



Con il patrocinio di



**CUADERNOS AISPI. Estudios de lenguas y literaturas hispánicas** es una revista científica con vocación internacional publicada por la Associazione Ispanisti Italiani, que acoge estudios sobre las lenguas, las culturas y las literaturas ibéricas e iberoamericanas. Se propone promover el debate académico entre investigadores, con la mayor profundidad científica, a partir de las tendencias emergentes del hispanismo internacional, tanto filológico-literario como lingüístico y traductológico.

**Cuadernos AISPI** tiene una periodicidad semestral. Cada número incluye una sección temática, alternativamente de Lengua o Traducción y de Literatura o Cultura, dedicada a temas de especial relevancia en las áreas interesadas, a cargo de un editor italiano y otro extranjero, de reconocido prestigio internacional.

La revista fue fundada en 2013 por la XIII Junta Directiva de la Asociación, compuesta por Pietro Taravacci, Maria Vittoria Calvi, Antonella Cancellier, Alessandro Cassol y Flavia Gherardi.

### **Associazione Ispanisti Italiani**

c/o Istituto Cervantes  
Via di Villa Albani 14/16  
00198 – Roma  
[www.aispi.it](http://www.aispi.it)

### **Direttrice**

Flavia Gherardi  
Dipartimento di Studi Umanistici (DSU)  
Università di Napoli Federico II

### **Progetto grafico e impaginazione**

LedizioniLediPublishing

© 2025 Associazione Ispanisti Italiani  
Registrazione presso il Tribunale di Milano n. 316 del 17/11/2015  
ISSN: 2283-981X

Edizione a stampa:  
© 2025 LedizioniLediPublishing  
Via Boselli 10, 20136 Milano - Italia

Informazioni sul catalogo, sulle ristampe e per acquistare copie cartacee ed elettroniche dell'editore: [www.ledizioni.it](http://www.ledizioni.it)  
Pubblicato secondo una licenza Creative Commons Attribuzione  
Non commerciale - Non opere derivate 4.0 Internazionale

# **CUADERNOS AISPI**

## **Estudios de lenguas y literaturas hispánicas**

Revista semestral de la Associazione Ispanisti Italiani

### **DIRECTORA / EDITOR IN CHIEF**

Flavia Gherardi (Università di Napoli "Federico II")

### **CONSEJO DE DIRECCIÓN / GOVERNING BOARD**

Giovanna Mapelli (Università di Milano)

Renata Londero (Università di Udine)

### **CONSEJO DE REDACCIÓN / EDITORIAL BOARD**

Sara Bani (Università "G. D'Annunzio" di Chieti-Pescara)

Federica Cappelli (Università di Pisa)

Alessandro Cassol (Università di Milano)

Florencio del Barrio de la Rosa (Università Ca' Foscari di Venezia)

Stefano Neri (Università di Verona)

Veronica Orazi (Università di Torino)

Gianluca Pontrandolfo (Università di Trieste)

Raffaella Tonin (Università di Bologna)

Katerina Vaiopoulos (Università di Udine)

### **SECRETARIOS DE REDACCIÓN**

Andrés Ortega (Università di Bergamo)

Francesca Coppola (Università de L'Aquila)

### **EDITORES DE RESEÑAS**

Felice Gambin (Università di Verona)

Luisa Chierichetti (Università di Bergamo)

### **COLABORADORES**

Giuliana Calabrese (Università di Milano)

Ida Grasso (Università della Calabria)

Sara Longobardi (Università Suor Orsola Benincasa di Napoli)

Angela Moro (Università di Pisa)

Valentina Paleari (Università di Milano)

## **CONSEJO CIENTÍFICO Y ASESOR / EDITORIAL ADVISORY BOARD**

Pedro Álvarez de Miranda (Universidad Autónoma de Madrid)  
Fausta Antonucci (Università di Roma Tre)  
Beata Baczyńska (Uniwersytet Wrocławski)  
Jean-François Botrel (Université Rennes 2)  
Antonio Briz (Universitat de València)  
Maria Vittoria Calvi (Università di Milano)  
Antonella Cancellier (Università di Padova)  
Giovanni Caravaggi (Università di Pavia)  
Rocío Elena Caravedo Barrios (Pontificia Universidad Católica de Perú)  
Ovidi Carbonell Cortés (Universidad de Salamanca)  
Anne Cayuela (Université de Grenoble Alpes)  
Guiomar Ciapuscio (Universidad de Buenos Aires)  
Enrico Di Pastena (Università di Pisa)  
Giuseppe Di Stefano (Università di Pisa)  
Ángela Di Tullio (Academia Argentina de Letras)  
Laura Dolfi (Università di Parma)  
Francisco Domínguez Matito (Universidad de La Rioja)  
Inés Fernández Ordóñez (Universidad Autónoma de Madrid)  
Teresa Ferrer Valls (Universitat de València)  
Isabel García Izquierdo (Universitat Jaume I – Castelló de la Plana)  
Antonio Gargano (Università di Napoli “Federico II”) †  
Giovanni Garofalo (Università di Bergamo)  
Folke Gernert (Universität Trier)  
David Gies (University of Virginia)  
Johannes Kabatek (Universität Zürich)  
Elena Liverani (Libera Università di Lingue e comunicazione IULM – Milano)  
María Antonia Martín Zorraquino (Universidad de Zaragoza)  
Francisco Matte Bon (Università degli Studi Internazionali di Roma – UNINT)  
José María Micó (Universitat Pompeu Fabra – Barcelona)  
Estrella Montolío Durán (Universitat de Barcelona)  
Gabriele Morelli (Università di Bergamo)  
Francisco Moreno Fernández (Universität Heidelberg)  
Veronica Orazi (Università Ca' Foscari di Venezia)  
Elide Pittarello (Università Ca' Foscari di Venezia)  
Marco Presotto (Università di Trento)  
José María Pozuelo Yvancos (Universidad de Murcia)  
Claudio Rodríguez Fer (Universidad de Santiago de Compostela)  
José Carlos Rovira Soler (Universidad de Alicante)  
Enrique Rubio Cremades (Universidad de Alicante)  
Aldo Ruffinatto (Università di Torino)  
Maria Caterina Ruta (Università di Palermo)  
Félix San Vicente (Università di Bologna – Forlì)  
Laura Silvestri (Università di Roma Tor Vergata)  
Pietro Taravacci (Università di Trento)  
Fernando Valls (Universitat Autònoma de Barcelona)

## **Diseño y maquetación**

Sarah Tafuro

---

## SECCIÓN MONOGRÁFICA

*Valorización del patrimonio lexicográfico hispánico y retrodigitalización de diccionarios*

### EDITORES

Carmen Castillo Peña

Alejandro Fajardo Aguirre

Introducción	9
<b>Amparo Alemany Martínez, Matteo De Beni, Alessandra La Manna,</b> La digitalización del <i>Vocabolario italiano e spagnolo (1620)</i> de Lorenzo Franciosini: retos metodológicos y nuevas oportunidades de estudio	19
<b>Florencia Ferrante, Chiara Valente, Ana Lourdes de Hériz,</b> El lemario italiano-español del <i>Diccionario de faltriquera</i> de Cormon y Manni (1805): transcripción automática y claves ecdóticas	45
<b>Maria Felisa Bermejo Calleja, Giulia Lanteri,</b> <b>María Joaquina Valero Gisbert, Maria Carmela Zaccone,</b> La transcripción digital del diccionario de L. Ambrozzi	73
<b>Hugo E. Lombardini,</b> El espacio de lo femenino en el área lexicográfica de <i>L'italiano istruito...</i> de F. Marín	97
<b>Natalia Peñín Fernández,</b> Estudios preliminares para la retrodigitalización del <i>Vocabulario Espanyol, Italiano y Tudesco</i> de Ignacio de Boria	119
<b>Cristian Díaz Rodríguez,</b> El <i>Diccionario de voces americanas</i> de Ayala (1777) en el <i>TLEAM</i> : sacrificar en aras de homogeneizar	139
<b>Francesca Dalle Pezze, Elisa Sartor,</b> El <i>Vocabolario italiano e spagnolo (1620)</i> de Lorenzo Franciosini: representación del discurso repetido para su digitalización	157
<b>Carmen Castillo Peña,</b> La digitalización de nomenclaturas bilingües español-italiano	173

## SECCIÓN GENERAL

<b>Maria Vittoria Ambrosini</b> , <i>¿Debo hacerlo o Tengo que hacerlo?</i> Reflexionar sobre la polisemia modal desde una perspectiva interlingüística	219
<b>Sonia Bailini</b> , <i>I diari della motocicletta</i> , un viaggio nella traduzione intersemiotica e interlinguistica	245
<b>Elena Liverani</b> , Lanzarote e Saramago nelle parole di una traduttrice d'eccezione: <i>La intuición de la isla</i> (2022) di Pilar del Río e i problemi di resa in lingua italiana	261
<b>Marisa Martínez Pérsico</b> , <i>Embodiment</i> y metáforas somáticas en la autotraducción bilingüe. El caso de la poeta vasca Miren Agur Meabe desde el enfoque de los <i>Cognitive Translation Studies</i>	277

