menudo tiene un valor negativo, sus efectos no siempre perjudiciales. La información que está presente puede considerarse una fuente de conocimiento utilizada para hacer inferencias sobre las personas cuando otros tipos de información no están disponibles, lo que puede ser útil si la información es correcta.

Los estereotipos son generalizaciones, y, a través de ellos, dividimos a las personas en grupos, naciones, o representantes de un cierto grupo étnico, profesión o cultura. El hablante tiene un cierto modelo en la mente sobre qué tipo de personas pertenecen a cada grupo y cuando se encuentra con una nueva persona que pertenece a un determinado grupo, puede dar por sentado que este individuo actúa de acuerdo con nuestro estereotipo marcado. Las generalizaciones asumen que los miembros de un grupo comparten los mismos valores y que se comportan de una manera predecible.

Según la investigación hecha en el campo de ELE en el contexto indio (Fuentes y Venkataraman, 2004), superar los estereotipos y prejuicios debe ser una prioridad absoluta en la enseñanza de una lengua extranjera con el propósito de fomentar el reconocimiento de los estudiantes del contexto sociocultural del idioma estudiado y prevenir una visión distorsionada que esté

lejos de la realidad. Lustig y Koester (1996) enumeran tres tipos de estereotipos en los que estos pueden ser imprecisos:

1. Los que se presupone que se aplican a todos los miembros de un grupo.
2. Los que no son correctos o son fuertemente exagerados.
3. Los que son fuertemente positivos y negativos solo para un cierto grupo de personas.

3. METODOLOGÍA E INFORMANTES

Desde el punto de vista metodológico, el presente trabajo es un estudio de campo. El carácter de la investigación realizada es descriptivo y el instrumento principal que hemos utilizado para la recogida de datos ha sido la encuesta. En dicha encuesta, se han combinado preguntas cerradas y escalas de valoración con unas preguntas abiertas al final.

No hay antecedentes de ningún estudio sobre el nivel de los estereotipos españoles existentes en la mente de los estudiantes indios en las aulas indias de ELE, así que el estudio se encontró un campo no tratado antes, donde las creencias y percepciones sobre los estereotipos del alumnado indio se desconocían completamente.

Este estudio se ha realizado en las universidades principales que ofrecen cursos de Grado en español. No se realizó ningún estudio de seguimiento o de *follow up*, dado que el objetivo era encontrar cuáles eran los estereotipos que tienen los estudiantes indios sobre España, sus nativos y su lengua; e indagar sobre el papel que juegan los profesores a la hora de la enseñanza de los aspectos culturales e interculturales en el aula. El análisis es de naturaleza cualitativa, y se analizarán los datos obtenidos en la encuesta de forma global, interesándonos más por las tendencias generales que por los matices sociolingüísticos.

3.1 Tipología de los participantes de la muestra

La investigación incluye a los estudiantes de las siguientes cuatro universidades públicas de la India que ofrecen curso de Grado.

- La Universidad Jawaharlal Nehru (Nueva Delhi);
- La Universidad de Delhi (Nueva Delhi);

- La Universidad del Inglés y otras Lenguas Extranjeras (Hyderabad); y
- La Universidad Doon (Dehradoon).

Los participantes son los alumnos que estudian español en los niveles principiante e intermedio. Para nuestro estudio, elegimos estos niveles porque se supone que ellos han estado menos expuestos a la cultura española, aunque no se puede decir esto de manera definitiva ya que un estudiante puede haber estado expuesto a la cultura española sin haber estudiado el idioma en las universidades. No obstante, partimos de la suposición de que un alumno de Grado que aprende español no habrá tenido mucho contacto con la lengua y cultura española.

Además de tratar de calibrar el nivel del conocimiento de la cultura y la competencia intercultural de los alumnos, hemos pretendido incluir a los profesores también para analizar los desafíos a los que se enfrentan a la hora de tratar los temas culturales en la clase y la importancia que otorgan a la cultura en la enseñanza de español como lengua extranjera. Elegimos a los profesores que enseñaban al nivel de Grado de las mismas universidades para realizar el estudio.

En cuanto al alumnado, no recurrimos a un muestreo, sino que pasamos los cuestionarios a todos los participantes presentes en la clase, con el fin de tener un resultado holístico de las opiniones de todos los alumnos que estudian español.

El grupo de encuestados consta de un total de 173 alumnos indios que aprenden español como lengua extranjera dentro del sistema educativo universitario. De los 173 alumnos totales, 98 son chicas y 75 son chicos. Generalmente, las edades de los alumnos oscilan entre los 17 y los 24 años.

Con respecto a los profesores, elegimos al azar a 15 profesores indios de entre 28-40 años que dan clases de español a los alumnos de Grado que recibieron el cuestionario. Intentamos incluir a más del 50 % de los profesores para tener una muestra representativa.

Universidad Jawaharlal Nehru	4
Universidad de Delhi	4
Universidad del Inglés y otras Lenguas Extranjeras	2
Universidad Doon	5
Total	15

Figura 1. Clasificación de la muestra de profesores.

4. ANÁLISIS DE LOS RESULTADOS

Después de elegir los participantes y la muestra, llevamos a cabo con éxito una encuesta que nos permitió recabar los resultados para su posterior análisis.

Tras preguntar el motivo que les acerca a estudiar español, los alumnos debían de elegir por lo menos cinco ítems para describir lo que pensaban sobre España o los españoles.

A continuación, presentamos las respuestas de los alumnos. De este modo, se pretende analizar la percepción estudiantil de los ítems que determinan la imagen de un país o su gente. Las respuestas reflejan la opinión de la mayoría de los alumnos en referencia a las siguientes esferas: saludos, hábitos gastronómicos, bebidas, música, fiestas, ropa, interacción social, cultura de trabajo, ciudades y otros. Dentro de estas áreas, los participantes tenían la opción de crear una respuesta semi-abierta y los resultados son:

Saludos

- Saludan a la gente de manera amistosa.
- Sus saludos son muy informales.
- Sus saludos son en ocasiones íntimos.

Hábitos gastronómicos

- En su mayoría comen alimentos no vegetarianos.
- Les gusta comer fuera de casa y comen mariscos en la paella y en otros platos tradicionales.
- Aman la comida picante como los indios.

Bebidas

- Toman mucho alcohol.
- Les encanta ir a los bares y tomar unas copas.
- Les encanta beber y visitar el pub regularmente.

Música

- Su música es fenomenal.
- Me gusta la música española.
- Su música es mundialmente famosa.

Fiestas

- España es conocida por sus manifestaciones culturales como la corrida de toros o la Tomatina.
- Tienen festivales muy bonitos como la Tomatina y las Fallas.
- La semana santa es el festival más importante de España.

Ropa

- Las chicas llevan ropa atrevida.
- Sus vestidos son muy coloridos.
- Tienen la libertad de vestir lo que quieran.

Interacción social

- La gente de España es muy abierta y amigable en sus interacciones.
- A las personas de España les gusta interactuar con otros.
- La gente de España es muy abierta, ruidosa y amigable.

Cultura de trabajo

- Son muy perezosos.
- La gente de España es muy trabajadora.
- Son muy sinceros y puntuales.
- Siempre necesitan echar una siesta por la tarde.

El análisis de las respuestas nos muestra que muchos de ellos tienen una percepción errónea o fragmentaria de la cultura española. Su respuesta nos dio la oportunidad de analizar algunos pensamientos y percepciones estereotipadas

que albergaban hacia la cultura española.

También, les preguntamos que enumeraran cualidades positivas y negativas del carácter español. Con esto, queríamos que el alumno transmitiera la percepción que tiene hacia la gente española.

En general, los participantes tienen una actitud positiva sobre el carácter español como muestra la siguiente figura.

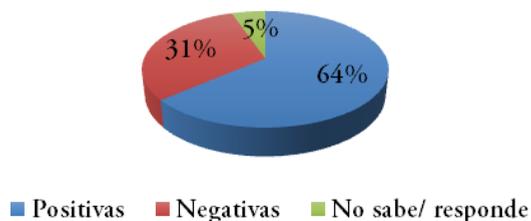


Figura 2. Valoraciones del carácter español por parte de los alumnos

A continuación, presentamos las valoraciones positivas y negativas del carácter español.

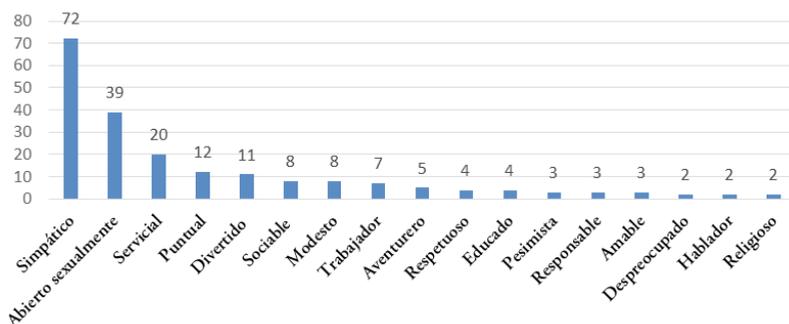


Figura 3. Valoraciones positivas del carácter español por parte de los alumnos

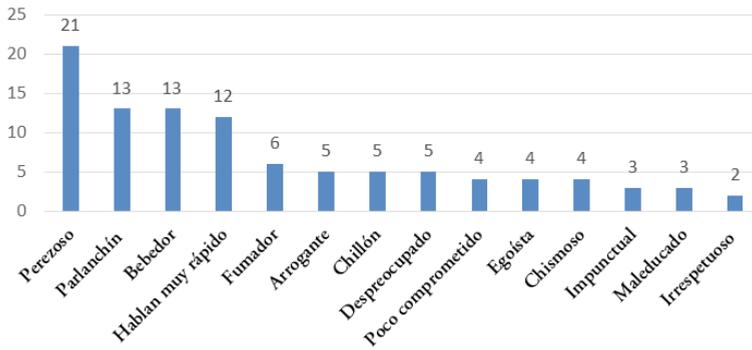


Figura 4. Valoraciones negativas del carácter español por parte de los alumnos.

En general, consideran que los españoles por su naturaleza son gente simpática, liberada sexualmente, servicial y puntual. En las valoraciones negativas, dicen que los españoles son perezosos, habladores, toman mucho alcohol, hablan muy rápido y fuman demasiado.

Se ve que los estudiantes a veces sobregeneralizan acerca del carácter de los españoles. Estas valoraciones por parte de los alumnos puede conducirlos a tener una visión estereotipada hacia la cultura meta.

En la siguiente pregunta, les proporcionamos los cuatro estereotipos más comunes de los españoles basados en el estudio de Fuentes y Venkataraman (2004) para que los alumnos los valoren. Esto nos ayudó a comprobar si estos pervivían todavía en la mente de los estudiantes de ELE en la India o no. Los estereotipos eran:

- Los españoles echan una siesta por la tarde.
- A los españoles les encanta ver la corrida de toros de vez en cuando.
- Los españoles van al bar todas las noches y beben cerveza.
- Los españoles son muy religiosos y van a la iglesia regularmente.

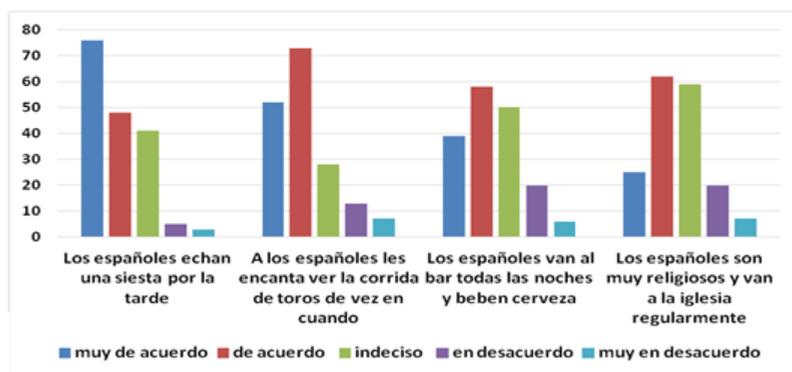


Figura 5. Valoración de la pervivencia de los estereotipos en los estudiantes indios de ELE

Estamos muy sorprendidos con los resultados obtenidos. La mayoría de los alumnos cree que estas afirmaciones son correctas.

A continuación, analizamos los típicos estereotipos españoles:

– *Los españoles echan una siesta por la tarde.*

El 72 % de los alumnos está muy de acuerdo o de acuerdo y el 23 % no está ni de acuerdo ni en desacuerdo mientras que el 5 % está en desacuerdo o muy en desacuerdo.

Este es el estereotipo más común de la gente española. La siesta es ahora tan ajena a la mayoría de los españoles como a los extranjeros que la integran en su imagen de España. Según el artículo de Jones (2017), casi el 60 % de los españoles nunca se echa una siesta. Este dato de nuestra encuesta muestra la solidificación de los estereotipos o la falta de conocimiento de la cultura meta por parte de los aprendientes de ELE.

– *A los españoles les encanta ver las corridas de toros de vez en cuando.*

El 72 % de los alumnos está muy de acuerdo o de acuerdo y el 16% no está ni de acuerdo ni en desacuerdo mientras que el 12 % está en desacuerdo o muy en desacuerdo.

Se les preguntó sobre las corridas y la mayoría respondió que a los españoles les encanta verlas de vez en cuando. No son conscientes de la polémica con respecto a las corridas de toros y a las nuevas legislaciones que prohíben y limitan dichas corridas en ciertas zonas de España.

– *Los españoles van al bar todas las noches y beben cerveza.*

El 57 % de los alumnos está muy de acuerdo o de acuerdo y el 28 % no está ni de acuerdo ni en desacuerdo mientras que el 15 % está en desacuerdo o muy en desacuerdo. La mayoría cree que los españoles pasan tiempo en los bares tomando copas.

– *Los españoles son muy religiosos y van a la iglesia regularmente.*

El 50% de los alumnos está muy de acuerdo o de acuerdo y el 34% no está ni de acuerdo ni en desacuerdo mientras que el 16% está en desacuerdo o muy en desacuerdo.

Esta creencia procede de su propio entorno, de la sociedad en la que habitan. En la India, ir al templo es algo común y a menudo los indios van al templo. De este modo, generalizan que la religión juega un papel muy importante en la sociedad española y los españoles son tan religiosos como los indios.

Les preguntamos mediante qué recurso o medio obtuvieron tal percepción hacia la cultura española, ya que queríamos comprobar si las percepciones se hacen dentro del aula o fuera del aula y hasta qué punto el profesor o los materiales influyen a la hora de crear estas percepciones en los alumnos.

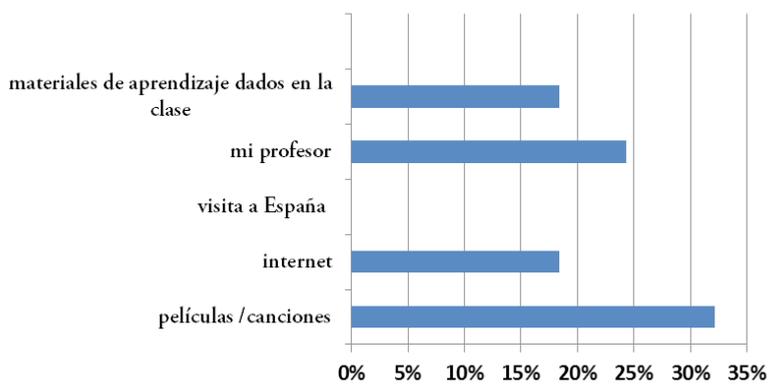


Figura 6. Fuentes de la percepción de la cultura española

El factor más importante sobre el que se tiene control es el de los materiales de aprendizaje y la actitud de los profesores indios, sumando un total de 42 %. Si bien es cierto que la proactividad de los estudiantes tiene una influencia

superior, con un total del 50 %, habría que buscar el modo de integrar estos aspectos en la clase y promover la formación de los profesores en este ámbito.

Según los profesores, los aspectos que crean problemas a la hora de enseñar la cultura española a los estudiantes indios son:

- Enseñarles a saludar con dos besos en las mejillas.
- Los hábitos alimentarios, el alcohol y la carne.
- Hay aspectos culturales que son nuevos para ellos. Por ejemplo, el encuentro en un bar.

Casi el 85% de los docentes encuestados afirma que los estudiantes indios de ELE tienen los siguientes estereotipos:

- Todos los españoles celebran el festival de la Tomatina.
- A todos les gusta participar en las fiestas de San Fermín.
- Todos los españoles van a ver las corridas de toros.
- Los españoles van al bar todos los días.

Además de estos, mencionan que:

- Los españoles (particularmente las chicas) entablan amistades rápidamente con cualquier persona que conocen.

Entonces, el problema proviene del hábito de estereotipar en general y de los estereotipos dominantes sobre la cultura española en particular. Ante esto, ¿qué deberían hacer los profesores de español? Los mismos profesores nos sugirieron medidas para superar la noción estereotipada que los alumnos tienen sobre España y su cultura.

Algunas de las sugerencias incluyen:

1. La lectura de libros y artículos sobre las tradiciones y costumbres españolas. El profesor debe sugerir a los alumnos algunos libros o proporcionar materiales relacionados durante el curso de la enseñanza.
2. La comparación de la cultura india con la española.
3. Ver películas que ofrecen diversas perspectivas culturales.
4. La presentación de posibles estereotipos con el fin de evitar cualquier malentendido.

5. La presentación de una actitud crítica frente a los estereotipos ya existentes sobre India y hacer entender que todo lo que se piensa sobre la India no es correcto. De la misma manera, enseñarles que todo lo que se dice o se piensa sobre España no corresponde con la realidad.
6. El aumento de la frecuencia de contacto entre los españoles y los alumnos indios.

Según ellos, estas sugerencias constituyen un marco que los profesores deben seguir a la hora de tratar temas culturales a lo largo del curso, pero el problema es que hay una diferencia entre saber lo que se debe hacer y realmente hacer lo que se debe hacer. La mayoría del profesorado tiene ideas muy claras cuando hablan de los aspectos culturales pero, este saber hacer no se refleja en la práctica docente como muestra la figura 5 de la valoración de la pervivencia de los estereotipos en los estudiantes indios de ELE.

Los docentes indios de ELE también son seres humanos con sus propias experiencias y pensamientos sobre otras culturas. No son siempre conscientes de estos sentimientos y de cómo los expresan, pero es posible que solo una breve observación por parte del profesorado en el aula sea recordada por los estudiantes durante muchos años después. Y dichas observaciones pueden ser negativas o positivas.

Las respuestas recibidas de los estudiantes muestran que existe una brecha entre el conocimiento y las creencias de los profesores y sus prácticas docentes. Seguramente, los profesores no hayan podido desarrollar su competencia intercultural plenamente o con el mismo rigor que desarrollan la competencia lingüística.

5. EL PAPEL DEL PROFESORADO EN EL CONTEXTO INDIO

En cuanto a la formación del profesorado indio encuestado, encontramos que 8 profesores (casi 53 %) no tienen ninguna formación específica en la enseñanza de lenguas extranjeras. Y de estos, 6 profesores responden que tampoco necesitan formación en cuanto a la enseñanza del componente cultural.

Kramersch (1996) afirma que el profesor en la clase de idiomas se ha visto durante mucho tiempo “[...] como el experto, que representa a una comunidad de habla nativa, y cuyo papel es impartir un cierto cuerpo de conocimiento a los hablantes no nativos para facilitar su integración en la comunidad del

habla”. Esta visión corresponde al ideal de la competencia lingüística de un hablante nativo, que ha sido y probablemente sigue siendo el objetivo del aprendizaje de idiomas para muchos.

Tradicionalmente, el papel del profesor se basa en las presunciones de que es la autoridad principal en todos los aspectos dentro del aula de idiomas. Se puede afirmar que es cierto en la enseñanza de un idioma extranjero en el contexto indio. En la India, se acepta que la dimensión pedagógica se limita a las iniciativas de los docentes y los estudiantes consideran que los materiales que se seleccionan son apropiados.

En el aula india, es el profesor quien decide el tipo de material que se usará para enseñar un idioma extranjero. Dicha autonomía de los profesores a la hora de proporcionar materiales de enseñanza establece una responsabilidad por su parte a la hora de incluir los aspectos multiculturales, pluriculturales e interculturales en la enseñanza de idiomas.

6. CONCLUSIÓN

Los resultados de este estudio manifiestan que los aprendices mantienen estereotipos tradicionales sobre la cultura española como se muestra en la figura 5. Sin embargo, esto no promueve una imagen negativa de la cultura meta, como aparece en la figura 2, ya que la mayoría presenta una actitud positiva hacia los españoles.

Por otro lado, los profesores indios necesitan ampliar su formación de ELE, a la par que desarrollar las sugerencias que nos han presentado en el aula para la implementación del componente cultural. De este modo, se logrará eliminar no solo los estereotipos anteriores (Fuentes y Venkataraman, 2004), sino también los nuevos estereotipos presentes en el alumnado como se ha mostrado en esta investigación.

REFERENCIAS BIBLIOGRÁFICAS

- BRISLIN R. W. 1986. "Prejudice and intergroup communication", *Intergroup communication.*, W. Gudykunst (ed.). London: Arnold. 74-85
- BRUNSCH C. 2005. *Stereotyping as a Phenomenon in Intercultural Communication.* Germany: GRIN Verlag, Norderstedt.
- FUENTES M., Venkataraman V. 2004. "Planteamientos interculturales del Marco Común Europeo de Referencia en el contexto sociocultural de la India y su influencia en el proceso enseñanza-aprendizaje del E-LE". *Encuentro Internacional de los Profesores del Español Como Lengua Extranjera del Asia-Pacífico*, Manila. Disponible en: https://cvc.cervantes.es/ensenanza/biblioteca_ele/publicaciones_centros/PDF/manila_2004/18_fuentes_v.pdf
- HILTON J. L., HIPPEL W. 1996. "Stereotypes", *Annual Review of Psychology* 47 (1): 237-71.
- JONES J. 2017. *It's time to put the tired Spanish siesta stereotype to bed.* Disponible en: <http://www.bbc.com/capital/story/20170609-its-time-to-put-the-tired-spanish-siesta-stereotype-to-bed>
- KRAMSCH C. 1996. "The cultural component of language teaching", *Language, Culture and Curriculum*, 8:2, 83-92.
- LIPPMANN W. 1922. *Public Opinion.* New York: Harcourt Brace
- LUSTIG M. W., Koester J. 1996. *Intercultural Competence. Interpersonal Communication Across Cultures.* 2a ed. London: Longman.
- SCHNEIDER D. J. 2004. *The Psychology of Stereotyping.* New York: Guilford Press.

LA POSICIÓN DEL VERBO EN EUSKERA: ¿VAN ESPECIALISTAS Y PROFESORADO EN LA MISMA DIRECCIÓN?

Julian Maia-Larretxea

Universidad del País Vasco-Euskal Herriko Unibertsitatea

RESUMEN

Este artículo versa sobre el orden de los elementos de la oración escrita en euskera. En el marco de la búsqueda del máximo nivel de eficiencia comunicativa en lengua vasca, en la actualidad se pueden distinguir dos tendencias diferenciadas en cuanto a la posición del verbo en oraciones afirmativas: una tiende a colocar el verbo en la parte final de la oración y la otra a adelantar su posición. En este estudio se analizan y se comparan las tendencias que se observan en un conjunto de especialistas (N=12) y en una muestra del profesorado de bachillerato de la Comunidad Autónoma Vasca (N=105). Para ello, se presenta a las personas participantes un conjunto de 25 pares de oraciones y se les solicita que elijan la ordenación que consideren preferente. Los resultados obtenidos apuntan a que el conjunto de especialistas presenta una mayor tendencia a preferir ordenaciones con el verbo situado en una posición más avanzada que la mostrada por el profesorado de bachillerato analizado.

Palabras clave: Euskera; orden de las palabras; posición del verbo; estudio empírico

ABSTRACT

This article tackles the sentences' element-order in Basque written prose. In the context of seeking the maximum level of communicative efficiency, at present time it exists two trends regarding the position of the verb in affirmative sentences: one tends to place the verb at the final part of the sentences whilst the other tends to put it in an advanced position. This piece of research analyses and compares those tendencies in a set of specialists (N=12) and in a sample of baccalaureate teachers (N=105). For that aim, participants chose their preferential element-order in a set of 25 pairs of sentences. Results point out that, unlike baccalaureate teachers, specialists mainly prefer sentences with the verb placed at an earlier position.

Keywords; Basque language; word-order; position of the verb; empirical study.

1. INTRODUCCIÓN

El euskera está viviendo, desde el último tercio del siglo XX, un proceso de normalización y revitalización en el que una de las cuestiones a debate es la posición del verbo en la oración afirmativa escrita de cierta extensión. En ese tema actualmente se pueden observar, esquemáticamente, dos diferentes tendencias o modelos.

Un modelo se inclina por colocar el verbo, elemento clave para el procesamiento del mensaje, en la parte final de la oración. Esta tendencia está basada en la obra de Altube titulada *Erderismos*, que fue publicada por primera vez en 1929 y posteriormente en 1975, y fue la referencia predominante durante los primeros años del proceso de normalización (Mitxelena, 1981).

En la última década del siglo XX, sin embargo, ya se empezó a detectar que la prosa construida siguiendo ese modelo presentaba dificultades de procesamiento, por lo que se empezaron a proponer otras formas de ordenación. Estos modelos, presentados bajo diferentes formulaciones de carácter más o menos general, coinciden en la preferencia por adelantar la posición del verbo con el fin de que la persona lectora pueda procesar el mensaje de manera gradual (p.ej. se pueden mencionar las contribuciones de Zubimendi & Esnal, 1993; Hidalgo, 1995, 2002; Rubio, 2002; Esnal, 2009; Euskaltzaindia, 2011; Maia, 2014; Sarasola, 2016). Mediante esta estrategia de adelantar el verbo se trata de obtener una mejora de la eficiencia comunicativa del mensaje evitando en lo posible la necesidad de releer toda la oración tras haber procesado el contenido informativo del verbo. Dicho de otra manera, se trata de reducir la denominada *retrocarga*, fenómeno que consiste en el efecto producido al colocar elementos decisivos para el procesamiento del mensaje (en este caso, el verbo) en la parte posterior de la oración, de manera que se dificulta su correcto procesamiento y comprensión en primera lectura.

En ese contexto, en el marco del debate para obtener el máximo de eficiencia comunicativa en la prosa lógico-discursiva escrita en euskera, es importante conocer las preferencias que sobre el tema presentan en la actualidad los y las usuarias de la lengua pertenecientes a ámbitos de carácter estratégico, como la administración, los medios de comunicación y difusión, y el sistema educativo. Este último merece una atención especial, debido a que tiene en nuestra sociedad la capacidad de acceder prácticamente a toda la población durante un período continuado de tiempo, lo que lo convierte en elemento clave para la difusión de las tendencias que puedan ser consideradas buenas prácticas y

que contribuyen a configurar las referencias canónicas. Esas referencias son necesarias también en el tema del orden de los elementos de la oración en euskera, en el que se observa la existencia de modelos diferenciados y ausencia de consenso (Maia, 2014; Maia y Badiola, 2015), debido a las características del proceso de normalización que está viviendo el euskera. En ese sentido, la investigación que presentamos a continuación tiene por objetivo general aportar información para colaborar en el debate acerca del orden más adecuado de los elementos de la oración en la prosa escrita en euskera.

2. LA INVESTIGACIÓN

El objetivo específico de este estudio es obtener información acerca de las preferencias que presenta un grupo de especialistas respecto a la posición del verbo en un conjunto de oraciones de cierta extensión, para después comparar sus resultados con los obtenidos con una muestra del profesorado del nivel de bachillerato. Se pretende ver qué tendencias existen en los colectivos analizados, y si ambos presentan características similares o divergentes.

2.1 Preguntas de investigación o hipótesis

A través de este estudio se pretende responder a las siguientes preguntas de investigación:

PI1: ¿Existe preferencia en los colectivos estudiados por ordenaciones de los elementos que colocan el verbo en una posición más retrasada en la oración?

PI2: En su caso, ¿con qué intensidad se manifiesta esa tendencia en el grupo de especialistas y en el del profesorado de bachillerato?

PI3: ¿Esa tendencia a colocar el verbo en una posición más retrasada es mayor en la muestra del profesorado que en el grupo de especialistas analizado?

PI4: ¿Qué criterios utilizan las personas participantes para justificar la opción elegida?

La formulación de esas PI tiene como base las hipótesis que se presentan a continuación:

H1: En los colectivos analizados existe efectivamente una cierta preferencia por ordenaciones más retrocargadas, lo que se atribuiría al canon teórico predominante a lo largo del siglo XX (sobre todo hasta la última década).

H2: La tendencia es mayor en el caso de la muestra de profesorado de bachillerato analizado que en el grupo de especialistas, porque este último estaría más actualizado respecto a las nuevas tendencias que están surgiendo en la línea de reducir la retrocarga con el fin de aumentar la eficiencia comunicativa de la prosa lógico-discursiva en euskera.

H3: Comparando con un estudio realizado en 2011 (Maia, 2017) los criterios simplistas (por ej. “en euskera el verbo se pone al final”) están perdiendo vigencia relativa, lo que se debería a la evolución en el plano teórico y a las diferentes características de la muestra (alumnado universitario en 2011 vs especialistas y profesorado en 2018).

2.2 Recogida y codificación de los datos

El procedimiento empleado para la recogida de datos es el siguiente: en 2018, se presenta a las personas de la muestra un conjunto de 25 pares de oraciones y se les solicita que elijan, en cada par de oraciones, la versión que consideran preferente. Las dos versiones de la oración que se presentan tienen el mismo contenido informativo pero en cada una de ellas los elementos de la oración están ordenados de forma diferente (respetando el criterio de gramaticalidad en todos los casos), de manera que en una de las versiones el verbo se presenta en una posición más retrasada que en la otra. Metodológicamente consideramos que la versión que tiene el verbo más adelantado posee una mayor eficiencia comunicativa, puesto que la posición en la que se presenta el verbo es un elemento determinante para poder procesar el contenido informativo del mensaje. Las personas encuestadas tienen, ante cada par de oraciones, las siguientes opciones: 1) elegir como preferente la versión con el verbo más retrasado; 2) elegir como preferente la versión con el verbo más adelantado; 3) dar ambas versiones como igualmente aceptables y dependientes de un contexto más amplio; 4) considerar ambas versiones de la oración como inaceptables; 5) no responder. A continuación, una vez que se ha mostrado la preferencia entre las dos versiones de la oración, se solicita a las personas encuestadas que señalen cuál es el criterio (o criterios) en que se basa su elección, para lo cual se puede elegir entre 5 respuestas tipificadas y una más de libre expresión. Las opciones son las siguientes: 1) criterio gramatical “duro”: elijo una oración porque la otra es agramatical; 2) criterio gramatical “relativo”: elijo una versión porque “es mejor euskera”; 3) criterio pragmático: elijo una versión porque se entiende más rápidamente/fácilmente; 4) elijo una versión porque en el código

escrito es más adecuada (en el uso oral también aceptaría la otra versión); 5) elijo una versión porque en euskera “el verbo se pone al final”; 6) elijo una versión por otros motivos.

Posteriormente, se codifican las respuestas recogidas, se introducen en la aplicación EXCEL, y se obtienen los valores que se señalan en el apartado de los resultados. También se obtienen datos referidos a cada uno de los ítems objeto de estudio, aunque por su especificidad y las limitaciones de espacio no pueden ser tratados en este trabajo.

2.3 Sobre los ítems analizados (y un ejemplo explicativo)

Se ha trabajado con 25 pares de oraciones. Cada par de oraciones constituye un ítem objeto de estudio. Se ha cuidado que las oraciones sean de cierta extensión, porque el fenómeno estudiado (posición del verbo y procesamiento de la información del conjunto de la oración) no resulta perceptible en las oraciones cortas, o, en otras palabras, se manifiesta más claramente cuanto más larga sea la oración. A continuación, con la intención de acercar el contenido del estudio también a personas no familiarizadas con la lengua vasca, mostramos como ejemplo uno de los ítems que han sido analizados, utilizando un nivel de detalle intermedio (no exhaustivo).

Supongamos las dos versiones de la oración en euskera que presentamos a continuación, **1a** y **1b** (señalamos en negrita los elementos más significativos para observar sus cambios de posición en las diferentes versiones; además, los subíndices empleados mantienen la referencia correspondiente a cada uno de los elementos que aparecen en los ejemplos, independiente de la lengua en que se presenten):

1a) [*Haren eginkizun nagusia*]₁ [*da*]₂ [*mundu guztian*]₃ [*babestea*]₄ [*alde guztietako eta bazter orotako jendearen giza-eskubideak eta demokrazia*]₅.

1b) [*Haren eginkizun nagusia*]₁ [*mundu guztian*]₃ [*alde guztietako eta bazter orotako jendearen giza-eskubideak eta demokrazia*]₅ [*babestea*]₄ [*da*]₂.

Esas dos versiones de la oración, ambas gramaticales en euskera, corresponderían a la oración **1c** en castellano estándar:

1c) [*Su tarea principal*]₁ [*es*]₂ [*proteger*]₄ [*en todo el mundo*]₃ [*los derechos humanos de toda la gente, y la democracia*]₅.

Ahora bien, si utilizamos un supuesto castellano cuyas ordenaciones de elementos fueran calcadas de las versiones **1a** y **1b** en euskera, tendríamos las siguientes secuencias, **2a** y **2b**:

2a) [Su tarea principal]₁ [es]₂ [en todo el mundo]₃ [proteger]₄ [de todos los lugares y rincones //de la gente los derechos humanos, y la democracia]₅

2b) [Su tarea principal]₁ [en todo el mundo]₃ [de todos los lugares y rincones//de la gente los derechos humanos, y la democracia]₅ [proteger]₄ [es]₂

Como se ve en **2a**, el verbo copulativo *es* está situado al comienzo de la oración, inmediatamente detrás del sujeto, por lo cual, una vez que el sujeto y el verbo han sido presentados, a partir de la posición de ese verbo copulativo, el/la lectora sabe que la información siguiente es el atributo, por lo que se puede disponer a procesar el resto del contenido informativo de la oración basándose en esa suposición.