## RESEÑAS CRÍTICAS

Sonia Bailini, Maria Vittoria Calvi, Elena Liverani (eds.); Roberto Ortí Teruel (coord.), <i>El español como lengua de mediación en contextos educativos y profesionales</i> (Alba Segura Lores)	229
Ana I. Codesido García, Carlos Hernández-Sacristán, Victoria Marrero-Aguiar (eds.), <i>Lingüística clínica en el ámbito hispánico:</i> <i>un panorama de estudios</i> (Alessandra Donati)	302
Veljka Ruzicka Kenfel, Raffaella Tonin, <i>Aproximación poliédrica</i> <i>a la cuestión del género en la literatura infantil y juvenil</i> (Chiara Albertazzi)	306
Natasha Leal Rivas, <i>La argumentación en la escritura académica</i> <i>de investigación</i> (Francisco José Sánchez García)	309
Nuria Pérez Vicente, <i>Los culturemas.</i> <i>Aprender a traducir con la literatura infantil y juvenil (LIJ)</i> (Rocío García Jiménez)	312
Isabel Santos Gargallo; Susana Pastor Cesteros (dirs.), <i>Metodología</i> <i>de la investigación en la enseñanza-aprendizaje del español</i> <i>como segunda lengua (L2)/ lengua extranjera (LE)</i> (Rosa María García Jiménez)	315
Milín Bonomi, Maria Vittoria Calvi, Marcella Uberti-Bona (eds.), <i>Demolingüística</i> <i>del español en Italia, con un anexo sobre Malta, San Marino</i> <i>y Ciudad del Vaticano</i> (Sonia Bailini)	318

---

# SECCIÓN MONOGRÁFICA

*Valorización del patrimonio lexicográfico hispánico y retrodigitalización de diccionarios*

## INTRODUCCIÓN

Carmen Castillo Peña

Alejandro Fajardo Aguirre

**Amparo Alemany Martínez, Matteo De Beni, Alessandra La Manna,** *La digitalización del Vocabolario italiano e spagnolo (1620)* de Lorenzo

Franciosini: retos metodológicos y nuevas oportunidades de estudio

**Florencia Ferrante, Chiara Valente, Ana Lourdes de Hériz,** El lecionario italiano-español del *Diccionario de faltriquera* de Cormon y Manni

(1805): transcripción automática y claves ecdóticas

**Maria Felisa Bermejo Calleja, Giulia Lanteri,**

**María Joaquina Valero Gisbert, Maria Carmela Zaccone,** La transcripción digital del diccionario de L. Ambruzzi

**Hugo E. Lombardini,** El espacio de lo femenino en el área lexicográfica de *L'italiano istruito...* de F. Marín

**Natalia Peñín Fernández,** Estudios preliminares para la retrodigitalización del *Vocabulario Espanyol, Italiano y Tudesco* de Ignacio de Boria

**Cristian Díaz Rodríguez,** El *Diccionario de voces americanas* de Ayala (1777) en el *TLEAM*: sacrificar en aras de homogeneizar

**Francesca Dalle Pezze, Elisa Sartor,** El *Vocabolario italiano e spagnolo* (1620) de Lorenzo Franciosini: representación del discurso repetido para su digitalización

**Carmen Castillo Peña,** La digitalización de nomenclaturas bilingües español-italiano





# INTRODUCCIÓN

**Carmen Castillo Peña, Università degli Studi di Padova**  
**Alejandro Fajardo Aguirre, Universidad de La Laguna**

A Felisa Bermejo, *in memoriam*<sup>1</sup>

El patrimonio lexicográfico, en su sentido más amplio, está constituido por todo texto que recoge unidades léxicas de una o más lenguas para dar cuenta de su significado, ya sea mediante definiciones, ejemplos o por medio de equivalencias intra o interlingüísticas. Esto implica que, junto a los diccionarios alfabéticos —monolingües, bilingües o multilingües, generales, terminológicos, de sinónimos o palabras afines, etimológicos, históricos, etc.—, forman también parte del patrimonio lexicográfico las nomenclaturas, los diccionarios conceptuales o cualquier listado léxico onomasiológico. Asimismo, se incluyen las listas de palabras que forman parte de los paratextos presentes en gramáticas y traducciones.

La preservación de este vasto y heterogéneo conjunto de textos es imperativa, pues los diccionarios —todos, desde los más importantes hasta los más humildes— son obras que conservan no solo el acervo léxico, sino también el patrimonio cultural, al ser testimonio de ideologías, visiones del mundo, conocimientos técnicos y científicos, así como denominaciones de objetos, especies, costumbres y hábitos de una comunidad en un momento histórico determinado.

Efectivamente, los diccionarios son productos metalingüísticos complejos, herramientas para la comprensión y el uso de las lenguas y, a la vez, expresión de la ideología dominante, de la cultura y de la historia de las comunidades lingüísticas. Además, la aproximación historiográfica a los textos lexicográficos permite observar la evolución del léxico y, junto a esta, la historia de los procesos y recursos metalingüísticos utilizados para representar el lenguaje, el devenir de las ideas sobre la lengua objeto y la trayectoria del discurso ideológico que inexorablemen-

---

1 Apenas concluidas las tareas de este monográfico, nos ha dejado Felisa. Consternados, los miembros del grupo TELEI, al que Felisa dedicó toda su pasión por los diccionarios, los amigos y compañeros que durante años hemos gozado de su saber, de su fina inteligencia, de su exquisita ironía y, sobre todo, de su gran humanidad, le rendimos homenaje.

te impregna el diccionario.

Al igual que ocurre en otros ámbitos relacionados con los bienes culturales, materiales o inmateriales, los avances tecnológicos relacionados con la digitalización han supuesto un gran impulso en la preservación de los diccionarios, abriendo el camino a una proficua línea de investigación basada en las metodologías de las humanidades digitales.

En este sentido, la investigación historiográfica en lexicografía se ha beneficiado de la posibilidad de acceder a textos menores o de difícil localización. Esto es posible gracias a las reproducciones digitalizadas de los fondos de las grandes bibliotecas, como, por ejemplo, la *Biblioteca Digital Hispánica* de la Biblioteca Nacional de España, o la *Biblioteca Digital* de la Real Academia Española, pero sobre todo, gracias a las bibliotecas virtuales que catalogan fondos digitales de múltiples procedencias, permitiendo la consulta —y en ocasiones el descubrimiento— de textos digitalizados que de otro modo estarían dispersos en la red, como es el caso de la magnífica *Biblioteca Virtual de la Filología Española* <<https://www.bvfe.es/es/>>, creada por Manuel Alvar Ezquerro y dirigida hoy por M.<sup>a</sup> Ángeles García Aranda o de la biblioteca de diccionarios del portal *Contrastiva* <<https://www.contrastiva.it>>, dirigido por Félix San Vicente.

Por otra parte, las humanidades digitales han transformado profundamente el propio concepto de diccionario, que ya no es solo un libro impreso, sino una base de datos, normalmente gratuita y de libre acceso, a la que se puede interrogar de varias formas y con diversas finalidades, adaptadas en cada ocasión a las necesidades específicas de distintos tipos de usuarios (Tarp 2019). Se trata de diccionarios proyectados como digitales, sin los vínculos macro y microestructurales que impone el papel. Con frecuencia, constituyen la única vía factible para proyectos que en versión impresa serían casi irrealizables: diccionarios de lenguas indígenas o en vías de extinción, de pares de lenguas con escasa representación, de nueva planta monolingües, terminológicos o para finalidades específicas, con nomenclaturas restringidas a un tipo particular de unidades léxicas o a variedades específicas. Son diccionarios de realización dilatada en el tiempo, que suelen estar sometidos a una actualización constante y que a menudo no cuentan con los recursos organizativos y económicos de una empresa comercial.

Sin pretensiones de agotar el listado posible, damos algunos ejemplos de diccionarios digitales del español, de distintas tipologías, alcances y finalidades:

- *Diccionario electrónico sáliba-español* (Estrada Ramírez) es un diccionario electrónico concebido para la “documentación de la lengua y la cultura” y la “protección del patrimonio idiomático” impulsado por el Instituto Caro y Cuervo.

- *Diccionario digital del español (DIDES)* (Fuertes Olivera), monolingüe, permanentemente actualizado, descrito por su autor como un “repositorio dinámico de datos lexicográficos (está sometido a un proceso de cambio constante)” y proyectado para el uso de la Inteligencia Artificial Generativa para describir “la realidad del español tal y como la misma se percibe analizando fuentes lexicográficas auténticas, actualizadas y fiables”.
- *Diccionario panhispánico de términos médicos* (Real Academia Nacional de Medicina de España 2023), dedicado al léxico científico de la medicina y con vocación panhispánica, “se presenta como una obra de acceso libre y gratuito” elaborada por un equipo integrado por las Academias Nacionales de Medicina de España y americanas, en el que han participado “traductores, informáticos, etimólogos, lexicógrafos y especialistas en codificación”.
- *Diccionario de colocaciones del español (DICE)*. Se trata de un diccionario realizado en el marco teórico de la lexicología explicativa y combinatoria, basado en el concepto de función léxica y con una nomenclatura por ahora limitada al campo léxico de los nombres del sentimiento (Alonso Ramos).
- *Diccionario de locuciones idiomáticas del español actual (DiLEA)*. Recoge el tipo de unidades léxicas indicadas en su título, en concreto 9075 entradas y 12 627 locuciones delimitadas diatópicamente en el español de España y de uso actual. El diccionario se actualiza periódicamente por lo que el número de entradas y locuciones aumenta progresivamente (Penadés Martínez 2019).
- *Diccionario de partículas discursivas del español*. Tiene una microestructura ideada para la descripción y ejemplificación del significado procedimental de las partículas. El corpus en el que se basa abarca sobre todo el español de España, pero se actualiza para ampliarlo a otras variedades (Briz; Pons; Portolés 2008).

Por último, está la digitalización de diccionarios, entendida no ya como una reproducción facsimilar digital, sino como la conversión y transformación de un diccionario impreso en un conjunto de datos legibles e interpretables por un sistema informático con la finalidad de obtener una edición digital interrogable. El producto final es similar al de un diccionario nacido digital, pero la base de partida es un diccionario impreso existente, tratado con el rigor historiográfico necesario en función de su datación, sus características lingüísticas y textuales y la finalidad misma de la edición digital, de ahí que se use el término *retrodigitalización*. Este neologismo designa el conjunto de operaciones informáticas,

ecdóticas y filológicas que permiten al investigador y al usuario interactuar con el diccionario editándolo, integrándolo con otros diccionarios, haciendo búsquedas complejas o, simplemente, consultándolo.

Son numerosos los diccionarios retrodigitalizados presentes en la red. Además de los que ha publicado la Real Academia Española, citamos los siguientes a título ilustrativo:

- *Diccionario del español de México* (en línea). Se trata de la versión electrónica de la segunda edición del diccionario (Lara 2024), con la posibilidad de búsquedas avanzadas y la presencia de recursos añadidos: listas de sufijos, conjugaciones, escritura de los numerales, el enlace a la presentación de otras investigaciones del Colegio de México y al Corpus del español de México contemporáneo<sup>2</sup>. Además, el usuario puede hacer propuestas o plantear dudas al equipo lexicográfico a través de un formulario.
- *Diccionario del español actual* (DEA). Esta tercera edición en versión electrónica es, a nuestro juicio, un caso híbrido, ya que parte de la retrodigitalización de las ediciones impresas del diccionario publicadas en 1999 y 2011 (Seco; Ramos; Andrés 1999), si bien el producto final se considera una edición nueva “notablemente aumentada y puesta al día” (Andrés 2023).
- *e-DRAE1884*. Con finalidad historiográfica, de sumo interés para el investigador en historia de la lexicografía y del léxico, consiste en una edición digital del DRAE de 1884, realizada a raíz de un proyecto dirigido por Gloria Clavería Nadal y Margarita Freixas Alás (Clavería Nadal; Freixas Alás).

Así pues, pueden distinguirse tres grandes tipos de productos, originados por la interacción entre las humanidades digitales y la lexicografía: las reproducciones facsimilares digitales, los diccionarios digitales y los diccionarios digitalizados o retrodigitalizados.

Con respecto a este último tipo, una de las aplicaciones que más se beneficia de la retrodigitalización de diccionarios es la elaboración de tesoros lexicográficos. En la tradición lexicográfica hispánica, el término *tesoro*, además de referirse a un cierto diccionario (como el de Sebastián de Covarrubias, que así lo indica en el

2 En el momento de la redacción de esta introducción, el enlace al corpus está inactivo porque envía a la primera versión del Corpus, que ha sido actualizada con una segunda versión publicada en una nueva URL <<https://cemcii.colmex.mx/index.html>>. Mencionamos la circunstancia, muy común por otra parte, no para disminuir el valor de este excelente diccionario en línea, sino como ilustración anecdótica de uno de los grandes problemas de las humanidades digitales, como es la necesidad de una actualización constante —en especial la de los enlaces—.

título: *Tesoro de la lengua castellana, o española*) se utiliza para denominar un *corpus glosariorum* (Nieto Jiménez; Alvar Ezquerra 2007: IX), esto es, un diccionario de diccionarios. Estos repertorios compilan y ordenan de forma exhaustiva, normalmente por criterios cronológicos, la información lexicográfica de una lengua. Su función principal es descriptivo-documental, con aplicaciones historiográficas y metalexicográficas que permiten rastrear tanto de las influencias de unos diccionarios en otros, como de las innovaciones. Entre los principales tesoros lexicográficos de la lengua española se encuentran los siguientes, los tres primeros, publicados solo en formato impreso, y los restantes, en versión digital:

- *Tesoro lexicográfico* (1492-1726) de Samuel Gili Gaya (1960). Obra de gran trascendencia para la historia de la lexicografía, a pesar de haber quedado inconclusa, pues solo llega hasta la letra E, de estar cronológicamente limitada a los diccionarios anteriores a la publicación del denominado *Diccionario de Autoridades* de 1726 y de no ser exhaustiva, pues son muchos los diccionarios no incluidos en el corpus del insigne lexicógrafo.
- *Nuevo Tesoro Lexicográfico del español* (s. XIV-1726) de Lidio Nieto Jiménez y Manuel Alvar Ezquerra (2007). Es una obra monumental, con un corpus de 145 repertorios, unas cien mil entradas, que llegan a casi un millón si se cuentan variantes según autor y más de seiscientas mil referencias de diferentes textos. Su corpus alberga diccionarios alfabéticos y conceptuales, monolingües, bilingües y multilingües, repertorios de ámbito diatópico, diccionarios de lenguas especiales, glosarios, etc.
- *Tesoro lexicográfico del español de Canarias* de Cristóbal Corrales Zumbado, Dolores Corbella Díaz y M.<sup>a</sup> Ángeles Álvarez Martínez (1992). Además de recoger las palabras de la tradición diccionarística del español de Canarias, incluye las aportaciones de estudios lexicológicos, especialmente la del *Atlas Lingüístico y Etnográfico de las Islas Canarias* (Alvar 1975-78).
- *Nuevo Tesoro Lexicográfico de la Lengua Española* (NTLE) de la Real Academia Española. Es, probablemente, el *Tesoro* más conocido. A pesar de sus indiscutibles méritos, su corpus es bastante limitado, al menos en relación con el *Tesoro* de Nieto y Alvar, solo se pueden buscar los lemas, no ofrece la posibilidad de la búsqueda avanzada y los resultados que proporciona consisten en las imágenes facsimilares de los diccionarios.
- *Tesoro lexicográfico médico* (*TèLeMe*) de Bertha Gutiérrez Rodilla. El Corpus comprende ocho diccionarios terminológicos de medicina originales, no traducidos, publicados entre mediados del XVIII y primeros años del XIX. Permite búsquedas avanzadas y, como el *NTLE* de la RAE, ofrece

como resultado el facsímil de la página en la que se encuentra la voz.

- *Tesoro lexicográfico del español de Puerto Rico en línea* de la Academia Puertorriqueña de la Lengua Española (2020). Actualmente contiene unas treinta mil palabras y frases puertorriqueñas, provenientes de un corpus de sesenta y cinco textos lexicográficos, pero también de investigación lexicológica, sobre el español de Puerto Rico escritos entre 1788 y 2022. Está basado en la retrodigitalización del *Tesoro lexicográfico del español de Puerto Rico* (Vaquero; Morales 2005), a la que se han añadido fuentes nuevas en un proceso de actualización constante.

El proceso de retrodigitalización pasa por dos grandes fases operativas: la primera consiste en la transcripción de los textos y la segunda en su codificación en un lenguaje de marcado XML. Para esta tarea se suele usar TEI, el estándar más extendido en el ámbito académico. Si bien la digitalización de un diccionario prevé unas normas de transcripción y codificación coherentes con su datación y del número de ediciones, la elaboración de un tesoro implica, además, un proceso de armonización y estandarización de criterios. Esto es esencial, ya que las normas adoptadas deben ser válidas para todos los diccionarios que lo conforman.

Los artículos de esta sección monográfica describen estos procesos —métodos para la transcripción automatizada, codificación en XML y armonización de esta para la construcción de un tesoro— ya que todos ellos plantean reflexiones y resultados en relación con dos proyectos de investigación en curso dedicados a la elaboración de tesoros digitales: el *Tesoro lexicográfico del español de América* (TLEAM) y el *Tesoro digital de la lexicografía bilingüe español-italiano* (TELEI). De estos trabajos se desprende una idea central: la retrodigitalización de diccionarios antiguos no consiste en una mera conversión de formatos (del papel a la pantalla). Al igual que cualquier otro texto sometido a un proceso de edición filológica, requiere un estudio crítico previo y, en el caso de los diccionarios, también un análisis metalexicográfico para determinar la naturaleza de los paratextos, las características de la macroestructura, la nomenclatura, y, sobre todo, la microestructura.

Varios de los artículos derivados del proyecto del TELEI se centran en la fase de transcripción. Alemany Martínez, De Beni y La Manna analizan con detalle las fases del trabajo necesario para la digitalización del *Vocabolario italiano e spagnuolo* (1620) de Lorenzo Franciosini, desde la conversión en formato texto con el uso del software de OCR Transkribus, hasta las decisiones editoriales relativas a la puntuación, la acentuación o el uso de las mayúsculas o el análisis interpretativo

del enunciado lexicográfico necesario para el etiquetado en XML TEI. Por su parte, Ferrante, Valente y de Hériz exponen con abundancia de detalles el minucioso proceso de análisis, ecdótico y metalexicográfico, previo a la digitalización del *Diccionario de faltriquera* de Cormon y Manni de 1805. Las autoras explican una serie de casos emblemáticos en los que se observan con claridad las dificultades que presenta la transcripción automática de un texto lexicográfico, incluso del siglo XIX, realizada con el software Transkribus. Asimismo, Bermejo Calleja, Lanteri, Valero Gisbert y Zaccone dedican una buena parte de su ensayo al uso de *Transkribus* en la transcripción del diccionario bilingüe de Ambruzzi y al análisis de fenómenos macroestructurales de este diccionario. Se enfocan en el uso de las abreviaturas, que no son lo uniformes que se esperaría en un diccionario de mediados del siglo XX, proponiendo soluciones para estandarizar la marcación.

La transcripción requiere un análisis crítico del texto que, como se apuntaba antes, no tiene solo un carácter estrictamente lingüístico (estado de lengua, grafemas, puntuación, erratas y errores, variantes morfológicas, etc.) sino también metalexicográfico (tipos de lematización, tipos de definición o de equivalencias, marcación, abreviaturas utilizadas, etc.). Este tipo de enfoque es el que ofrecen los artículos de Lombardini y de Peñín Fernández, ambos dedicados a la lexicografía bilingüe no alfabética. Lombardini estudia la marcación del género gramatical en la nomenclatura incluida en *L'italiano istruito nella cognizione della lingua spagnuola* de Francisco Marín (1833), ya que, como se sabe, las nomenclaturas tienen una microestructura mínima, normalmente reducida al lema y a la equivalencia interlingüística, por lo que, la información gramatical, si la hay, es irregular y de expresión variable, ciertamente no normalizada con una abreviatura. Peñín Fernández, por su parte, dedica su artículo a un análisis crítico completo del *Vocabulario Espanyol Italiano y Tudesco* de Ignacio de Boria (1719), una nomenclatura trilingüe español-italiano-alemán que aporta la peculiaridad de su condición inédita.