Sin embargo, en la organización de la secuencia **2b**, el verbo copulativo (cast. *es*; eus. *da*) no es presentado hasta el final de la oración, por lo cual la persona lectora, hasta la aparición de ese verbo, no sabe a qué atenerse respecto a la información que está recibiendo, cuyo sentido completo se percibe una vez que el verbo copulativo está a disposición de la persona que está leyendo, y no antes. La tardía aparición del verbo es el motivo por el cual es muy posible que la persona lectora tenga que releer la oración completa una vez que ha procesado el contenido del verbo, que es un elemento determinante para establecer el universo predicativo y procesar progresivamente los otros elementos que constituyen el contenido informativo completo.

En cuanto al resultado que se ha obtenido en la encuesta acerca del ítem que estamos comentando, consideramos que es ciertamente significativo: el conjunto de especialistas prefirió la ordenación de mayor retrocarga, la que hemos denominado **1b** y **2b**, sólo en un 8,3 % de los casos, mientras que el profesorado prefirió, como media, la versión de retrocarga elevada en un 61,9 % de los casos. Como se ve, en ese ejemplo, el conjunto de especialistas tiene una tendencia mucho mayor a preferir ordenaciones de retrocarga reducida (en otras palabras, con el verbo situado en una posición anterior) que el colectivo de docentes de bachillerato analizado.

2.4 Participantes

En esta investigación se ha contado con la generosa colaboración de un total de 105 docentes del nivel de Bachillerato de diferentes disciplinas, tanto lingüísticas como no lingüísticas, pertenecientes a 21 centros de la Comunidad Autónoma Vasca, 15 de ellos públicos y 6 concertados (entre ellos, 4 ikastolas).

Además, también han colaborado desinteresadamente 12 especialistas que trabajan en los ámbitos de la traducción, los medios de comunicación y la universidad. Se ha asegurado que esas personas respondieran al criterio de especialización en el tema de estudio, si bien no pueden ser consideradas como representativas del conjunto global de las y los especialistas, dado que, para su elección, no se han utilizado criterios estadísticos sino que se ha tenido que atener al criterio de accesibilidad hacia las personas encuestadas. Aun así, consideramos que los resultados de este grupo de especialistas constituyen una referencia comparativa que informaría acerca de la tendencia más actualizada sobre el tema de este estudio.

2.5 Resultados

En el **cuadro 1** pueden verse los resultados comparativos entre la muestra de especialistas y del profesorado en valores medios porcentuales. Como dato significativo señalamos la clara diferencia existente entre los resultados del grupo de especialistas y el del profesorado de bachillerato: este último prefiere las ordenaciones de alta retrocarga en un 55,2 % de los casos, mientras que la muestra de especialistas sólo opta por esas ordenaciones en el 21% de los casos. Por el contrario, el conjunto de especialistas opta por las opciones menos retrocargadas en un 52,7 % de los casos mientras que esa opción es preferida por la muestra del profesorado solamente en un 17 % de las respuestas. Consideramos significativo ese diferente comportamiento entre el grupo de docentes y el de especialistas.

Es también digno de reseñar que tanto en el colectivo de especialistas como en el del profesorado hay un porcentaje importante de respuestas que aceptan como válidas las dos ordenaciones (19,7 % en el grupo de especialistas y 23,3 % en el del profesorado de bachillerato), que en definitiva dependerían de las características de un contexto lingüístico más amplio. Si se suman los porcentajes de preferencias por una menor retrocarga con aquellos que han

aceptado las dos ordenaciones, los porcentajes ascienden, en el grupo de especialistas, hasta un 72,4 % de “aceptación” de las ordenaciones de baja retrocarga, frente al 40,3 % en el caso del profesorado de bachillerato (una diferencia de 32,1 %).

	Preferencia por alta retrocarga (%)	Preferencia por baja retrocarga (%)	Ambas ordenaciones válidas (%)	Ambas ordenaciones incorrectas (%)	Ns/nc
Especialistas (N=12)	21 % (dt: 13,4)	52,7 % (dt: 28,8)	19,7 % (dt: 16,8)	6 % (dt:9,1)	0,7 %
Docentes bachillerato (N=105)	55,2 % (dt: 22,1)	17 % (dt: 16,2)	23,3 % (dt:18,9)	3,5 % (dt:4,9)	1 %

Cuadro 1. Resultados comparativos, % de valores medios (especialistas y profesorado bachillerato)(dt=desviación típica)

En el mismo **cuadro 1** se puede ver también que la desviación típica (d.t.) es elevada, lo que indica una notable dispersión en las respuestas.

	Especialistas (N=12)	Profesorado (N=105)
Quien más veces elige retrocarga alta (=elige verbo retrasado)	44 % de los ítems	100 %
Quien menos veces elige retrocarga alta (=elige verbo retrasado)	4 %	0 %
Quien más veces elige retrocarga baja (=elige verbo adelantado)	96 %	84 %
Quien menos veces elige retrocarga baja (=elige verbo adelantado)	8 %	0 %

Cuadro 2. Perfiles correctores extremos. Valores extremos obtenidos entre los/las participantes. Resultados (%)

Por otra parte, en el **cuadro 2** se pueden ver los datos referidos a los perfiles

correctores extremos, que presentan asimismo una enorme dispersión (hasta el punto de alcanzar valores casi completamente extremos), tanto si se consideran los valores correspondientes a la opción con el verbo más retrasado (es decir, la versión más retrocargada) como a la preferencia por la ordenación que tiene el verbo más adelantado (en otras palabras, la opción con menor retrocarga). Esa diferencia tan notable en los valores extremos de los dos colectivos es un indicador de la divergencia de criterios existente entre las personas encuestadas. Aun así, es notable el hecho de que en el conjunto de especialistas la persona que más preferencia manifiesta por las ordenaciones más retrocargadas “solamente” utilice esa opción en el 44 % de los ítems, alejándose así de los valores extremos y manifestando un rechazo relativo (y llamativo) hacia las ordenaciones más retrocargadas.

En el **cuadro 3** se pueden ver los datos referidos a las razones de las preferencias expresadas. La razón predominante entre las personas especialistas corresponde al criterio pragmático/comunicativo (60,8 %), mientras que en el caso del profesorado esa razón iguala al criterio gramatical relativo (“es mejor euskera”), alcanzando prácticamente el 30 % de las respuestas. Entre el profesorado tiene un notable seguimiento el criterio gramatical “duro” (es decir, se elige una ordenación de los elementos porque la otra se considera gramaticalmente incorrecta), que alcanza el 22,3 %, mientras que esa razón es menos elegida por el grupo de especialistas (7,2 %). Así pues, el criterio pragmático-comunicativo es seguido en mayor medida por los y las especialistas que por el profesorado, lo que sugiere la posibilidad de una evolución en los criterios, que estaría en un estadio más avanzado en el caso de los y las especialistas. Apuntamos que el criterio más normativista del profesorado esté probablemente influenciado por su función docente, especialmente (y comprensiblemente) preocupada por transmitir a los y las estudiantes un modelo canónico de lengua “claro”, factor que le llevaría a optar por las tendencias predominantes más conservadoras, con menos margen para el seguimiento a propuestas innovadoras (y, por tanto, más discutibles).

Criterio	Especialistas (N=12)	Profesorado (N=105)
Corrección vs incorrección (criterio gramatical “duro”)	7,2	22,3
Mejor modelo de lengua (criterio gramatical “relativo”)	23,2	29,5
Más fácilmente “legible/comprensible” (criterio pragmático-comunicativo)	60,8	29,1
Más adecuado para código escrito	8,8	9,2
“en euskera el verbo se pone al final”	0	5,9
Otros criterios	0	4,0

Cuadro 3: Criterios aducidos para explicar las opciones elegidas (%)

3. REFLEXIÓN Y CONCLUSIONES

Existe entre los y las participantes cierta tendencia a preferir ordenaciones con más alta retrocarga (el verbo hacia el final de la oración), tanto en el grupo de especialistas como en el de docentes de bachillerato, lo que sería consecuencia de la influencia del *modelo idealista-idiosincrático*, antes predominante y todavía en vigor (Maia, 2014; Maia y Badiola, 2015; véase el **cuadro 1**).

Sin embargo, el grupo de especialistas muestra una tendencia mucho menor hacia las ordenaciones más retrocargadas: 21 % vs 55,2 %. Sugerimos que esa diferencia sería debida al hecho de que el conjunto de especialistas utilizaría criterios más actualizados, basados en propuestas de ordenaciones correspondientes al *modelo de adaptación progresiva y ponderada* (Maia y Badiola, 2015), que discuten la necesidad y/o la idoneidad de colocar el verbo en la parte final de las oraciones afirmativas.

También se ha detectado un impulso hacia la preferencia por ordenaciones con el verbo más adelantado (lo que significa una menor retrocarga). Entre los y las especialistas analizadas, esa tendencia alcanza el 52,7 % de las respuestas, frente a un 17 % recogido en la muestra del profesorado (ver **cuadro 1**). Así pues, la tendencia a preferir ordenaciones de menor retrocarga es notablemente mayor entre los y las especialistas que entre el profesorado de bachillerato.

Existe asimismo una notable tendencia, alrededor del 20% de las respuestas, hacia la aceptación de las dos ordenaciones planteadas, haciéndolas depender de un entorno de enunciación más amplio (19,7 % para los y las especialistas, y 23,3% en la muestra de docentes de bachillerato). Ello parece indicar una postura precavida ante casos que se consideran dudosos, lo que podría indicar cierto conocimiento de la complejidad del tema y el alejamiento de soluciones demasiado simplificadoras.

Respecto a las razones de las preferencias, la más mencionada por el grupo de especialistas es la pragmática (“es más fácilmente legible/comprendible”: 60,8%). Así, el criterio “pragmático” de “comunicatividad” es con mucho el más elegido por los y las especialistas. Ese criterio “pragmático” es notablemente secundado también por el profesorado (29,1 %).

El criterio menos frecuentemente elegido entre los y las especialistas es la razón simplista formulada como “en euskera el verbo se pone al final” (0%). Llama la atención que en un estudio realizado con alumnado universitario en 2011 (Maia, 2017), dicha razón superficial fue la más elegida por los y las estudiantes (28 %).

En cuanto a las razones preferidas por el profesorado de bachillerato, el criterio simplista sólo es secundado en un 5,9 % del total de las respuestas, por lo que, si comparamos los dos estudios de 2011 y de 2018, ese criterio ha perdido fuerza de manera notable (desde 28 % hasta el 5,9 %), en beneficio de razones más matizadas, como el criterio “gramatical relativo” (que ha sido la respuesta mayoritaria entre el profesorado de bachillerato: 29,5 %), y el criterio “pragmático-comunicativo”, que llega también hasta el 29,1 %.

El criterio gramatical “duro” (“una forma es correcta y la otra es incorrecta”), en los ítems objeto de este estudio, tiene poco valor relativo entre las personas especialistas (7,2 % de las respuestas; antes 20,5 % entre los y las estudiantes, en 2011). Es de señalar que este criterio sigue teniendo entre el profesorado un seguimiento notable (22,3 %).

En definitiva, los datos obtenidos permiten sugerir una evolución que va desde posiciones más simplistas y más tendentes a altas retrocargas hacia unos criterios más matizados, basados en aspectos pragmáticos de comunicatividad y en criterios lingüísticos más cuidadosamente ponderados, que en su conjunto apuntan a una mayor aceptación de ordenaciones con menores retrocargas.

Así pues, respondiendo a la pregunta que encabeza este artículo, parece que especialistas y profesorado van en la misma dirección (adelantar la posición del verbo en la oración), pero el grupo de especialistas iría más avanzado, puesto

que acepta en mayor medida unas ordenaciones de reducida retrocarga.

En ese contexto, consideramos crucial la labor del sistema educativo en la tarea de extender e intensificar la reflexión sobre el orden de los elementos de la oración que resulte más adecuado en cada caso para alcanzar una mayor eficiencia comunicativa en la prosa escrita en euskera. En ese sentido, aquellos ítems analizados en este estudio que han sido ampliamente preferidos por el grupo de especialistas en su versión con retrocarga reducida son, a nuestro juicio, ejemplos de mejora pragmático-comunicativa, y por lo tanto sería positivo promocionar esas ordenaciones en lugar de otras que colocan el verbo en una posición posterior.

Finalmente, a la hora de proponer ordenaciones preferentes de los elementos de la oración también hay que tener en cuenta las tendencias y recursos adoptados por las poderosas lenguas que coexisten con el euskera (español/francés, inglés...), así como la mayor o menor dificultad de procesamiento de las diferentes ordenaciones por parte de los y las hablantes plurilingües de euskera, que en número creciente utilizan tres o más lenguas en su actividad diaria.

REFERENCIAS BIBLIOGRÁFICAS

- ALTUBE S. 1975 [1929]. *Erderismos*. Bilbao: Cinsa.
- ESNAL, P. 2009. “Testu-arkitektura eta gramatika, ekintza komunikatiboaren baitan”, *Hizpide*, 71: 3-160.
- EUSKALTZAINDIA 2011. *Hitz-ordena. Erabilera estrategikoa*. Bilbao: Euskaltzaindia.
- HIDALGO, B. 1995. *Hitzen ordena euskaraz*. Tesis doctoral. UPV-EHU.http://www.euskara.euskadi.net/appcont/tesisDoctoral/PDFak/bittor_hidalgo.pdf (10-05-2019)..
- HIDALGO, B. 2002. “Hitzen ordena esaldian”, *Senex*, 25: 75-109.
- MAIA, J. 2014. “Euskararen hitz-ordenaren norabidea: idealismo ideologikotik pragmatismo komunikatiborantz”, *Fontes Linguae Vasconum*, 117: 121-168.
- MAIA, J. 2017. “Erranaldiko elementuen ordena euskaraz: Lehen Hezkuntzako irakasle-gaiek zer hobetsi duten eta zergatik”, *Tantak*, 29 (2): 69-97.
- MAIA-LARRETxea, J.; Badiola-Urbe, N. 2015. “Word-order and Basque logical-discursive prose development: Different approaches”, *Dominated Languages in the 21st Century*, Barbara Schrammel-Leber, Christina Korb (Eds.). Graz: Karl-Franzens-Universität Graz. 134-147.
- MITXELENA, K. 1981. “Galdegai eta mintzagaia euskaraz”, *Euskal linguistika eta literatura: bide berriak*. Bilbao: Publicaciones de la Universidad de Deusto.
- RUBIO, J. 2001. *Euskararen garabideak*. Irun: Alberdania.
- SARASOLA, I. 2016. *Bitakora kaiera*. Donostia-San Sebastián: Erein.
- ZUBIMENDI, J. R., ESNAL, P. 1993. *Idazkera-liburua*. Vitoria-Gasteiz: Gobierno Vasco-Consejería de Cultura.

MEDIDAS PARA CALCULAR LA COMPLEJIDAD RELATIVA DE LAS LENGUAS NATURALES

M. Dolores Jiménez López
Universitat Rovira i Virgili

RESUMEN

Este trabajo se enmarca en el ámbito de los estudios sobre complejidad lingüística, un área que en los últimos años ha adquirido especial relevancia en las investigaciones sobre el lenguaje natural. Como alternativa a las medidas propuestas para calcular la complejidad de las lenguas naturales, proponemos basar el cálculo de la complejidad relativa en el concepto de “aprendibilidad”. Para ello presentamos un modelo basado en técnicas de aprendizaje automático y determinamos la complejidad lingüística (en términos de coste/dificultad de adquisición) en función del número de interacciones que necesita el sistema para llegar a un buen nivel de dominio de la lengua.

Palabras clave: complejidad lingüística, aprendizaje automático, aprendibilidad

ABSTRACT

This work focuses on the studies on linguistic complexity, an area that in recent years has acquired special relevance in the research on natural language. As an alternative to the measures proposed to calculate the complexity of natural languages, we propose to base the calculation of relative complexity on the concept of “learnability”. For this, we present a model based on machine learning techniques and determine the linguistic complexity (cost/difficulty of acquisition) in terms of the number of interactions that the system needs to reach a good level of language proficiency.

Keywords: linguistic complexity, machine learning, learnability

1. INTRODUCCIÓN

En este trabajo, nos centramos en el ámbito de los estudios sobre complejidad lingüística y proponemos medidas basadas en modelos de aprendizaje automático para calcular la complejidad relativa de las lenguas naturales.

El número de trabajos publicados en los últimos años sobre complejidad tanto en el ámbito de la lingüística teórica como en el de la lingüística aplicada (Baechler y Seiler, 2016; Baerman et al., 2015; Coloma, 2017; Conti, 2018; Di Domenico, 2017; Kortman y Szmrecsanyi, 2012; La Mantia et al., 2017; McWhorter, 2012; Newmeyer y Preston, 2014; Ortega y ZhaoHong, 2017) pone de manifiesto el interés por encontrar un método para calcular la complejidad lingüística e intentar responder a la pregunta de si todas las lenguas son iguales en lo que a complejidad se refiere o, si por el contrario, difieren en sus niveles de complejidad.

Las medidas propuestas para calcular la complejidad lingüística son muy variadas y dependen del tipo de complejidad que se esté analizando. Los estudios sobre complejidad absoluta utilizan medidas como el número de categorías o de reglas, la ambigüedad, la redundancia o la longitud de la descripción (Miestamo, 2008). Quienes se ocupan de complejidad relativa proponen diferentes medidas en función de la tarea (aprendizaje, adquisición, procesamiento) y el tipo de agente (hablante, oyente, niño, adulto) que consideren. La complejidad de procesamiento (Hawkins, 2009) o la complejidad de aprendizaje de segundas lenguas en adultos (Trudgill, 2001; Kusters, 2003) son ejemplos de medidas que se han propuesto en términos de dificultad/coste. En muchos casos, se utilizan medidas procedentes de otras disciplinas. La teoría de la información con formalismos como la entropía de Shannon o la complejidad de Kolmogorov (Dahl, 2004; Juola, 2008; Bane, 2008; Miestamo, 2008); los modelos computacionales basados en gramáticas de restricciones (Blache, 2011); o la teoría de sistemas complejos (Andrason, 2014) son ejemplos de áreas que han proporcionado medidas para calcular la complejidad lingüística.

A diferencia de lo que se ha venido haciendo en los estudios sobre complejidad lingüística, nuestra investigación se centra en el proceso de adquisición y considera al niño que adquiere su lengua materna para calcular la complejidad relativa de las lenguas. Para medir la complejidad utilizamos un modelo de aprendizaje automático y determinamos el coste/dificultad de adquisición en función del número de interacciones que necesita el sistema para llegar a un

buen nivel de dominio de la lengua.

En este trabajo, tras revisar algunas de las medidas que se han propuesto en el ámbito de los estudios sobre complejidad lingüística, presentamos nuestro modelo para calcular la complejidad relativa en términos de adquisición de primeras lenguas; mostramos las medidas consideradas para evaluar el rendimiento de nuestro sistema de aprendizaje (exactitud y completitud); discutimos los diferentes problemas que plantean estas medidas, entre ellos la dificultad de disponer de un “gold standard” para evaluar el modelo; e introducimos el modelo de lenguaje que hemos diseñado para generar este estándar.

2. ¿CÓMO SE MIDE LA COMPLEJIDAD LINGÜÍSTICA?

Las medidas y criterios que se han propuesto para calcular la complejidad lingüística varían y dependen de los intereses específicos de investigación y de la definición de complejidad adoptada.

Los principales tipos de complejidad que aparecen en la bibliografía son los siguientes:

- *Complejidad absoluta vs. complejidad relativa* (Miestamo, 2008)
 - La *complejidad absoluta* se define como una propiedad objetiva del sistema y se calcula en términos de número de partes del sistema, número de interrelaciones entre las partes o longitud de la descripción de un fenómeno. Esta aproximación es habitual en los estudios de tipología y está representada por trabajos como McWhorter (2001) y Dahl (2004).
 - La *complejidad relativa* tiene en cuenta a los usuarios del lenguaje. Se identifica con la dificultad o coste de procesamiento, aprendizaje o adquisición. Es habitual en los estudios de sociolingüística y psicolingüística y está representada por trabajos como el de Kusters (2003).
- *Complejidad global vs. complejidad local* (Miestamo, 2008).
 - La *complejidad global* calcula la complejidad total del sistema lingüístico. Se trata de una tarea difícil que se enfrenta al problema de la representatividad —no es posible considerar de forma exhaustiva todos los aspectos relevantes de la gramática de una lengua— y al problema de la comparabilidad. La contribución

de los diferentes dominios gramaticales a la complejidad global es inconmensurable.

- La *complejidad local* analiza la complejidad de subdominios particulares de la lengua y se presenta como una tarea abordable.
- *Complejidad del sistema vs. complejidad estructural* (Dahl, 2004).
 - La *complejidad del sistema* hace referencia a las propiedades de una lengua, mide el número de distinciones dentro de una categoría y calcula el contenido de la competencia del hablante.
 - La *complejidad estructural* calcula la cantidad de estructura de un objeto lingüístico, analiza la estructura de las expresiones.

A partir de estas diferencias en la definición de complejidad, las medidas que se han propuesto pueden clasificarse en dos grandes grupos:

1. *Medidas basadas en la gramática.* Se trata de medidas de complejidad absoluta y, por tanto, objetivas. Suelen calcular y comparar el grado de complejidad de cada componente gramatical, considerando el número de categorías, la longitud de la descripción, la ambigüedad, la redundancia, etc.
2. *Medidas basadas en el usuario.* Son medidas subjetivas en las que se calcula la complejidad en términos de coste y dificultad desde el punto de vista del usuario: 1) adquisición del lenguaje: ¿se tarda más en adquirir unas gramáticas que otras?; 2) aprendizaje de segundas lenguas: ¿se tarda más en aprender unas gramáticas que otras?; 3) uso del lenguaje: ¿son algunas gramáticas más difíciles de usar que otras?

En lo referente al primer bloque —medidas basadas en la gramática—, podemos hacer referencia a las medidas de complejidad local que Kortmann y Szmrecsanyi (2012) destacan: *complejidad fonológica*: dimensión del inventario fonológico, distinciones tonales, fonología suprasegmental, restricciones fonotácticas, etc.; *complejidad morfológica*: morfología flexiva, alomorfía, etc.; *complejidad sintáctica*: número de reglas sintácticas, recursividad, subordinación, etc.; *complejidad semántica y léxica*: homonimia, polisemia, distinciones en los pronombres personales, etc.; *complejidad pragmática*: grado de inferencia impuesta por la lengua.

La entropía de Shannon o la complejidad de Kolmogorov son medidas que se han considerado para calcular la complejidad lingüística en términos absolutos. En el primer caso, se considera que un mensaje es complejo si

tiene un gran contenido de información, y una lengua es compleja si enviar el mensaje en esa lengua requiere mucho más ancho de banda que el contenido de información del mensaje. La complejidad de Kolmogorov, por su parte, mide la informatividad de una cadena como la longitud del algoritmo necesario para describir/generar esa cadena. Calcula la longitud de la descripción necesaria para especificar un objeto. Cuanto más grande es la descripción, mayor es la complejidad.

En lo referente al segundo bloque —medidas basadas en el usuario—, Menn y Duffield (2014), por ejemplo, proponen medidas basadas en el esfuerzo de procesamiento y utilizan métodos empíricos basados en análisis psicolingüísticos (tiempo de procesamiento; precisión a la hora de procesar la estructura; posibilidad de realizar otras tareas mientras se procesa la estructura, etc.) o neurolingüísticos (miden el esfuerzo cerebral a la hora de procesar una determinada estructura) que permiten calcular la complejidad relativa y estructural. En lo referente a la complejidad relativa en el aprendizaje de segundas lenguas, Trudgill (2001) defiende que “linguistic complexity equates with difficulty of learning for adults” y Kusters (2003) define la complejidad como “the amount of effort an outsider has to make to become acquainted with the language in question [...]. An outsider is someone who learns the language in question at a later age, and is not a native speaker”.

Concluimos este breve repaso a algunas de las medidas de complejidad utilizadas, con la propuesta de Kortmann y Szmrecsanyi (2012). Estos autores consideran las siguientes medidas de complejidad. En primer lugar, lo que ellos denominan “*absolute-quantitative complexity*” que se puede resumir en la idea de “más significa más complejo” y en la que se aplica un criterio eminentemente cuantitativo en el que lo único que interesa es el número de elementos en los distintos componentes de una lengua. En segundo lugar, la denominada “*redundancy-induced complexity*” —considerada una medida que abarca tanto la complejidad absoluta como la relativa— que considera como indicador de complejidad la redundancia. En tercer lugar, la “*irregularity-induced complexity*” —que combina la complejidad absoluta y relativa— y que tiene en cuenta la presencia de irregularidad y la relaciona con altos grados de complejidad. Y, por último, lo que ellos denominan “*L2 acquisition complexity*” que mide la complejidad calculando la dificultad para aprender una lengua como lengua extranjera. A estas cuatro medidas añaden tres más: “*processing complexity*”; “*morphological naturalness*” e “*information-theoretic complexity*.”

3. COMPLEJIDAD EN TÉRMINOS DE “APRENDIBILIDAD”

Como alternativa a las medidas propuestas en los estudios sobre complejidad lingüística y con el objetivo de medir la complejidad relativa teniendo en cuenta el proceso de adquisición del lenguaje, en este trabajo proponemos el uso de herramientas computacionales procedentes del ámbito del aprendizaje automático para calcular la complejidad de las lenguas naturales. Estas herramientas nos permiten proponer como medida para determinar la complejidad lingüística el concepto de *aprendibilidad*. El concepto de aprendibilidad es habitual en los estudios de aprendizaje automático. Aplicado al ámbito del aprendizaje automático de lenguas, podemos definir la teoría de la aprendibilidad como un modelo que tiene como objetivo “*to logically and mathematically formulate the possible ways that language can be learned from the input of the surrounding speech community*” (Gierut, 2007).

En este trabajo, aplicamos al cálculo de la complejidad relativa de las lenguas el modelo propuesto en Becerra *et al.* (2015, 2016a, 2016b). Este modelo se presenta como un sistema artificial que, sin ningún conocimiento lingüístico previo, puede aprender una lengua a partir de pares que consisten en una oración y el contexto en el que esta se ha producido.

De las dos aproximaciones propuestas por Becerra *et al.* (2015, 2016a, 2016b), en este trabajo usamos el modelo que aprende a partir del “*abstract scenes dataset*” (Becerra *et al.*, 2016a, 2016b). Se trata de una base de datos creada utilizando la herramienta Amazon’s Mechanical Turk (AMT). La base de datos contiene imágenes/escenas de clip art en las que se representan niños jugando al aire libre y oraciones que describen estas imágenes. Para generar esta base de datos se pidió a los trabajadores de AMT que a partir de 80 piezas de clip art crearan escenas que representarían a un niño y una niña con diferentes poses y expresiones faciales, y algunos objetos, como juguetes, árboles, animales, sombreros, etc. Luego, a un nuevo grupo de trabajadores se les pidió que describieran las escenas usando tres oraciones; en las descripciones debían usar palabras básicas que pudieran aparecer en un libro para niños. En total, la base de datos contiene 10.020 imágenes y 30.060 oraciones.

El sistema aprende a partir de pares (O, I) que consisten en una oración (O) y una imagen (I), donde la oración O (parcialmente) describe la imagen I. Una oración se representa como una secuencia de palabras (n-grama). Una fase de pre-procesamiento transforma la información proporcionada por la base de datos en un contexto. Los contextos están formados por un conjunto

de elementos que describen propiedades y relaciones entre los objetos en la imagen. Es importante poner de manifiesto que el contexto solo describe lo que el aprendiz puede percibir. En contraste con otros enfoques, el significado no está representado explícitamente, el aprendiz tiene que descubrirlo. Las imágenes se proporcionan al sistema a través de una representación basada en lógica de primer orden. Usando técnicas de programación lógico-inductiva, el sistema computa el significado asociado con cada una de las frases (n-gramas) de la base de datos. Es decir, el sistema aprende una relación/función entre un n-grama y su representación semántica. El significado de un n-grama es lo que hay en común en todas las situaciones en las que el n-grama es utilizado. Los experimentos muestran que el sistema es capaz de aprender una serie de relaciones n-gramas/significados y usar ese conocimiento para una variedad de propósitos. De hecho, el sistema es capaz de aprender el significado de las palabras y de generar frases relevantes para una escena determinada.

Proponemos utilizar este sistema artificial para estudiar la complejidad lingüística desde un punto de vista relativo porque consideramos que este modelo presenta analogías importantes con el proceso de adquisición de primeras lenguas. En el modelo de aprendizaje automático proporcionamos a la máquina una serie de frases producidas todas ellas en un contexto y la máquina debe identificar el lenguaje subyacente a partir de estos datos, en definitiva, la máquina debe ser capaz de aprender la lengua. De manera similar, en el proceso de adquisición del lenguaje, los niños reciben datos lingüísticos y a partir de ellos aprenden su lengua materna. El *input* que recibe el sistema tiene propiedades similares al que recibe el niño en su entorno de aprendizaje y el sistema no tiene conocimientos previos sobre la lengua a aprender. Además el sistema permite realizar análisis croslingüísticos, ya que se utiliza un algoritmo único para aprender cualquier lengua, y esto podría ser equivalente a la capacidad innata que permite a los humanos adquirir una lengua.

¿Cómo calculamos el coste/dificultad para adquirir una lengua utilizando este sistema artificial? En términos de aprendibilidad o dificultad para adquirir el lenguaje. Contamos la cantidad de ejemplos necesarios para que el sistema logre aprender una lengua determinada. Es decir, determinamos el coste/dificultad de adquisición en función del número de interacciones que necesita el sistema para llegar a un buen nivel de dominio de la lengua.

4. EVALUACIÓN DEL MODELO: “GOLD STANDARD”

El modelo que utilizamos calcula la cantidad de interacciones que son necesarias para lograr un buen nivel de rendimiento en una lengua determinada mediante el uso de un algoritmo único para aprender cualquier lengua. Por lo tanto, este modelo nos permite calcular el *coste* —en términos del número de interacciones— para alcanzar un buen nivel de actuación en una lengua dada. Ahora bien, ¿cómo sabemos que el sistema ha llegado a un “buen” dominio de la lengua? Utilizamos distintas medidas para evaluar la actuación del aprendiz como son la precisión de los significados referenciales aprendidos o la capacidad del sistema para comentar imágenes no mostradas, es decir la capacidad de generar frases relevantes encadenando n-gramas. Pero, sobre todo evaluamos el rendimiento de nuestro sistema de aprendizaje en términos de:

- Exactitud (*Correctness*): número frases correctas que el aprendiz es capaz de producir.
- Completitud (*Completeness*): capacidad del sistema de producir todas las frases correctas posibles.

A la hora de evaluar en términos de exactitud y completitud el rendimiento del sistema nos encontramos con un importante problema: dado un contexto, no es trivial especificar el conjunto total de frases posibles. Por tanto, es difícil disponer de un “gold standard” para evaluar el modelo. Para solucionar este problema, hemos definido un *modelo de lenguaje* para generar un “gold standard” que sirva para evaluar el sistema de aprendizaje.

Para definir el *modelo de lenguaje* se utilizan gramáticas semántico-contextuales y se combina razonamiento de tipo deductivo y razonamiento de tipo abductivo. El modelo se implementa mediante tecnología de programación lógica estándar e integra reglas gramaticales y una base de conocimiento semántico contextual. Las reglas gramaticales se especifican con la notación de las *Definite Clause Grammars* (para Prolog) y el programa se utiliza de forma deductiva. La base de conocimiento semántico contextual —conjunto de hechos lógicos que representan información sobre los objetos de las imágenes, sus propiedades y sus relaciones— está representada por *Constraint Handling Rules* (Christiansen y Dahl, 2005) y realiza un análisis abductivo de oraciones dadas.

El *modelo de lenguaje*, combinando competencia gramatical y competencia semántica, es capaz de:

- Extraer información de las frases que conforman el corpus;
- y generar nuevas frases para construir un “gold standard”.

Nuestro objetivo es, para cada imagen del corpus, extender la base de conocimiento contextual con el conocimiento que contienen las oraciones sobre la imagen, de manera que podamos generar nuevas oraciones coherentes con la imagen.

A partir de la base de datos de la que disponemos (imágenes + frases), se desarrolla una gramática en un proceso iterativo que combina conocimiento general sobre la lengua y las construcciones y el vocabulario utilizado en el corpus. En una primera fase, la gramática *extrae* información de la pequeña base de conocimiento de la que disponemos inicialmente (contexto para cada imagen) y del análisis de las frases y produce una *base de conocimiento ampliada*. En una segunda fase, utilizando los contextos, las frases y la base de datos ampliada, la gramática *genera* todas las frases posibles para una imagen determinada.

En resumen, el *modelo de lenguaje* definido permite:

- Mejorar la base de conocimiento semántico contextual, mediante una interpretación abductiva de las oraciones que permite extraer el conocimiento incrustado en oraciones.
- Generar nuevas frases, mediante un análisis deductivo y utilizando la base de conocimiento semántico contextual mejorada.
- Construir parte del conocimiento del mundo, analizando una gran colección de oraciones para diferentes imágenes.

Obtenemos así un “gold standard” que nos permite evaluar el rendimiento del sistema de aprendizaje. Ahora, comparando el output del sistema con el “gold standard” generado, podemos calcular el número frases correctas que el sistema de aprendizaje es capaz de producir (exactitud) y determinar si tiene la capacidad de producir *todas* las frases correctas posibles (completitud).

De esta forma, el cálculo de la complejidad puede realizarse de dos maneras diferentes (considerando tanto la exactitud como la completitud):

1. Podemos considerar un número fijo de interacciones (igual para cualquier lengua). Si el sistema presenta mayor exactitud y/o

completitud en una lengua que en otra, esto significa que la lengua en la que consigue mayor exactitud y/o completitud es menos compleja.

2. Podemos contar el número de interacciones necesarias para conseguir una exactitud y/o completitud óptima. La lengua que más interacciones necesite para llegar a ese nivel de exactitud y/o completitud óptima será más compleja.

5. CONCLUSIONES

El trabajo que hemos presentado se enmarca en un proyecto de investigación (FFI2015-69978-P) —que pretende proporcionar herramientas para determinar el nivel de complejidad relativa de las lenguas— y tiene como objetivo rebatir el dogma de la equicomplejidad considerando dos de las objeciones que sus defensores han planteado a quienes han pretendido determinar la complejidad de las lenguas: 1) Las lenguas son inconmensurables en lo que a complejidad se refiere. 2) Todas las lenguas tienen el mismo nivel de complejidad.