La fase de codificación y cómo resolver la necesaria homogeneización para la constitución de un tesoro digital son los puntos clave de los cuatro artículos siguientes; el primero está dedicado a un diccionario del TLEAM y los otros tres, a diccionarios del corpus de TELEI.

Díaz Rodríguez, tras reseñar brevemente la finalidad, las características y las fases de elaboración del TLEAM, se concentra en el proceso de homogeneización de los materiales que constituyen el corpus de este Tesoro, a partir del estudio de caso del *Diccionario de voces americanas* de Manuel José de Ayala (1777), tratando en particular la normalización de la ortografía, la lematización geminada, la du-

plicidad de lemas y la codificación de las unidades poliléxicas. A este último tema, está también dedicado el trabajo de Dalle Pezze y Sartor, quienes analizan con detalle una de las cuestiones más relevantes del *Vocabolario* de Franciosini, como es la fraseología, o el discurso repetido, siguiendo la terminología coseriana utilizada por las autoras. El artículo da cuenta de la dificultad que presenta el etiquetado metalexicográfico de este tipo de estructuras, difícilmente sistematizables en los diccionarios antiguos. Para ello, las autoras trabajan con un subcorpus piloto en el que analizan la técnica lexicográfica empleada por Franciosini para lematizar y establecer equivalencias, así como las condiciones que debe cumplir su etiquetado en XML-TEI Lex0, que es el estándar de codificación utilizado en TELEI. En la misma línea, García Jiménez y Pérez Vázquez describen las características específicas del *Vocabulario italiano-español* del *Diccionario marítimo* de O'Scanlan publicado en 1831 y discuten las propuestas de codificación para los lemas pluriverbales, lemas geminados o las equivalencias definicionales de este diccionario de especialidad, bilingüe, aunque no bidireccional. Por último, también Castillo Peña aborda con detalle los problemas de codificación que presentan las nomenclaturas en TELEI, en cuanto diccionarios conceptuales cuya microestructura es difícilmente homologable a la de los diccionarios alfabéticos.

En conclusión, estos trabajos ponen de manifiesto, por un lado, la gran variedad genérica y cronológica de los textos estudiados: diccionarios del español hablado en América, diccionarios bilingües, diccionarios de especialidad, nomenclaturas, glosarios incluidos en gramáticas, textos impresos de gran trascendencia, textos menores y textos inéditos. Por otro lado, ponen de relieve la necesidad de un conjunto específico de recursos metodológicos y tecnológicos propios de las humanidades digitales. Los procesos de digitalización, retrodigitalización y armonización en la construcción de tesoros lexicográficos evidencian que la edición crítica de diccionarios requiere no solo un tratamiento filológico riguroso, sino también una reflexión metalexicográfica. Esta reflexión es fundamental para garantizar su coherencia y su utilidad para la investigación. Por tanto, estos proyectos en curso confirman la oportunidad de este enfoque para asegurar la preservación y la proyección futura de los repertorios lexicográficos, consolidando así un campo de estudio que consideramos esencial para la historiografía lingüística y para la comprensión de la tradición lexicográfica en su conjunto.



## Referencias bibliográficas

- Academia Puertorriqueña de la Lengua Española, (2020), *Tesoro lexicográfico del español de Puerto Rico en línea*. [29/03/2025] <<https://tesoro.pr/>>
- Alonso Ramos, Margarita; Grupo DICE (dir.<sup>a</sup>) (en línea), *Diccionario de colocaciones del español (DICE)* [26/06/2025] <<http://www.dicesp.com>>.
- Alvar, Manuel (1975–78), *Atlas Lingüístico y Etnográfico de las Islas Canarias*, Las Palmas de Gran Canaria, Cabildo Insular de Gran Canaria.
- Andrés, Olimpia (dir.<sup>a</sup>) (2023), *Diccionario del español actual* (edición electrónica). Fundación BBVA [14/06/2025] <<https://www.fbbva.es/diccionario>>.
- Briz, Antonio; Pons, Salvador; Portolés, José (coords.) (2008), *Diccionario de partículas discursivas del español*. [04/04/2025] <[www.dpde.es](http://www.dpde.es)>.
- Corrales Zumbado, Cristóbal José; Corbella Díaz, Dolores; Álvarez Martínez, María Ángeles (1996), *Tesoro lexicográfico del español de Canarias*, Madrid, Real Academia Española.
- Clavería Nadal, Gloria; Freixas Alás, Margarita (en línea) *e-DRAE 1884* [15/04/2025] <<https://edrae1884.uab.cat>>.
- Estrada Ramírez, Hortensia (coord.) (en línea), *Diccionario electrónico sáliba-español: una propuesta de documentación de la lengua y la cultura sálibas*, Bogotá, Instituto Caro y Cuervo [24/06/2025] <<https://saliba.caroycuervo.gov.co>>
- Fuertes Olivera, Pedro (en línea), *Diccionario Digital del Español (DIDES)*. [24/06/2025] <<https://diesgital.com/>>.
- Gili Gaya, Samuel (1960), *Tesoro lexicográfico (1492-1726)*, Madrid, Consejo Superior de Investigaciones Científicas.
- Gutiérrez Rodilla, Bertha M. (dir.<sup>a</sup>) (en línea) *Tesoro lexicográfico médico (TeLeMe)* [en línea]. [28/04/2025] <<http://teleme.usal.es>>
- Lara, Luis Fernando (dir.) (2024), *Diccionario del español de México (DEM)*, México DF, El Colegio de México. Versión en línea [15/05/2025] <<https://dem.colmex.mx>>.
- Nieto Jiménez, Lidio; Alvar Ezquerro, Manuel (2007), *Nuevo Tesoro Lexicográfico del español (s. XIV-1726)*, Madrid, Arco Libros.
- Penadés Martínez, Inmaculada (2019), *Diccionario de locuciones idiomáticas del español actual (DiLEA)* [03/04/2025] <[www.diccionariodilea.es](http://www.diccionariodilea.es)>.
- Real Academia Española (en línea), *Nuevo tesoro lexicográfico de la lengua española (NTLLE)*, [04/06/2025] <<https://apps.rae.es/ntlle/SrvltGUISalirNtle>>.
- Real Academia Nacional de Medicina de España (2023), *Diccionario panhispánico de términos médicos* [25/05/2025] <<https://dptm.es/>>.
- Seco, Manuel; Andrés, Olimpia; Ramos González, Gabino (1999), *Diccionario del español actual*, Madrid, Aguilar.
- Tarp, Sven (2019), “La ventana al futuro: despidiéndose de los diccionarios para abrazar la lexicografía”, *RILEX. Revista sobre investigaciones léxicas*, 2: 5-36.
- Vaquero, María; Morales, Amparo (2006), *Tesoro Lexicográfico del Español de Puerto Rico*, San Juan, Academia Puertorriqueña de la Lengua Española.



**AMPARO ALEMANY MARTÍNEZ  
MATTEO DE BENI**

**ALESSANDRA LA MANNA\***

# **LA DIGITALIZACIÓN DEL *VOCABOLARIO ITALIANO E SPAGNOLO* (1620)**

## **DE LORENZO FRANCIOSINI: RETOS METODOLÓGICOS Y NUEVAS OPORTUNIDADES DE ESTUDIO**

**Università degli Studi di Verona**

### **Resumen**

El presente artículo aborda la digitalización con codificación XML-TEI del *Vocabolario italiano e spagnolo* (1620) de Lorenzo Franciosini —una obra central de la lexicografía hispanoitaliana de la edad moderna— en el marco de un proyecto de investigación actualmente en curso. Además de presentar las fases de trabajo, en la contribución se mostrará, gracias a ejemplos pertinentes, cómo el proceso de digitalización permite ahondar en el conocimiento de la obra lexicográfica estudiada.

palabras clave: lexicografía bilingüe italiano-español, Lorenzo Franciosini, codificación XML-TEI

### **Abstract**

***The digitisation of Lorenzo Franciosini's Vocabolario italiano e spagnolo (1620): methodological challenges and new opportunities for its study***

*This article aims to examine the digitisation with XML-TEI encoding of Lorenzo Franciosini's Vocabolario italiano e spagnolo (1620), a central work of Hispano-Italian lexicography of the Modern Age, within the framework of a research project currently underway. In addition to presenting the workflow and its phases, the contribution will show, thanks to relevant examples, how the digitisation process allows us to deepen our knowledge of the lexicographical work under study.*

keywords: Spanish-Italian bilingual lexicography, Lorenzo Franciosini, XML-TEI encoding

## 1. Introducción

El *Vocabolario italiano e spagnolo* (Roma, 1620) de Lorenzo Franciosini es, sin duda alguna, una joya de la historia de la lexicografía bilingüe entre español e italiano (cfr., entre otros, Gallina 1959: 263-284; Martínez Egido 2002; Valencia y Peña 2003; San Vicente 2010: 63-80). Antes de la publicación de esta obra la producción de diccionarios bilingües italiano-español era bastante pobre: además del pequeño repertorio manuscrito de Nicolao Landucci de 1562, que presenta las palabras en italiano con sus equivalentes en español, destaca el *Vocabulario de las dos lenguas toscana y castellana* de Cristóbal de las Casas de 1570, que se considera el primer verdadero diccionario bilingüe español-italiano, después sustituido como herramienta de referencia por el de Franciosini. El éxito del diccionario del lexicógrafo italiano queda ratificado por sus muchas ediciones<sup>1</sup> y estriba, entre otras razones, en la voluntad de su autor de producir una obra guiada por prioridades prácticas<sup>2</sup>.