Hemos presentado un modelo para calcular la complejidad relativa en términos de adquisición de primeras lenguas y una herramienta para generar un “gold standard” que permita evaluar el rendimiento del sistema de aprendizaje.

El modelo de aprendizaje automático que utilizamos permite calcular la complejidad de las lenguas naturales. Utiliza técnicas de programación lógico inductiva para que el sistema aprenda un lenguaje a partir de frases que se producen en un contexto y analiza una serie de frases y el contexto en el que estas se producen para aprender una lengua. Permite contar la cantidad de interacciones (secuencias imagen/frase) que necesita el sistema para aprender la lengua, haciendo posible determinar la complejidad de la lengua considerada.

Entre las ventajas que este modelo presenta para medir la complejidad relativa de las lenguas destacamos las siguientes: utiliza datos reales; se centra en el proceso de aprendizaje; no requiere ningún conocimiento previo sobre la lengua y aprende de forma incremental.

REFERENCIAS BIBLIOGRÁFICAS

- ANDRASON, A. 2014. Language complexity: An insight from complex-system theory. *International Journal of Language and Linguistics*, 2/2: 74-89.
- BAECHLER, R. y SEILER, G. 2016. *Complexity, Isolation, and Variation*. Berlin: Mouton de Gruyter.
- BAERMAN, B., BROWN, D. y Corbett, G.G. 2015. *Understanding and Measuring Morphological Complexity*. Oxford: Oxford University Press.
- BANE, M. 2008. "Quantifying and measuring morphological complexity". En Ch. Chang y H. Haynie (eds.), *Proceedings of the 26th West Coast Conference on Formal Linguistics*. Somerville: Cascadilla Proceedings Project, 69-76.
- BECERRA-BONACHE, L., BLOCKEEL, H., GALVÁN, M. y JACQUENET, F. 2015. "A first-order-logic based model for grounded language learning". En *Advances in Intelligent Data Analysis XIV*. Berlin: Springer, 49-60.
- BECERRA-BONACHE, L., BLOCKEEL, H., GALVÁN, M. y JACQUENET, F. 2016a. "Learning language models from images with regll". En *Machine Learning and Knowledge Discovery in Databases -European Conference 2016*, 55-58.
- BECERRA-BONACHE, L., BLOCKEEL, H., GALVÁN, M. y JACQUENET, F. 2016b. "Relational grounded language learning". En *22nd European Conference on Artificial Intelligence*. The Hague, 1764-1765.
- BLACHE, Ph. 2011. "A computational model for linguistic complexity". En G. Bel-Enguix, V. Dahl y M.D. Jiménez-López (eds.), *Biology, Computation and Linguistics. New Interdisciplinary Paradigms*. Amsterdam: IOS Press, 155-167.
- CHRISTIANSEN, H. y DAHL, V. 2005. "HYPROLOG: A new logic programming language with assumptions and abduction". En M. Gabbrielli and G. Gupta (eds.), *Logic Programming, 21st International Conference*. Berlin: Springer, 159-173.
- COLOMA, G. 2017. *La Complejidad de los Idiomas*. Bern: Peter Lang.
- CONTI JIMÉNEZ, C. 2018. *Complejidad Lingüística: Orígenes y Revisión Crítica del Concepto de Lengua Compleja*. Bern: Peter Lang.
- DAHL, O. 2004. *The Growth and Maintenance of Linguistic Complexity*. Amsterdam: John Benjamins.

- DI DOMENICO, E. 2017. *Syntactic Complexity from a Language Acquisition Perspective*. Newcastle upon Tyne: Cambridge Scholars Publishing.
- GIERUT J. A. 2007. Phonological complexity and language learnability. *American Journal of Speech-Language Pathology*, 16(1): 6-17.
- HAWKINS, J.A. 2009. "An efficiency theory of complexity and related phenomena". En G. Sampson, D. Gil y P. Trudgill (eds.), *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press, 252-268.
- JUOLA, P. 2008. "Assessing linguistic complexity". En M. Miestamo, K. Sinnemäki y F. Karlsson (eds.), *Language Complexity: Typology, Contact, Change*. Amsterdam: John Benjamins, 89-108.
- KORTMANN, B. y SZMRECSANYI, B. 2012. *Linguistic Complexity: Second Language Acquisition, Indigenization, Contact*. Berlin: Mouton de Gruyter.
- KUSTERS, W. 2003. *Linguistic Complexity: The Influence of Social Change on Verbal Inflection*. Utrecht: LOT.
- LA MANTIA, F., LICATA, I. y PERCONTI, P. (eds.) 2017. *Language in Complexity. The Emerging Meaning*. Berlin: Springer.
- MCWHORTER, J. 2001. "The world's simplest grammars are creole grammars", *Linguistic Typology*, 6: 125-166.
- MCWHORTER, J. 2012. *Linguistic Simplicity and Complexity: Why do Languages Undress?* Berlin: Mouton de Gruyter.
- MENN, L. y DUFFIELD, C.J. 2014. "Looking for a "Gold Standard" to measure language complexity: what psycholinguistics and neurolinguistics can (and cannot) offer to formal linguistics". En Newmeyer, F.J. y Preston, L.B. (eds.), *Measuring Grammatical Complexity*. Oxford: Oxford University Press, 281-302.
- MIESTAMO, M. 2008. "Grammatical complexity in a cross-linguistic perspective". En M. Miestamo, K. Sinnemäki y F. Karlsson (eds.), *Language Complexity: Typology, Contact, Change*. Amsterdam: John Benjamins, 23-42.
- NEWMAYER, F.J. y PRESTON, L.B. (eds.) 2014. *Measuring Grammatical Complexity*. Oxford: Oxford University Press.

- ORTEGA, L. y ZHAOHONG, H. 2017. *Complexity Theory and Language Development*. Amsterdam: John Benjamins.
- TRUDGILL, P. 2001. Contact and simplification: historical baggage and directionality in linguistic change. *Linguistic Typology*, 5 (2/3): 371-374.

**CODIFICACIÓN Y ETIQUETADO EN LOS CORPUS
DE APRENDICES Y SU APLICACIÓN DIDÁCTICA: LA
PROPUESTA DEL *CORPUS DE INTERLENGUA
ESPAÑOLA DE APRENDICES SINOHABLANTES*
(CINEAS)¹**

M^a Ángeles Calero Fernández
Maribel Serrano Zapata
Universidad de Lleida
M^a Begoña Gómez-Devís
Universidad de Valencia

RESUMEN

Los criterios con los que se diseña, se cataloga y se etiqueta un corpus lingüístico determinan las aplicaciones que dicho corpus podrá tener (Leech, 1993). En esta publicación se revisa la codificación y el etiquetado de distintos corpus de aprendices de español como lengua extranjera (ELE), así como algunos programas informáticos utilizados para ello, con el fin de evaluar el tipo de información de estos corpus que puede ayudar a caracterizar los estadios de la interlengua y a determinar los factores lingüísticos y extralingüísticos que intervienen en la formación, estructura y progreso de la interlengua.

Palabras clave: lingüística del corpus, corpus de aprendices, enseñanza de ELE

ABSTRACT

The criteria for designing, encoding and annotating a learner corpus determine its potential as a research tool and as a pedagogical resource. This paper reviews the criteria employed in coding and annotating a series of written learner corpora of Spanish. It also examines the possibilities these corpora offer for the following tasks: 1) identifying the features that characterize the different stages of the interlanguage of learner Spanish; 2) determining which linguistic and extralinguistic factors intervene in the formation, structure and progress of the interlanguage; and 3) determining which factors intervene in the teaching-learning process of a foreign language.

¹Este trabajo se ha realizado dentro del proyecto *Elaboración y catalogación de un corpus de textos escritos en ELE producidos por estudiantes sinohablantes*, que se está desarrollando en la Universidad de Lleida y que está financiado por el Ministerio de Economía y Competitividad (Nº de Referencia: FFI2016-80280-R) con la participación de Fondos Feder.

Keywords: corpus linguistics, corpus of apprentices, teaching of ELE

1. INTRODUCCIÓN

La pujanza actual de las investigaciones basadas en corpus no se concibe sin el notable desarrollo de los corpus textuales en lengua inglesa, iniciado por el estructuralismo americano de la segunda mitad del siglo XX. A partir de entonces la evolución ha sido muy rápida pues los avances originados por la necesidad de estudiar la lengua en uso no solo han alcanzado al desarrollo de la lingüística como disciplina sino también a la informática, que ha facilitado cada vez más el manejo de ingentes cantidades de datos; asimismo han sido cada vez más las lenguas sobre las que se han ido construyendo corpus y estos han ido derivando en toda una tipología.

Una definición básica de corpus lingüístico fácilmente reconocible podría ser: conjunto de textos informatizados producidos en situaciones reales que se han seleccionado siguiendo una serie de criterios lingüísticos explícitos los cuales garantizan que dicho corpus pueda ser usado como muestra representativa de la lengua (EAGLES, 1996). Ciñéndonos exclusivamente a consideraciones didácticas, las competencias de los corpus son numerosas. Para el alumno, el corpus constituye una base sólida para identificar las estructuras lingüísticas más frecuentes en las producciones reales de los hablantes nativos de una lengua; le otorga la autonomía de elegir por sí mismo qué aprender, cómo aprenderlo y en qué orden, y, además, es un instrumento para encontrar respuestas a una tipología muy variada de dudas concretas y de profundizar en ellas por medio del acceso a amplios contextos reales. Al profesor no nativo, el corpus le permite basarse no exclusivamente en su intuición y en ejemplos elaborados *ad hoc*, sino en una fuente amplia y fiable de recursos lingüísticos, así como seleccionar un *input* suficiente y de calidad al que enfrentar a sus alumnos (Pérez, 2007: 11). Los corpus de aprendices son un tipo de corpus lingüístico que recoge muestras de habla (oral y/o escrita) producidas por hablantes no nativos; su finalidad es registrar la forma en que estos se expresan en la lengua extranjera y, en consecuencia, aporta al docente información muy valiosa para la planificación de la enseñanza.

Pero no se puede acceder a la información de un corpus si este no ha sido preparado para buscar datos de manera rápida, sistemática y automática. Además de digitalizarlo, es necesario codificarlo y etiquetarlo, y la forma en

que se realicen estas tres tareas condicionará indefectiblemente las aplicaciones que dicho corpus pueda tener y el tipo de explotación que pueda acometerse (Leech, 1993).

Nuestro trabajo va a centrarse en la forma en que han sido diseñados los corpus de aprendices y su efecto en la enseñanza-aprendizaje de la lengua extranjera. Los objetivos que perseguimos son: 1) describir la forma en que se ha codificado y etiquetado los corpus de aprendices de ELE a partir de los metadatos que se tienen en cuenta, 2) presentar los programas informáticos utilizados para anotarlos, incluido uno recientemente creado en la Universidad de Lleida, y 3) por último, evaluar si los metadatos y la anotación que se han empleado pueden ayudar a caracterizar las etapas de la interlengua y los elementos lingüísticos y extralingüísticos que intervienen.

Así pues, daremos primero un repaso a este tipo de corpus y a los metadatos que utilizan. Seguidamente veremos distintos anotadores. Finalmente, concluiremos la utilidad didáctica que tiene todo ello.

2. LOS CORPUS DE APRENDICES DE ELE

Los corpus de aprendices no son desconocidos en ELE. Varias autoras (Penadés 2005, Baralo 2010, Ainciburu 2012, Buyse, Fernández y Verweckken 2016) han señalado la necesidad de bancos de datos longitudinales de ELE en diferentes contextos de aprendizaje de la lengua que puedan servir de punto de referencia para planificar su enseñanza. Y en los últimos 15 años han ido surgiendo corpus que recogen textos orales o escritos producidos por estudiantes de español. La diversidad que presentan estos corpus hace imprescindible describirlos para que docentes e investigadores puedan conocer las posibilidades de explotación de cada uno de ellos y, en consecuencia, puedan saber cuáles son los más adecuados a sus intereses (Rojo, 2010). Alonso (2016), resumiendo las principales características de estos corpus, destaca que la mayoría son escritos, transversales, con textos que desarrollan temas prefijados, especialmente narraciones y textos argumentativos, redactados por informantes habitualmente anglófonos adultos que están aprendiendo español y que tienen distintos niveles de competencia. Pasemos, pues, a relacionarlos y a retratarlos brevemente.

Los corpus orales de ELE que son consultables son:

- a. el *Corpus Oral de Español como Lengua Extranjera*² (CORELE) (http://cartago.lllf.uam.es/corele/home_es.html),
- b. el *Fono.ELE* (<http://www3.uah.es/fonoele/corpus-muestra.php>),
- c. el *Spanish Corpus Proficiency Level Training* (SPT) (<http://www.laits.utexas.edu/spt/intro>) y
- d. el *Spanish Learner Language Oral Corpora* (SPLLOC1 y SPLLOC2) (<http://www.splloc.soton.ac.uk/>)³.

El SPT y el SPLLOC recogen muestra de anglófonos de todos los niveles de dominio de la lengua, mientras que CORELE y *Fono.ELE* contienen muestras de informantes de diferentes lenguas cuyos niveles son de A2 y B1, en el primer caso, y desde A2 hasta C1, en el segundo. Los textos del SPLLOC y de CORELE son entrevistas, descripciones y relatos; los del SPT son diálogos a partir de una serie de preguntas, y *Fono.ELE* contiene grabaciones tanto de breves conversaciones estructuradas como de textos, frases y palabras leídas en voz alta. El SPLLOC y el CORELE cuentan con un corpus de control de hablantes nativos.

Por su parte, los corpus escritos de aprendices de ELE que son consultables son:

- a. el *Corpus de aprendices de español* (CAES) (<http://galvan.usc.es/caes>),
- b. el *Corpus del español de los italianos* (CORESPI) (<http://corespiyorite.altervista.org/>) y
- c. el *Corpus para el análisis de errores de aprendices de ELE* (CORANE) (Cestero y Penadés, 2009)⁴.

²En su versión inglesa se llama *Spanish Learner Oral Corpus* y puede consultarse en http://cartago.lllf.uam.es/corele/home_en.html.

³El CORELE fue realizado por Leonardo Campillos y subvencionado por la Consejería de Madrid y el Fondo Social Europeo. Por su parte, al frente del equipo Fono.ELE están Ana Blanco y M^a Ángeles Álvarez. El SPT fue realizado en el College of Liberal Arts de la Universidad de Texas en Austin bajo el liderazgo de Dale Koike. El *Spanish Learner Language Oral Corpora* fue financiado por el Consejo de Investigación Económica y Social del Reino Unido y llevado a cabo por un equipo integrado por 5 profesoras de las Universidades de York, Newcastle y Southampton.

⁴El CAES es el resultado de un proyecto financiado por el Instituto Cervantes y desarrollado por un equipo de la Universidad de Santiago de Compostela, liderado por Guillermo Rojo e Ignacio Palacios. La autora del CORESPI es Sonia Bailini. Por su parte, el CORANE fue financiado por el Consejo Social de la Universidad de Alcalá y su investigadora principal fue

El CAES y el CORANE contienen textos de aprendices de diversas lenguas maternas (7, en el primer caso, y 23, en el segundo), mientras que el CORESPI es específico de aprendices italo hablantes. Los niveles de competencia son distintos en cada corpus: en el CAES encontramos textos desde A1 hasta C1; en el CORESPI, desde A1 hasta B2; y, en el CORANE, desde A2 hasta C1. La tipología textual también varía de un corpus a otro: el CAES es el que presenta mayor diversidad con cartas, postales, correos electrónicos, reservas de hotel, solicitudes de admisión, biografías, narraciones, ensayos y reseñas críticas, todo recogido en línea. En cambio, el CORESPI y el CORANE se circunscriben a un tipo concreto de texto: en el primer caso, interacciones por email entre aprendices y hablantes nativos, y, en el segundo, textos ensayísticos redactados en el aula y fuera del aula. Los tres corpus son longitudinales, pues acogen muestras de los mismos aprendices en varios momentos de su proceso de aprendizaje.

También es accesible el LANGSNAP (<http://langsnap.soton.ac.uk/>), que ofrece tanto textos orales como escritos de estudiantes anglófonos. La oralidad se obtiene a través de entrevistas semiestructuradas y narraciones, y las muestras escritas, mediante breves textos argumentativos. Se trata de otro corpus longitudinal, puesto que se recogen 6 textos distintos de cada informante.

Por otro lado, se han confeccionado corpus de aprendices de español, tanto orales como escritos, de los que se conocen algunas de sus características pero que no son de acceso libre, sino que, al menos de momento, son o han sido explotados únicamente por los equipos que los han elaborado. Entre los orales, tenemos el *Corpus Oral de Interlengua Bilingüe Español-Italiano* (CORINÉI), con interacciones nativo/no nativo entre estudiantes de Traducción españoles e italianos a través de Skype, Facebook, Dropbox y similares, tratando diversos temas a lo largo del curso académico (González et al., 2016). Entre los corpus escritos, contamos con el *Anglia Polytechnic University Learner Spanish Corpus* (APU), el corpus *Aprender a Escribir en Lovaina* (Aprescrllov), el *Corpus de Aprendices de Español como Lengua Extranjera* (CAELE), el *Corpus Escrito del Español como L2* (CEDEL2), el *Japanese Learner Corpus of Spanish* (JLCS) y el *Tartu Learner Corpus of Spanish as a L3* (Tartu)⁵. Los corpus que toman muestras

Ana M^a. Cestero.

⁵Para más información sobre corpus de aprendices, pueden consultarse el *Learner corpora around the world*, de la Universidad Católica de Lovaina (<https://uclouvain.be/en/research-institutes/ilc/cecl/learner-corpora-around-the-world.html>), el *Indexador de Corpus de Aprendices de Español*, desarrollado en la Universidad Complutense de Madrid, (http://repositorios.fdi.ucm.es/corpus_aprendices_espa%C3%B1ol/view/paginas/view_paginas.php?id=1) y el buscador de corpus de la revista electrónica *LinRed* (<http://www.linred.es/corpus.html>), de la Universidad

de aprendices de una lengua materna específica son el Aprescilov (holandés), el CEDEL2 (inglés), el JLCS (japonés), el Tartu (estonio). En cambio, el APU y el CAELE recogen textos procedentes de estudiantes de español con lenguas maternas diversas. En cuanto a los niveles de competencia, el APU y CEDEL2 contemplan todos los niveles; el CAELE se mueve entre textos de A2 y B1, y el Aprescilov contiene textos desde A1 hasta C1.

3. LOS METADATOS EN LOS CORPUS DE APRENDICES DE ELE

Para entender las posibilidades de explotación que tiene un corpus, además de conocer qué tipos de textos incluye, cómo se han obtenido estos, qué características tienen los aprendices que han proporcionado las muestras orales y/o escritas, es necesario saber cuál es la anotación no lingüística (codificación) que emplea dicho corpus, cuál es su anotación lingüística (etiquetado) y cómo se recuperan los datos.

Los metadatos que tienen en cuenta los corpus de aprendices de ELE que ofrecen aplicativos de consulta automática, metadatos que deben ser codificados y que permiten filtrar la información en las búsquedas automáticas, son de dos tipos: sociolingüísticos y textuales. Los primeros constituyen las habituales variables sociales adscritas o adquiridas que se utilizan en los estudios sociolingüísticos: así, se contempla el sexo o género, la edad y la lengua materna, como variables adscritas, y, como variables adquiridas, podemos encontrar el curso, el centro de estudios, el nivel de dominio de ELE (subjeto o no), los años de estudio del español, las estancias en países hispanohablantes, el nivel de contacto con la cultura hispana (subjeto), el nivel de dominio de otras LE (subjeto o no), la lengua hablada en casa, o las experiencias de aprendizaje. Por su parte, los metadatos textuales incluyen el tipo de tarea que realiza el aprendiz y que genera su muestra de habla oral o escrita, si esta tarea se ejecuta en un contexto controlado o no, o el tipo textual que se requiere al aprendiz.

Pues bien, el corpus *FonoEle* permite extraer datos atendiendo al sexo o género, a la edad, al nivel de dominio de ELE, al nivel de competencia en otra lengua extranjera, y al contacto con el español. El CAES proporciona búsquedas por sexo o género, edad, lengua materna, nivel de competencia de ELE y país de residencia. El CORESPI posibilita filtrar a partir del sexo o género, la edad, el nivel de competencia o dominio de la lengua, otras segundas lenguas conocidas y el tipo textual. El LANGSNAP ofrece los datos de Alcalá de Henares.

por tarea, por el momento de recogida de la muestra (previa a la estancia en el extranjero, durante la estancia o posteriormente a ella) y por informante; aunque se da información del sexo, la edad, las otras lenguas extranjeras que conoce el aprendiz, los años que lleva estudiando ELE y la condición que haya tenido durante la estadia (estudiante de movilidad, lector, profesor de apoyo), tal cosa se ofrece mediante la descripción que se hace de cada informante, pero no es posible realizar búsquedas automáticas a partir de estas variables sociolingüísticas.

En cuanto a la anotación lingüística, el SPLLOC, el CAES y el CORESPI han sido anotados morfosintácticamente, razón por la cual no solo se pueden buscar palabras, sino también categorías gramaticales y/o estructuras lingüísticas. El CORELE y el CORANE han marcado los tipos de errores, por lo que estos se pueden localizar en el corpus.

4. LOS PROGRAMAS INFORMÁTICOS PARA ANOTAR CORPUS DE APRENDICES DE ELE

En la actualidad, existen diversas aplicaciones informáticas tanto de software libre como comerciales que ayudan en la anotación de corpus lingüísticos. Veamos algunas de ellas y en qué tipo de corpus han sido usadas. La *UAM Corpus Tool* (de 2012, en su versión 2.8) (<http://www.corpustool.com/index.html>) permite diseñar en árbol las etiquetas específicas que vaya a necesitar el investigador o docente, enlazarlas a los distintos fragmentos que se quieren etiquetar y almacenar el texto etiquetado en XML. Una vez etiquetados los textos, es posible realizar búsquedas aplicando distintos filtros y obtener estadísticas comparativas entre los diferentes subgrupos resultantes de los filtros (frecuencias absolutas y relativas, prueba t de *student* y *chi*²): por ejemplo, saber en qué porcentaje se da el mismo error en cada nivel de dominio de la lengua extranjera. Es una herramienta fácil de utilizar, amigable, de software libre (Figura 1), pero tiene la limitación de que no puede trabajarse de manera holística los errores, sino un tipo concreto de error cada vez (por ejemplo, la confusión entre SER y ESTAR). Ha sido utilizado en estudios de C. Lozano y A. Mendikoetxea⁶ a partir del CEDEL2 y con el corpus CAELE (Ferreira y Elejalde, 2017).

⁶Para una relación de las investigaciones de estos autores basadas en la explotación del CEDEL2 y en el uso de *UAM Corpus Tool*, consúltese <https://www.uam.es/proyectosinv/woslac/publications.htm>.

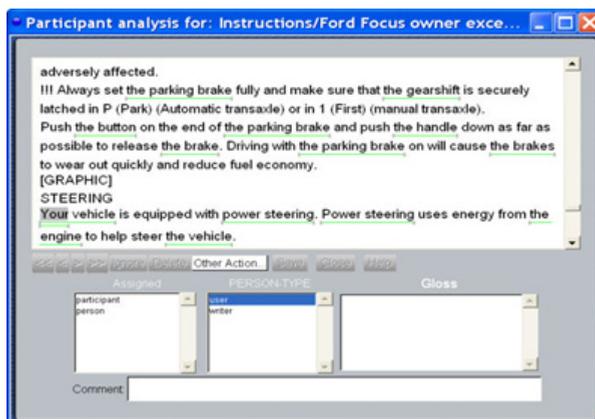


Figura 1. Vista del anotador *UAM Corpus Tool*

El programa *Knowtator* (<http://knowtator.sourceforge.net/>) (Figura 2) es una herramienta de anotación de texto de software libre desarrollada en la Universidad de Colorado que utiliza el *backend* basado en archivos Protégé (CLIPS) para guardar la información. Proporciona un mecanismo muy flexible para definir esquemas de anotación, incluida la anotación semántica. Ha sido utilizada por Alonso (2013) aprovechando un conjunto de textos del CEDEL2.

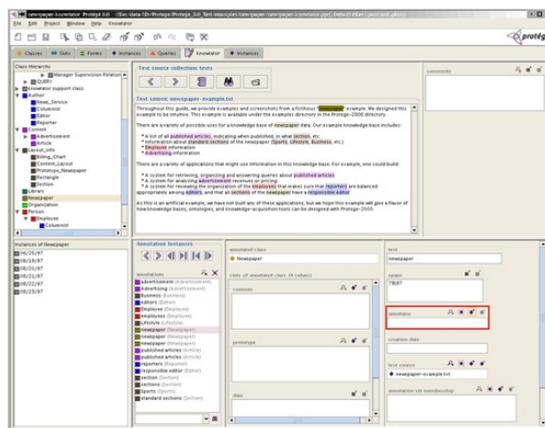


Figura 2. Vista del anotador *Knowtator*

El etiquetador *Markin* (<https://www.cict.co.uk/markin/>), también de software libre, facilita la anotación y corrección de textos electrónicos. Puede

importar el texto de un alumno para marcarlo pegándolo desde el portapapeles o directamente desde un archivo de documento. Una vez que se ha importado el texto, *Markin* proporciona un conjunto completo de herramientas que permiten al profesorado marcar y anotar el texto (Figura 3), que queda guardado como un documento XHTML o RTE, que es devuelto al alumno.

Markin tiene un uso muy pedagógico porque, cuando el aprendiz accede al texto anotado, ve claramente los errores, que quedan resaltados en rojo y, situando el cursor en las marcas, puede obtener más detalles sobre la anotación o el comentario del profesor. Esta herramienta ofrece estadísticas del error. Ha sido usado con el corpus *Aprender a Escribir en Lovaina* (Aprescrllov) (Buyse y González, 2013).

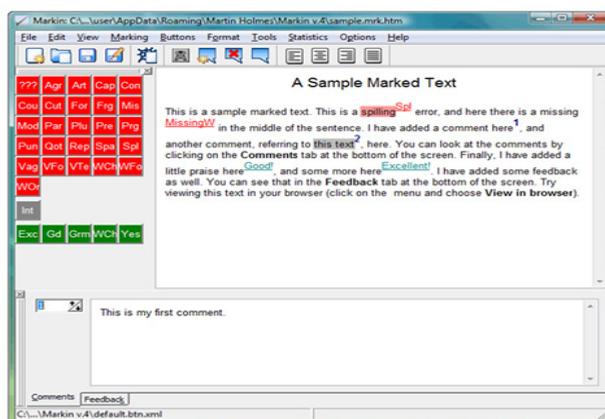


Figura 3. Vista del anotador *Markin*

La interfaz CATMA (<https://catma.de/>), también de código abierto, ha sido desarrollada por la Universidad de Hamburgo (Figura 4). Marca y analiza textos en web con etiquetas fácilmente definibles. Permite búsquedas interactivas en lenguaje natural, ofrece funciones analíticas estadísticas y no estadísticas automatizadas. Tiene un uso intuitivo y puede utilizarse incluso en idiomas con sistemas de escritura propios (árabe, hebreo, chino). Se utiliza en el CORESPI (Bailini, 2016).

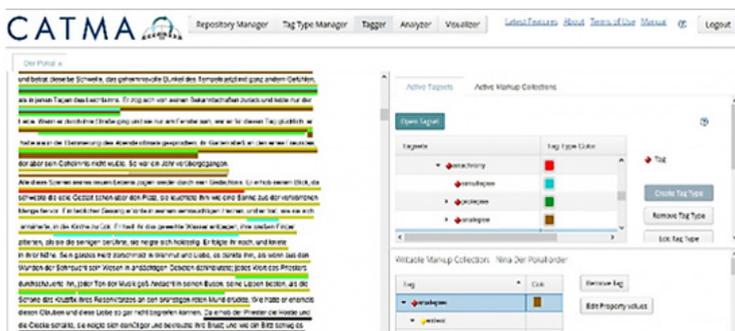


Figura 4. Vista del anotador CATMA

Por último, el programa *Nvivo* (Figura 5) ha sido diseñado por Qualitative Software Research (QSR) para investigaciones basadas en datos cualitativos que procesa de forma cuantitativa y automática (<http://www.qsrinternational.com/nvivo-spanish.>).

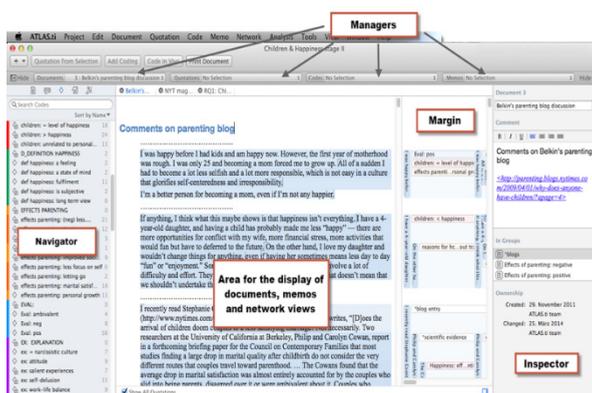


Figura 5. Vista del anotador *Nvivo*

Ayuda en la compilación de textos orales y/o escritos, en su organización y en la creación de la taxonomía de etiquetas para identificar fenómenos. Permite visualizar el contexto de etiquetado, analizar ocurrencias, cruzar variables, aplicar análisis estadísticos. Contiene una opción, 'contexto de etiquetamiento', que muestra el error inserto en el texto. Este anotador lo usan Ferreira, Elejalde y Vine (2014) en un corpus preliminar del CAELE (Ferreira y Elejalde, 2017: 518).

5. EL ETIQUETADOR *TEXTANNOT*

Codificar textos, pero sobre todo etiquetarlos, exige un trabajo colaborativo entre lingüistas e informáticos. El lingüista establece para qué ha de servir el corpus, qué información se quiere obtener, cuáles han de ser los filtros, por tanto, debe tomar decisiones que influirán en la explotación futura que tenga el corpus y en el propio resultado de las investigaciones que se deriven de dicho corpus.

Como hemos visto, la lingüística computacional ha dado ya varios programas informáticos que permiten al investigador etiquetar textos. Existen, incluso, anotadores automáticos, como el *FreeLing*, que, sin embargo, no evitan que haya que aplicar una desambiguación manual. Los etiquetadores existentes están pensados para trabajar un conjunto limitado de textos o una serie limitada de fenómenos. Y los etiquetadores automáticos, basados en la gramática normativa, no acaban de ser adecuados para identificar los errores propios de la interlengua de los aprendices de ELE.

Por estas razones, en el proyecto *Elaboración y catalogación de un corpus de textos escritos en ELE producidos por estudiantes sinohablantes* financiado por el Ministerio de Economía y Competitividad —hoy Ministerio de Ciencia, Innovación y Universidades— (Nº de Ref. FFI2016-80280-R) y desarrollado en la Universidad de Lleida con un equipo de investigadores de esta universidad, de la Universidad de Valencia, de la Universidad de Salamanca, de la Universidad de Estudios Extranjeros de Tianjin, de la Universidad de Changzhou y de la Universidad de Soochow, ha sido necesario diseñar una herramienta específica para poder catalogar y etiquetar de manera semiautomática las muestras escritas y los errores cometidos por aprendices chinos de ELE, el *TextAnnot* (Figura 6). Se trata de una aplicación Web desarrollada con Spring Boot, que es un *framework* o entorno de trabajo proporcionado por API REST (un tipo de interfaz de programación de aplicaciones), que funciona como una arquitectura de *software* que incluye —además de un lenguaje de programación— bibliotecas, programas y otros elementos que sirven para construir y entrelazar los diferentes componentes del proyecto, en este caso, el corpus y el anotador. El código fuente ha sido creado a partir de GitHub, una web que almacena y ayuda a gestionar códigos fuente públicos.

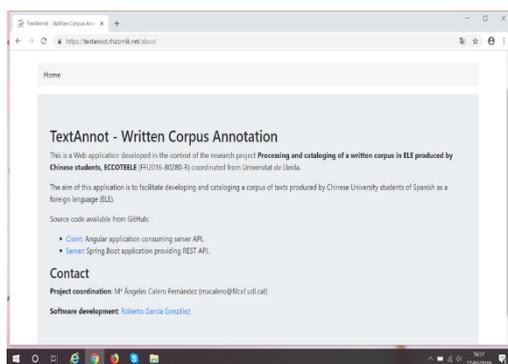


Figura 6. Página de inicio de *TextAnnot*

TextAnnot permite la personalización tanto de los metadatos utilizados para la catalogación de las muestras, como de la jerarquía de etiquetas utilizada para la anotación lingüística. El proyecto es de código abierto y acceso libre (<https://github.com/rhizomik/TextAnnot>), tanto para su uso como para colaborar en su desarrollo y mantenimiento. No requiere ningún tipo de instalación: puede accederse a él desde cualquier navegador. Ha sido diseñado por dos miembros del equipo de investigación, que son especialistas en comunicación hombre-máquina. Esta aplicación está en estos momentos en prueba y comenzará a utilizarse en breve para la catalogación, el etiquetado y la consulta del corpus que se está construyendo en el marco del proyecto FFI2016-80280-R, el *Corpus de Interlengua Española de Aprendices Sinohablantes* (CINEAS). En fases posteriores, *TextAnnot* quedará a disposición de los investigadores que necesiten etiquetar los errores de corpus escritos (y orales, si se acompañan de transcripciones) de aprendices, sean de lengua extranjera o de lengua materna.

El CINEAS (<https://cineas.udl.cat/>) es un corpus longitudinal constituido por textos redactados por alumnos chinos de Filología Hispánica que cursan sus estudios en las Universidades que participan en el proyecto y que se sitúan entre los niveles A1 y C1, incluidos. Los textos son descriptivos, narrativos, expositivos y argumentativos. Han sido escritos a mano en el aula y han sido recogidos en distintos momentos de cada año académico, desde 2016-2017 hasta 2019-2020. Los metadatos que se recogen son el sexo/género, el curso, la fecha, la universidad en la que se recoge la muestra, el nivel de inglés acreditado, si ha visitado un país hispanohablante y durante cuánto tiempo, así como el género discursivo (Calero, Terrado y Gómez-Devís en prensa). En estos momentos alcanza casi 4000 textos, que están siendo digitalizados en

HTML, catalogados y preparados para su etiquetado en el anotador.