Además, el “maestro de lenguas” toscano llevó a cabo otras tareas de interés para las dos lenguas en cuestión: realizó la primera traducción al italiano de *El Quijote* de Cervantes —cuyas dos partes se publicaron en 1622 y 1625— y publicó una conocida *Grammatica spagnola e italiana* (1624) y unos *Diálogos apazibles* junto con el *Nomenclator o Registro de algunas cosas curiosas y necessarias de saberse*

\* Este trabajo se ha realizado en el marco del proyecto “Un nuovo ambiente digitale per il recupero del patrimonio lessicografico: il *Tesoro digitale della lessicografia bilingue spagnolo-italiano*”, financiado por la Unión Europea – Next Generation EU, Missione 4 Componente 1 CUP C53D23004010006 / PRIN 2022 MUR 20229W73WR. Participan las universidades de Padua, Bolonia, Génova, Pisa, Turín y Verona; esta última es la que se encarga del diccionario de Franciosini. Aunque el presente ensayo se ha elaborado de forma conjunta, a efectos académicos Amparo Alemany Martínez es responsable de los apartados 3 y 4.2, Matteo De Beni de los apartados 1, 2 y 5 y Alessandra La Manna de los apartados 4.1, 4.3 y 4.4.

1 Se cuentan trece ediciones de 1620 a 1796. Otras dos, “que varios bibliógrafos mencionan, una de 1633 (Ginebra, P. Marcello) y otra de 1776 (Venecia, Baglioni)”, no tienen localización conocida (San Vicente 2010: 64).

2 El diccionario de Franciosini sirvió como herramienta en las labores de traducción y como instrumento de aprendizaje lingüístico durante más de dos siglos. En este sentido, son significativas las propias palabras del autor en la portada de la edición de 1620: “Vocabolario Italiano, e Spagnolo non piu dato in luce. Nel quale con la facilità, e copia, che in altri manca, si dichiarano, e con proprietà convertono tutte le voci Toscane in Castigliano, e le Castigliane in Toscano. Con le Frasi, & alcuni Proverbi, che in ambe due le lingue giornalmente occorrono; con una chiara, e breve Regola per leggere, e scriuere, & una succinta introduzione, con Avvertimenti di molte cose notabili. Opera vtilissima, e necessaria a' Predicatori, Segretari, e Traduttori, che con legittimo senso e vero fondamento le voglion tradurre, o imparare”.

a los estudiosos de lengua española (1626)<sup>3</sup>.

Los aspectos aludidos bastarían para justificar la inclusión de la obra de Franciosini en un repertorio lexicográfico digital como el que, con la sigla *TELEI*, se está creando en el seno del proyecto “Un nuovo ambiente digitale per il recupero del patrimonio lessicografico: il *Tesoro digitale della lessicografia bilingue spagnolo-italiano*”. El proyecto prevé la creación de una constelación de ediciones digitales de obras lexicográficas emblemáticas para el marco italoespañol, producidas entre el siglo XVI y el XX, que se codificarán con el etiquetado XML-TEI y se podrán consultar a través de una herramienta diseñada *ad hoc*. Esto permitirá superar la lectura convencional de los documentos metalingüísticos (cfr. Castillo Peña 2020): entre otras ventajas, al codificar el texto, se lematizan voces, se resaltan características semánticas y se organiza la información de forma jerárquica (de los *elementos* penden *atributos*), brindando la oportunidad de recuperar toda la información a través de filtros. El estándar de codificación TEI es un sistema de etiquetado y marcación textual que, por su flexibilidad, permite adaptarse a la estructura de una obra lexicográfica. De hecho, se trata de un sistema creado desde sus inicios para “formulate and disseminate guidelines for the encoding and interchange of machine-readable texts intended for literary, linguistic, historical, or other textual research” (TEI Consortium 1988 *apud* Castro Prieto 2022: 222). La guía TEI, además, tiene un estándar para diccionarios, que se rige en una amplia flexibilidad:

First, because the structure of dictionary entries varies widely both among and within dictionaries, the simplest way for an encoding scheme to accommodate the entire range of structures actually encountered is to allow virtually any element to appear virtually anywhere in a dictionary entry. It is clear, however, that strong and consistent structural principles do govern the vast majority of conventional dictionaries, as well as many or most entries even in more ‘exotic’ dictionaries; encoding guidelines should include these structural principles. (TEI Consortium 2024, cap. 10)

Cabe recalcar que el proceso de digitalización previsto en nuestro proyecto permitirá ahondar en el conocimiento del diccionario de Franciosini, gracias a las posibilidades de búsqueda y conexión entre los datos que brinda, ya en una primera fase, el texto plano y aún más, en una segunda fase y a través de la plataforma de *TELEI*, el texto digital con etiquetado.

3 Los *Diálogos* y el *Nomenclator* aparecen junto a la *Gramática* a partir de la edición de esta de 1638.

## 2. Objetivos y estructura del trabajo

El presente estudio pretende exponer y justificar las fases y estrategias seguidas para transformar el *Vocabolario* en un documento codificado en XML-TEI y, asimismo, presentar una caracterización formal de la obra. En concreto, se resalta el proceso de segmentación, transcripción y digitalización y se describen las fases metodológicas que sustentan dichas tareas, aprovechando las capacidades de aprendizaje del programa utilizado, Transkribus, un sistema de reconocimiento de caracteres y transcripción automática (§ 3). Se trata, así pues, de las fases previas a la codificación XML-TEI con estándar Lex-0 que, posteriormente, se realizará a partir de una plataforma informática creada *ad hoc*, cuyo principal cometido será el de anejar el corpus de diccionarios con el que conformar el ya mencionado *TELEI*. Asimismo, también presentaremos brevemente las normas de edición que hemos fijado gracias al detallado conocimiento del diccionario que nos ha brindado la transcripción íntegra de su texto. A continuación, se desarrolla una exposición de las características y los problemas que, a partir de la transcripción y su posterior revisión, se han detectado en el texto lexicográfico estudiado (§ 4). Para ello, se usarán ejemplos de diversas letras, tanto del primer volumen (italiano-español) como del segundo (español-italiano). El trabajo con el documento ha permitido detectar incongruencias dentro del *Vocabolario* que señalamos para que, a partir de ellas, se puedan exponer las soluciones a las que hemos llegado y las ventajas que supone su hallazgo para comprender y analizar la obra. Finalmente, el trabajo termina con unas conclusiones (§ 5).

## 3. La digitalización del *Vocabolario*: metodología y fases

A continuación, nos centramos en la hoja de ruta establecida para digitalizar el *Vocabolario* y acondicionarlo hacia su inserción en un repositorio más amplio en el que se permitan, entre otras muchas opciones, la consulta, navegación e interacción con otros repertorios lexicográficos. Se procede, por lo tanto, a desglosar las fases metodológicas que se han llevado a cabo para tener la obra transcrita y revisada.

### 3.1 *La segmentación del diccionario*

El *Vocabolario* se divide en dos partes: en la primera la lengua de entrada es el italiano y la de salida el español, esto es, se dan los equivalentes en español de un lema en italiano; en la segunda, el lema se da en español, acompañado por sus equivalentes en italiano. Los dos tomos constan de un total de 677 y 792 páginas, respectivamente. El trabajo con el volumen total de información no habría sido eficaz, dado que el programa habría tenido que efectuar muchas operaciones (codificación, formalización de los constituyentes del escaneado, discriminación de los caracteres leídos, etc.) en un documento muy extenso y con una enorme cantidad de datos. Por consiguiente, se han separado los dos volúmenes y, al mismo tiempo, se han hecho distintos documentos por cada una de las letras que dividen ambos tomos, esto es: un documento de la letra A, otro de la B y así sucesivamente para las dos versiones. Naturalmente, el número de páginas entre una letra y otra es muy distinto: varía entre las 105 páginas de la C del segundo volumen y unas poquísimas páginas de letras como la Ñ, la X o la Z.

### 3.2 *Transkribus: un HTR con Inteligencia Artificial*

Transkribus es un *software* basado en OCR (*Optical Character Recognition*), es decir, un sistema de reconocimiento de caracteres que permite codificar material de texto escaneado y, asimismo, discriminar la información fija para segmentarla. Por lo tanto, subdivide bloques de contenido, facilitando así su lectura. En concreto, en cuanto OCR, transcribe las partes identificadas de una imagen en *.txt*, *.docx*, HTML, entre otros<sup>4</sup>. Además, Transkribus también es un HTR (*Handwritten Text Recognition*):

Attraverso Transkribus è possibile, sulla base di trascrizioni fornite da personale esperto (*ground truth*), sviluppare modelli di riconoscimento automatico di testo scritto da una determinata grafia.

In questo modo a fronte della trascrizione manuale di un *ground truth* di circa 10.000 parole è possibile “allenare” un modello di riconoscimento HTR, supportato da un algoritmo basato sui principi del calcolo delle probabilità e sul funzionamento delle reti neurali, in grado di decifrare quel tipo di grafia in percentuali oscillanti tra il 92% e il 98%. (Zappulli, Iorio 2018: 51)

4 El programa, en su versión de pago, también es capaz de etiquetar con TEI.

Por lo tanto, este programa resulta beneficioso para las investigaciones con imágenes escaneadas de textos escritos a mano y de impresos con tipografías que para un OCR convencional resultarían de difícil lectura. No es de extrañar que este entorno virtual se haya utilizado en el ámbito de las humanidades digitales, puesto que asegura un reconocimiento rápido y eficaz de textos de distinta índole y de características muy variadas. Además, Transkribus se enmarca en el Aprendizaje Automático, dado que el programa aprende a medida que recibe más *input* del mismo objeto. Al aplicar este modelo a la lexicografía, se abandona la idea de que un diccionario es un “sistema estático de información” (Wiegand 2007: 38), dado que, gracias a la sucesiva fase de etiquetado, el texto se vuelve dinámico. Por este motivo, la IA de este *software* resulta beneficiosa para las investigaciones de carácter diacrónico, dado que permite crear un modelo de transcripción a partir de documentación equivalente o relacionada.