Lo que permite *TextAnnot* es cargar textos individuales o en paquetes, registrar metadatos que faciliten su gestión y procesamiento, señalar en cada texto el fragmento que contiene un error, anotar el tipo de error del que se trata seleccionando del abanico que se ofrece en unos desplegables en forma de árbol —por tanto, en jerarquía—, marcar cada palabra o secuencia incorrecta tantas veces cuantos errores distintos haya en ella, ampliar y modificar la tipología de errores prefijados, revisar las anotaciones, establecer filtros, hacer búsquedas y contabilizar casos (en frecuencias absolutas y relativas). Para establecer un listado preliminar de etiquetas con las que se va a poner a prueba la anotación de errores, se han tenido en cuenta los sistemas de etiquetas utilizados en diferentes corpus, así como los establecidos en estudios sobre el error en aprendices de lenguas extranjeras, algo que no podemos tratar aquí por falta de espacio.

6. CONCLUSIONES

Tras esta revisión de los corpus de aprendices de ELE, puede concluirse que la destreza que mejor puede caracterizarse es la expresión escrita, al menos así lo entendemos dado que la mayoría de los corpus son escritos. Los metadatos que se utilizan y que, en algunos casos, sirven para filtrar la información susceptible de extraerse del corpus, son especialmente el sexo o género, la edad, la lengua materna (cuando son corpus multilingües), el nivel de dominio del español o el conocimiento de otra(s) lengua(s) extranjera(s). Este tipo de metadatos hace posible analizar los elementos lingüísticos distinguiendo a los informantes por rasgos sociales, lo que facilitaría la detección de eventuales influencias de estos factores en el aprendizaje de ELE, como sería previsible si atendemos a los resultados de los estudios sociolingüísticos tanto de comunidades monolingües como bilingües. Por otro lado, la anotación morfosintáctica que han recibido el SPLLOC, el CAES y el CORESPI y el etiquetado de tipos de errores del CORANE y el CORELE, pueden ayudar a identificar fenómenos de la interlengua.

El CEDEL2, el CAELE, el CORESPI y el Aprescrilov han utilizado programas de software libre distintos, más o menos intuitivos y flexibles, que sirven para diferentes tipos de etiquetado (de categorías gramaticales, de errores, semántico...) y que proporcionan cálculos estadísticos de aquello que se busca. El *TextAnnot* está pensado para el análisis de errores, y aporta una

tipología de errores prefijada (no hay que crearla) que puede ser aprovechada en su diseño original o puede ser adaptada a las necesidades de investigación o docentes y a las características de las muestras que forman el corpus en el que se aplique. Se está empleando por primera vez en el CINEAS, pero podrá utilizarse en cualquier corpus de aprendices, tanto de lengua extranjera como de lengua materna. Un anotador de estas características posibilitaría la comparación de datos entre diferentes corpus y aumentaría el impacto de los resultados obtenidos en la planificación didáctica de ELE, en general o destinada a colectivos específicos de aprendices.

REFERENCIAS BIBLIOGRÁFICAS

- AINCIBURU, M^a. C. 2012. "Recursos para estudiar el español a partir de corpus", *Revista Nebrija de Lingüística Aplicada a la Enseñanza de Lenguas*, 12: 209-215.
- ALONSO RAMOS, M. 2013. "Colocaciones, diccionarios y corpus de aprendices", *Eugenio Coseriu, in mermoriam, XIV Jornadas de Lingüística*, M. Casas Gómez (dir) y R. Vela Sánchez (ed.). Cádiz: Universidad de Cádiz. 57-71.
- ALONSO RAMOS, M. 2016. "Spanish learners corpus research. Achievements and challenges", *Spanish Learner Corpus Research. Current trends and future perspectives*, M. Alonso-Ramos (ed.). Amsteerdam/Philadelphia: John Benjamins. 3-32.
- BAILINI, S. 2016. *La interlengua de lenguas afines. El español de los italianos, el italiano de los españoles*, Milano: Edizioni Universitarie di Lettere Economia Diritto.
- BARALO OTTONELLO, M. 2010. "La investigación en español como lengua segunda: necesidad de un corpus de español no nativo". *V Congreso Internacional de la Lengua Española (CILE)*, Valparaíso (Chile). Recuperado a partir de: http://congresosdelalengua.es/valparaiso/ponencias/lengua_educacion/baralo_marta.htm
- BUYSE, K. 2011. "¿Qué textos en línea utilizar para qué fines en el aula de ELE?", *Del texto a la lengua: La aplicación de los textos a la enseñanza-aprendizaje del español L2-LE*, vol I, J. de Santiago Guervós, H. Bongaerts, J. J. Sánchez Iglesias, M. Seseña (coords). Madrid: Arco-Libros. 277-288.
- BUYSE, K. y GONZÁLEZ MELÓN, E. 2013. "El corpus de aprendices Aprescrliv y su utilidad para la didáctica de ELE en la Bélgica multilingüe", *Plurilingüismo y enseñanza de ELE en contextos multiculturales. XXIII Congreso Internacional ASELE*, en B. Blecua, S. Borrell, B. Crous y F. Sierra (eds.). Girona: ASELE y Universidad de Girona. 247-261.
- BUYSE, K., FERNÁNDEZ PEREDA, L., VERVECKKEN y K. 2016. "The Aprescrliv corpus, or broadening the horizon of Spanish language learning in Flanders", *Spanish Learner Corpus Research. Current trends and future perspectives*, M. Alonso-Ramos (ed.). Amsteerdam/Philadelphia: John Benjamins. 143-168.
- CALERO FERNÁNDEZ, M^a. Á., TERRADO PABLO, F. J. y GÓMEZ-DEVÍS, M-B.

- en prensa. “Análisis de errores y variación: la importancia de los corpus lingüísticos”, *Variación lingüística y su didáctica. Actas del Congreso Internacional sobre Variación Lingüística en Español: Cómo trabajarla en el aula* (Salamanca, 28 y 29 de junio de 2017). Universidad de Salamanca.
- CESTERO MANCERA, A. M^a., PENADÉS MARTÍNEZ, I. 2009. *Corpus de textos escritos para el análisis de errores de aprendices de E/LE (CORANE)*. Alcalá de Henares: Universidad de Alcalá de Henares. CD-ROM.
- EAGLES. 1996. *Text Corpora Working Group Reading Guide*. Eagles (Expert Advisory Group on Language Engineering). Recuperado de <http://www.ilc.cnr.it/EAGLES/corpintr/corpintr.html>
- FERREIRA CABRERA, A., ELEJALDE GÓMEZ, J., VINE JARA y A. 2014. “Análisis de Errores Asistido por Computador basado en un Corpus de Aprendices de Español como Lengua Extranjera”, *Revista Signos. Estudios de Lingüística*, 47(86): 385-411.
- FERREIRA CABRERA, A. y ELEJALDE GÓMEZ, J. 2017. “Análisis de errores recurrentes en el Corpus de Aprendices de Español como Lengua Extranjera, CAELE”, *Revista Brasileira de Lingüística Aplicada*, 17: 509-538.
- GONZÁLEZ ROYO, C., CHIAPELLO, S., MARTÍN SÁNCHEZ, T., MURA, G. Á., PASCUAL ESCAGEDO, C., PUIGDEVALL BAFALUY, N. y REGAGLILO, A. 2016. “CORINÉI (Corpus Oral de Interlengua Bilingüe Español-Italiano): Elaboración, análisis y aplicación a la enseñanza/aprendizaje de la interacción con las TICs”, *Innovaciones metodológicas en docencia universitaria: resultados de investigación*, por J. D. Álvarez Teruel, S. Grau Company, M^a. T. Tortosa Ybáñez (coords.). Universidad de Alicante. 1951-1958.
- LEECH, G. 1993. “Corpus Annotation Schemes”, *Literary and Linguistic Computing*, 8(4): 275-281.
- PENADÉS MARTÍNEZ, I. 2005. “Corpus para el análisis de errores en el aprendizaje de ELE. Presentación”, *Linred*, 3: 1-5.
- PÉREZ ÁVILA, E. 2007. “El corpus lingüístico en la didáctica del léxico del español como LE”, *Boletín de la Asociación para la Enseñanza del Español como Lengua Extranjera*, 37: 11-27.
- ROJO, G. 2010. “Sobre codificación y explotación de corpus textuales: otra comparación del Corpus del Español con el CORDE y el CREA”,

Lingüística, 24: 11-50. Ardila A. 2011. "Neuropsicología del lenguaje", *Manual de Neuropsicología*, J. Tirapu Ustárroz, M. Rios Lago, F. Maestú Unturbe (eds.). Barcelona: Viguera. 99-121.

DISEÑO Y CREACIÓN DE UN CORPUS DE APRENDICES DE ESPAÑOL EN JAPÓN (CELEN)

Pilar Valverde

Universidad Kansai Gaidai, Japón

RESUMEN

Presentamos CELEN, el primer corpus de aprendices dedicado en exclusiva al análisis de la expresión escrita de los estudiantes japoneses de español. Se trata de un corpus con vocación pedagógica, concebido como una suma de varios subcorpus locales, que son una muestra representativa de los textos que los estudiantes escriben durante un año académico en las clases de ELE de una universidad determinada. En este artículo se describe el primer subcorpus, correspondiente a la Universidad Kansai Gaidai, formado por 1.840 textos de 459 aprendices, unas 140.000 palabras. Se trata de un corpus abierto que está siendo ampliado con textos de otras instituciones. Está disponible para descarga bajo una licencia CC-NY-NC y puede ser consultado en línea en la plataforma Sketch Engine.

Palabras clave: corpus de aprendices, español en Japón, aprendices japoneses de español.

ABSTRACT

We present CELEN, the first learner corpus geared to the analysis of the written expression of Japanese students of Spanish. It is a corpus with a pedagogical aim, conceived as a sum of several local subcorpora, which constitute a representative sample of the texts that students write during one academic year in the Spanish classes of a given university. This article describes the Kansai Gaidai University subcorpus, made up of 1,840 texts from 459 learners, about 140,000 words. CELEN is an open corpus and is being extended with texts coming from others institutions. It is available for download under a CC-NY-NC license and can be consulted online in the Sketch Engine platform.

Keywords: learner corpora, Spanish in Japan, Japanese learners of Spanish

1. EL ESPAÑOL COMO LENGUA EXTRANJERA EN JAPÓN

En Japón el español es una lengua de estudio minoritaria, con una popularidad semejante a la de otras lenguas europeas como el francés o el alemán, y muy por detrás del coreano, el chino y la lengua extranjera por excelencia, el inglés. El Instituto Cervantes en su último informe (2018) calcula que hay unos 60.000 estudiantes de español en el país y, en la clasificación de países por número de estudiantes, Japón ocupa el puesto número 20 —en el número 1 se encuentra Estados Unidos (más de ocho millones de estudiantes) y en el 2, Brasil (más de seis millones)—.

En cuanto a los contextos de instrucción, en la enseñanza reglada predomina el nivel universitario, seguido a mucha distancia por la secundaria y bachillerato¹, mientras que la presencia del español en la escuela primaria es prácticamente nula. En la enseñanza no reglada encontramos las academias de idiomas², entre las que podemos incluir el Instituto Cervantes de Tokio³, los centros culturales y los que estudian español de forma independiente (véase Sanz *et al.*, 2015 y Ugarte, 2012).

Dado el reducido número de estudiantes y su concentración en ciertos contextos de enseñanza, por el momento el corpus que presentamos aquí recoge exclusivamente textos del ámbito universitario, donde suelen distinguirse dos modalidades de estudio: el español como asignatura de segunda lengua extranjera y el español como carrera.

En el contexto universitario, los estudiantes de carreras muy diversas tienen que estudiar durante un año académico (30 semanas) una segunda lengua extranjera, esto es, una lengua extranjera distinta del inglés, y el español es una de las opciones ofrecidas⁴. Normalmente se imparten 2 clases (de una

¹Según los datos del Ministerio de Educación, Cultura, Deporte, Ciencia y Tecnología de Japón (MEXT), de 2003 a 2009, el número de centros de bachillerato que impartían otras lenguas extranjeras además del inglés pasó de 454 a 2.027, y el español se ofrecía solo en un centenar de ellos. Asimismo, el número de estudiantes de bachillerato que estudiaban español se triplicó de 2003 a 2009, siendo 2.763 en 2009 (Ugarte, 2012).

²Según los datos del Ministerio de Economía, Comercio e Industria de Japón, en 2005 había 1.144 academias y el 12,5% (esto es, unas 150) ofrecía clases de español (Ugarte, 2012).

³En 2016 tenía 914 alumnos en sus cursos (Instituto Cervantes, 2018), una cifra semejante a la de los centros de Austria, Bélgica, Bulgaria, Túnez o República Checa.

⁴En 2009 el español ocupaba el sexto lugar en la lista de idiomas que ofrecen las universidades, tras el inglés, el chino, el francés, el alemán y el coreano.

hora y media de duración cada una) por semana, una a cargo de un profesor japonés y otra a cargo de un profesor hispanohablante. Se trata por lo tanto de un curso de corta duración (90 horas) en el que generalmente se prioriza el conocimiento teórico, razón por la cual al terminar el curso, la mayoría de estudiantes tiene un nivel básico y no superaría un examen de nivel A1.

En cuanto al contexto universitario especializado, el español puede estudiarse de forma intensiva, como parte fundamental de la licenciatura, en un departamento de español. Las clases de esta lengua (5 o más a la semana, esto es, 7,5 horas) suelen ser obligatorias durante los dos primeros años y optativas en los dos últimos, durante los cuales se priorizan los estudios de área (economía, política, relaciones internacionales, historia, literatura, etc.). Como resultado, al terminar la carrera los estudiantes alcanzan un nivel que puede ir desde A2, en el caso de los estudiantes de mediano rendimiento que se quedan en Japón, hasta B2, en el caso de los alumnos más aventajados que realizan estancias en el extranjero.

De las 780 universidades del país, más de 200 ofrecen el español como lengua extranjera optativa (Ugarte, 2012), mientras que solo una quincena ofrece la especialidad de español⁵. Esto significa que la gran mayoría de estudiantes universitarios de español se encuentran por debajo del nivel A1 y solo unos 4500 estudian “Filología Hispánica” (Sanz, 2013).

En el corpus que presentamos aquí se incluirán textos procedentes de ambos contextos, pero prestando especial atención al contexto especializado, con el objetivo de recoger textos correspondientes a varios niveles de dominio, al menos desde A1 hasta B2 del MCER.

En cuanto al perfil de los estudiantes, la gran mayoría no tiene conocimientos previos de español al empezar sus estudios universitarios, ni contacto con esta lengua fuera de clase, y apenas usará el español en su trabajo después de graduarse, pues en este país la especialidad cursada es prácticamente irrelevante para la vida laboral. Según los resultados de la encuesta realizada por el grupo GIDE (2012) en 40 universidades, la mayoría estudia español porque es una de las lenguas más habladas del mundo y porque le interesa la cultura del mundo hispano (la música, el deporte, el baile, etc.), y su principal objetivo es usar la lengua para comunicarse con la gente cuando viajen a países de habla española.

⁵En algunas universidades también existe la posibilidad de cursar cuatro o más clases de español semanales, pero no con la categoría de especialidad.

Como la mayoría de universidades del país —tres de cada cuatro— son privadas⁶, los métodos de enseñanza, la formación del profesorado y las directrices de cada una varían considerablemente. Mientras que para graduarse en algunas se requiere alcanzar un nivel suficiente para escribir un resumen de la tesina de graduación y superar una entrevista sobre la misma en español, en otras se exige solamente que los estudiantes hayan superado las asignaturas obligatorias y obtenido los créditos necesarios, independientemente de su nivel de dominio del español en el momento de graduarse. Para saber más sobre la situación de la enseñanza de ELE en el contexto universitario japonés y las propuestas de reforma, véase Sanz (2013), Sanz et al. (2015), Moreno (2015) y Escandón (2017).

2. LOS CORPUS DE APRENDICES DE ESPAÑOL

El desarrollo de corpus de textos escritos por aprendices de una lengua extranjera, denominados *corpus de aprendices*, empezó en los años ochenta del siglo XX. El primer corpus de aprendices a gran escala fue el *International Corpus of Learner English* (ICLE) (Granger et al., 2009), coordinado desde la Universidad de Lovaina, cuya primera versión fue publicada en CD-ROM en 2002 y que en la versión 2, publicada en 2009, contiene 4,5 millones de palabras procedentes de aprendices de inglés con once lenguas maternas distintas.

En el ámbito hispánico el desarrollo de corpus de aprendices se encuentra todavía en una fase inicial (Alonso-Ramos, 2016). Si acudimos al repositorio digital de recursos lingüísticos del proyecto europeo CLARIN⁷, en el apartado dedicado a corpus de aprendices⁸, encontramos 24 corpus monolingües de aprendices de nueve lenguas (inglés, finés, sueco, alemán, árabe, checo, francés, húngaro y noruego), pero ninguna de ellas es el español. Tenemos que llegar al apartado de “Corpus que se encuentran fuera de la infraestructura de CLARIN” para encontrar solamente un corpus de español disponible para consulta y descarga, el *Spanish Learner Language Oral Corpus* (SPLLOC)⁹, un

⁶ Según el MEXT en 2018 había 780 universidades: 604 privadas, 86 nacionales y 90 locales. En línea: http://www.mext.go.jp/b_menu/toukei/002/002b/1403130.htm.

⁷ *Common Language Resources and Technology Infrastructure* (CLARIN) es un proyecto europeo que tiene como objetivo desarrollar y reunir recursos tecnológicos en relación al uso y aplicación del lenguaje y de la cultura.

⁸ <https://www.clarin.eu/resource-families/L2-corpora>.

⁹ <http://www.splloc.soton.ac.uk/>.

corpus oral destinado a la investigación sobre la adquisición del español como lengua extranjera.

Existen otros listados en línea más amplios, como el *Learner corpora around the world*¹⁰, mantenido por los autores del ICLE de la Universidad de Lovaina. En dicha lista, en el momento de escribir este artículo, figuran 19 corpus de aprendices de español. Sin embargo, la mayoría de los que aparecen en ella no están disponibles para ser descargados o consultados online por otros investigadores, ya sea porque son de acceso restringido, están todavía en una fase de desarrollo o simplemente quedaron inacabados y nunca fueron publicados. Así pues, a la escasez de corpus se suma la dificultad para acceder a la mayoría de ellos, lo que limita su utilidad.

Si centramos nuestra atención exclusivamente en los corpus escritos, como el que aquí presentamos, y sin tener en cuenta corpus de reducido tamaño compilados para llevar a cabo investigaciones particulares¹¹, existen en este momento dos que pueden ser consultados libremente, el Corpus Escrito del Español L2 (CEDEL2)¹² y el Corpus de Aprendices de Español como Lengua Extranjera (CAES)¹³, y cinco mediante registro o compra¹⁴. Cuatro contienen datos de aprendices con una lengua materna determinada —el inglés (CEDEL2), el taiwanés (CATE), el italiano (CORESPI) y el portugués (COMET)— y el resto, lenguas maternas variadas.

El primer corpus de aprendices para el español fue CORANE (Cestero et al., 2001), que reúne 957 composiciones escritas por estudiantes de ELE de los cursos de español para extranjeros de la Universidad de Alcalá de los niveles A2, B1, B2 y C1, y de varias lenguas maternas. Los datos fueron obtenidos en el año 2000 y publicados en formato CD-ROM en 2009.

¹⁰<https://uclouvain.be/en/research-institutes/ilc/cecl/learner-corpora-around-the-world.html>.

¹¹Como los publicados en el portal de la revista Linred: <http://www.linred.es/corpus.html>.

¹²<http://cedel2.learnercorpora.com/>.

¹³<https://galvan.usc.es/caes>.

¹⁴Son los siguientes:

- El corpus *Aprender a Escribir en Lovaina* (Aprescilov).
- Corpus para el Análisis de errores de aprendices de E/LE (CORANE).
- El Corpus Multilíngue para Ensino e Tradução (COMET).
- El Corpus de Aprendices Taiwaneses de Español (CATE).
- El Corpus del Español de los Italianos (CORESPI).

El corpus de mayor tamaño es *Aprescrllov* (Buyse, 2012), otro de los pioneros, que cuenta con aproximadamente un millón de palabras recogidas desde 2004 en la Universidad de Lovaina.

Le sigue en tamaño CEDEL2 (Lozano y Mendikoetxea, 2013), que desde 2006 recoge textos escritos por hablantes de inglés L1 a través de un formulario electrónico, llegando actualmente a las 800.000 palabras. Este es además el único que permite la descarga de los datos.

El corpus CAES (Rojo y Palacios, 2016), promovido por el Instituto Cervantes, contiene muestras de hablantes de seis lenguas maternas y, a pesar de su reducido tamaño —575.000 palabras en total— fue concebido como un corpus abierto y en el futuro aspira a constituirse en un corpus representativo del español, al estilo del corpus ICLE para el inglés.

Entre todos los corpus publicados, solamente CORANE contiene algunos textos escritos por hablantes de japonés, por lo que podemos afirmar que CELEN, el corpus que presentamos aquí, es el primero dedicado en exclusiva a aprendices japoneses de español.

3. DISEÑO DEL CORPUS

El desarrollo de corpus de aprendices, a diferencia del desarrollo de corpus de nativos, se ve fuertemente condicionado por el campo del que proceden los investigadores implicados y el uso que se le pretende dar. Así, podemos diferenciar principalmente tres tipos de corpus según los destinatarios principales, que pueden ser:

1. Profesionales de la enseñanza como profesores, creadores de materiales didácticos, evaluadores, etc. Se trata del tipo más común, al que pertenecen los corpus CORANE, *Aprescrllov* y CAES.
2. Investigadores sobre adquisición de lenguas, como el corpus CEDEL2 o SPLLOC.
3. Investigadores en el campo del procesamiento del lenguaje natural o la lingüística computacional, como los corpus multilingües de Lang-8 (Mizumoto *et al.*, 2011) y WordReference (Berdicevskis, 2018).

CELEN nace con una vocación pedagógica, ya que su objetivo principal es facilitar la aplicación de la lingüística de corpus a los profesores de ELE que trabajan en Japón. No obstante, uno de los principios fundamentales que

seguimos al desarrollar el corpus es el de favorecer su reutilización por parte de investigadores con intereses diversos. Para conseguirlo se han tomado las siguientes medidas:

1. Cada texto incluye una gran cantidad de información sobre el aprendiz y sobre la situación en la que se escribió, para que cada investigador pueda seleccionar los textos que cumplen ciertos requisitos.
2. El corpus puede ser descargado y manipulado con fines de investigación bajo una licencia CC BY-NC 4.0 y consultado en línea en la plataforma Sketch Engine (véase el apartado 5).

En cuanto a su uso en el ámbito de la enseñanza, si bien es cierto que algunos aspectos de la enseñanza de ELE deben ser revisados y actualizados usando metodología de corpus (o metodología científica en general) (Llorián, 2017), también lo es que la investigación en este campo suele estar alejada de la realidad que se vive en el aula de lenguas extranjeras (Tono, 2016): mientras que los investigadores suelen estar más interesados en crear una descripción del aprendiz “arquetípico”, los profesores se muestran más interesados por lo que hacen realmente sus estudiantes particulares en las clases (Seidlhofer, 2002, Mukherjee y Rohrbach, 2006).

Para acortar esa distancia entre el ámbito de la investigación y la realidad del aula creemos que es importante disponer de subcorpus *locales*, esto es, representativos de una institución en particular, para que los profesores puedan investigar sobre las producciones de sus alumnos, en textos que son relevantes en su contexto de enseñanza, y llevar a cabo proyectos de investigación-acción.

Empezamos a recoger datos en la Universidad Kansai Gaidai, donde cerca de 300 estudiantes se matriculan cada año en la carrera de español. Se trata de un corpus abierto y tras la experiencia acumulada en la recogida de datos en esta universidad durante el año académico 2018, esperamos ampliarlo con la colaboración de otras universidades japonesas que ofrecen la licenciatura en español.

En cuanto los criterios lingüísticos, el corpus está formado exclusivamente por textos escritos, recogidos a lo largo de un año académico. Se trata, por lo tanto, de un corpus longitudinal, ya que registra la evolución de un grupo de aprendices a lo largo de un periodo de tiempo, a la manera de un Portafolio. Concretamente, en la Universidad Kansai Gaidai recogimos aproximadamente la mitad de los textos escritos por los estudiantes durante un año académico:

entre 1 y 8 textos por aprendiz (entre 3 y 6 textos del 78 % de aprendices).

En cuanto a las tareas escritas, estas son las que los estudiantes llevan a cabo en el marco de las asignaturas impartidas por profesores nativos como parte de exámenes parciales o finales, actividades de clase o tareas para escribir en casa. Por lo tanto, están basadas en un tema propuesto por el profesor, y la mayoría son guiadas en el sentido de que incluyen algunos puntos que los estudiantes deben incluir en su texto, al estilo de las tareas de expresión escrita de los exámenes DELE, o semi-guiadas, a partir de un tema propuesto por el profesor. Algunas han sido escritas con limitaciones de tiempo, mientras que otras han sido escritas en casa sin restricciones de tiempo. Asimismo, se indica si el aprendiz ha tenido la posibilidad de acceder a ayuda externa (diccionario, Internet, etc.) o no.

Sobre los aprendices, se trata como hemos comentado antes de jóvenes universitarios, ya que la enseñanza de esta lengua es prácticamente inexistente en etapas educativas anteriores y nuestro corpus tampoco contiene por el momento textos procedentes de academias de idiomas, centros culturales u otros contextos de enseñanza. Se incluyen solamente textos de aprendices cuya lengua materna es, al menos, el japonés. Se indica también el nivel de español e inglés según certificados oficiales, si tiene alguno.

El nivel de español de los textos según el MCER es determinado según el nivel de la asignatura en la cual se escribió el texto, aunque este nivel no se corresponda siempre con el nivel “real” de los textos en términos de competencia lingüística, o según el número de meses u horas de clase recibidas. De esta forma, durante el primer año se usa un libro de texto de nivel A1, y durante el segundo año, un libro de nivel A2, aunque el grado de dominio lingüístico de cada estudiante puede variar dependiendo de factores individuales como la motivación o las horas de estudio, entre otros. Los textos del corpus, por lo tanto, reflejan el nivel de dominio que los profesores de ELE encuentran en sus clases.

4. RECOGIDA DE DATOS

4.1 *La institución: solicitud al comité de ética*

Cada vez más universidades han actualizado sus políticas con respecto a la investigación que involucra el uso de sujetos humanos, y por ello fue necesario entregar una solicitud al Comité Universitario de Evaluación de la Investigación antes de comenzar la investigación, para garantizar la anonimidad de los participantes.

4.2 *Los aprendices: cuestionarios y formularios de consentimiento*

Asimismo, fue necesaria la colaboración y coordinación de los profesores del departamento de español, ya que se trata de una de las universidades con mayor número de estudiantes de especialidad, casi 300 por curso. Durante las primeras semanas del curso los profesores de las clases de Gramática entregaron un cuestionario y un formulario de consentimiento a los estudiantes de sus clases. En total, 13 profesores recogieron 508 formularios en 27 grupos de estudiantes.

Para el diseño del cuestionario tuvimos en cuenta los datos incluidos en otros corpus de aprendices de español y en corpus de referencia de otras lenguas, como el ICLE (Granger *et al.*, 2009) o MERLIN (Boyd *et al.*, 2014), para que el corpus pueda adaptarse a las necesidades de la mayor cantidad de investigadores. Mediante el cuestionario obtuvimos información sobre: edad, sexo, año académico, lengua(s) materna(s) del aprendiz, de la madre y del padre, lengua(s) hablada(s) en casa, contactos personales en países de habla española, estancias en países de habla española, edad de inicio en el estudio del español, títulos oficiales de español y de inglés y conocimientos de otras lenguas.

Aproximadamente el 85 % de los estudiantes completaron el cuestionario. El resto no lo hizo porque estuvieron ausentes ese día o prefirieron no participar. Los perfiles de los aprendices que dieron su consentimiento y que completaron los cuestionarios en su totalidad fueron introducidos en la base de datos. Después de descartar aquellos participantes cuya lengua materna no incluye el japonés o de los cuales no recibimos ningún texto durante el curso

académico, el número total de aprendices incluidos en el corpus es de 459.

Aproximadamente dos tercios son mujeres, como suele ocurrir en las facultades de lenguas extranjeras, y tienen entre 18 y 22 años. El 99 % son hablantes monolingües de japonés y el 1 % hablan japonés y otra lengua. Para el 98% de ellos la lengua materna de la madre, la del padre y la lengua hablada en casa es el japonés, y solo tres viven en una familia con algún miembro cuya lengua materna es el español. Aproximadamente el 80 % pertenece al programa regular, y el resto al programa intensivo (en el que a partir del segundo año reciben más clases de lengua). El 45 % son estudiantes de primer año, el 45 % de segundo, el 9 % de tercero y el 1 % de cuarto. El 90 % empezó a estudiar español a partir de los 18 años. Tres cuartas partes no tienen contactos personales (familia, amigos, etc.) en países de habla hispana y el 88 % no ha estado nunca en un país de habla hispana. Entre el 12 % que sí ha estado, la mayoría ha estado menos de un mes, y los países preferidos son España, Perú y México. En cuanto al nivel de español certificado, un 13 % tiene algún título de español: el más popular es el *Supeingo Kentei (Evaluación del conocimiento de la lengua española)*, organizado por la Sociedad Hispánica del Japón y que tiene seis niveles, desde el 6, el más básico, hasta el 1, el más avanzado. Los niveles más comunes entre los aprendices son el 4 (35 %), 5 (27 %) y 6 (28 %). En lo que respecta al nivel de inglés, el 68 % manifiesta tener uno o más títulos (EIKEN, TOEIC, TOEFL ITP, GTEC u otros), y en su equivalencia con el nivel del MCER lo más común es el nivel B1 (el 56 %), seguido por A2 (30 %) y A1 (9 %).

4.3 Los textos

Nueve profesores a cargo de las clases de Conversación de primer y segundo año del programa regular recopilaron dos textos por semestre (de los cuatro que escriben en total como parte de los exámenes parciales), en total cuatro durante el año académico, en 21 grupos de estudiantes. Cinco profesores de las clases del programa intensivo de segundo a cuarto año recopilaron varios textos en seis grupos, dependiendo de la disponibilidad del profesor. El 77 % de textos proceden de exámenes parciales, el 17 % de tareas, el 4 % actividades de clase y el 2 % de exámenes finales, por lo que el 78 % fue escrito sin acceso a ayuda externa como el libro de texto, el diccionario o Internet.

En total el corpus contiene 1.840 textos, unas 140.000 palabras, y cada texto está relacionado con un gran número de metadatos sobre las características del

aprendiz y del texto en cuestión: fecha, nivel de español de la clase (Apéndice, Figura 1), curso, semestre, meses de clase cursados hasta el momento de escribir el texto, macrofunción predominante (Apéndice, Figura 2), género textual (Apéndice, Figura 3), título, destinatario (Apéndice, Figura 4), autenticidad, noción específica predominante¹⁵ (Apéndice, Figura 5), libro de texto y unidad que se ha estudiado antes de escribir el texto, modo de asignación del tema, situación en la que se escribió el texto, límite de tiempo, acceso a ayuda externa, número de palabras requerido y medio de escritura.

La mayoría de textos fueron escritos a mano por los aprendices, y por lo tanto han sido transcritos con ayuda de OCR o manualmente, revisados, anonimizados e ingresados en la base de datos. Durante la transcripción se ha intentado reflejar lo más fielmente posible el original, incluyendo los errores. Cuando hay ambigüedad en el texto, como suele hacerse en otros corpus de aprendices, evitamos introducir errores de más y hacemos una asunción positiva, esto es, suponemos que el aprendiz no cometió un error. Solo hemos modificado el original para ocultar la información personal que pueda aparecer: nombres propios, números de teléfono, direcciones postales, direcciones de correo electrónico, etc.

5. CONCLUSIONES Y TRABAJO FUTURO

El corpus puede ser consultado en Sketch Engine¹⁶, una herramienta para consultar y crear corpus usada por numerosos profesionales de la lengua como lexicógrafos, traductores, profesores, etc. Para acceder a la plataforma, es necesario registrarse. Para tener acceso al corpus, hay que solicitarlo a la autora, indicando el correo electrónico, para poder darle de alta como usuario.

Con esta herramienta, es posible obtener varios tipos de resultados: concordancias (*Concordance*), listas de palabras (*Wordlist*), colocaciones y combinaciones de palabras (*Word Sketch*), sinónimos o palabras semejantes (*Thesaurus*), comparaciones entre las colocaciones de dos palabras (*Word Sketch Difference*) y expresiones multipalabra (*N-grams*).

En la Figura 1, pueden verse por ejemplo las líneas de concordancia resultado de la búsqueda de la colocación del verbo *ser* seguido por el adverbio *bien*, un error común entre los aprendices de español.

¹⁵Según la lista del *Plan Curricular* del Instituto Cervantes (2007).

¹⁶<https://www.sketchengine.eu>.

Details	Left context	KWIC	Right context
1 0143-H-KG-A22-375h-s	mica. Me gustaba paella. Siempre quería comer a la paella. Primera escuela mio	era muy bien. Pero segunda escuela mio era muy mal. Me gustaba mi profesor. Cua	
2 0244-M-KG-A21-270h-s	as tapas de queso, croquetas y pollo frito. todos con muy ricos. Interior del tienda	es muy bien. Los camareros son simpáticos y guapos. Pero menu de bebida poco, i	
3 0154-M-KG-A21-270h-s	acellunas, paella. Ha dicho que son muy muy bien. No sé otras comidas pero han	sido bien que nos comidos. Los camareros han sido muy simpáticos. Quiero ir al resta	
4 0400-H-KG-A11-60h-s	sta es Nara. Nara está izquienda de Osaka. Nara es antiguo pero algadoble. Nara	es muy bien. Nara hay muchos templos. Por ejemplo, Todai y Kiyomizu. Nara hay ui	
5 0221-M-KG-A22-375h-s	La comida que mi gustaba más es parte, merocotón y sopa de vegetales. Mi vida	es muy bien. Vivía en Nishinomiya. Mi casa era cerca de "Koushien". Mi casa es no	
6 0205-M-KG-A22-405h-s	n! Muchos besos, #NombreM. ¡Hola! Me llamo #NombreM. ¿Cómo estás? Japón	es muy bien. Por que hay muchas bien comidas. "Sushi, Sukiyaki, Mattucha". Me er	
7 0166-H-KG-A21-270h-s	e se ofrece comida francesa. He ido con mi amiga. Hemos sentido que el comida	es muy bien pero los camareros son muy antipáticas. Por ejemplo, he contado la cu	
8 0030-M-KG-B12-585h-s	al. Además el sueldo es un poco más alto que normal y el condición de mi trabajo	es muy bien. Son una de las razones porque me encanta mi trabajo. Creo que la mx	
9 0347-M-KG-A12-210h-s	locar el kimono y pasar en Kioto. Buenas tardes. #NombreM. ¡Hola! ¿Qué tal? Yo	soy muy bien. Quiero jugar con mis amigos en vacaciones de primavera. ¿Por qué ni	
10 0174-H-KG-A21-338h-s	n para mi. Salgamos para cenar. Cenamos en un restaurante de indio. La comida	es muy bien. Porque fue importante que mi cumpleaños de una vez al año. gracias.	
11 0186-M-KG-A22-375h-s	taba jugar en el parque y jugar con bol. Me gustaba Sushi que es arroz con pez. ¡	Es muy bien! Cuando era pequeña vivía en Wakayama con mi familia y dos perros.	
12 0217-M-KG-A21-270h-s	de está restaurante es muy tranquila y amable. La comida es muy bueno, el lugar	es bien. Está restaurante es mi favorito. He yendo a nuevo restaurante. Se llama "vi	
13 0228-M-KG-B12-585h-s	obre limpiador. Ojalá que encuentre buen trabajo. Un saludo, #NombreH Hola, yo	soy bien. ¿y? Muchas gracias Celebras sobre he encontrado un trabajo estupend	
14 0118-M-KG-A21-270h-s	te. Otro amiga ha pedido zumo de uva ensaladilla de marisco y paella. La comida	es bien. El lugar es poco lejos de estación de Umeda. Los camareros son amables :	
15 0107-M-KG-A22-305h-c	ague a la casa. Me perdí todas las esperanzas. Dormí todos un día. ¿Qué tal? Yo	soy bien. He oído que estás un poco cansado. ¿Qué te pasa? La resulta no es impor	
16 0336-M-KG-A11-60h-s	ico en Chatatown. Chacha town está norte de Kitakyushu. Gente de Kitakyushu	es muy bien. Ciudad Kobe está al norte de Awaji. Ciudad Kobe hay muchas casas, i	
17 0356-H-KG-A11-60h-s	iy un castelo. Se llama Osaka castelo. Hay un torre. Se llama Tsutenkaku. Osaka	es muy bien. Hay un parque. Se llama Universal Stadio Japan. Universal Stadio Jap	
18 0358-M-KG-A12-210h-s	las vacaciones de primavera. ¿Vamos a Tokio? Harajuku se come bien. Shibuya	es muy bien, porque yo quiero comprar ropa. Luego, ¿quieres ir al cine? La película	
19 0129-M-KG-A22-420h-s	nijos y cenar con mi familia. A veces comeré con mi hermana y su novio. Mi futuro	será muy bien. Cuando tengo 22 años, trabajaré en Tokyo y ganaré mucho dinero. Y, i	
20 0209-H-KG-A22-375h-s	migos por lo usado. A mi me gustaba el curry porque eso que mi madre cocinaba	era muy bien. Vivía en Higashiosaka. Higashiosaka es centro de Osaka. Higashiosal	

Figura 1. Líneas de concordancia del verbo *ser* seguido por el adverbio *bien*.

También es posible descargar el corpus completo y manipularlo directamente, bajo una licencia CC BY-NC 4.0. Véase la página web del proyecto: <https://sites.google.com/view/celen/>.

En el futuro, seguiremos ampliando el corpus con textos de otras instituciones, lo que permitiría, a partir de un muestreo de los datos, crear un corpus de referencia del español en Japón. También estudiamos la posibilidad de incluir textos procedentes de contextos de práctica informales como los blogs y las redes sociales para el aprendizaje de lenguas extranjeras y aumentar la precisión de la anotación morfosintáctica automática revisando los errores más frecuentes.

AGRADECIMIENTOS

Esta investigación ha sido posible gracias a todos los profesores de español de la Universidad Kansai Gaidai que colaboraron amable y desinteresadamente en la recogida de cuestionarios y textos y a los estudiantes que generosamente aceptaron participar cediendo sus datos.

La investigación ha sido financiada por *kakenhi* (17H07270), *Grant-in-Aid for Scientific Research* de la Japan Society for the Promotion of Science.

APÉNDICE

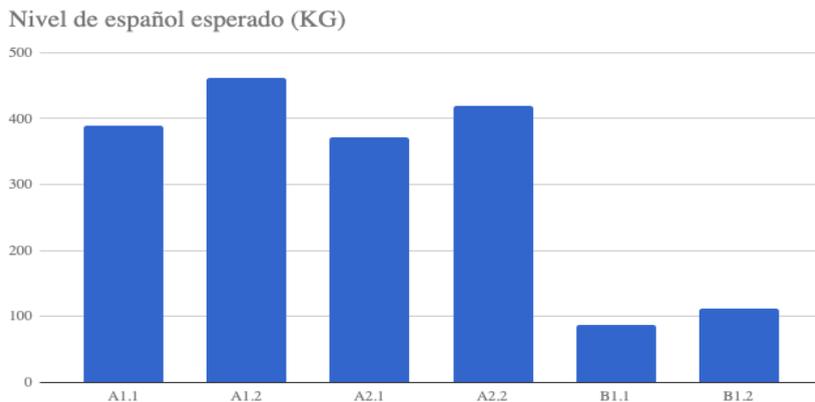


Figura 1. Número de textos según el nivel del MCER esperado

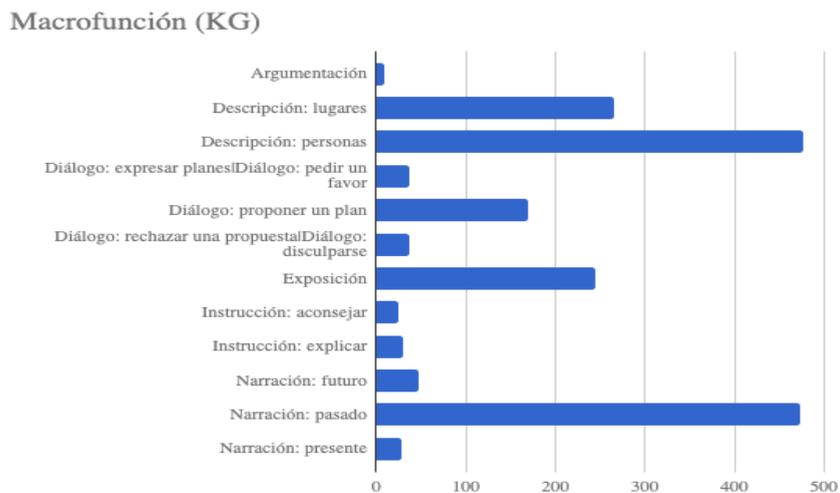


Figura 2. Número de textos según su macrofunción predominante

Género textual (KG)

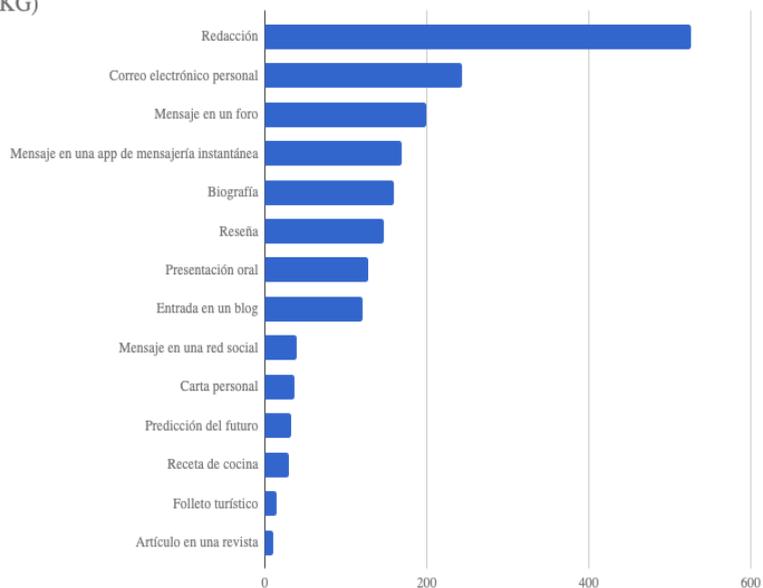


Figura 3. Número de textos según su género textual

Destinatario (KG)

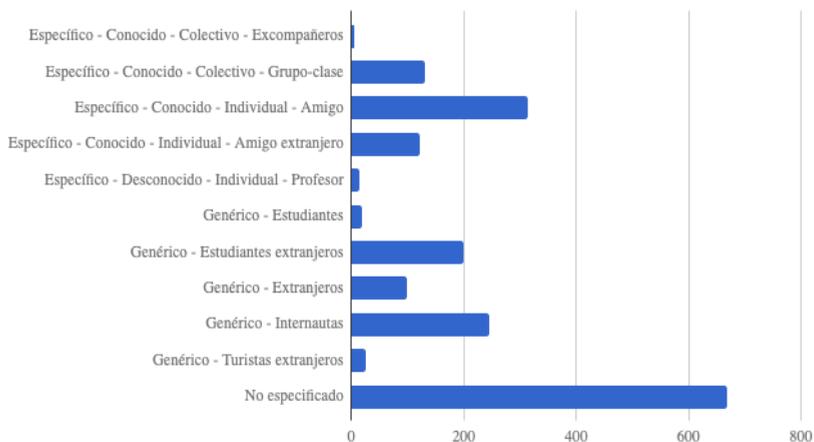


Figura 4. Número de textos según el tipo de destinatario

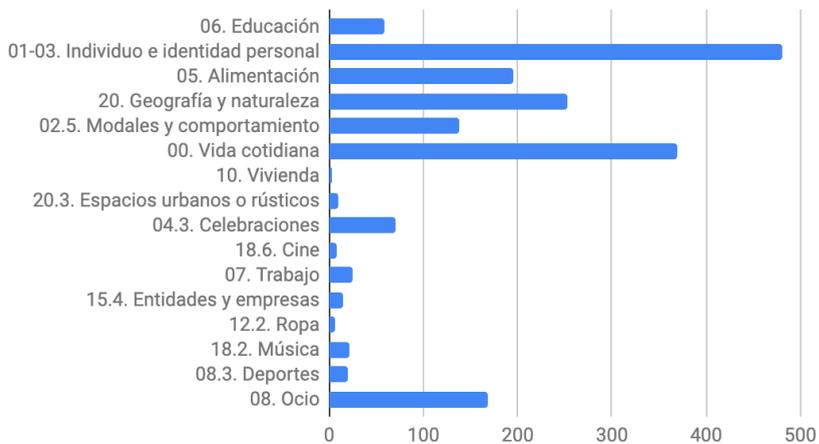


Figura 5. Número de textos según la noción específica predominante

REFERENCIAS BIBLIOGRÁFICAS

- ALONSO-RAMOS, M. 2016. Spanish Learner Corpus Research. Current trends and future perspectives. John Benjamins.
- BERDICEVSKIS, A. 2018. “Do non-native speakers create a pressure towards simplification? Corpus evidence”. En Cuskley, C., Flaherty, M., Little, H., McCrohon, L., Ravignani, A. & Verhoeft, T. (Eds.): *The Evolution of Language: Proceedings of the 12th International Conference (EVLANGXII)*. 41-43.
- BOYD A., HANA J., NICOLAS L., MEURERS D., WISNIEWSKI K., ABEL A., SCHONE K., STINLOVA B. y VETTORI C. 2014. “The MERLIN corpus: Learner language and the CEFR”, *Proceedings of LREC 2014*. Reykjavik, Iceland: European Language Resources Association.
- BUYSE K. 2012. “El corpus de aprendices Aprescrilov y su utilidad para la didáctica de ELE en la Bélgica multilingüe”, *Actas de ASELE 2012*.
- CESTERO MANCERA A.M., PENADÉS MARTÍNEZ I., BLANCO CANALES A., CAMARGO FERNÁNDEZ L. y SIMÓN GRANDA J.F. 2001. “Corpus para el análisis de errores de aprendices de E/LE (CORANE)”, *Actas de ASELE 2001*. 527– 534.
- ESCANDÓN A. 2017. “La enseñanza universitaria de ELE/EL2 en Japón, tiempo de reformas”, *Boletín de la Asociación para la Enseñanza del Español como Lengua Extranjera*, ASELE, 56. 25-29.
- GIDE. 2012. *Cuestionario sobre Análisis de Necesidades Aplicado a los Alumnos Universitarios Japoneses de Español. Informe*. GIDE, Grupo de Investigación de la Didáctica del Español. En línea: <http://gide.curhost.com/publicacionesCuest.html>.
- GIDE. 2015. *Un modelo de contenidos para un modelo de actuación. Enseñar español como segunda lengua extranjera en Japón*. GIDE, Grupo de investigación de la didáctica del español. En línea: <http://gide.curhost.com/archivos/201601Modelo.pdf>.
- GRANGER S., DAGNEAUX E., MEUNIER F., PAQUOT M. 2009. *International Corpus of Learner English v2* (Handbook + CD-Rom). Presses universitaires de Louvain, Louvain-la-Neuve.
- INSTITUTO CERVANTES. 2007. *Plan Curricular del Instituto Cervantes. Niveles de*

- referencia para el español*. Madrid: Biblioteca Nueva. En línea: https://cvc.cervantes.es/ENSEÑANZA/biblioteca_ele/plan_curricular/indice.htm
- INSTITUTO CERVANTES. 2018. *El español: una lengua viva 2018*. En línea: https://cvc.cervantes.es/lengua/espanol_lengua_viva/.
- LLORIÁN GONZÁLEZ S. 2017. “Claves de una revisión de los Niveles de Referencia para el Español, basada en metodología de corpus”. *Marcoele Revista de Didáctica de ELE*, 25.
- LOZANO C. y MENDIKOETXEA A. 2013. “Learner corpora and second language acquisition: the design and collection of CEDEL2”, *Automatic Treatment and Analysis of Learner Corpus Data*, N. Ballier, A. Díaz-Negrillo y & P. Thompson (eds.). Amsterdam: John Benjamins. 65–100.
- MIZUMOTO T., KOMACHI M., NAGATA M. y MATSUMOTO Y. 2011. “Mining Revision Log of Language Learning SNS for Automated Japanese Error Correction of Second Language Learners”, *Proceedings of the 5th International Joint Conference on Natural Language Processing (IJCNLP)*, Chiang Mai, Thailand.147-155.
- MORENO GARCÍA C. 2015. “Necesidades en la enseñanza del español en universidades japonesas. Reflexiones y sugerencias”, *Actas del II Congreso Internacional sobre el español y la cultura hispánica del Instituto Cervantes de Tokio*. Tokio: Instituto Cervantes de Tokio. 66-77.
- MUKHERJEE J. y ROHRBACH J-M. 2006. “Rethinking applied corpus linguistics from a language-pedagogical perspective: New departures in learner corpus research”, *Planning, Gluing and Painting Corpora: Inside the Applied Linguistics Workshop*, B. Kettemann y G. Marko (eds.). Frankfurt:Peter Lang. 205-232.
- PADRÓ L. y STANILOVSKY E. 2012. “FreeLing 3.0: Towards Wider Multilinguality”, *Proceedings of the Language Resources and Evaluation Conference (LREC 2012)* ELRA.
- ROJO G. y PALACIOS MARTÍNEZ I. 2016. Learner Spanish on computer, *Spanish Learner Corpus Research: Current trends and future perspectives*, M. Alonso (ed.), John Benjamins. 55-87.
- SANZ YAGÜE M. 2013. “Situación presente y futura del español en Japón. Elucubraciones de una profesora dirigidas a otros profesores en Asia (y a las autoridades)”, Jornadas organizadas por la editorial Edinumen y

la Universidad de Hong Kong, con la colaboración de la Consejería de Educación en China y el Instituto Cervantes de Pekín.

- SANZ YAGÜE M., ESCANDÓN GODOY A., ROMERO DÍAZ J., RAMÍREZ GÓMEZ D. y CIVIT I CONTRA R. 2015. *Enseñar español en Japón. Algunos aspectos de la enseñanza a japoneses*. Kobe City University of Foreign Studies Annals of Foreign Studies, 89. En línea: <http://id.nii.ac.jp/1085/00001825/>.
- SEIDLHOFER B. 2002. "Pedagogy and local learner corpora. Working with learning-driven data". *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*, S. Granger, J. Hung y S. Petch-Tyson (eds.). Philadelphia: John Benjamins. 213–234.
- TONO Y. 2003. "Learner corpora: Design, development and applications", *Proceedings of the 2003 Corpus Linguistics Conference* [ICREL Technical paper 16]. Lancaster University. 800-809.
- TONO Y. 2016. "What is missing in learner corpus design?", *Spanish Learner Corpus Research: Current trends and future perspectives*, M. Alonso (ed.). John Benjamins. 33-52.
- UGARTE V. 2012. "El español en Japón", *Anuario Cervantes 2012*. Instituto Cervantes. En línea: http://cvc.cervantes.es/lengua/anuario/anuario_12/ugarte/p01.htm.

¿ES POSIBLE LA DETECCIÓN AUTOMÁTICA DE LA VIOLENCIA LINGÜÍSTICA EN LAS REDES SOCIALES?

Susana Campillo Muñoz

Universitat Rovira i Virgili, Tarragona

RESUMEN

La violencia lingüística es un fenómeno presente en la actualidad. En las redes sociales ve modificada la forma en que se materializa y esto ha impulsado la creación de métodos computacionales para su análisis automatizado. Esta investigación se centra en el análisis de enunciados violentos en el contexto virtual. Observamos que en los desarrollos computacionales actuales no se limitan las categorías de violencia, se emplean diferentes clases según los autores y se consideran categorías discretas. Además, son escasos los trabajos que consideran la información extralingüística y los recursos propios del mundo virtual. Por ello, defendemos la inclusión de la información contextual, tanto sociocultural como de la red y del usuario, para desarrollar una categorización en términos difusos y en diferentes niveles lingüísticos. Creemos necesario el trabajo conjunto de informáticos y lingüistas.

Palabras clave: violencia lingüística, detección automática, redes sociales

ABSTRACT

Verbal violence is a phenomenon with a strong impact on contemporary society. In social media, however, the way it emerges differs, propelling thus the formulation of computational analysis methods to detect it and prevent it. This investigation focuses on violent statements analysis in a virtual context. The categories used in current computational detection systems are not limited, show differences across authors and are treated as discrete entities. In addition, there are few studies that include extralinguistic data and virtual communication features. Therefore, in this paper I advocate the inclusion of sociocultural, network and user contextual data, in order to develop a classification system which uses fuzzy categories and various linguistic levels. This calls for the collaboration of linguists and computer engineers.

Keywords: linguistic violence automatic detection, social media

1. INTRODUCCIÓN

La violencia verbal y, en concreto, en las redes sociales es un fenómeno presente en nuestra sociedad. Las redes sociales han conformado y forman un entorno virtual en el que las relaciones sociales se ven modificadas.

Con relación a la violencia verbal, cada día son más numerosas las noticias de denuncias a cantantes, humoristas y otros profesionales por enunciados emitidos en el contexto virtual. Estos hechos derivan en diversidad de opiniones de los hablantes sobre estos sucesos y sobre las condenas que se imponen. Asimismo, la variabilidad de consecuencias legales que conllevan los enunciados violentos en general, y aún más en internet, refleja la importancia del componente contextual y sociocultural en sus valoraciones.

El interés y la actualidad del fenómeno han impulsado a los investigadores a desarrollar herramientas para el análisis de estos enunciados, catalogándolos bajo etiquetas diferentes: *hate speech*, *ciberacoso*, *lenguaje racista*, *sexista u ofensivo*, entre otros. No obstante, nos seguimos encontrando con casos difíciles de analizar y de clasificar, como los enunciados implícitos.

A partir de lo dicho, nos preguntamos si es posible la detección automática de estos enunciados atendiendo a la variabilidad inherente del fenómeno.

Nuestra investigación parte de la pragmática y se centra en los enunciados violentos en un contexto determinado, en este caso, en el entorno virtual.

El objetivo principal es proponer una formalización que dé cuenta de la información relevante para el análisis de los mensajes violentos. Para ello, nos fijamos los siguientes subobjetivos:

- Definir la violencia verbal y establecer una clasificación en términos difusos de las distintas categorías de violencia verbal.
- Delimitar los patrones recurrentes en los enunciados violentos en las redes sociales.
- Concretar la información lingüística y sociocultural que influye en la comunicación de la violencia verbal.
- Establecer el modo más adecuado de introducir la información lingüística y sociocultural en su análisis.

Así pues, por un lado, estudiaremos tanto la bibliografía existente sobre violencia verbal como el conocimiento metapragmático de los hablantes; por otro lado, analizaremos los enunciados violentos de un corpus extraído de las

redes sociales, y las investigaciones y desarrollos computacionales referentes a dicho fenómeno.

2. MARCO TEÓRICO

El fundamento teórico de nuestro estudio parte de cuatro teorías que se toman como base y de diferentes estudios en relación a la violencia verbal. Así, nuestra investigación parte, en primer lugar, de la teoría de los actos de habla (Austin, 1955), puesto que afirmamos que al hablar se realiza un tipo particular de acción; de la teoría de la cortesía (Brown y Levinson, 1987) y de su contraparte, la teoría de la descortesía, (Culpeper, 1996, 2010; Culpeper, Haugh, Sinkeviciute), ya que la violencia verbal combina la parte lingüística y la social; y de la teoría de la relevancia (Sperber y Wilson, 1994), que nos permite calcular el grado de implícito y explícito atendiendo a la diferente información contextual.

Partimos de las siguientes premisas teóricas en relación con la violencia verbal. El fundamento de la violencia es cultural y, por tanto, el grado de variabilidad determina la diversidad de marcos contextuales desde los que se evalúa. Además, la violencia verbal se puede realizar de manera explícita o implícita [el 34% de los tweets irónicos son ataques, ridiculizan o dañan a alguien (Van Hee et al., 2018)]. Por lo que se refiere a la evaluación y valoración de la violencia verbal, esta es contextual, es decir, la información situacional y del contexto comunicativo es imprescindible para analizarla.

Por otro lado, las redes sociales suponen un cambio de paradigma en los estudios de comunicación (Herring, 2004). Tanto la información disponible para el usuario, como el contexto de comprensión y los recursos comunicativos son distintos a los de la comunicación cara a cara; han surgido, pues, nuevos géneros textuales con características formales diferentes que hay que considerar para un análisis automático eficaz.

Resumiendo, las redes sociales son un espacio en el que la violencia lingüística tiene lugar con características y significado social modificados respecto a su realización en el mundo físico (Miró, 2016). La lingüística computacional ofrece herramientas y perspectivas desde las que comprobar las diversas teorías sobre la comunicación mediada por ordenador. El procesamiento del lenguaje natural para las redes sociales requiere adaptar los métodos tradicionales o crear nuevos (Farzindar, 2018).

Parece claro que, para la detección automática de la violencia verbal en las redes sociales, es necesario un cambio de perspectiva en referencia al paradigma tradicional del estudio tanto de la comunicación como de la violencia verbal.

3. PROBLEMAS EN EL ANÁLISIS DE LA VIOLENCIA VERBAL

Tras realizar un estudio de la bibliografía existente y de las investigaciones actuales sobre el tema, hemos detectado un conjunto de problemas, tanto teóricos como de la aplicación práctica en la detección de la violencia verbal. Clasificamos estos problemas en dos bloques: problemas teóricos y problemas computacionales, aunque tienen una clara relación y se retroalimentan.

3.1 Problemas teóricos

Entre las investigaciones actuales encontramos varios trabajos de detección de mensajes racistas, *hate speech*, ciberacoso y lenguaje ofensivo. Así, observamos distintas materializaciones del fenómeno, es decir, varios subtipos que se engloban bajo un término genérico como puede ser *descortesía* o *agresión verbal* (Culpeper, 2017), y encontramos también una escasa delimitación entre estas clases. No se distinguen las etiquetas empleadas en las categorías y, de hecho, se emplean clasificaciones distintas según el ámbito de estudio del que se parta (Schmidt y Wiegand, 2017).

Asimismo, el contexto comunicativo en el mundo virtual posee unas características diferentes que afectan a la calidad y cantidad de información contextual (Yus, 2018). Surgen nuevos tipos de textos o géneros textuales de las diferentes herramientas de comunicación en línea, y estos textos poseen unos rasgos recurrentes y unas características lingüísticas propias. Por tanto, parece lógico que se tenga en cuenta, en el análisis de la violencia verbal, la información contextual compartida, que dependerá de la red social en cuestión y de la información que incorpore, y el género discursivo que se investiga. Esta información es imprescindible a la hora de determinar qué rasgos considerar en su análisis automático (Brenes Peña y Fuentes Rodríguez, 2017) y también cómo procesarlos para una posterior detección automática. Estudios recientes empiezan a incorporar esta información (Schmidt y Wiegand, 2017); sin embargo, no se aprecia en general una

inclusión de las características del contexto comunicativo en la mayoría de las investigaciones.

De este modo, los problemas teóricos expuestos se convierten en metodológicos. El hecho de la falta de caracterización clara y uniforme de la violencia verbal conlleva distintas metodologías de análisis (Ajdari y Ismaili, 2017) según se conciba como un fenómeno gradual o con categorías cerradas, por ejemplo. Las clasificaciones, además de ser varias, parten de distintos parámetros (la intención, el grado de lesión...), hecho que dificulta un análisis minucioso con las herramientas computacionales disponibles.

Todo esto aplicado al análisis automático conlleva una concepción simplista de la violencia verbal, centrada en el análisis del mensaje atendiendo básicamente a la parte textual (Ajdari y Ismaili, 2017) y, por tanto, resulta ser un análisis parcial de dicho fenómeno.

3.2 Problemas computacionales

En los desarrollos computacionales las categorías de violencia verbal que se emplean también varían según los autores (Waseem, 2016; Razavi et al., 2010), no se definen claramente y se consideran categorías discretas.

Se analiza, pues, la presencia o ausencia de palabras consideradas obscenas, racistas u ofensivas. Además, se emplean métodos estadísticos y la mayoría de los estudios consideran enfoques generales, sin atender a rasgos particulares para la violencia verbal como sucede, por ejemplo, en el discurso de odio con la apelación de un «vosotros» frente a un «nosotros» (Fortuna, 2017; Biere, 2018).

A esto hay que añadir las diferencias entre la información que se considera y los niveles de análisis. Son pocos los trabajos que incluyen la información extralingüística (información del usuario, de la red social...) y los recursos de estas plataformas (puntuación, emoticonos...), y abundan los análisis léxico-sintácticos (Burnap y Williams, 2016): la minería del texto basada en frecuencias de tokens, bolsas de palabras (Nahar, Li y Pang, 2013) o analizadores sintácticos.

No obstante, si se pretende identificar estos mensajes es esencial considerar las condiciones contextuales (Whitten, 2018). Algunos estudios recientes empiezan a incluir más niveles de análisis e información extralingüística,

tales como los estados sociales, las emociones (Douce Munezero, 2017) o el análisis de opinión y la polaridad; otros introducen también el conocimiento del mundo y la metainformación del usuario (Schmidt y Wiegand, 2017). Cabe destacar la importancia de los límites contextuales que subrayan algunos autores para identificar el lenguaje grosero como recurso expresivo (Biere, 2018), sin considerarlo violencia verbal.

Con relación al procesamiento de los mensajes, en la fase de normalización del texto se excluyen las marcas de puntuación propias de recursos estilísticos, emoticonos, mayúsculas..., con lo que consideramos que se pierde información relevante. Las clasificaciones son binarias (violento o no a partir de la asignación de violencia inherente en el término) y se aplican clasificaciones de otros contextos socioculturales o traducciones literales de estudios en otros géneros y entornos culturales.

Por último, observamos variedad de algoritmos y de distinto tipo de aprendizaje: supervisado (Douce Munezero, 2017) y no supervisado (Gordeev, 2016); métodos clásicos y aprendizaje profundo (Badjatiya et al., 2017).

Consideramos, así, que los lingüistas e investigadores deberían ir a la par con la evolución de las tecnologías digitales y de las relaciones sociales que se establecen en ellas (Androutsopoulos y Beißwenger, 2008).

4. PROPUESTA PARA EL ANÁLISIS DE LA VIOLENCIA VERBAL

Tras resumir los distintos problemas teóricos y prácticos para el análisis de la violencia verbal en las redes sociales, proponemos lo siguiente.

En primer lugar, es necesaria una descripción y caracterización clara del fenómeno para poder desarrollar una formalización que dé cuenta de la información relevante para su detección automática. Respecto a las clasificaciones, consideramos que es necesario atender a la variabilidad contextual, tanto en lo referente a la parte sociocultural como a la de la situación comunicativa. Para ello, hay que tener en cuenta tanto la información bibliográfica de las teorías tradicionales como la de las investigaciones actuales, así como el conocimiento de los hablantes con relación al fenómeno. Se ha de considerar el marco contextual y sociocultural en el caso del español, como ya hacen algunos estudios sobre la comprensión de la ironía, atribuyendo, por ejemplo, marcos a palabras, a cadenas de palabras y a situaciones (Yus, 2013).

Cabe considerar, además, la posible diferente percepción del mismo

fenómeno en el contexto virtual y el contexto físico: atendiendo a la mayor exposición de la imagen positiva en el mundo virtual, ¿un mismo acto realizado en internet y cara a cara es violento en la misma medida en ambos contextos? Es necesario también tener en cuenta estos factores y clasificar los mensajes según el grado de violencia en su emisión a través de las redes sociales.

Como hemos expuesto, las redes sociales conforman un entorno en el que la información disponible en el proceso de comunicación se ofrece de manera distinta. Es pertinente, pues, tener en cuenta qué información se tiene al alcance al emitir y recibir estos enunciados violentos y cómo se puede formalizar. Se ha de atender al contexto compartido en el mundo virtual estableciendo, por ejemplo, condiciones de *tipicalidad* que caractericen el prototipo del fenómeno [algunos autores, como Utsumi (2007), emplean esta metodología en el estudio de la ironía].

Finalmente, la violencia verbal se puede comunicar de manera implícita o explícita; por tanto, se han de incluir diferentes niveles lingüísticos tanto en la extracción de patrones como en el análisis de enunciados. Creemos necesario incluir diferentes rasgos en diferentes niveles [puntuación, emoticonos, mayúsculas... (Robinson et al., 2018)] y aquellos fenómenos de *citacionalidad* (Butler, 1997) que evocan situaciones violentas.

Por tanto, proponemos:

- Caracterizar de manera clara, no ambigua, la violencia verbal en el contexto español. Se analizarán las definiciones tradicionales de descortesía realizadas en el ámbito lingüístico y las del ámbito legal, fijando una caracterización concreta para el contexto cultural español.
- Establecer categorías difusas en las clasificaciones de violencia verbal. Las diferentes clases que forma parte del fenómeno son graduales, por lo que se han de fijar en un continuum. Estas categorías se extraerán de los estudios realizados y del conocimiento metapragmático de los hablantes mediante encuestas.
- Realizar un análisis en diferentes niveles lingüísticos. Se extraerá un corpus de datos y se estudiarán los rasgos lingüísticos (fonéticos, léxicos, sintácticos y semánticos) propios de las categorías anteriormente fijadas, así como los recursos más comunes (metáforas, comparaciones...).
- Considerar la información disponible en la red social (textual, del usuario, de la red...), siendo conscientes de las limitaciones por motivos

de privacidad. Analizaremos los patrones situacionales que se repiten en los enunciados de violencia verbal para fijar las variables estables en estos enunciados.

- Extraer patrones lingüísticos y pragmáticos (hablantes, tema, contexto comunicativo, recursos) de la violencia verbal. Una vez investigados los rasgos lingüísticos y situacionales, formalizaremos estas características en las categorías de violencia verbal para posibilitar su análisis computacional.

Para ello, será necesario analizar un corpus de datos que se extraerá de las redes sociales (con herramientas como la API de Twitter) y analizar también el discurso de los hablantes sobre el fenómeno (mediante estudios psicolingüísticos) y los estudios realizados en el ámbito académico.

5. CONCLUSIONES

En este trabajo nos hemos centrado en la violencia verbal, en concreto en los problemas que hemos detectado en su análisis y hemos propuesto algunas modificaciones para el análisis del fenómeno en las redes sociales. Destacamos tres puntos fundamentales de nuestra investigación.

En primer lugar, la violencia verbal es cultural y es necesario incluir la información sociocultural en su análisis. En segundo lugar, las categorías de violencia verbal difieren entre culturas y suponen diferentes grados de violencia, por lo que es necesaria una formalización de estas categorías en términos difusos y enmarcadas en su contexto sociocultural. En tercer lugar, las redes sociales ofrecen recursos y herramientas diferentes a las de la comunicación cara a cara. Parece lógico modificar la perspectiva de análisis e incorporar estos recursos lingüísticos en la detección de mensajes violentos en las redes sociales.

Los métodos de detección automática deberían considerar las tres premisas anteriores y ampliar la información que incorporan en el análisis de estos mensajes: incluyendo la información sociocultural y contextual, la del género textual en cuestión y la variabilidad en la percepción del fenómeno.

Creemos necesario revisar y actualizar las bases teóricas de la violencia verbal y de la información que incluyen los desarrollos computacionales. Para poder desarrollar un análisis computacional de la violencia verbal es crucial examinar la información sociocultural relevante y los diferentes niveles lingüísticos, teniendo en cuenta el contexto que suponen las redes sociales. Es necesario, pues, el trabajo conjunto de lingüistas e informáticos.

Por todo lo dicho, el trabajo que estamos realizando se centra, por un lado, en la investigación de las variables situacionales o *condiciones idóneas* (Utsumi, 2007) constantes en la emisión de estos enunciados para delimitar el marco situacional y, por otro lado, en el estudio de los diferentes métodos computacionales existentes y aplicados en la detección del fenómeno.