Transkribus es capaz de leer y transcribir cualquier documento en formato PDF, pero el documento tiene que cumplir con una serie de características para que la transcripción sea lo más precisa y limpia posible, en particular, se necesita una digitalización nítida, sin excesivos cortes o pliegues en las páginas escaneadas o juegos de luces y sombras<sup>5</sup>. Además, para un trabajo de transcripción de gran envergadura como el que aquí se expone, es necesario, en aras de crear un modelo de transcripción, que los documentos de trabajo estén relacionados: de este modo, se agiliza el procesamiento del texto por parte de la IA —la cual puede focalizarse en separar las diferencias y acercar las semejanzas en los distintos documentos que se pretenden transcribir— y se crea un modelo basado en un prototipo y no en textos periféricos.

En suma, Transkribus propicia un mayor alcance de la investigación, dado que transpone el texto escaneado a una versión más maleable que habilita nuevas funciones textuales (búsqueda de información, facilidad lectora, acceso a distinguir información, etc.) o sintáctico-semánticas (acerca la semejanza y discrimina la diferencia en las partes fragmentadas, por lo tanto, favorece no solo la búsqueda sino también el etiquetado y segmentación de los datos). En el punto que sigue, se mostrará cómo se ha preparado el *Vocabulario* con el fin de convertir la imagen en texto.

---

5 En nuestro caso, hemos trabajado con la versión escaneada de los ejemplares de la primera edición del *Vocabulario* conservados por la Staatliche Bibliothek Regensburg (Alemania): signatures 999/Ling.620 (1), para el primer volumen, y 999/Ling.620 (2), para el segundo.



### 3.3 Creación de un modelo de transcripción propio y criterio editorial

Como todas las IA, Transkribus necesita de cierto entrenamiento para lograr que el *input* se convierta en el *output* deseado. Esto se traduce en que, al inicio del trabajo, será mayor el coste que el beneficio en términos de tiempo y esfuerzo, pero cuando el programa reconozca la suficiente cantidad de información y la discrimine se volverá provechoso.

Antes de proceder con el reconocimiento del texto, es necesario preparar la estructura de la página que se quiere transcribir. En nuestro caso, estamos ante una obra lexicográfica del siglo XVII, por lo tanto, la preparación deberá tener en cuenta las características de nuestra pieza, a saber, su disposición en dos columnas, que se corresponden con el elenco jerarquizado alfabéticamente de las entradas lexicográficas, y la presencia de una cabecera y de un pie de texto. En la primera se muestran las dos letras iniciales de los lemas que se consignan en cada una de las dos columnas que dividen la página, en el segundo se presenta el reclamo, esto es, el lema (o el comienzo de este) que inicia la página siguiente, además de otras marcas tipográficas. En este sentido, en el caso del *Vocabolario* de Franciosini, para que Transkribus lea el documento conforme deseamos, debemos seleccionar por separado —con la opción de *región*— las dos columnas que conforman el texto, ordenadas de izquierda a derecha, y discriminar las partes que no deseamos de la hoja (cabecera y pie de página).

Hay modelos variados de reconocimiento textual en este *software*, dado que existen muchos grupos de investigación implicados en la transcripción de documentos de distintas épocas, lenguas, tipologías textuales, etc. En nuestro caso, el modelo seleccionado para llevar a cabo la primera parte de la transcripción es *Transkribus Print M1*, que se describe así en la página del propio *software*:

Extended multi-language Transkribus print model, including antiqua and blackletter prints, typewriter, computer print outs and decorative fonts includes more languages than print 0.3. The CER in M1 is higher than in 0.3 which is due to a more varied validation set. For languages that were already included in 0.3 the new M1 usually performs equally well or slightly better than 0.3. Curated by the Transkribus team, this model is occasionally updated with community data for continuous improvement. (READ-COOP SCE 2024, en línea)

Dicha elección se debe a que *Print M1* cubre el periodo cronológico del diccionario, es decir, es capaz de transcribir textos de un amplio periodo que abarca tam-

bién el siglo XVII, y su sistema de reconocimiento incluye el español<sup>6</sup>. A partir del texto creado desde el modelo M1, los transcriptores revisaron y corrigieron la transcripción para, posteriormente, crear un modelo que cometiera menos errores (imagen 1). El siguiente paso, por lo tanto, ha sido crear un modelo de transcripción propio del diccionario, que tenga en cuenta las correcciones asumidas. La creación del modelo se sustenta mediante la transcripción de la letra más larga (la letra C de cada volumen) y se ha efectuado a partir de tener la letra más del 35 % transcrito, corregido y revisado. Transkribus permite crear modelos sobre el 10 % del texto que ya está en su versión definitiva, sin embargo, teniendo en cuenta el volumen de trabajo que supone transcribir una obra de esta cantidad de páginas, hemos creído conveniente rebasarlo.



Imagen 1. Texto transcrito con el modelo propio en la plataforma de Transkribus

Para llevar a cabo la corrección del texto transcrito, desde el grupo de investigación hemos creado unas pautas para llevar a cabo un modelo de transcripción ideal, cuyos criterios fundamentales han sido *a)* intervenir lo menos posible el texto para respetar las características del original y de su época y *b)* homogeneizar los signos ortográficos utilizados con función metalingüística, que son el punto y el corchete, manteniendo la siguiente estructura para la entrada: *lema. [ equivalente 1. equivalente 2. equivalente n.* Por lo tanto, conservamos o, si falta,

6 Resulta de gran importancia la elección del modelo de transcripción, dado que, si el modelo reconoce fácilmente los caracteres, el entrenamiento a partir de la corrección de la transcripción resulta más fácil para el reconocimiento automático posterior.

reponemos el punto y seguido que cierra el lema y el punto final que concluye la entrada, así como el punto como separador entre equivalentes; de la misma manera, conservamos o, si falta o se ha sustituido, reponemos el corchete que marca la transición del lema de la lengua origen a los equivalentes de la lengua meta.

Además, a pesar de la poca intervención en el texto, sí que se han introducido algunos otros retoques: en primer lugar, se han eliminado las separaciones de palabras entre una línea y la siguiente del original (tanto con guion para señalar el salto de línea como sin ello). En segundo lugar, se ha normalizado el uso de *u* con respecto a *v*. En tercer lugar, se han corregido las erratas textuales evidentes. Esta última operación se ha podido llevar a cabo gracias a la redacción de una bitácora de erratas: de esta manera, los fallos que se han encontrado durante el proceso están reseñados, localizados y analizados. En este sentido, en el texto editado se pretende efectuar una limpieza de los desajustes que hemos detectado.

### 3.4 Codificación en XML-TEI

La última fase del trabajo es la codificación del texto a partir de la creación de un administrador creado *ad hoc* para *TELEI*. Una muestra la podemos ver en la imagen 2.

No obstante, cabe decir que, hasta llegar a la versión definitiva de la transcripción, se han podido observar distintas características peculiares del diccionario, incluyendo varios desajustes o incongruencias. En el siguiente apartado se reseñará una caracterización formal del diccionario, haciendo hincapié en estos aspectos y las soluciones que se han aportado.

XML

```

<entry type="mainEntry" xml:id="Francio1620.nunzio" xml:lang="it">
  <form type="lemma">
    <orth>Nunzio, o Ambasciador del Papa</orth>
  </form>
  <gramGrp>
    <gram norm="sustantivo" type="pos"/>
    <gram norm="masculino" type="gender"/>
  </gramGrp>
  <metamark function="variantLemmaDelimiter">o</metamark>
  <form type="variant">
    <orth>Ambasciador del Papa</orth>
  </form>
  <metamark function="translationDelimiter">.<[/metamark>
  <sense xml:id="Francio1620.nunzio.s.1.">
    <cit type="translationEquivalent" xml:lang="es">
      <form>
        <orth>Nunzio</orth>
      </form>
    </cit>
  </sense>
</entry>

```

Imagen 2. La entrada *Nunzio* etiquetada según las normas TEI Lex-0 en *TELEI*

#### 4. Caracterización formal del *Vocabolario*

La organización del *Vocabolario* resulta, en general, homogénea y se estructura bajo una serie de principios que se replican a lo largo de los dos volúmenes. En esta sección se expondrá, en primer lugar, el conjunto de características comunes a la microestructura (§ 4.1); en segundo lugar, las divergencias que se han encontrado con respecto a las reglas generales, a saber, los inconvenientes localizados en el contenido que complican la tarea de realizar una transcripción y una posterior edición crítica de la obra (§ 4.2). Hemos dividido aquí en dos grupos los signos

ortográficos: los que tienen carácter metalexicográfico (§ 4.2.1) y los que carecen de este valor (§ 4.2.2). Del mismo modo, desarrollaremos también los problemas que atañen a otros aspectos formales (§ 4.3) y destacaremos una serie de características singulares de esta que repercuten en su codificación digital (§ 4.4).

#### 4.1 *Características generales*

El artículo lexicográfico del *Vocabolario* de Franciosini se sistematiza, como ya se ha advertido, mediante la estructura siguiente: *lema*. [ *equivalente 1. equivalente 2. equivalente n.*

La mayoría de las entradas ocupan entre una y tres líneas, aproximadamente. El lema aparece en italiano para el primer tomo y en español para el segundo. Tras el lema en el idioma de entrada, se utiliza un punto, seguido por un corchete que introduce los equivalentes, los cuales, en general, no cuentan con un corchete de cierre. De este modo, se establece una definición mediante equivalencia o sinonimia entre el idioma origen y el meta. Cada uno de estos equivalentes se yuxtapone con punto y seguido y el artículo finaliza con un punto y aparte. Por ejemplo:

(1) edificador. [ edificatore. fabbricatore.