Asimismo, estamos trabajando en la recopilación del corpus: considerando el método más adecuado, atendiendo a diferentes niveles lingüísticos, a la información disponible (de la red, del usuario, del conocimiento del mundo) y las restricciones legales. Recopilado el corpus, llevaremos a cabo la extracción de rasgos y patrones de estos enunciados que nos permitirá proponer una formalización de la información a considerar en la detección automática de la violencia verbal en las redes sociales.

REFERENCIAS BIBLIOGRÁFICAS

- AJDARI, J., ISMAILI, F. 2017. "Automatic Hate Speech Detection in Online Contents Using Latent Semantic Analysis". 2nd World Conference on Technology, *Innovation and Entrepreneurship*. 12-14 de Mayo, Estambul, Turquía.
- ANDROUTSOPOULOS, J.; Beißwenger, M. 2008. "Introduction. Data and Methods", Computer-Mediated Discourse Analysis. *Language@Internet*. 5(9): 1-7.
- AUSTIN J. 1955. *Cómo hacer cosas con palabras: palabras y acciones*. Barcelona: Paidós, 1971.
- BADJATIYA, P., GUPTA, S., GUPTA, M. 2017. "Deep Learning for Hate speech detection in tweets", *Proceedings of the 26th International Conference on World Wide Web Companion*. 759-760.
- BIERE, S. 2018. "Hate Speech Detection Using Natural Language Processing Techniques", Department of Mathematics, Faculty of Science. Vrije Universiteit Amsterdam.
- BRENES PEÑA, E., FUENTES RODRÍGUEZ, C. 2017. "Impoliteness and Thematic Variation in Spanish Television Interviews", *Studies in Media and Communication*, Vol. 5, 1. 50-62.
- BROWN P., LEVINSON, S. 1987. *Politeness. Some Universals in Language use*. Cambridge: Cambridge University.
- BUTLER, J. 1997. *Excitable speech: A Politics of the Performative*. Londres: Routledge.
- BURNAP, P., WILLIAMS, M. L. 2016. *Us and them: identifying cyber hate on Twitter across multiple protected characteristics*. EPJ Data Science, 5: 11.
- CULPEPER, J. 1996. "Toward an Anatomy of Impoliteness", *Journal of Pragmatics* 25. 349-367.
- CULPEPER, J. 2010. "Conventionalised Impoliteness Formulae", *Journal of Pragmatics* 42. 3232-3245.
- CULPEPER, J., HAUGH, M., KÁDÁR, D. 2017. *The Palgrave Handbook of Linguistic (im)politeness*. Londres: Palgrave Macmillan.
- CULPEPER, J. 2017. "Impoliteness metalinguistic labels and concepts in

- English”, *Doing Pragmatics Interculturally. Cognitive, Philosophical and Sociopragmatic Perspectives*. Berlín: Mouton de Gruyter. 135-147.
- DOUCE MUNEZERO, M. 2017. *Leveraging emotion and Word-based features for antisocial behavior detection in user-generated content*. University of Eastern Finland, n. 264.
- FARZINDAR, A. 2018. “Natural Language Processing for Social Media”, *Synthesis Lectures on Human Language Technologies*, 10, 2. 1-195. California: Morgan & Claypool Publishers.
- FORTUNA, P. 2017. *Automatic detection of hate speech in text: an overview of the topic and dataset annotation with hierarchical classes*. Trabajo de Fin de Máster. Faculdade de Engenharia, Universidade do Porto, Oporto.
- GORDEEV, D. 2016. “Automatic verbal aggression detection for Russian and American imageboards”, *Procedia-Social and Behavioral Sciences*, 06/12/2015. Moscú, Rusia.
- HERRING, S. C. 2004. “Computer-mediated discourse analysis”, *Designing for Virtual Communities in the Service of Learning*, S. A. Barab; R. Kling; J. H. Gray (eds.). Nueva York: Cambridge University Press. 338-376.
- MIRÓ LLINARES, F. 2016. “Taxonomía de la comunicación violenta y el discurso del odio en Internet”, *Revista de los Estudios de Derecho y Ciencia Política*, Universitat Oberta de Catalunya. 93-118.
- NAHAR, V., LI, X., PANG, Ch. 2012. “An Effective Approach for Cyberbullying Detection”, *Communication in Information Science and Management Engineering*, 3. 238-247.
- RAZAVI, A., INKPEN, D., URITSKY, S., MATWIN, S. 2010. “Offensive language detection using multilevel classification”, *Advances in Artificial Intelligence*, 6085/2010. 16-27.
- SCHMIDT, A., WIEGAND, M. 2017. “A Survey on Hate Speech Detection using Natural Language Processing”, *Proceedings of the Fifth International Workshop on Natural Language Processing for Social Media*. Valencia, España, Abril 3-7.
- SPERBER D., WILSON D. 1986. *La Relevancia. Comunicación y procesos cognitivos*. Madrid: Visor, 1994.
- UTSUMI, A. 2007. “Verbal Irony As Implicit Display of Ironic Environment:

- Distinguishing Ironic Utterances From Nonirony”, *Irony in Language and Thought. A cognitive Science Reader*, R. Gibbs; H. Colston (eds.). Nueva York/Londres: Lawrence Erlbaum Associates. 499-529.
- YUS, F. 2013. “An inference-centered analysis of jokes”, *Irony and Humor. From pragmatics to discourse*, L. Ruiz Gurillo, M. B. Alvarado Ortega (eds.). Ámsterdam/Filadelfia: John Benjamins Publishing Company. 56-82.
- YUS, F. 2018. “The interface between pragmatics and internet-mediated communication”, *Pragmatics and its Interfaces*, C. Illie, N. R. Norrick. Ámsterdam: John Benjamins. 267-290.
- VAN HEE, C., LEFEVER, E., HOSTE, V. 2018. “Exploring the fine-grained analysis and automatic detection of irony on Twitter”. *Lang Resources & Evaluation*. 52. 707-731.
- WASEEM, Z., DAVISON, T., WARMSLEY, D. WEBER, I. 2017. *Understanding Abuse: A Typology of Abusive Language Detection Subtasks*, Proceedings of the First Workshop on Abusive Language Online.
- WHITTEN, S. 2018. “A recognition-sensitive phenomenology of hate speech”, *Critical Review of International Social and Political Philosophy*. Belfast, Irlanda.

PALABRAS COGNADAS EN EL VOCABULARIO ACADÉMICO DEL INGLÉS, ESPAÑOL Y ESTONIO

Mari Kruse

Universidad de Tartu, Estonia

RESUMEN

La lengua estonia, con su procedencia uralaltaica, se considera bastante diferente de las lenguas indoeuropeas, tanto por su sintaxis y morfología como por su léxico. Sin embargo, en el estonio también existe un amplio abanico de palabras «internacionales». Hasta el momento, esta parte del léxico no ha recibido atención notable por parte de los investigadores. El presente estudio analiza la proporción de palabras cognadas (las que tienen forma reconocible y significado análogo) dentro del New Academic Vocabulary List (NAVL, Gardner y Davies 2014) y su traducción al español y al estonio, concluyéndose que la mitad del vocabulario estudiado es “internacional” en estonio. No obstante, nuestros resultados asimismo ponen de manifiesto que la riqueza del vocabulario en el L1 (estonio) es un factor determinante a la hora de buscar palabras reconocibles en otras lenguas.

Palabras clave: vocabulario académico, cognados, adquisición de lenguas extranjeras, escritura en lenguas extranjeras

ABSTRACT

The Estonian language, being part of the Uralic family, is generally considered to differ significantly from Indo-European languages in its syntax, morphology and lexis. Nevertheless, there are also numerous “international” words. This part of the Estonian lexicon has not received notable attention from researchers. The present study analyses the proportion of cognate words (those which share a recognizable form and an analogous meaning) between the New Academic Vocabulary List (NAVL, Gardner y Davies 2014) and its translation into Spanish and Estonian, coming to the conclusion that roughly half of the analysed vocabulary is “international” in Estonian. That being said, it was also patent that L1 (Estonian) lexical richness is a determining factor in finding recognizable words in other languages.

Keywords: academic vocabulary, cognates, foreign language acquisition, foreign language writing

1. INTRODUCCIÓN

Es de conocimiento general que lenguas emparentadas, como las romances o las germánicas, tienen abundante vocabulario en común. Sin embargo, en los siglos XX-XXI, también el contacto entre lenguas históricamente no emparentadas se ha multiplicado e intensificado, resultando en un vocabulario «internacional» que tiene uso especialmente abundante en contextos más formales, como puede ser el ámbito académico. Sobre todo, este vocabulario internacional tiene una base latina o griega, como en estos ejemplos (que se podrían ampliar a otras lenguas):

ENG	radio	analysis	professor
ESP	radio	análisis	profesor
EST	raadio	analüüs /analy:s/	professor

Los dos primeros ejemplos son relativamente sencillos, pero con el tercero, *profesor*, ya se puede ver que hay margen para el error: no solo es que signifique puestos diferentes en cada lengua, sino que, incluso en la propia lengua inglesa, también varía el significado que tiene: puede referirse a «catedrático» (igual que en estonio) o simplemente a una persona que enseña en la universidad. Así pues, este ejemplo funciona como una buena ilustración de que el solapamiento semántico entre formas cognadas puede ser total, parcial o nulo (Chamizo Domínguez 2009, Granger 1996, Franzen 1998).

En este artículo, voy a seguir la argumentación ofrecida en Kruse 2018 y emplear el término «cognados» para este tipo de palabras, igual que lo hacen, por ejemplo, Carroll 1992, Frantzen 1998, Chacón Beltrán 2000, Cameron Scarcella y Boyd Zimmermann 2005. Así, se definiría un **cognado** como **una palabra que tiene forma parecida y significado igual en lenguas diferentes** (Lightbown y Spada, 2006: 99), resultando en una clasificación tripartita para el léxico, semejantemente a lo que propone Jiang (2000: 67):

- desconocidos (la forma no se reconoce),
- amigos (la forma permite inferir el significado) y
- falsos amigos (la forma lleva a conclusiones equivocadas sobre el significado), que, a su vez, pueden ser
 - totales (siempre engañan) o

— parciales (dependiendo del contexto, pueden significar o no lo que sugiere la forma familiar).

Esta definición contrasta con las clásicas que requieren raíces etimológicas comunes para que se pueda considerar que dos palabras son cognadas o falsos amigos. Sin embargo, en una situación comunicativa habitual, el parentesco histórico no es un factor que un usuario naïf pueda tomar en cuenta con facilidad. A su vez, se sigue el “principio de entidad razonada” (Kellerman, 1983) y la “hipótesis de equivalencia” (Swan, 1993): si se encuentra un parecido formal, se supone que conlleva un parecido semántico (o funcional, habría que añadir) si no hay razones de peso para pensar lo contrario. Por lo tanto, compartimos la opinión de Chacón Beltrán (2000: 225) de que la definición que ofrecemos es “más pragmática y pedagógica”.

2. COGNADOS ENTRE EL INGLÉS, ESPAÑOL Y ESTONIO

La cantidad de cognados varía entre lenguas diferentes. Generalmente, se considera que una proporción mayor supone una ventaja para aprendientes de lengua extranjera, mientras que su cantidad limitada sería un factor inhibitorio (p.ej., Ringbom 1986, 1992). Se ha estimado que, en un ámbito académico, más del 80% del vocabulario del inglés es cognado para hablantes de lenguas romances (Coxhead y Byrd, 2007: 132), con lo cual el conocimiento de una de estas lenguas mejora notablemente la posición para aprender la otra, especialmente en la modalidad receptiva. Mientras tanto, los hablantes de otras lenguas pueden encontrarse en posiciones mucho menos ventajosas. Mi trabajo con estudiantes estonios del español me ha llevado a preguntarme cuál, exactamente, es esa posición inicial para ellos, porque, por una parte, se podría suponer que su L1 no les ofrece mucho apoyo para ampliar el vocabulario; pero, por otra, hay que considerar que el español suele ser la L3 o L4 para ellos, y que ya tienen conocimientos del inglés al empezar sus estudios de esta lengua. En este contexto, se han planteado dos **preguntas de investigación** principales:

- ¿En qué grado es «internacional» (=con bases grecolatinas cognadas) el vocabulario académico en estonio?
- A la hora de aprender español, ¿en qué grado aumenta la cantidad de cognados para un estonio si tiene buen dominio del inglés?

2.1. *Trasfondo y estudios previos*

Aunque las relaciones léxicas entre el español y el inglés se han estudiado ampliamente, no existen comparaciones con el estonio, especialmente en cuanto al vocabulario académico. El estonio es una pequeña lengua uralaltaica, concretamente fino-báltica, que tiene aproximadamente 1,1 millones de hablantes, la gran mayoría de los cuales reside en Estonia. Históricamente, ha tenido **una fuerte influencia del alemán**, igual que del **ruso**, pero la lengua más parecida es el **finlandés**. Desde 1991, cuando Estonia recuperó su independencia de la URSS, los procesos de democratización, emancipación y desarrollo tecnológico han estimulado cambios rápidos en el léxico, especialmente por influencia del inglés. Sin embargo, aunque el estonio cuenta con un diccionario etimológico (Metsmägi *et al.* 2012), este no recoge los préstamos de las últimas décadas y, por lo tanto, el peso de las entradas con etimología inglesa (y también grecolatina) es muy reducido, mientras que, en realidad, ha sido la influencia imperante en el pasado reciente. En el ámbito académico existen tanto una lista de vocabulario inglés-estonio (TERMEKI, 2016) como un glosario académico del estonio (Metslang y Kibar, 2012), pero ninguno de los dos presta atención a los cognados, por lo cual se puede afirmar que el presente estudio explora un terreno desconocido en el campo de contactos entre lenguas. Como avance al estudio se puede comentar que búsquedas esporádicas de algunos de los 500 términos incluidos en la lista investigada (NAVL, Gardner y Davies, 2014) no dieron ningún resultado en el diccionario etimológico estonio (Metsmägi *et al.* 2012), mientras que, en realidad, aproximadamente la mitad de ellos tienen raíces cognadas en estonio.

2.1.1 Academic Word List (Coxhead 1998)

Como hay muchas personas que manejan las lenguas española e inglesa a diario, estudios comparativos entre los dos idiomas ya han producido una variedad de frutos. En contextos generales tenemos listas y diccionarios de cognados, igual que materiales de estudio (por nombrar algunos, Gallego 2014 o las publicaciones de la editorial Velázquez Press¹).

¹Su selección de diccionarios bilingües y materiales de estudio, tanto para el ámbito académico como más generales, se puede consultar en <https://velazquezpress.com/>. En el marco del estudio presente, se puede destacar la serie *Common Core* (“Núcleo común”) para estudiantes.

En el campo académico, encontramos estudios enfocados en su vocabulario específico. Desde su publicación en 1998, *Academic Word List* de Averil Coxhead se convirtió en la lista de referencia para el inglés y, como consecuencia, los estudios comparativos con el español también la usaban como base: se ha estudiado la cantidad de cognados que incluye (Bushong, 2010) y su transparencia para hispanohablantes (Lubliner y Hiebert, 2011). En investigaciones empíricas, Cameron Scarcella y Boyd Zimmermann (2005) trataron de determinar si los hispanohablantes efectivamente tenían alguna ventaja en comparación con hablantes de lenguas asiáticas, llegando a la conclusión de que podían tener una ventaja al leer, pero no al escribir. Su recomendación era proporcionar una instrucción explícita para sacar provecho de los cognados.

2.1.2 New Academic Vocabulary List (Gardner y Davies, 2014)

Aunque, en su momento, la AWL de Coxhead siguió unas pautas de compilación rigurosas y fundamentadas, con el avance de los estudios basados en macrodatos le empezaron a llegar críticas (resumidas en Gardner y Davies, 2014: 307-308). En comparación con la AWL, la NAVL:

- se basa en un corpus más amplio y más actual, el subcorpus académico del Corpus of Contemporary American English (COCA, Davies 2008) (120/425 millones de palabras, respectivamente, en el momento de compilar la NAVL);
- contiene los **500 lemas**² más frecuentes de dicho corpus, y no se organiza en familias de palabras como en la AWL, donde las 570 entradas, en realidad, se bifurcaron en una cantidad mucho mayor de formas conjugadas y derivadas;
- no excluye las palabras de uso general si se emplean frecuentemente en el ámbito académico, pues se basa completamente en frecuencias.

Por la rigurosidad y la actualidad de sus principios de compilación, se consideró a la NAVL como la lista más adecuada para llevar a cabo el presente estudio. Además, todavía no se han publicado investigaciones contrastivas basadas en la NAVL. Así, el presente estudio da un paso más allá de la línea investigativa

²En realidad, son 499 lemas porque el adjetivo *general* aparece dos veces, marcado como adjetivo y como adverbio, mientras que el adverbio *generally* también se incluye aparte.

iniciada con los estudios de la AWL.

3. MÉTODO Y PREGUNTAS DE INVESTIGACIÓN

Se llevó a cabo un análisis manual de la NAVL, empleando varios diccionarios de referencia en línea: *Collins English-Spanish Dictionary*, *WordReference English-Spanish Dictionary*, diccionario inglés-estonio de Aare Vesi (Vesi 2011, basado en los diccionarios de Eesti Keele Instituut), diccionario inglés-estonio de FiloSoft, *Oxford English Dictionary* (ahora *Lexico.com*), *Diccionario de la lengua española* (Real Academia Española, 23ª ed., 2014) y *Eesti keele seletav sõnaraamat* (Eesti Keele Instituut, 2009). Para cada lema de la NAVL, se tomó nota de lo siguiente:

- traducción al español y al estonio;
- si era cognado en español y/o estonio;
 - en caso afirmativo, si el cognado era primario [o sea, la forma (más) habitual o no];
 - en caso afirmativo, el grado de solapamiento semántico del cognado: total/parcial/falso amigo.

Por ejemplo, las primeras tres entradas de la lista, las más frecuentes, quedaron como sigue:

ENG	study (n)	group (n)	system (n)
ESP	estudio, investigación, reflexión	grupo	sistema, método, organismo, cuerpo, aparato, régimen, formación
cognado	+	+	+
primario	+	+	+
solapamiento	ENG más amplio	total	total
EST	uurimus, uuring	grupp, rühm	süsteem
cognado	-	+	+
primario	N/A	(+)*	+
solapamiento	N/A	total	total

*(+ para el estonio indica que se trata de un equivalente primario que alterna libremente con otro no cognado (en el caso del ejemplo, “rühm”).

Tabla 1. Ejemplo de la categorización empleada

La categorización empleada asume el punto de vista de un usuario sin conocimientos etimológicos, por lo cual parejas como “discovery” - “descubrimiento” o “design” - “diseño”, que tienen un origen común se clasificaron aparte.

Concretamente, se buscó responder a las siguientes preguntas de investigación:

- En la NAVL, ¿cuántas de las 500 palabras son cognadas para hablantes de español y estonio? (En otras palabras, ¿en qué parte es internacional el vocabulario académico estonio?)
- ¿Cuántos de los cognados son demostrablemente primarios? En tal caso, la palabra inglesa no tendría acepciones múltiples que se expresaran con palabras no cognadas en español, y el cognado en español sería el equivalente más habitual y natural.

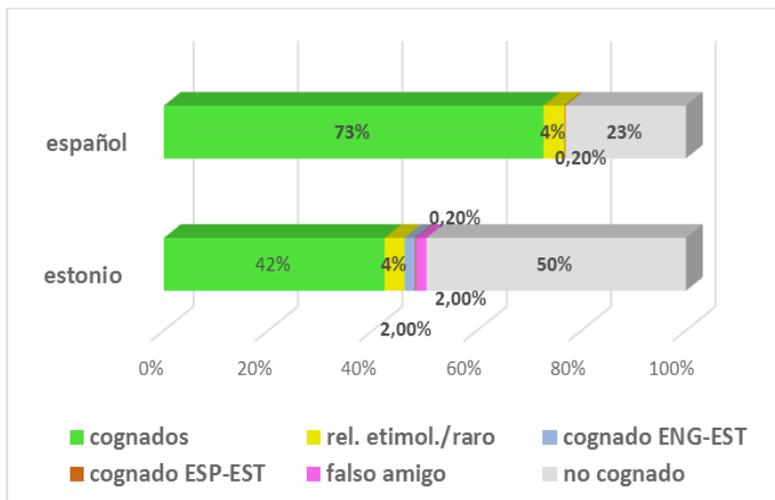
4. RESULTADOS

Para el español, NAVL tenía

- 363 cognados;
- 20 palabras de relación etimológica que no necesariamente se reconocerían como cognados;
- 1 palabra que no es cognada en el inglés, pero sí entre el español y el estonio;
- 115 no cognados.

Para el estonio, contenía

- 211 cognados;
- 19 cognados que tenían un significado muy específico o poco frecuente;
- 10 cognados ENG-EST (no con ESP; 1 de los 10 es muy raro y específico);
- 1 cognado ESP-EST (no con ENG);
- 10 falsos amigos ENG-EST;
- 249 no cognados.



A continuación, se presentarán algunos ejemplos de palabras de cada grupo.

- Cognados entre las tres lenguas:

ENG	system
ESP	<u>sistema</u> , método, organismo, cuerpo, aparato, régimen, formación
EST	<u>süsteem</u>
ENG	social
ESP	<u>social</u> , sociable
EST	<u>sotsiaalne</u> , ühiskondlik
ENG	process
ESP	<u>proceso</u> , procedimiento, protuberancia; requerimiento de comparecencia, citación
EST	<u>protsess</u> , kulg, jätke

Por otra parte, había 110 entradas que eran cognadas o tenían etimología común, pero que no se podían considerar como la opción más natural o frecuente en español, o sea, no eran cognados primarios. Por ejemplo:

ENG	actual
ESP	real, exacto, textual, propiamente dicho; <u>actual</u> (DLE: “1. Dicho del tiempo en que se está: presente; 2. Perteneciente o relativo al tiempo actual”.)
EST	tóeline, tegelik; nüüdne, praegune; päevakajaline, <u>aktuaalne</u> (EKSS: “actual, importante en el momento dado, de interés actual”)
ENG	conduct
ESP	llevar, llevar a cabo, <u>conducir</u> , dirigir, gestionar, presentar, comportarse
EST	läbi viima, juhtima (ka soojust, elektrit); saatma; dirigeerima

Algunos cognados se podían considerar muy específicos y/o infrecuentes, como, p.ej., ESP *centuria* = EST *tsentuuria* (DLE: “en la milicia romana, compañía de 100 hombres”; = ENG *century*), o EST *žurnaal* (“revista, diario, registro”, poco usado = ENG *journal*). Esto era especialmente evidente para el estonio, ya que, en muchas ocasiones, el cognado se consideraría “extranjero”, por así decir, aunque para algunas palabras pueda ser un equivalente perfectamente habitual. Son casos como:

eksisteeriv (“existente”, pero normalmente se usaría *olemas olev* que significa lo mismo; = ENG *existing*).

En general se podría inferir que si un texto estonio optara siempre por los cognados, correría el riesgo de sonar afectado y rebuscado. Por otra parte, también significa que los alumnos estonios que tienen un amplio vocabulario en su lengua materna, o sea, que conocen también las palabras “extranjeras”, encontrarán una cantidad mucho más amplia de cognados tanto en inglés como en español.

En resumen, de los 363 cognados, 272 eran cognados primarios en español:

- 124 eran “fiables”, en el sentido de que coincidían los campos semánticos de la palabra inglesa y el cognado en español;
- 135 tenían un campo semántico más amplio en inglés o solapaban en parte (p.ej., ENG *institute* (*n*) - ESP *instituto*; ENG *express* - ESP *expresar/exprimir*);
- 10 entradas tenían 2 cognados en competencia (p.ej. ENG *influence*

(v) - ESP *influenciar / influir*);

- 3 eran falsos amigos (p.ej., que no existe el sustantivo *individual* sino es *individuo*).

Para el estonio, se llegó a los siguientes resultados:

- entre el inglés y el estonio, había 121 cognados primarios (además, *grado-kraad*, que es primario en ESP y EST, pero con *degree* se reconoce difícilmente). De estos:
 - 68 eran “fiables” y se podrían usar sin problemas semánticos ni estilísticos;
 - eran primarios para uno de los sentidos que tiene la palabra inglesa;
 - eran primarios, pero tenían un campo semántico más amplio en inglés.
- 119 no se podían considerar cognados primarios, y de estos:
 - 74 tenían un campo semántico más amplio en inglés que en estonio;
 - 33 eran palabras raras o específicas.

5. CONCLUSIONES

Como era de esperar, la cantidad de cognados con el inglés es notablemente más amplia en el español que en el estonio. A grandes rasgos, se puede concluir que, en un contexto académico, **un hispanohablante podría esperar reconocer hasta tres de cada cuatro palabras** (74-77 %, en función de si se incluyen o no las palabras con relación etimológica), **mientras un estoniohablante contaría con dos** (48 %). Al mismo tiempo, tanto en estonio como en español existe una multitud de alternativas para las palabras cognadas con el inglés y, en numerosas ocasiones, probablemente son más adecuadas y/o exactas. En concreto, solo un 24,8 % de la lista eran cognados primarios en el español, y un 13,6 % en el estonio. El resto puede desviar al aprendiente con sentidos paralelos o solapamientos parciales. En otras palabras, aunque **una base fuerte en inglés puede resultar útil para aprender español, en la práctica, apoyarse demasiado en esta lengua puede resultar en un vocabulario empobrecido y extranjerizado.**

Como segundo resultado, se ha visto que la mitad del vocabulario académico

estonio tiene raíces grecolatinas y/o se ha tomado prestado del inglés. A la luz de todo ello, es un defecto importante que el diccionario etimológico estonio no incluya préstamos recientes, ya que las proporciones de la procedencia etimológica del léxico han cambiado drásticamente. Como conclusión importante se destaca el hecho evidente de que **estonios con un vocabulario amplio (“educado”) en L1 tienen más ventajas para reconocer cognados**. En el futuro, tenemos la intención de validar esta idea empíricamente realizando un experimento de reconocimiento en que estonios de diferentes trasfondos lingüísticos, pero sin conocimientos del español, busquen palabras reconocibles en un texto en español. Como han subrayado diferentes investigadores —p.ej., Cameron Scarcella y Boyd Zimmermann (2005), Gardner (2007), Wesche y Paribakht (2010)—, sin recibir instrucción en el reconocimiento de cognados y en las relaciones derivacionales, y sin experiencia con el aprendizaje de lenguas extranjeras, el potencial para reconocer un significado familiar detrás de una forma familiar no siempre se llega a realizar. Por lo tanto, los resultados de tal experimento ayudarían a arrojar luz sobre la capacidad de reconocer cognados en contextos realistas.

REFERENCIAS BIBLIOGRÁFICAS

- BUSHONG R.W. II. 2010. The Academic Word List reorganized for Spanish-speaking English language learners. Tesis de máster. Orlando: University of Central Florida.
- CAMERON SCARCELLA R., BOYD ZIMMERMANN C. 2005. "Cognates, cognition and writing: An investigation of the use of cognates by university second-language learners", *Language in Use: Cognitive and Discourse Perspectives on Language and Language Learning*. A. Tyler et al. (eds.). Georgetown, WA: Georgetown University Press. 123-136.
- CARROLL, S. E. 1992. "On cognates". *Second Language Research* 8/2: 93-119.
- CHACÓN BELTRÁN M. R. 2000. *La enseñanza del vocabulario en inglés como L2: el efecto del énfasis en la forma lingüística de cognados falsos*. Tesis doctoral. Sevilla: Universidad de Sevilla, Facultad de Filología.
- CHAMIZO DOMÍNGUEZ P. J. 2009. "Los falsos amigos desde la perspectiva de la teoría de conjuntos". *Applied Linguistics Now. Understanding Language and Mind = La lingüística aplicada actual: comprendiendo el lenguaje y la mente*, C. M. Bretones Callejas et al. (eds.). Almería: Universidad de Almería. 1111-1126.
- COLLINS SPANISH ONLINE DICTIONARY. Disponible en <https://www.collinsdictionary.com/> (consultado el 10/3/2019).
- COXHEAD A. J. 1998. *An academic word list (English Language Institute Occasional Publication No. 18)*. Wellington, New Zealand: Victoria University of Wellington.
- COXHEAD A., BYRD P. 2007. "Preparing writing teachers to teach the vocabulary and grammar of academic prose", *Journal of Second Language Writing* 16/3: 129-147.
- DAVIES M. 2008- *The Corpus of Contemporary American English (COCA): 560 million words, 1990-present*. Provo, UT: Brigham Young University.
- EESTI KEELE INSTITUUT 2009. *Eesti keele seletav sõnaraamat* (en línea). Disponible en <https://www.eki.ee/dict/ekss/> (consultado el 15/7/2019).
- FILOSOFT 2007-. *Inglise-eesti sõnastik* (en línea). Editores A. Tedresalu y K. Traks. Disponible en https://www.filosoft.ee/dict_enet/ (consultado el 15/7/2019).

- FRANTZEN D. 1998. "Intrinsic and extrinsic factors that contribute to the difficulty of learning false cognates", *Foreign Language Annals* 31/2: 243-254.
- GALLEGO L. 2014. *Spanish English Cognates: Thousands of Spanish English Cognates to Help You Expand Your Vocabulary Instantly*. Lexiconer.
- GARDNER D. 2007. "Validating the construct of *word* in applied corpus-based vocabulary research: a critical survey", *Applied Linguistics* 28/2: 241-65.
- GARDNER D., DAVIES M. 2014 "A new academic vocabulary list", *Applied Linguistics* 35/3: 305-327.
- GRANGER S. 1996. "Romance words in English: from history to pedagogy", *Words. Proceedings of an International Symposium*, J. Svartvik (ed.). Stockholm: Almqvist and Wiksell International. 105-121.
- JIANG N. 2000. "Lexical representation and development in a second language", *Applied Linguistics* 21: 47-77.
- KELLERMAN E. 1983. "Now you see it, now you don't", *Language Transfer in Language Learning*, S. Gass, L. Selinker (eds.). Rowley: Newbury House. 112-134.
- KRUSE M. 2018. *La transferencia en personas plurilingües: los falsos amigos como un obstáculo y una oportunidad en la enseñanza y aprendizaje de lenguas extranjeras* (tesis doctoral). Tartu: Tartu Ülikooli Kirjastus.
- LEXICO (en línea; hasta la primavera de 2019, *Oxford English Dictionary*). Disponible en <https://www.lexico.com/en> (consultado el 15/7/2019).
- LIGHTBOWN P. M., SPADA N. 2006. *How Languages Are Learned*. 3ª edición. Oxford: Oxford University Press.
- LUBLINER S., HIEBERT E. H. 2011. "An analysis of English-Spanish cognates as a source of general academic language", *Bilingual Research Journal* 34/1: 76-93.
- METSLANG H., KIBAR, T. 2012. *Üldakadeemiline sõnavara. Abivahend eesti keele õppeks kõrgkoolis*. Tallinn: Tallinna Ülikool.
- METSMÄGI I., SEDRIK M., SOOSAAR S.-E. 2012. *Eesti keele etimoloogia sõnaraamat*. Tallinn: Eesti Keele Sihtasutus.

- MORANTE VALLEJO R. 2005. *El desarrollo del conocimiento léxico en segundas lenguas*. Madrid: Arco/Libros.
- REAL ACADEMIA ESPAÑOLA. 2014. *Diccionario de la lengua española*, 23ª ed. (en línea). Disponible en <https://dle.rae.es/> (consultado el 10/3/2019).
- RINGBOM H. 1986. "Crosslinguistic influence and the foreign language learning process", *Crosslinguistic Influence in Second Language Acquisition*, E. Kellerman, M. Sharwood Smith (eds.). New York, etc.: Pergamon Press. 150-162.
- RINGBOM H. 1992. "On L1 transfer in L2 comprehension and L2 production", *Language Learning* 42/1: 85-112.
- TERMEKI 2016- *Akadeemilise väljendusoskuse terminid* (en línea). Disponible en <https://term.eki.ee/termbase/view/8068645/> (consultado el 10/3/2019)
- SWAN M. 1997. "The influence of the mother tongue on second language vocabulary acquisition and use", *Vocabulary: Description, Acquisition and Pedagogy*, N. Schmitt, M. McCarthy (eds.) Cambridge: Cambridge University Press. 156-180.
- VESI A. 2011. *Eesti-inglise-eesti sõnaraamat* (en línea). Disponible en <https://aare.edu.ee/dictionary.html> (consultado el 15/7/2019).
- WESCHE M. B., PARIBAKHT T. S. 2010. *Second Language Acquisition: Lexical Inferencing in a First and Second Language: Cross-linguistic Dimensions*. Bristol, Buffalo, Toronto: Multilingual Matters.
- WORDREFERENCE ENGLISH-SPANISH DICTIONARY (en línea). Disponible en <https://www.wordreference.com/> (consultado el 15/7/2019).

LA INFORMACIÓN ETIMOLÓGICA EN LOS DICCIONARIOS GENERALES DE LENGUA DEL ESPAÑOL Y DEL ITALIANO DEL SIGLO XX

Maria Iannotti

Liceo Galilei Civitavecchia

Università Niccolò Cusano

RESUMEN

La etimología tiene su mayor proyección en la lexicografía. En el diccionario etimológico encuentra su natural e inevitable desarrollo, con el fin de registrar y fundamentar de manera científica el origen de las palabras. No obstante, la información etimológica está presente también en algunos diccionarios generales sincrónicos, sean descriptivos, sean normativos, y obedece mayoritariamente al concepto de etimología-origen, que aporta noticias sobre la raíz léxica, la procedencia y la composición de la palabra. La motivación del presente estudio se sustenta principalmente en la convicción de que el papel de las obras lexicográficas no puede limitarse solo a la simple consulta, sino que debería proporcionar al lector una información más completa y exhaustiva del hecho lingüístico.

Palabras clave: etimología, lexicografía

SINTESI

L'etimologia trova nel dizionario etimologico il suo naturale ed inevitabile strumento, con il fine di registrare e giustificare scientificamente l'origine delle parole. Nonostante ciò, è presente anche in alcuni dizionari generali sincronici, sia descrittivi che normativi, ed obbedisce al concetto di etimologia-origine, apportando notizie sulla radice lessicale, l'origine e la formazione della parola. Il motivo del presente studio si basa essenzialmente sulla convinzione che il ruolo delle opere lessicografiche non può limitarsi alla semplice consultazione, ma deve fornire al lettore una informazione più completa ed esaustiva del fatto linguistico.

Parole chiave: etimologia, lessicografia

1. LA INFORMACIÓN METALINGÜÍSTICA: LA ETIMOLOGÍA

El discurso etimológico es de tipo metalingüístico, sirve para hablar de la lengua misma, es decir, su definición explicita la función de la unidad léxica definida. Se trata de un comentario sobre el signo y no de una paráfrasis del significado.

Un análisis de la información etimológica en el artículo lexicográfico permite constatar que el comentario etimológico en los diccionarios generales de lengua entrega informaciones del étimo en cuanto significante. Ahora bien, para que el comentario etimológico sea útil y no se quede como algo marginal dentro de la microestructura de los diccionarios de lengua, es necesario ampliarlo significativamente. De hecho, la etimología, además de interesarse por el origen de la palabra, analiza también su incorporación a un idioma, su fuente, su edad, las variaciones de forma (productividad morfológica) y de significados (productividad sémica) que las palabras sufren a medida que las lenguas evolucionan y se adaptan a un determinado momento histórico.

El criterio etimológico, según el cual, en la escritura de las palabras, debe respetarse en alguna medida la forma gráfica de su étimo, ha operado en la configuración del sistema ortográfico del español (Real Academia Española, 2010: 37). Este criterio funciona en oposición al fonológico y explica el motivo por el que no hay correspondencia, en algunas palabras, entre grafía y pronunciación. Asimismo, el criterio etimológico sigue operando hoy a la hora de fijar la grafía de aquellas palabras que contienen fonemas que admiten varias presentaciones gráficas, pues lo habitual es respetar, en esos casos, los grafemas etimológicos.

En el plano de la significación, en su etimología, es en donde podemos encontrar las pistas perdidas de la existencia de las palabras. Por eso la información etimológica debería representar un hecho funcional en el artículo lexicográfico, un acto útil, ventajoso y apto para llegar a un conocimiento más profundo de las lenguas y del lenguaje; en efecto, la indicación etimológica, reflejando directa o indirectamente la realidad histórica del período que vive, establece vínculos significativos entre la historia de la lengua y la historia de la cultura.

«La etimología no es una curiosidad erudita de interés puramente histórico, sino que es la base misma de la propiedad idiomática» (Ramón Menéndez Pidal, 1945: XXII).

2. LA LEXICOGRAFÍA ACADÉMICA

2.1 *La Accademia della Crusca y la etimología*

La relación entre la Academia italiana y la etimología fue muy difícil. Hubo un total desinterés en las dos primeras ediciones del *Vocabolario*, en donde los étimos se indicaban muy raramente y solo en función de la definición. En los prolegómenos a la tercera edición, la etimología se consideraba «fuori affatto dell'intenzione della presente Opera» (*Crusca*, 1691: 14) y en el prefacio a la cuarta edición se leía: «ci siamo astenuti in tutto, e per tutto dall'assegnare l'etimologie, e l'origine di qualsivoglia Voce essendo per lo più incerte, e dubbie».

La *Crusca* solo se limitó a decir «*Voce greca, Voce latina*» (Monti, 1819: 42) para legitimar las palabras italianas, justificando la relación estrecha entre las tres lenguas. Solo en la quinta edición los académicos decidieron abolir la práctica de la doble traducción en latín y en griego para sustituirla con la indicación de la etimología.

Desgraciadamente, los principios del *Prefacio* no se aplicaron con el mismo rigor y equilibrio a las etimologías de todas las voces. Faltó un grupo de especialistas que se ocupara de etimología, así que esa labor se quedó al margen del trabajo principal que siguió consistiendo en las definiciones y en la selección de los ejemplos.

2.2 *Real Academia Española*

El interés por la etimología de la Academia Española fue bastante inconstante. Empieza en 1726, fecha de la publicación del primer tomo del *Diccionario de autoridades*, en el que la Academia afirma que todos los buenos diccionarios tienen que incluirla entre sus informaciones.

En la segunda edición de *Autoridades*, de 1770, se incluyeron solo las etimologías que parecían propias y fundadas. Sucesivamente, en la nueva versión de 1780, que hoy se conoce como la primera edición del *Diccionario de la Real Academia Española (DRAE)*, se eliminaron todas las etimologías de las voces y también las citas de autoridades, dejando solo la voz, la definición y la correspondencia latina, con la intención de realizar una obra más manejable

y económica. Desde entonces han aparecido 23 ediciones, y habrá que esperar a la undécima, la de 1869, para que la Academia suprima también las correspondencias latinas.

Con la duodécima edición, de 1884, comienzan a incluirse las etimologías (excepto en las voces de origen incierto o desconocido o de palabras de etimologías obvias) hasta que, en 1914, la inclusión se hizo definitiva.

3. LA LEXICOGRAFÍA NO ACADÉMICA

3.1. Los diccionarios generales italianos y españoles: análisis y comparación

Analicemos la información etimológica en los diccionarios italianos y españoles del siglo XX y todo lo que está íntimamente relacionado con ella¹: el recorrido lleva, así, a la presentación de la etimología, derivados y compuestos, extranjerismos, criterio del orden de las acepciones, fuentes y, finalmente, representación de la homonimia y polisemia.

Podemos observar que todos los diccionarios generales analizados, tanto los italianos como los españoles, tienen características comunes.

1. Clasificación: lingüístico, monolingüe, general, selectivo excepto el *GRADIT* (al que se considera exhaustivo), pancrónico, descriptivo, alfabético.
2. Dimensión: se sitúan entre 70 000 y 120 000 lemas o 2000 o 3000 páginas (excepto el *GRADIT*, obra monumental de mayor alcance).
3. Destinatarios: siguiendo las indicaciones de los mismos diccionarios, el público potencial a quien se dirigen es muy amplio.

El étimo de la misma voz analizada en los diccionarios generales aporta un tipo de información *stricto sensu*, es decir, en el sentido propio y más restrictivo de la palabra. No hay reflexión alguna acerca del origen de la palabra ni de su cambio a lo largo del tiempo. En algunos diccionarios italianos, sobre todo los más recientes, la información aparece más detallada y, además del étimo, se incluye su procedencia y la primera fecha de atestiguación. En los españoles, excepto el de María Moliner que incluye el étimo del étimo, solo se da el origen latino.

¹De la lexicografía italiana se analizan *Zingarelli*, *Devoto-Oli*, *VOLIT*, *DISC*, *GRADIT*; de la española, *VOX*, *DUE*, *GDLE*, *CLAVE*.

La etimología está separada por símbolos (paréntesis, corchetes, tipos diferentes de letras), constituye *de facto* un tipo más de información del signolema como significante (Bugueño Miranda, 2004: 179) y su representación cambia conforme a los diccionarios. Las formas se registran en el caso nominativo (a veces con genitivo) o acusativo. Si la palabra no está documentada, aparece con un asterisco.

La fecha de atestación, adquirida por varias fuentes, generalmente escritas, aparece con la fecha exacta, con la especificación de un período de tiempo o con la indicación de un siglo. En los diccionarios italianos, las palabras procedentes de otras lenguas se incluyen con la fecha de acogida en la lengua italiana. Las fuentes pertenecen principalmente a dos grandes diccionarios de la editorial UTET: el *Dizionario de la lengua italiana* de Nicolò Tommaseo, publicado en 1861, y el *Grande dizionario de la lengua italiana*, diccionario histórico publicado a partir de 1961, nacido como actualización del primero y fundado por Salvatore Battaglia, bajo la dirección de Giorgio Bàrberi Squarotti. En poquísimos casos no se ha encontrado documentación alguna útil para la datación.

Los diccionarios españoles, generalmente, no incluyen la datación; este es el motivo por el que Menéndez Pidal, en el Prólogo al *Diccionario general ilustrado de la lengua española*, está convencido de que «[...] se cometen frecuentes errores en las investigaciones etimológicas» (DGILE, 1945: XXII).

Por lo que concierne a las abreviaturas, todos los diccionarios italianos (excepto el *Zingarelli*) introducen *Lat.*, en caso de tradición directa, a través del uso hablado y popular; *Dal lat.*, en caso de recuperación docta, a través de las fuentes escritas; asimismo, frecuentemente, tanto en los diccionarios italianos como en los españoles, se añaden especificaciones con referencia al período de pertenencia del término: *lat.* (latín clásico), *lat. tardo*, *lat. volg.*, *lat. mediev.*, *lat. scient.*

Generalmente, la reconstrucción etimológica se limita al directo antecedente cuando la base encontrada en la lengua de origen ya es portadora del contenido semántico suficiente para explicar el significado del lema etimologizado. Precisamente ese es el límite de las indicaciones etimológicas en los diccionarios generales: no fijar como objetivo remontarse a un pasado más remoto, no plantearse reconstruir un pasado efectivamente muerto, situaciones propias de otros tiempos, con caracteres formales, fonéticos, morfológicos o valores semánticos propios de otros tiempos, que convertiría a la palabra en una fuente de historia. Y el límite del antecedente directo se supera solo en caso de que el

ulterior retroceso aporte una integración importante, como sucede cuando se remonta desde el latín hasta el griego o el hebreo, o desde una lengua extranjera moderna o un dialecto italiano hasta la lengua que proporciona el vocablo.

3.1.1. Derivados y compuestos

Por lo que concierne a los derivados, en todos los diccionarios italianos las fórmulas *der. di* (o *del*) o *da* indican que el paso de la palabra originaria al lema se ha producido mediante modificaciones morfológicas; generalmente, la derivación es vertical, a menos que el significado del lema no imponga una solución alternativa.

Un problema que podría nacer frente a distintas palabras que proceden de la misma base puede ser su orden cronológico: en estos casos es útil la consulta de un diccionario etimológico, porque es el único que incluye la forma base y trata todos los derivados o los compuestos relativos. Una ulterior confirmación nos la da la datación de la primera atestación de la palabra: generalmente, la forma con la fecha más antigua será la base sobre la que se han formado los derivados y los compuestos². Está claro que la formación de las palabras no siempre se puede atribuir a un procedimiento formal de unión de elementos a una base de partida. Aquí la competencia del hablante y sus conocimientos previos permiten reconocer la relación semántica que une un derivado o un compuesto con las palabras base de partida: así es inmediatamente transparente que *mattonne* no es una alteración de *matto*, ni *magone* tiene nada que ver con *magò*.

Como es evidente, en estos diccionarios generales, normalmente se hace referencia a la base latina, se fija la última etapa de la evolución de la palabra, sin indicar la razón del verdadero significado, ni determinar el significado primitivo de la palabra a través de su análisis formal.

En los diccionarios generales españoles analizados, el étimo de los derivados no aparece, sobre todo si son voces claramente deducibles de la palabra base.

²En la palabra *carta* aparece la fecha de 1294; en *cartolaio*, la de 1387; *cartolina* es de 1476; *incartare*, de 1598; *incarto*, de 1812; *incartamento* se remonta a 1855: así, se puede reconstruir fácilmente la cronología del conjunto de las palabras derivadas de la base *carta*.

Para los compuestos, generalmente la fórmula etimológica usada en los diccionarios italianos es *comp. di.*; en los españoles no hay una fórmula específica, sino más bien una conjunción o un signo de adición (en el *DUE* no se especifica la composición).

3.1.2. Los extranjerismos

Los diccionarios acogen también los extranjerismos que se lematizan. No se consideran extranjerismos las palabras que se han adaptado a la lengua italiana, aunque aparezca en todos los diccionarios la procedencia.

Los diccionarios españoles incluyen los extranjerismos no adaptados, introduciéndolos con la voz de procedencia sin el étimo.

Generalmente, los latinismos, así como los helenismos, usados en el italiano y en el español contemporáneo en su forma originaria, se tratan como los extranjerismos.

3.1.3. El criterio de orden de las acepciones

Una de las mayores preocupaciones para un lexicógrafo en la redacción de un artículo léxico es decidir los criterios de orden de las distintas acepciones de referencia del término.

¿Cuál es la ordenación idónea para las diferentes acepciones? Generalmente, la norma prevé la alternancia de un criterio histórico a uno lógico a los que se añade, en los repertorios modernos, una ordenación *frequenziale* (Rey-Debove, 1971: 88). Las tres diferentes modalidades basan su elemento de atención en el primer significado históricamente atestiguado, o en un significado concreto del que se originan los sentidos trasladados, o en la acepción que más frecuentemente aparece en el uso. En cambio, el criterio etimológico, que consiste en poner como primer semema el documentado primero en la sincronía del sistema lingüístico objeto de la descripción, no se utiliza mucho en los diccionarios pancrónicos; solo el *VOX* y el *DUE* dan la precedencia a las acepciones más próximas a la etimología y, luego, ordenan las siguientes según su mayor proximidad a las que las preceden.

No siempre es posible una coherencia total en la aplicación de los criterios indicados: muy a menudo se llega a soluciones combinadas adoptadas por

el lexicógrafo, con el fin de obviar resultados poco claros o anacrónicos. Los autores casi nunca consiguen seguir rigurosamente el criterio adoptado de manera coherente para toda la nomenclatura y en las frecuentes combinaciones se intenta conciliar las opuestas exigencias de diacronía y sincronía.

3.1.4. Fuentes italianas y españolas

La base de datos y el leuario de los diccionarios generales analizados se ha fundado en la exploración y utilización de múltiples fuentes.

Sin duda alguna, un cambio radical en la investigación etimológica italiana fue la publicación del *Dizionario etimologico della lingua italiana (DELI)* de Manlio Cortellazzo e Paolo Zolli, publicado en cinco volúmenes entre 1979 y 1988. El panorama etimológico italiano antes de la llegada del *DELI* comprendía el *Dizionario etimológico italiano (DEI)* de Carlo Battisti y Giovanni Alessio, en cinco tomos publicados entre 1950 y 1957; el *Avviamento alla etimologia italiana* de Giacomo Devoto, publicado a partir de 1966; el *Lessico Etimologico Italiano*, fundado y dirigido por Max Pfister y todavía incompleto. Esta última obra se inspira directamente en el *FEW* de Von Wartburg y, junto a ella, aspira a incluir todo el léxico italiano, también el de los dialectos.

De cualquier manera, todos los que nacieron después del *DELI* sufrieron directa o indirectamente su influencia. El *Devoto-Oli* se remonta más allá, tomando como ejemplo el *DEI*, sin duda alguna la obra etimológica más significativa y completa, con un leuario muy amplio que incluye no solo el léxico de uso, sino también numerosos dialectalismos, tecnicismos y vocablos antiguos³. Los autores se centran tanto en las definiciones como en la reconstrucción de la historia de las palabras, introducidas con exactitud gracias a los numerosos textos analizados. Otra obra de referencia, con un leuario heterogéneo que incluye palabras dialectales y jergales, es el *Vocabolario etimologico italiano (VEI)* de Angelico Prati, cuyas hipótesis etimológicas son originales y, por eso, acogidas por la tradición lexicográfica sucesiva, aunque Prati insiste excesivamente sobre explicaciones de tipo onomatopéyico y expresivo-imitativo (Baglioni, 2016: 115).

³Una aportación significativa de la última edición del *Devoto-Oli* es el *Etimologico* de Alberto Nocentini, publicado en 2010. El *Etimologico* contiene hipótesis nuevas sobre la etimología de algunas palabras controvertidas, aun presentando un leuario más reducido con respecto al de los otros diccionarios etimológicos de aquel período.

Finalmente, la publicación de la *Accademia* desde hace más de cuatrocientos años es la obra de referencia de la lengua italiana y de toda la lexicografía moderna.

No siempre se conocen las fuentes de las etimologías en los diccionarios generales españoles. El autor denuncia que la mayoría de las etimologías están tomadas del francés, traducidas de Littré; incluso algunas palabras castellanas y francesas de étimo común son explicadas como galicismos (Toro y Gisbert, 1909: 3). Por supuesto, a la fuente francesa tan acreditada hay que añadir la influencia que tuvo el *Dizionario etimologico delle lingue romanze* (*Etymologisches Wörterbuch der romanischen Sprachen*) de Diez, de 1853.

Es cierto que, con la publicación del *Diccionario crítico etimológico de la lengua castellana* de Corominas, editado por primera vez en 1954, la Academia tuvo una enorme base de datos para revisar etimologías ya existentes o introducir nuevas, aunque la Academia, como es notorio, reconocerá su deuda hacia Corominas solo en 1984.

3.1.5. El tratamiento de la homonimia

Un discurso aparte, pero directamente relacionado con la materia etimológica, merece el tratamiento de la polisemia y la homonimia en los diccionarios generales, dos tipos distintos de relaciones sincrónicas entre significantes y significados. En la polisemia, los significados están relacionados, tanto etimológica como semánticamente: eso significa que tienen una historia común y una afinidad semántica que depende del criterio lingüístico de la economía (cuando hablamos asociamos al mismo lema significados similares a los originarios). En la homonimia, en cambio, la igualdad entre dos palabras se debe a una serie de controversias etimológicas, es decir, por una serie de fenómenos y acontecimientos los dos términos han asumido la misma grafía o el mismo sonido, a pesar de seguir indicando referentes diferentes. En resumen, lo que distingue la homonimia de la polisemia es la etimología: el criterio etimológico reconoce que existe polisemia en los casos en que a distintos significados correspondan significantes idénticos que, desde el punto de vista diacrónico, tienen un idéntico origen; se trata de homonimia, al contrario, cuando a diferentes significados corresponden significantes idénticos que, desde el punto de vista diacrónico, se remontan a significantes diferentes.

4. CONCLUSIONES

El trabajo que hasta aquí se ha desarrollado pone de manifiesto que la presencia de la información etimológica en los diccionarios generales de uso no coloca a la etimología en su natural perspectiva diacrónica, recorriendo la historia de la palabra, sino que la analiza desde el punto de vista sincrónico, es decir, estudiando los elementos constitutivos y los principios fundamentales de la lengua en un determinado momento. Para que la investigación etimológica se realice de manera satisfactoria, la documentación debería estar dirigida a averiguar cuándo se encuentra una palabra por primera vez, dónde, en qué forma y con qué significado, y si su uso se ha visto interrumpido desde el momento de su primera aparición. Por este motivo, cada indagación etimológica científica no puede prescindir de la adquisición de toda la documentación disponible relativa a la palabra o a la familia de palabras en examen.

Es preciso señalar que estas informaciones específicas y detalladas solo se pueden recuperar gracias a la consulta de los diccionarios históricos y etimológicos, ya que no la reducen a un probable étimo entre paréntesis, sino que la analizan considerando varios criterios: la cronología, la evolución fonética y semántica, la relación con otras lenguas, la cantidad de datos que proporciona, o el ámbito de acción.

Además, en los diccionarios estudiados, la indicación etimológica obedece al concepto de etimología-origen, fuente de noticias sobre la raíz léxica, la procedencia y, a veces, la composición de la palabra, sin ofrecer aclaración alguna, de ahí que la etimología tenga un papel propiamente de consulta en estos tipos de diccionarios. Por eso, para que el comentario etimológico resulte útil, hay que ampliarlo significativamente. De hecho, la etimología, además de centrarse en el origen de las palabras, debería analizar también las variaciones de formas y de significados que estas sufren a lo largo del tiempo. Así, suponiendo que se conozca el étimo de una palabra, se podría operar no solo en dirección regresiva, es decir, desde la palabra al étimo, sino por vía progresiva, partiendo del étimo y llegando a los resultados modernos en diferentes lenguas. Conocer las dinámicas con las que las palabras cambian su forma y significado es el cometido fundamental para quienes quieran ocuparse de etimología. Una etimología sin el contexto que la ha generado no tiene mucho interés porque, aunque arroje luz parcialmente sobre el aspecto de la lingüística contrastiva, deja al lector completamente desorientado respecto a su uso.

En resumidas cuentas, ¿para qué necesitamos la información etimológica en los diccionarios generales?

La indicación etimológica en los diccionarios sincrónicos solo se puede justificar plenamente en caso de que ayude al usuario a comprender el significado o la formación de una palabra, a informarle sobre su uso, a mejorar la ortografía y a orientar sobre la distinción entre homonimia y polisemia. En caso contrario, habría una sobrecarga informativa que se podría emplear mejor para ampliar la microestructura de los artículos.

El objetivo, también de los diccionarios generales que acogen la etimología, debería ser el de remontarse a un pasado más remoto, plantearse reconstruir un pasado efectivamente muerto, situaciones propias de otros tiempos, con caracteres formales, fonéticos, morfológicos o valores semánticos de otros tiempos, para convertir la palabra en una fuente de historia. En caso contrario, el comentario etimológico seguirá siendo concebido como un elemento secundario dentro de la microestructura de los diccionarios generales de lengua.

REFERENCIAS BIBLIOGRÁFICAS

- BAGLIONI, D., 2016. *L'Etimologia*. Roma: Carocci.
- BUGUEÑO MIRANDA, F.V., 2004. «La etimología en el diccionario de la lengua», *Revista Letras* (Curitiba), 64, 2004: 173-188.
- MONTI, V., 1819. *Proposta di alcune correzioni ed aggiunte al Vocabolario della Crusca*. Milano: Fontana.
- REY-DEBOVE, J., 1971. *Étude linguistique et sémiotique des dictionnaires français contemporains*. Paris: Mouton.
- TORO Y GISBERT, M., 1909. *Enmiendas al Diccionario de la Academia*. París: Ollendorff.

DICCIONARIOS

- BATTISTI, C. y ALESSIO, G., 1950-57. *Dizionario etimologico italiano (DEI)*. Firenze: Barbera.
- BATTISTI, C. y ALESSIO, G., 1996. *Gran diccionario de la lengua española (GDLE)*. Barcelona: Larousse Planeta.
- BATTISTI, C. y ALESSIO, G., 1997. *CLAVE. Diccionario de uso del español actual*. Madrid: Ediciones SM.
- CORTELLAZZO, M., ZOLLI, P., 1979. *Dizionario etimologico della lingua italiana (DELI)*. Bologna: Zanichelli.
- DE MAURO, T., 1999-2000. *Grande dizionario italiano dell'uso (GRADIT)*. Torino: UTET.
- DEVOTO, G. y OLI, G.C., 1971. *Dizionario della lingua italiana*. Firenze: Le Monnier. Edición sucesiva de 2004.
- DEVOTO, G. y OLI, G.C., 1996. *Diccionario de la lengua española*. Barcelona: Larousse Planeta.
- DURO, A., 1986-1997. *Vocabolario della lingua italiana (VOLIT)*. Roma: Istituto dell'Enciclopedia Italiana.
- MENÉNDEZ PIDAL, R., 1945. «El diccionario que deseamos», prólogo a *VOX. Diccionario general ilustrado de la lengua española (DGILE)*. Barcelona:

Bibliograf.

- MOLINER, M., 1966. *DUE, Diccionario de uso del español*, [1.^a ed.]. Madrid: Gredos. Ediciones posteriores en 1998, 2007.
- NOCENTINI, A., con la collaborazione di Parenti, A., 2010. *l'Etimologico*. Milano: Le Monnier.
- PFISTER, M., 1979. *Lessico etimologico italiano (LEI)*. Mainz: Akademie der Wissenschaften und der Literatur.
- PRATI, A., 1951. *Vocabolario etimologico italiano*. Torino: Garzanti.
- REAL ACADEMIA ESPAÑOLA, (172639 [2013]): *Diccionario de la lengua castellana en que se explica el verdadero sentido de las voces, su naturaleza y calidad, con las frases o modos de hablar, los proverbios o refranes y otras cosas convenientes al uso de la lengua*, Madrid: Imprenta de Francisco del Hierro y otras.
- REAL ACADEMIA ESPAÑOLA, 2010. *Ortografía de la lengua española*. Madrid: Espasa Libros.
- SABATINI, F., COLETTI, V., 1997. *Dizionario italiano Sabatini (DISC)*. Coletti. Firenze: Giunti.
- SABATINI, F., COLETTI, V., *Vocabolario degli Accademici della Crusca. III impressione, 1691*. Firenze: Accademia della Crusca.
- SABATINI, F., COLETTI, V., *Vocabolario degli Accademici della Crusca. IV impressione, 1729-1738*. Firenze: Domenico Maria Manni.
- ZINGARELLI, N., 1922. *Vocabolario della lingua italiana*. Bologna: Zanichelli. Ediciones posteriores en 1970, 2016.

ENTORNO PERSONAL DE APRENDIZAJE Y PORTAFOLIOS ELECTRÓNICOS EN ESTUDIOS DE TRADUCCIÓN

Marcos Cánovas

Universitat de Vic - Universitat Central de Catalunya

RESUMEN

La incorporación del concepto de entorno personal de aprendizaje o PLE (del inglés personal learning environment) a los estudios de traducción favorece el aprendizaje cooperativo y los estudiantes tienen acceso a nuevos recursos profesionales. Nuestra propuesta unifica el PLE y el portafolios (en el que el estudiante presenta muestras significativas del progreso formativo y propone reflexiones en torno al proceso), por medio de la plataforma Mahara. Aquí presentamos los resultados de un estudio y un proyecto. En el estudio recogimos datos a partir del uso del entorno Mahara por parte de estudiantes de grado que no lo habían empleado antes. Los datos muestran un aumento significativo del acceso de los estudiantes a recursos electrónicos del ámbito de la traducción. El proyecto propone un modelo para un máster en traducción que combina los entornos abiertos Moodle y Mahara. Este modelo integra los recursos del portafolios, visibiliza los entornos personales de aprendizaje y potencia la cohesión de todo el máster de manera transversal.

Palabras clave: portafolios, entorno personal de aprendizaje, PLE, Mahara, formación de traductores

ABSTRACT

The incorporation of the Personal Learning Environment (PLE) concept into translation courses encourages cooperative learning and students gain access to new professional resources. Our proposal brings together PLE and portfolios (in which each student compiles relevant evidence of their personal progress and reflects on the learning process), through the Mahara platform.

Here we present the results of a study and a project proposal. In the study we collected data on first-time Mahara use by undergraduate students. The data shows a significant increase in student access to online translation resources. The project proposal is an implementation strategy for a master's degree in translation, combining the open source Moodle and Mahara platforms. This tandem integrates portfolio resources, making personal learning environments visible and enhancing coherence across the whole master's course.

Keywords: portfolio, personal learning environment, PLE, Mahara, translator training

1. ENTORNO PERSONAL DE APRENDIZAJE Y PORTAFOLIOS EN ESTUDIOS DE TRADUCCIÓN

La incorporación del concepto de entorno personal de aprendizaje o PLE (del inglés *personal learning environment*) a los estudios de Traducción lleva a algunas consideraciones que conectan directamente con competencias fundamentales en la formación de traductores.

Los medios electrónicos permiten hacer visibles recursos que no ofrece la institución académica y que el estudiante se procura porque le ofrecen información complementaria para su formación. Estos elementos constituyen el entorno personal de aprendizaje de los estudiantes. Hacer visibles estos entornos personales favorece la interacción entre los miembros del grupo, algo que es equivalente a la interacción en entornos profesionales reales. Como consecuencia de esta interacción se potencia el aprendizaje cooperativo y los estudiantes tienen acceso a nuevas fuentes y recursos que proporcionan otros miembros del grupo y que son determinantes para potenciar su autonomía y tomar decisiones adecuadas a la hora de resolver los retos que plantean los proyectos formativos.

El PLE se enlaza con el constructivismo social de Vigotsky (1978): los estudiantes construyen el conocimiento a partir de la reflexión y la interacción. La tecnología permite unos desarrollos específicos que han dado lugar a teorías como el conectivismo Siemens (2005): el aprendizaje surge de las conexiones entre diferentes fuentes. Igualmente, la teoría LaaN (Learning as a Network) (Chatti, Schroeder y Jarke, 2012; Chatti, 2013) añade al concepto del aprendizaje como conexión la consideración de los elementos vinculados al contexto social y la capacidad del estudiante para gestionar este entorno y obtener rendimiento en su formación.

En cualquier caso, la idea de PLE va más allá de las teorías concretas. Entra en juego el papel del docente como facilitador, de acuerdo con los presupuestos del socioconstructivismo, pero en este caso lo que facilita es la capacidad para establecer redes virtuales. El estudiante recombina las fuentes de información y los contactos, y la institución educativa le proporciona la plataforma donde puede recoger y plasmar toda esta información y compartirla.

Por otro lado, el entorno propuesto emula las circunstancias de la traducción profesional. Los profesionales están constantemente acudiendo a fuentes de información y documentación, y tienen algunas que son de consulta habitual. Igualmente, dialogan entre sí en foros o listas de correo para intercambiar

opiniones o resolver problemas específicos de terminología, estructuras sintácticas, equivalencias, etc. Y también hay en el mundo profesional revistas especializadas, blogs y publicaciones donde hay un espacio para la reflexión. De hecho, el mundo de la traducción profesional es inseparable de los entornos virtuales: los traductores trabajan con dispositivos electrónicos, se comunican con dispositivos electrónicos y se informan con dispositivos electrónicos.

La propuesta que se presenta vincula el concepto de entorno personal de aprendizaje con el de portafolios electrónico. El portafolios tiene un recorrido importante en la docencia, y los medios tecnológicos le han proporcionado mecanismos óptimos para la actualización digital. La convergencia entre PLE y portafolios electrónicos se refleja en estudios como los de Attwell (2007), Abrami et al. (2008), Barret y Garrett (2009), Cambridge (2009 y 2010), Drexler (2010), Tur y Urbina (2012), *Journal of Literacy and Technology* (2014).

Es la plataforma elegida se mostrarán fuentes de información, se reflexionará sobre esta información y se encontrarán herramientas de interacción entre los propios estudiantes o abiertas a Internet (v. Adell y Castañeda, 2013).

La plataforma de portafolios electrónicos que se ha usado es Mahara. Mahara facilita la convergencia a que nos hemos referido entre portafolios y PLE: permite visibilizar los resultados del aprendizaje, permite reflexionar (punto en común entre los portafolios y los PLE) y permite visibilizar recursos externos (en el campo de la traducción pueden ser glosarios, textos paralelos o especializados, artículos académicos, opiniones de profesionales, foros de debate, etc.). Teniendo presentes estas consideraciones, incorporamos a algunas materias del grado en Traducción e Interpretación de la Universidad, de forma experimental, el acceso a la plataforma Mahara desde las aulas Moodle de la institución. A partir de esta experiencia, elaboramos un modelo de implementación de la plataforma de manera transversal en un máster (el modelo está pensado para en Máster en Traducción Especializada de la Universidad, pero es aplicable a otros estudios).

2. ACCESO A PORTAFOLIOS ELECTRÓNICO EN ASIGNATURAS DE TRADUCCIÓN E INTERPRETACIÓN (GRADO)

2.1 Preparación y objetivos de la propuesta

Se trataba de que los estudiantes tuvieran acceso a un espacio propio en la plataforma Mahara, como complemento a las tareas e indicaciones que se encontraban en las aulas virtuales de la institución, basadas en Moodle.

De una manera sintética, apuntaremos las características de Mahara. Se trata de un espacio que tiene como eje fundamental las páginas web que genera el propio estudiante (llamadas “vistas”). Estas páginas las puede estructurar el estudiante de la manera que le parezca adecuada e insertar los contenidos que crea convenientes (incluyendo las páginas agrupadas o colecciones). Entre estos contenidos, destaca el llamado “diario” (en realidad, un blog, en terminología más convencional) en el que se plasman las reflexiones. Una vez creada una página, se puede elegir con quién se comparte (si se comparte): con todo el grupo, con alguna persona o personas en concreto, o se puede dejar abierta a Internet.

De esta manera, mientras el entorno Moodle está controlado principalmente por el profesorado, en Mahara son los propios estudiantes los que editan las páginas, deciden con quién las comparten, realizan las interacciones y seleccionan la información. Por lo tanto, las páginas creadas podían desempeñar funciones como:

- Gestionar la información.
- Compartir datos (lo que incluye, por ejemplo, enlazar otras páginas con novedades de la industria de la traducción).
- Interactuar con colegas para resolver aspectos vinculados a un encargo profesional.
- Reflexionar sobre aspectos específicos de la profesión y compartir las reflexiones en un entorno abierto.

Sin embargo, una plataforma que da tanta autonomía necesariamente requiere cierta orientación antes de ser utilizada a pleno rendimiento. De esta manera, dar formación a los estudiantes se convierte en un reto que hay que superar necesariamente si se quiere lograr que estos estudiantes operen de manera satisfactoria.

2.2 Formación a los estudiantes y desarrollo de la propuesta

La formación se desarrolló teniendo en cuenta aspectos generales (la plataforma) y específicos del proyecto. De entrada, fue fundamental considerar las modalidades de estudio, puesto que una parte de los estudiantes eran presenciales y otra parte estudiaba en modalidad online. A partir de aquí, se realizaron las acciones siguientes:

- Se creó una guía de uso de Mahara en formato de texto, adaptada específicamente al proyecto.
- Se realizó una sesión de videoconferencia sobre el entorno, dirigida a los estudiantes online, que quedó grabada como tutorial y disponible también para los estudiantes presenciales.
- Con los estudiantes presenciales se realizó una sesión de formación en clase.

Al mismo tiempo, y precisamente por la gran cantidad de opciones que ofrece Mahara, se optó por ofrecer una página-modelo que fue el punto de partida para cada estudiante. La página se copiaba y después se personalizaba de acuerdo con las preferencias de cada persona. De entrada, mostraba tres columnas: una orientada a la recopilación de información, otra a la reflexión de los estudiantes (a partir de un diario) y, la tercera, a la interacción (por ejemplo, se podían enlazar foros externos sobre temas de traducción).

Los estudiantes desarrollaron sus páginas durante un semestre y pudieron visibilizar y comentar las páginas de las demás personas del grupo. En paralelo, al comienzo y al final del semestre rellenaron un cuestionario.