Del mismo modo, se han delimitado también funciones frecuentes para la coma, a saber: división de un segmento A de otro B unido por conjunción copulativa o disyuntiva (2, 3); separación de reformulaciones (*cioè*) o remisiones internas (*vedi*) (4, 5); uso de la coma con oraciones explicativas o especificativas (6):

(2) apurar, o llevar al cavo alguna cosa.

(3) Apurado verse. [ vedersi stracco, e consumato.

(4) abbaimento, cioè l'abbaiare. [ Ladrado.

(5) Abbattacchiare, vedi abbacchiare.

(6) appetitoso. [ appetitoso, che sveglia, o incita l'appetito.

Los lemas y los equivalentes suelen tener minúscula inicial (excepto algunas excepciones, cfr. § 4.3). Además, se ha podido determinar que Franciosini utiliza las letras capitales de manera sistemática en una serie de casos para resaltar diacríticamente la importancia de ciertos elementos. Más allá, como es obvio, de elementos de la onomástica, topónimos y otros nombres propios, también lo hace con fórmulas de tratamiento, cargos, títulos religiosos, etc. (7), materias, saberes y ramas del conocimiento (8), gremios, artes y oficios (9) y otros elementos de los

que quiere resaltar su relieve (10):

- (7) embiar Embaxadores. [ mandare, o spacciare Imbasciatori.
- (8) Astrologia. [ Astrologia.
- (9) Escrivano legal, ò Real. [ Notaio publico, e fide digno.
- (10) Astrolabio. [ Astrolabio. strumento matematico per conoscer il moto de Ciel.

A pesar de que la ordenación resulta en la mayoría de los casos sistemática, se han encontrado una serie de incongruencias en la microestructura que se vinculan con las características mencionadas, puesto que en muchas ocasiones se cometen olvidos o se transgreden normas que parecían fijas y, en adición a ello, existen confusiones en el uso de ciertos elementos. Los obstáculos para llevar a cabo una transcripción homogénea no solo guardan relación con el nivel ortotipográfico, sino también con otros elementos que se reseñan en los apartados siguientes.

#### 4.2 *Los signos ortográficos: variación y desajustes*

A continuación, se reseñan los casos que contradicen la sistematización expuesta en el apartado anterior y se explican las soluciones que se han aplicado durante el proceso de transcripción de la obra.

##### 4.2.1 Signos ortográficos con carácter metalexícográfico

Los signos que tienen mayor relevancia a lo largo de la obra son los que cumplen con un papel estructural. Como ya se ha advertido, el punto señala las divisiones internas y externas de las entradas y por regla general abarca tres funciones metalexigráficas: la separación entre el lema y sus equivalentes; la separación de los distintos equivalentes; la finalización de la entrada. Sin embargo, a veces se utilizan otros signos de puntuación en lugar del punto, como la coma; otras veces el punto de entrada a los equivalentes falta:

- (11) carrasco, vedi carrasca, [ leccio.
- (12) compito, cortese. [ cortés, discreto.
- (13) habituado [ habituato, vedi abitur.
- (14) dorado. [ indorato, dorador. [indoratore,

La sustitución del punto por una coma sucede con cierta frecuencia y es posible que se deba a una confusión del cajista. La solución a la que se ha llegado durante la fase de transcripción es la reposición del punto solo en las tres funciones meta-lexicográficas enunciadas arriba.

El otro signo con carácter metalexicográfico es el corchete. Para este se ha advertido la función, que ya hemos comentado arriba, de marcar el vínculo entre el lema y sus equivalentes. Uno de los problemas con este signo, aunque poco frecuente, es su olvido. Por ejemplo, en (15) la expresión española (“es hombre de mucha autoridad”) no introduce su traducción mediante el corchete que tendría que colocarse entre la lengua origen y el equivalente en la lengua meta:

(15) Es hombre de mucha autoridad. E persona grave e di molta autorità.

Durante el trabajo, se ha comprobado también la repetición del corchete, que, al ser una clara equivocación, se ha corregido en la transcripción. Una muestra de esto es (16):

(16) dolerse. [ dolersi. [ lamentarsi.

Por último, se han encontrado casos en los que este signo de cierre aparece más de una vez para resaltar el cambio de código entre los idiomas que conforman el *Vocabolario*. Esto es, marca la múltiple alternancia del español y el italiano en una misma entrada. Su función se reseña más detalladamente en el apartado dedicado a la información lingüística complementaria (§ 4.4.2).

En suma, se han expuesto aquí los signos que solo cumplen con funciones metalexicográficas. En contraposición, en la sección que sigue, se abordarán los que cumplen con otras funciones.

#### 4.2.2 Otros signos ortográficos

Si bien no tienen una función estructural como los anteriores, otros signos ortográficos introducen informaciones relevantes sobre cómo la obra expone ciertos lemas, contextos oracionales, definiciones o modismos. En primer lugar, el paréntesis se suele utilizar con cierta frecuencia para añadir información accesorio, de manera coherente con su función ortográfica:

(17) a todo ruedo. [ modo di dire furbesco, e vale da far ogni cosa, da bosco, e da riviera (come si suol dire).

(18) coger à palabras. [ pigliare uno allà parola, cioè se uno (come si usa nelle Corti) fa mille offerte e promesse, il credergli, che è cosa da semplici.

Estas informaciones adicionales sobre modismos o unidades fraseológicas suelen ser notas aclaratorias acerca del significado de estas. Por lo tanto, las acotaciones como “come si suol dire”, “come si usa in...”, “come usa...”, etc. son más habituales en el volumen italiano-español, lo cual se explica por la voluntad de Franciosini de dar cuenta de ciertos significados convencionalizados en el ámbito del español, pero que pueden resultar ajenos a los italo hablantes.

Otra de las funciones del paréntesis es la de completar la información de una entrada que, debido al espacio de las dos columnas que dividen los tomos, no cabía dentro de la entrada a la que pertenece. Se trata, así pues, de un caso de entradas dislocadas, como se aprecia en la imagen que sigue:

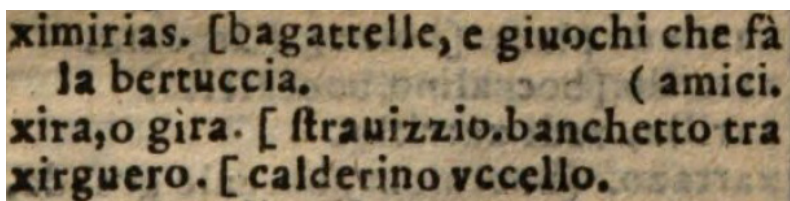


Imagen 3. Uso del paréntesis en entradas dislocadas

Como se puede observar, “amici.” pertenece a la entrada *xira, o gira*:

(19) xira, o gira. [ stravizzio. banchetto tra amici.

En este sentido, la forma en la que se ha procedido con el paréntesis depende de la función que cumpla: obviamente, los casos donde el paréntesis presenta información accesoria se han mantenido, mientras que el paréntesis que marca que una entrada está dislocada por falta de espacio se ha suprimido y el fragmento de texto se ha añadido a la entrada que corresponde, tal y como se muestra en (19).

Existen otros casos de paréntesis que se han interpretado como confusiones del propio Franciosini o de los cajistas que realizaron los volúmenes. Por ejemplo, en (20) se yuxtapone un paréntesis con un corchete, donde claramente se da una información que comúnmente se dispone de forma parentética:

(20) A otro perro con esse guesso. [ Modo di dire che inferisce che uno conosce l'inganno, o la burla, e significa che la vadia à fare a un'altro, perche à lui non gliela ficcherà. [come si suol dire.)



En casos como (20), se ha suprimido el corchete de apertura y se ha cambiado por un paréntesis, en tanto que resulta ser una información complementaria, como los ejemplos anteriores. Del mismo modo, también aparecen paréntesis colocados por equivocación, como sucede en (21), donde también se ha optado por su supresión, puesto que se ha interpretado como errata:

(21) ancora strumento di ferro, col quale gittato ne fondi dell'acque ) si fermano i Navili. [ ancora.

Otro de los casos que nos ocupa en esta sección se refiere a los ejemplos encuadrados entre signos de interrogación o exclamación. En ocasiones, Franciosini, para remarcar énfasis, interrogación o duda cambia la modalidad oracional, pero solo en segmentos determinados. Un ejemplo es el caso que tenemos en (22):

(22) Dò, o a donde. [ dove?

Como se puede apreciar, la entrada no tiene ningún signo de interrogación, pero sí resulta ser una interrogación indirecta, por lo tanto, Franciosini lo reporta en la equivalencia por medio de la modalidad interrogativa. De esta forma, aproxima el significado de la entrada en la lengua origen, en este caso el español al italiano. No es así en todas las entradas. Por ejemplo, en (23) o (24) las dos oraciones son interrogativas, dado que se traduce y se lematiza el lema de la entrada en su forma interrogativa y no la declarativa:

(23) por donde? [ di dove? o per dove?

(24) Por ventura es ahi? [ e egli cosi dauero? è egli davvero? modo di dire.

En este apartado se ha querido mostrar cómo se ha trabajado con la transcripción de los signos ortográficos en los dos volúmenes que componen el *Vocabolario* de Franciosini. En este sentido, se ha intentado dejar el texto lo más fiel y cercano al original, dado que las intervenciones que se han realizado se han efectuado con el fin de otorgar homogeneidad a la transcripción y nunca una actualización que rompa con la índole y la época de la obra.