2.3 Cuestionarios: datos obtenidos y discusión

El objetivo de los cuestionarios fue recoger la percepción de los estudiantes respecto al uso que hacían de recursos virtuales relacionados con el objeto de su formación, la traducción (recursos susceptibles de ser visibilizados en las páginas compartidas de Mahara). Se preguntaba, por ejemplo, sobre la consulta de diccionarios en línea, consulta de sitios especializados, comunicación de voz sobre IP (como Skype) o uso de redes sociales, siempre desde la perspectiva de la profesión. El cuestionario inicial lo respondieron 40 estudiantes y el final, 27. Tenían que indicar si consideraban que su acceso a determinados recursos

era inexistente, poco frecuente, frecuente o muy frecuente. La última pregunta era sobre el uso global de los recursos en línea. Los resultados, comparando ambos cuestionarios y agrupando las respuestas frecuente y muy frecuente, determinaron los siguientes porcentajes:

	Cuestionario inicial	Cuestionario final
	% Frecuente o muy frecuente	% Frecuente o muy frecuente
Diccionarios	80	100
Sitios especializados	83	75
Artículos especializados	69	75
Redes sociales	49	71
Wiki	63	64
Voz sobre IP	40	57
Portafolios	3	50
Fotos	31	43
Blogs	34	39
Microblogging	29	36
Foros	34	36
RSS	14	21
Uso global	89	100

Tabla 1

Por otra parte, la pregunta de cuáles son los recursos preferidos, desde la perspectiva de la resolución de encargos de traducción, dio lugar a los porcentajes de respuesta siguientes:

Cuestionario inicial	%	Cuestionario final	%
Wiki	46	Diccionarios	92
Diccionarios	74	Redes sociales (Facebook)	43
Microblogging (Twitter)	14	Voz sobre IP (Skype / Hangouts)	25
Blogs (Blogger)	20	Blogs (Blogger)	18
Voz sobre IP (Skype / Hangouts)	26	Google	18
Sitios web especializados	11	Wiki	14
		e-Portafolios (Mahara)	14
		Microblogging (Twitter)	14

Tabla 2

Respecto a la tabla 1, de los datos recogidos se desprende un aumento significativo del acceso y uso de los estudiantes a recursos electrónicos relacionados con el ámbito de la traducción después de visibilizar y compartir con el resto del grupo los recursos de su entorno personal de aprendizaje. Excepto el acceso a sitios especializados, que disminuye, pero no de manera significativa (por un lado, quizás porque llega más información de otras fuentes, pero, al mismo tiempo, la menor frecuencia de uso se ve compensada por el incremento de los artículos especializados), todos los otros recursos ven incrementada su frecuencia. Destaca, lo que no sorprende, dadas las características de la experiencia, el incremento en el uso del portafolios. Se puede señalar, asimismo, el aumento del empleo de voz sobre IP o de redes sociales. Por otra parte, los diccionarios electrónicos en línea pasan a ser una preferencia unánime, y también el uso de estos recursos en general.

En lo que respecta a la tabla 2, los estudiantes tenían que señalar un máximo de cinco recursos preferidos. Entre el cuestionario inicial y el final se refleja el aumento que se ha comentado en la tabla 1, pero no necesariamente se establece una correlación entre los recursos más utilizados y los preferidos. Destaca, en todo caso, un desplazamiento de las preferencias entre el primer

y el segundo cuestionario: aparecen con fuerza en el segundo las redes sociales (centradas en Facebook), que no eran relevantes en el primer cuestionario, disminuye considerablemente la preferencia por los entornos wiki (Wikipedia), se puede suponer que compensada por el acceso a más recursos, y aumenta la preferencia por otras herramientas (como los diccionarios online). En el segundo cuestionario, aparece el buscador Google (de hecho, las búsquedas afinadas con Google son una poderosa ayuda para resolver problemas de traducción). Finalmente, la preferencia por otros recursos se mantiene más o menos igual antes y después de la experiencia con el portafolios (por ejemplo, la consulta de blogs o el uso de Twitter).

3. PROPUESTA DE IMPLEMENTACIÓN DEL PORTAFOLIOS EN UN MÁSTER EN TRADUCCIÓN

Después de la experiencia de la incorporación del portafolios a algunas aulas del grado, se propone un proyecto para implementar de manera generalizada en un máster en traducción un modelo de multiplataforma virtual que incluye los recursos abiertos Moodle y Mahara. Así, en las aulas virtuales institucionales basadas en Moodle se introduce como complemento el acceso a Mahara. Se pretende que la plataforma Mahara funcione de manera transversal: los contenidos no solo se comparten en las diferentes materias, sino que el estudiante tiene un repositorio de trabajos y recursos que facilitan una reflexión global sobre todo el proceso formativo. Esta reflexión tiene especial relevancia en determinados momentos, como en la realización de prácticas y, especialmente, del trabajo final de máster. Se trata, de una manera general, de potenciar la autonomía de los estudiantes con la vista puesta en el mundo profesional, y de reforzar la coherencia de los estudios y las conexiones entre las materias. De una forma más específica, se plantean los objetivos siguientes:

- Presentar muestras seleccionadas y significativas del proceso formativo, junto con una reflexión sobre este aprendizaje.
- Conectar con los recursos del entorno personal de aprendizaje de los estudiantes (redes sociales, blogs, sitios web especializados, etc.), que se harán visibles en el contexto educativo.
- Realizar el seguimiento de las prácticas.
- Realizar el seguimiento del trabajo final de máster.
- Conectar las materias del máster (por ejemplo, con evaluaciones

compartidas).

- Preparar un portafolios orientado a la inserción laboral.

Por lo tanto, además de recoger los aspectos que se han comentado en el apartado anterior (formación de estudiantes y páginas-modelo que se despliegan en las asignaturas), hay que tener en cuenta cuestiones adicionales:

- La formación se tiene que extender al profesorado, por medio de documentos de texto y vídeos tutoriales. Ya no se trata de profesorado implicado en el proyecto que trabaja en asignaturas concretas y que conoce el recurso Mahara, sino de profesorado que usa por primera vez la plataforma.
- Además de la página-modelo para las asignaturas, se deben determinar páginas que sirvan de base para desarrollar las prácticas y el trabajo final de máster. A partir estas páginas, los contenidos se pueden orientar a un portafolios profesional. Por otra parte, el diario será transversal en todo el máster.
- Se proponen diferentes niveles de implicación de las asignaturas, con la recomendación de que cada asignatura cree un grupo de Mahara y proponga actividades que supongan compartir páginas con información y reflexión en el propio grupo-clase.

A partir de todos estos aspectos, se puede comparar la estructura de un máster convencional (figura 1) con la estructura que se describe aquí, que incluye el modelo de portafolios transversal (figura 2). En el primer caso, la combinación una serie de asignaturas obligatorias y optativas puede llevar a la consecución de determinadas especialidades (dependiendo del máster). Paralelamente, se pueden realizar prácticas y las competencias adquiridas en todo el proceso se reflejan en la elaboración del trabajo final de máster. En todo caso, este modelo no asegura la cohesión entre las diversas asignaturas o la incorporación de tareas o contenidos compartidos:



Figura 1. Estructura estándar de un máster

Por su parte, el portafolios transversal introduce una herramienta que establece conexiones entre las asignaturas y que facilita compartir contenidos dentro de una materia y entre materias diferentes. Asimismo, facilita igualmente que la información recogida en asignaturas y en prácticas se reoriente hacia el trabajo final de máster y a la generación de un portafolios profesional:

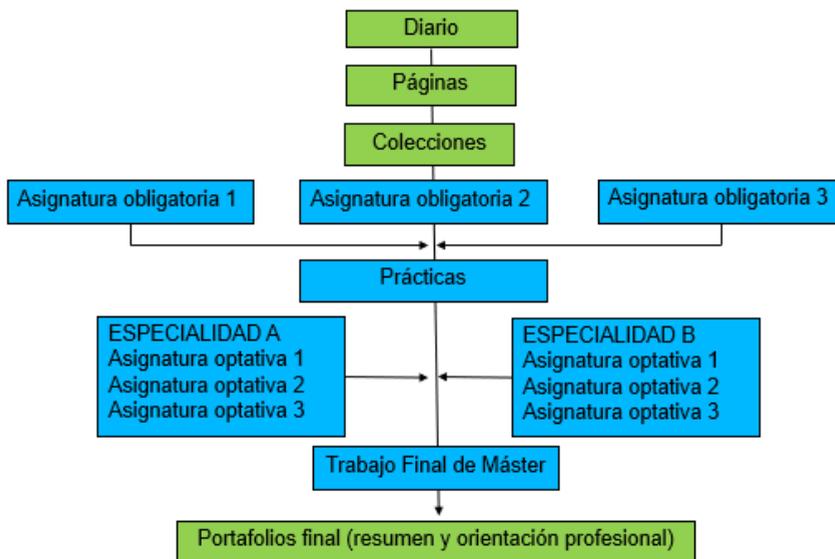


Figura 2. Estructura de un máster con portafolios transversal

4. CONCLUSIÓN

Uno de los retos de la formación de traductores es la orientación hacia la realidad de la industria de la traducción profesional. Para ello se deben tener en cuenta herramientas especializadas de traducción, pero también elementos de interacción y documentación relevantes para resolver los encargos profesionales.

El portafolios electrónico, que recoge muestras del aprendizaje y reflexión, y que al mismo tiempo permite compartir recursos e información en el contexto de un grupo-clase o más allá, constituye la base conceptual que facilita cumplir estos objetivos. En todo caso, toma sentido pleno si se desarrolla de manera general en todo un ciclo de estudios, como aquí se propone en el caso de un máster en traducción. Cuando es así, puede contribuir de manera decisiva a cohesionar todo el ciclo y reforzar la trayectoria de los estudiantes, así como la visibilización de unos objetivos globales y compartidos en todo el contexto de los estudios.

No se pueden obviar, en todo caso, las dificultades. En primer lugar, la tendencia tradicional a la fragmentación de los estudios en asignaturas aisladas, contraria a la permeabilidad de contenidos y actividades que se propone aquí. Esta tendencia a la fragmentación se puede reflejar, por ejemplo, en opiniones del profesorado: no solo se tratará de formarlo sino, en ocasiones, de convencerlo. Los propios alumnos vienen igualmente de esta tradición, por lo que también será conveniente reorientar sus hábitos. Finalmente, está la parte puramente tecnológica: casi todas las instituciones educativas tienen campus virtual con aulas de Moodle o de otro tipo, pero no es habitual que se disponga de Mahara, y menos integrado al sistema institucional, de manera que, por ejemplo, no se tengan que gestionar manualmente los registros de los estudiantes.

Todos estos elementos nos sitúan ante un modelo que puede dar la vuelta a una visión tradicional de la enseñanza universitaria pero, que, precisamente por eso, se tiene que implementar con convicción, unos recursos mínimos y todo el apoyo institucional.

REFERENCIAS BIBLIOGRÁFICAS

- ABRAMI, P., WADE, A., PILLAY, V., ASLAN, O., BURES, E., BENTLEY y C. 2008. "Encouraging self-regulated learning through electronic portfolios". *Canadian Journal of Learning and Technology*, 24 (3).
- ATTWELL, G. 2007. "E-Portfolios- the DNA of the Personal Learning Environment?" *Journal of eLearning and Knowledge Society*, 3 (2): 39-61.
- BARRETT, H; GARRETT y N. 2009. "Online Personal Learning Environments: Structuring Electronic Portfolios for Lifelong and Life Wide Learning", *On the Horizon*, 17 (2): 142-152.
- CAMBRIDGE, D. 2009. "Two faces of integrative learning online". *Electronic Portfolios 2.0. Emergent Research on Implementation and Impact*, Cambridge, D.; Cambridge, B. y Yancey, K. (eds.). Virginia: Stylus. 41-49.
- CAMBRIDGE, D. 2010: *Eportfolios for Lifelong Learning and Assessment*. San Francisco: JosseyBass.
- CASTAÑEDA, L. y ADELL, J. (eds.). 2013. *Entornos Personales de Aprendizaje: claves para el ecosistema educativo en red*. Alcoy: Marfil.
- CHATTI, M., SCHROEDER, U. y M. JARKE y M. 2012. "LaaN: Convergence of Knowledge Management and Technology-Enhanced Learning". *IEEE Transactions on Learning Technologies*, 5 (2).
- DREXLER, W. 2010. "The networked student model for construction of personal learning environments: Balancing teacher control and student autonomy", *Australasian Journal of Educational Technology*, 26 (3): 369-385
- JOURNAL OF LITERACY AND TECHNOLOGY. SPECIAL EDITION: Personal Learning Environments: Current Research and Emerging Practice*. 2014. Volume 15, Number 2.
- SIEMENS, G. 2005. "Connectivism: A learning theory for the digital age", *International Journal of Instructional Technology and Distance Learning*, 2 (1): 3-10.
- TUR, G. y URBINA, S. 2012. "PLE-based ePortfolios: Towards Empowering Student Teachers' PLEs through ePortfolio Processes". *Proceedings of the The PLE Conference 2012*.
- VYGOTSKY, L. 1978. *Mind in Society*. London: Harvard University Press.

UNA APROXIMACIÓN A LA TRADUCCIÓN DE LA VARIACIÓN LINGÜÍSTICA EN *GOMORRA – LA SERIE*

Matias Vedaschi

Laura Cruz García

Universidad de Las Palmas de Gran Canaria

RESUMEN

El estudio que aquí se ofrece se centra en la variación lingüística presente en la versión original de “Gomorra – La serie” y su traducción a otras lenguas. La mayor parte de los diálogos de la serie se expresan en dialecto napolitano. Para llegar al público de todo el país, ha sido subtitulada en su versión italiana. Precisamente, la peculiaridad de utilizar el dialecto como lengua vehicular de la serie ha despertado nuestras dudas sobre la transmisión de la realidad y de las temáticas que se representan en la serie hacia culturas diferentes, que son, en este caso, la española y la británica. El objetivo principal, por tanto, es determinar de qué forma se refleja la variación lingüística de dicha serie televisiva en su traducción al español y al inglés. Para ello, hemos analizado cuatro capítulos y ofreceremos los resultados de un estudio contrastivo, que mostrarán cuáles han sido las opciones elegidas para transmitir dicha variación lingüística en su versión doblada al español y subtitulada al inglés desde un enfoque teórico-práctico.

Palabras clave: variación lingüística, análisis, subtitulado, dialecto

ABSTRACT

This study focuses on the linguistic variation found in the original version of Gomorra: The Series and its translation into other languages. The majority of the series' dialogue is conducted in the Neapolitan dialect, with the Italian version subtitled to enable a country-wide audience to be reached. It is precisely this peculiarity, of employing a dialect as the vehicular language of the series, which led the authors of this study to question how the reality of the situation and the themes portrayed in the series are conveyed to different cultures, specifically Spanish and British cultures. Therefore, the main objective of this study is to determine the way in which the linguistic variation within this series is reflected in its Spanish and English translations. Consequently, four episodes have been analysed and the results of the contrastive study are presented. From a theoretical and practical perspective, these results illustrate the strategies selected in order to convey the aforementioned linguistic variation in the dubbed Spanish version and subtitled English version.

Keywords: language variation, analysis, subtitling, dialect

1. INTRODUCCIÓN

Con este trabajo pretendemos determinar de qué forma se traslada la variación lingüística de *Gomorra – La serie*, donde predomina el dialecto napolitano, a los distintivos textos meta que se contemplan en este estudio: 1) italiano estándar en la versión subtitulada, que ha sido emitida en toda Italia a través del canal de pago Sky Italia y en los DVDs disponibles en el país, 2) español en la versión doblada para España y 3) inglés en la versión subtitulada que encontramos en los DVDs disponibles en Reino Unido y Estados Unidos.

La peculiaridad de utilizar el dialecto como lengua vehicular de la serie ha despertado nuestras dudas sobre la transmisión de la realidad y de las temáticas que se representan en la serie hacia culturas diferentes, en este caso la española y la británica.

El dialecto, según Marcató (2002), De Mauro y Lodi (1993), es un sistema autónomo con respecto a la lengua estándar utilizado a nivel local. Así, Pellegrini (1997) identifica cinco sistemas lingüísticos en Italia: el ladino, las variedades de los dialectos alto-italianos, las variedades de los dialectos toscanos, las variedades de los dialectos centro-meridionales y el sardo. El dialecto empleado en la serie objeto de este estudio es el dialecto centro-meridional.

Tal y como afirman Grassi, Sobrero y Telmon (1998), es necesario tener en cuenta que los dialectos no son sistemas cerrados, sino que varían; presentan siempre un grado de variabilidad de tipo diatópico (existen variantes en la forma de hablar una misma lengua debido a la distancia geográfica que separa a los hablantes), diacrónico (cambios a lo largo de un periodo extenso de tiempo), diafásico (cambio de registro) y diastrático (caracterizado por el nivel social o cultural del hablante). Todos estos elementos presentan una dificultad añadida para la traducción, puesto que establecer equivalencias entre lenguas/culturas distintas, cuando se trata de aspectos propios de determinadas variedades lingüísticas, no siempre da lugar a una traducción adecuada.

2. ANALISIS DE LAS ESTRATEGIAS TRADUCTOLÓGICAS EMPLEADAS

Gomorra – La serie se estrenó en mayo de 2014 en la plataforma de satélite Sky Italia, en su canal Sky Atlantic, y está basada en la novela *Gomorra*, de Roberto Saviano (2006). Trata sobre la vida de la familia mafiosa Savastano, de sus

enemigos, en particular, de Salvatore Conte y de cómo Pietro Savastano, el patriarca, encarga a su mano derecha, Ciro Di Marzio, que prepare a su hijo, Gennaro Savastano, para convertirse en el nuevo *Capo della famiglia*, es decir, el nuevo jefe de los barrios del norte de Nápoles y de la familia Savastano.

Para el análisis se ha empleado una muestra de cuatro capítulos de la serie, concretamente, los capítulos 1, 2, 3 y 4 de la primera temporada, que se emitieron en Italia en 2014. A modo de estudio preliminar, nos hemos centrado en cómo se han trasladado al español y al inglés algunas expresiones peculiares típicas del área del sur de Italia.

Para el estudio, se ha tomado como base la clasificación de estrategias de Martí Ferriol (2013: 119-122), de la cual han sido de utilidad para nuestro análisis las siguientes:

- a) Traducción literal: la traducción representa el texto origen, pero el número de las palabras y el orden de la frase del texto origen no coinciden con los del texto meta.
- b) Amplificación: para que el receptor meta interprete correctamente el referente cultural, el traductor introduce en el texto meta explicaciones o información extra que no está presente en el texto origen.
- c) Omisión: el traductor omite por completo en el texto meta algún elemento presente en el texto origen. Se suele emplear para solventar dificultades derivadas de la velocidad de elocución o incapacidad de encontrar un equivalente adecuado.
- d) Adaptación cultural: el traductor reemplaza un elemento cultural del texto origen por otro término de la cultura meta que cumpla con una función similar o igual.
- e) Préstamo: el traductor integra una palabra o una expresión en el texto meta sin modificarla.
- f) Compresión: el traductor sintetiza los elementos lingüísticos del texto origen.

Para cada elemento analizado se ofrecen tablas que muestran su uso en la versión original y en las distintas versiones meta.

Las tablas que reflejan los diálogos y los subtítulos están divididas en cuatro filas, en las cuales hemos utilizado 4 acrónimos:

TO: texto en versión original en napolitano

TM1: versión con subtítulos en italiano estándar

TM2: versión doblada al español

TM3: versión con subtítulos en inglés

En el caso de las versiones subtituladas en italiano estándar e inglés, para reflejar cómo aparecen segmentados los subtítulos, hemos utilizado dos símbolos:

/ cuando el subtítulo aparece en la pantalla en dos líneas;

// cuando el subtítulo termina y empieza uno nuevo.

2.1 El vocablo “Zio”

Describiremos tres ejemplos diferentes del uso de esta expresión hallados en los capítulos 2 y 3 de la serie. El primer ejemplo se encuentra en el capítulo 2, que lleva por título “Ti fidi di me?”¹. El contexto en el que se emplea es el siguiente: la policía está inspeccionando la mercancía, unas latas de piña, de un buque que acaba de atracar en el puerto de Nápoles. Pocos segundos después, se observa que las fuerzas armadas encuentran droga en las latas. Zecchinetta, un miembro del clan Savastano, llama a Pietro para informarle de lo ocurrido.

TO

Zecchinetta: **Zio**, abbiamo un problema con mamma.

TM1

Zecchinetta: **Zio**, abbiamo un problema con mamma.

¹En la versión española se empleó la expresión “¿Confías en mí?” y en la versión inglesa, “Do you trust me?”.

TM2

Zecchinetta: Don P, tenemos un problema con mamá.

TM3

Zecchinetta: Uncle, // here we have a problem with Mom.

La palabra *zio*, además de su significado habitual “hermano de padre o de madre”, de acuerdo con el *GRADIT – Grande Dizionario Italiano dell’uso*², se emplea en las regiones meridionales de Italia como expresión de respeto hacia un cura o hacia un familiar mayor con el cual se tiene mucha diferencia de edad. En muchos casos se convierte de manera abreviada en *zi*³.

En italiano estándar se ha optado por un préstamo, manteniendo el vocablo en la versión original; en la versión doblada al español se ha optado por una amplificación que clarifica el TO en el sentido de que reproduce el tono de respeto de la expresión original, mientras que en inglés el vocablo “zio” se ha traducido de manera literal como “uncle”, alejándose con ello del sentido del original.

El segundo ejemplo se encuentra en el episodio 3, que lleva por título “L’omm ‘e casa”³. Don Pietro está en la cárcel, porque al detenerlo por conducción temeraria la policía ha encontrado, en su coche, una bolsa llena de dinero y, en su chaqueta, una dosis de cocaína, ambas robadas a Bolletta, miembro de su clan. Gracias a un policía corrupto, Don Pietro logra obtener un móvil y llama a Ciro y su clan para preguntarles cómo van los negocios.

²<http://www.accademiadellacrusca.it/en/reference-bibliographic/dictionaries/gradit-grande-dizionario-italiano-delluso>.

³En la versión española “El hombre de casa” y en la versión inglesa “The Mother and Home”.

TO

Pietro: 'Na passeggiata di salute. L'appalto del terz' munno comm' è iuto, tutt' buono?

Zecchinetta: Tutt' a post, **Zí.** De Rosa vi ringrazia assaje.

Pietro: [...] O fum'?

Zennichetta: Oh **Zí,** lu canale co' di Vaio e Carluccio si è aperto, ma chist' stanno cagato sotto. [...]

Malammore: Statemi buon', **Zio.**

TM1

Pietro: una passeggiata di salute. L'appalto del terzo mondo, com'è andato, bene?

Zecchinetta: Tutto a posto, **Zio.** De Rosa vi ringrazia molto.

Pietro: [...] Il fumo?

Zennichetta: **Zio,** il canale con / Di Vaio e Carluccio si è aperto, // ma si stanno cagando sotto. [...]

Malammore: Statemi bene, **Zio.**

TM2

Pietro: Como en unas vacaciones. Las licitaciones en Tercer Mundo, ¿qué tal? ¿Todo en orden?

Zecchinetta: Todo en orden, **zio.** De Rosa os da las gracias.

Pietro: [...] ¿Y el costo?

Zennichetta: **Zio,** el canal con Di Vaio y Carluccio se ha abierto, pero ahí están bastante acojonados. [...]

Malammore: Cuidese, **zí.**

TM3

Pietro: A piece of cake. / And that third world contract?

Zecchinetta: It's all okay, De Rosa thanks you.

Pietro: [...] Smoke?

Zennichetta: A channel's opened with Di Vaio and Carluccio, / but they are shittin' bricks. [...]

Zecchinetta: Take care.

Se observa, una vez más, que en la versión italiana estándar subtitulada se utiliza el préstamo. Lo mismo ocurre en la versión en español, a diferencia del ejemplo anterior, lo que hace que el vocablo pueda resultar incomprensible para los espectadores españoles. Por su parte, en la versión inglesa subtitulada se ha optado por omitir el vocablo. Un motivo muy probable sería que los miembros están hablando con Pietro a través del altavoz y al despedirse lo hacen cuatro hombres a la vez. En el doblaje es fácil reflejar la superposición de voces, pero en el subtitulado es muy difícil. Aquí la función principal de la comunicación se cumple, es decir, con *take care* el espectador inglés entiende que se están despidiendo.

El tercer ejemplo también se ha encontrado en el capítulo 3. En este caso, Don Pietro entra en su celda y todos los presos lo acogen con mucho respeto. Es aquí donde conoce a Pasqualino, un joven encarcelado por atraco y secuestro que está esperando su juicio. Pasqualino cuenta su historia.

<p>TO</p> <p>Pasqualino: Giudicabile pure io, Zio. Rapina e sequestro, a Mergellina. [...] Belle e buone sono arrivate le volanti per tutt' 'e parti, si sono portati pure l'elicottero, zio, si pareva nu film americano.</p>
<p>TM1</p> <p>Pasqualino: Sono giudicabile anch'io, Zio. // Rapina e sequestro, a Mergellina. [...] Belle e buone sono arrivate / le volanti da tutte le parti // si sono portati anche l'elicottero, / zio, sembrava un film americano.</p>
<p>TM2</p> <p>Pasqualino: Yo también estoy esperando el juicio. Robo y secuestro en Mergellina. [...] De golpe llegaron patrullas de todas partes, hasta vinieron en helicóptero. Parecía una peli americana.</p>
<p><u>Versión con subtítulos en inglés</u></p> <p>Pasqualino: I'm up too, uncle. // Robbery and kidnapping in a jewelry in Naples. / [...] Suddenly there were cops, helicopters, / just like an American movie.</p>

En la versión subtitulada al italiano estándar se utiliza el préstamo, mientras que, en el doblaje español, en este caso, se opta por la omisión y en los subtítulos en inglés se recurre de nuevo, erróneamente, a una traducción literal. En la versión italiana estándar subtitulada, el préstamo es la técnica ideal ya que todos los italianos entienden perfectamente que el uso de “zio” en este contexto es una demostración de respeto; en español, en este caso, el traductor ha optado por una omisión y en inglés por una traducción literal.

2.2 La expresión “Terzo Mondo”

En el capítulo 3, Pietro, desde la cárcel, está manteniendo una conversación telefónica con un miembro de su clan para saber cómo van los negocios.

La expresión “Terzo Mondo”, que literalmente significa “Tercer Mundo”, se refiere a un área de Nápoles que comprende los barrios marginales de Scampia, Miano, Piscinola, San Pietro a Paterno y Secondigliano.

TO

Pietro: ‘Na passeggiata di salute. L’appalto del terz’ munno comm’ è iuto, tutt’ buono?

Zecchinetta: Tutt’ a post, **Zí.** De Rosa vi ringrazia assaje.

TM1

Pietro: una passeggiata di salute. / L’appalto del terzo mondo, com’è andato, bene?

Zecchinetta: Tutto a posto, **Zio.** De Rosa vi ringrazia molto.

TM2

Pietro: Como en unas vacaciones. Las licitaciones en tercer mundo, ¿qué tal? ¿Todo en orden?

Zecchinetta: Todo en orden, **zio.** De Rosa os da las gracias.

TM3

Pietro: A piece of cake. / And that third world contract?

Zecchinetta: It’s all okay, De Rosa thanks you.

En todos los casos se ha optado por una traducción literal que puede llevar a confusión al espectador porque la familia Savastano está haciendo negocios con unos delincuentes senegaleses que también venden drogas. La expresión “tercer mundo” se podría relacionar erróneamente en este caso con los delincuentes senegaleses. Consideramos que se podría haber optado por una traducción explicativa, como podría ser “barrios bajos”, ya que es una expresión utilizada en un registro informal.

2.3 Las expresiones “Le Case dei Puffi” y “Sette Palazzi”

“Le case dei Puffi⁴” es un conjunto de bloques de viviendas que constituye el núcleo del tráfico de droga de la Camorra. Cada día acuden a este lugar decenas de drogodependientes para comprar sus dosis. Para la Camorra, el negocio de estos edificios supone una facturación de varios millones de euros al año. Se les denomina así porque al haber sido construidas abusivamente, los techos de las viviendas son más bajos de lo que establece la normativa urbanística y pueden vivir ahí solo personas de baja estatura.

Por su parte, los “Sette Palazzi⁵”, conocidos como *le vele*⁶ de Scampia, son edificios residenciales de forma triangular que recuerdan una vela latina. Tal y como afirma Banini (2013), *le vele* son el núcleo de la degradación y del tráfico de droga. Simbolizan el mal, la ilegalidad y la derrota del Estado Italiano.

Estas expresiones se emplean por primera vez en el capítulo 2:

TO

Bolletta: Malamò e Zecchinetta erano già assaje no? Pe' nu strunz come Parisi. E poi tenevo che fa. Dovevo andare alle **case de Puffi** e ai **sette palazzi** a pijà i sordi.

⁴Traducción literal: “Las casas de los pitufos”.

⁵Traducción literal: “Los Siete Edificios”.

⁶Traducción literal: “las velas”.

TM1

Bolletta: Malamò e Zecchinetta erano piu che sufficienti / per uno stronzo come Parisi, no? // E poi avevo da fare. // Dovevo andare alle **case de Puffi** / e ai **sette palazzi** a prendere i soldi.

TM2

Bolletta: Malamó y Zecchinetta se bastaban, ¿no? Para un mindundi como Parisi. Además, tenía que hacer. Quería ir a las **casas de los Pitufos en los siete edificios** a por el dinero.

TM3

Bolletta: Malamò and Zecchinetta were enough / for an asshole like Parisi. // Besides I was busy, / I had to go to the **Smurf house** and get the money.

En italiano estándar se mantienen estas expresiones a través del préstamo. En español se traducen literalmente, pero con un error de comprensión ya que se ubican las casas de los pitufos en los Siete Edificios, cuando en realidad son dos sitios diferentes.

En inglés se omiten los Siete Edificios y se refieren a la casa de los Pitufos en singular, como si fuera solo una casa. En este caso, la traducción literal para las casas de los pitufos cumple la función del TO ya que, si leemos la definición del verbo “smurf” en el *Urban Dictionary*, “Someone who smurfs is usually hired by drug runners to move a large amount of cash to launder it”, vemos que con esta acepción del término la traducción literal es adecuada y mantiene el rasgo cultural del TO.

2.4 El vocablo “Cazzimmoso”

De acuerdo con el *GRADIT: Grande Dizionario Italiano dell'uso*⁷, “cazzimmoso” es un vocablo propio del dialecto napolitano que describe una actitud firme y también agresiva, interpretada favorablemente como expresión de fuerza y personalidad. En italiano estándar no hay traducción ya que el término engloba un concepto. En el capítulo 4, “Sangre africano”⁸, Pietro está hablando con su abogado sobre la posibilidad de salir de la cárcel.

TO

Abogado: Pietro, io ho fatto di tutto, non c'è stato niente da fare. U giudice è **cazzimmoso**. Ma se ci pensa un altro po', decorrono i termini e ti devono pure scarcerare.

TM1

Abogado: Pietro, io ho fatto di tutto, / non c'è stato niente da fare. // Il giudice è “**cazzimmoso**”. // Ma se ci pensa un altro po', / decorrono i termini / e ti devono pure scarcerare.

TM2

Abogado: Yo he hecho lo posible, pero no ha habido manera. El juez **me da mala espina**, pero como se lo siga pensando se pasan de fecha y te tienen que excarcelar.

TM3

Abogado: Pietro, I did all could, / still nothing, // the judge is **smart**, // but they can't hold you much longer, / they'll have to release you.

⁷<http://www.accademiadellacrusca.it/en/reference-bibliographic/dictionaries/gradit-grande-dizionario-italiano-delluso>.

⁸En la versión española nos encontramos con “Sangre africana” y en la versión inglesa, con “African Blood”.

Gracias al contexto, se entiende que el juez no es sobornable. En italiano estándar se ha utilizado un préstamo. En español se ha omitido el adjetivo para utilizar una amplificación a través de la cual, a pesar de no comunicar lo mismo que el adjetivo del TO, el espectador español entiende que el juez no es sobornable.

En la versión subtitulada en inglés se ha utilizado la técnica de compresión para poder explicar el adjetivo, pero este cambio no reproduce fielmente el sentido del TO.

3. CONCLUSIONES

Entendemos que trasladar las expresiones peculiares típicas del sur de Italia descritas en este estudio preliminar no es tarea simple para los traductores. A la hora de enfrentarse a la traducción de la variación lingüística, un traductor tiene que llevar a cabo una labor importante de documentación que, lamentablemente, por cuestiones de tiempo dictadas por las productoras o las agencias de doblaje y subtítulo, no siempre se realiza adecuadamente.

La versión subtitulada al italiano estándar opta, de manera lógica, por la traducción palabra por palabra puesto que ningún italiano tendría, en principio, dificultades a la hora de entender las denominaciones peculiares analizadas. Estas expresiones están a la orden del día en los telediarios y en la prensa.

En inglés se ha optado por una traducción literal, lo que ayuda a los espectadores a entender un poco más la vida de los personajes; esto no ocurre en la versión doblada al español, donde se opta por utilizar préstamos, alejando al público del día a día de los protagonistas de la serie.

Como ya se ha señalado más arriba, en el estudio aquí presentado únicamente se han analizado las denominaciones peculiares del dialecto napolitano que hemos encontrado en los primeros cuatro capítulos de la primera temporada de *Gomorra – la serie*. Por un lado, resulta imprescindible ampliar la muestra de vocablos con el visionado de los 8 restantes capítulos con el fin de incrementar la representatividad del análisis.

Por otro lado, y con vistas a un desarrollo a medio o largo plazo de nuestra investigación, además de las denominaciones peculiares, nos centraremos en la traducción de los apodos de los miembros del clan mafioso protagonista de la serie y la importancia que éstos tienen en la cultura origen.

REFERENCIAS BIBLIOGRÁFICAS

- BANINI, T. 2013. *Identità Territoriali. Questioni, metodi, esperienze a confronto*. Milano: Franco Angeli.
- GRADIT: *Grande dizionario italiano dell'uso* [consultado el 01/2019 en <http://www.accademiadellacrusca.it/en/reference-bibliographic/dictionaries/gradit-grande-dizionario-italiano-delluso>].
- DE MAURO T., LODI M. 1993. *Lingua e dialetti*. Roma: Editori Riuniti.
- DIZIONARIO TRECCANI. [consultado el 12/2018 en <http://www.treccani.it/vocabolario/>].
- GRASSI C., SOBRERO A., TELMON T. 2003. *Introduzione alla dialettologia italiana*. Bari: Laterza.
- MARCATO C. 2002. *Dialecto, dialetti e italiano*. Bologna: Il Mulino.
- MARTÍ FERRIOL, J.L. 2013. *El método de traducción: Doblaje y subtitulación frente a frente*. Castelló de la Plana: Publicacions de la Universitat Jaume I.
- PELLEGRINI, G. B. 1977. *Profilo dei dialetti italiani. Carta dei dialetti d'Italia*. Pisa: Pacini Editore.
- SAVIANO, R. 2006. *Gomorra*. Milano: Mondadori.
- URBAN DICTIONARY [consultado el 12/2018 en <https://www.urbandictionary.com>].