En resumen, en este apartado se ha tenido en cuenta el uso de la puntuación como medio para estructurar funciones metalexigráficas o guiar el discurso contenido en los volúmenes y, en la siguiente sección, se hará referencia a las circunstancias de la forma de los lemas, esto es, cómo se han transcrito y editado las letras capitales y la acentuación del *Vocabolario* de Franciosini.

### 4.3 *Los problemas en la forma*

En este apartado se analizarán los problemas encontrados en el *Vocabolario* en cuanto a las oscilaciones entre mayúscula/minúscula inicial (§ 4.3.1) y a la acentuación (§ 4.3.2).

#### 4.3.1 Alternancia entre mayúscula y minúscula: casos diferenciales

El interés de este apartado radica en la falta de homogeneidad del nivel ortográfico, dado que tanto en el primer volumen como en el segundo existen discordancias en el uso de las mayúsculas sobre las minúsculas. Los lemas suelen ir con minúscula (excepto en los casos apuntados en § 4.1), pero la tendencia es la de utilizar mayúscula en colocaciones, frases hechas, modismos, refranes, etc., aunque no de forma sistemática. Véase (25), *s. v. aprieto*, donde la expresión española sí se encuentra en letra capital, de la misma forma que su primer equivalente en italiano, sin embargo, no ocurre lo mismo con sus otros equivalentes:

(25) Veome en grande aprieto. [ Mi veggo a cattui termini. a mali partiti. mi veggo in vn gran pericolo.

Ocurre algo similar en los casos que siguen, pero existe cierta divergencia entre la fraseología del idioma origen y la del idioma meta:

(26) a priessa. [ con fretta. a fretta.

(27) No ay priessa. [ Ci è tempo, cioè questo si può far con agio.

(28) Da os priessa. [ fate presto. finitela. sollecitate. studiateui.

Vuelve a no ser sistemático en un ejemplo como (29):

(29) a puros ruegos lo alcançamos. [ l' hauiamo conseguito, o l'ottenemmo, perche ci valle il pregare, o per via di preghi.

En el caso del refranero, nos encontramos con la misma situación. Tal y como se puede comprobar en (30):

(30) Quien mucho abarca, poco aprieta [ Chi molto abbraccia, poco strigne.

En general, las formas de discurso repetido se escriben con mayúscula inicial para resaltar el carácter diferencial de estas construcciones con respecto a los lemas convencionales. Sin embargo, esta tendencia no se ve reforzada sempre, dado que

no se aprecia una sistematización sobre este aspecto a lo largo de los dos volúmenes.

En cuanto a la alternancia entre mayúscula y minúscula, advertimos del caso diferencial de la letra A del volumen italiano-español: al comienzo de la sección, todo se encuentra en letras capitales<sup>7</sup>. Los lemas que se recogen en las primeras páginas con mayúscula son expresiones o colocaciones en las que se usa *a* con función de preposición (31, 32, 33) o bien casos en que se explica el uso de *a* cuando se encuentra sola en un determinado contexto gramatical (34, 35):

- (31) A uno a uno. [ Uno a Uno.
- (32) A branchi. [ A manadas.
- (33) A corpo à corpo. [ Cara a cara.
- (34) A tal volta serve di particola riempitiva. Dolente, e povero à me. [ Pobre, negro, o desdichado de mi.
- (35) A, ogni volta che è con i verbi di luogo, ò stato significa in. Quando questo accade io e. o [sic] à Siviglia. [ Quando esto acontecio yo estaba en Sevilla.

En concreto, se puede comprobar que en las primeras 20 páginas de la letra A en el volumen italiano-español la situación es la que sigue:

	Mayúsculas	Minúsculas
Entradas en italiano	281	710
Equivalentes en español	573	418

Tabla 1. Primeras 20 páginas de la letra A en italiano

En este sentido, en las entradas en italiano, se puede ver como en las páginas iniciales (en concreto de la 35 a la 37), los lemas se escriben solamente en mayúscula (se trata de los casos reseñados anteriormente). No obstante, esta tendencia cambia paulatinamente a partir de la página 38, puesto que se empieza a usar la primera letra de la entrada en minúscula. También con respecto a los equivalentes en español, a pesar de encontrarse en mayúscula en las primeras páginas, la tendencia general del diccionario es la de escribirlos en minúscula: el uso de la mayúscula se difumina a lo largo del volumen, aunque se dé en ciertos momentos puntuales<sup>8</sup>.

7 En el volumen español-italiano, en cambio, se detecta el uso de la mayúscula inicial solamente en las primeras 17 entradas, antes del comienzo de la subsección “AB”.

8 En el intento de comprender si Franciosini les da un valor especial a estos elementos ortotipográficos, se han buscado las voces *maiuscola* y *minuscola* en el cuerpo textual del diccionario, con el objetivo de averiguar si se ofrece al respecto una definición o alguna otra información. Sin embargo,

A pesar de la heterogeneidad descrita, se concreta que mayoritariamente los volúmenes cuentan con una lematización en letra minúscula. Por lo tanto, excepto los casos ya reseñados en § 4.1, en nuestra edición los lemas se editarán de manera armónica con minúscula inicial para el conjunto de la obra.

#### 4.3.2 Acentuación

Atendiendo al criterio de no modernizar la lengua y su representación, en nuestra edición mantenemos la acentuación del texto original. Asimismo, la conservación de la acentuación que exhibe el *Vocabolario* no dificulta en absoluto la interpretación del texto por parte del lector/usuario de *TELEI*. Sin embargo, existen ciertas particularidades de la acentuación que serán objeto de un aparato crítico para la futura edición anotada.

Como se puede apreciar en (36) y (37), se acentúa *figliastro* pero no *figliastra*. En este sentido, aquí en vez de sistematizar señalamos con una nota crítica que se respeta la ortografía del original:

(36) antenado, o alnado. [ figliastro.

(37) antenada, o alnada. [ figliastra.

De la misma forma, existen otros ejemplos donde Franciosini en un lapso muy breve acentúa o no acentúa ciertos lemas:

(38) assentarse o sentarse. [ porsì à sedere. sedère.

Sin embargo, como ya se ha dicho, hemos creído conveniente mantener el diccionario conforme se encuentra en origen, esto es, acentuado unas veces y no acentuado en otros casos análogos. Con todo, se anotará en la edición, a partir de la transcripción efectuada, lo que resalta particularmente de la microestructura del diccionario.

---

no hay rastro de las dos palabras en ninguno de los dos volúmenes. Tampoco se documentan en la primera edición del *Vocabolario degli Accademici della Crusca* (1612), que ve la luz pocos años antes del *Vocabolario* de Franciosini y que nuestro autor tuvo la posibilidad de consultar. De hecho, hay que esperar la tercera edición del repertorio de la Crusca (1691) para que una de estas voces —*maiuscola*— se consigne en la nomenclatura.

#### 4.4 Otras características singulares: unas calas

En este apartado nos focalizamos en la clasificación de algunos elementos singulares cuyo hallazgo resulta funcional no solamente en la creación de criterios de transcripción, sino también en la fase de etiquetado.

##### 4.4.1 Ejemplos de uso

Con el intento de detectar un esquema común entre las dos partes del *Vocabolario*, funcional a la fase sucesiva del proyecto, esto es, el etiquetado, se ha investigado sobre el uso del ejemplo en los dos volúmenes. En términos generales, el esquema que Franciosini sigue en la presentación de los ejemplos es el siguiente en ambas partes del diccionario<sup>9</sup>:

*Lema L1 → equivalente L2*  
*Ejemplo L1 → traducción ejemplo L2*

Italiano-español

(39) abbacinato. [ cegado, o ciego. encandilado.  
specchio abbacinato, o appannato.  
[ espexo apañado.

Español-italiano

(40) abatirse. [ avvilirsi. perdersi d'animo.  
calarsi così.  
abatiose el aguila al suelo. [ l'aguila si calò in  
terra.<sup>10</sup>

Sin embargo, como en otro caso expuesto anteriormente (§ 4.3.1), las primeras páginas de la letra A constituyen un caso diferencial, puesto que en ellas encontramos una divergencia sustancial entre el resto de los dos tomos, esta vez en relación con los ejemplos de uso. En cuanto al volumen italiano-español, las primeras 122 entradas hacen referencia a ocurrencias de la voz *a* con función de preposición, en un determinado contexto gramatical o en expresiones o colocaciones. En todos estos casos, la fórmula general que acabamos de presentar no resulta perfectamen-

9 El esquema que aquí se presenta se centra en los ejemplos de usos, dejando de lado todos los demás elementos que pueden encontrarse en la microestructura e interpolarse entre el lema y el ejemplo o seguir este último (remisiones, informaciones gramaticales, etc.).

10 Los ejemplos (39) y (40) reproducen las sangrías del original, aunque no la división en líneas; ambos elementos carecen de representación en la versión digital, en la que la disposición del texto no se reproduce gráficamente, sino a través de la relación de dependencia entre entrada principal y subentrada.

te aplicable, ya que carecen de uno de los cuatro elementos que la constituyen: el equivalente en español (41, 42, 43). Por tanto, en dichas entradas del primer volumen el esquema general es sustituido por el siguiente: *lema it.* → *información metalingüística del valor del lema* → *ejemplo de uso it.* → *traducción del ejemplo al esp.*<sup>11</sup>:

(41) A, in vece di á suo. Quelle donne lo dissero a mariti. [ Aquellas mujeres lo dijeron a sus maridos.

(42) A, in vece di preposizione del caso accusativo. Io ho gran voglia d'andare à Belmonte. [ Tengo grande gana de yr à Belmonte.

(43) A, in vece di sopra. Montai a cavallo. [ Subì à cavallo.

A diferencia de las entradas anteriores, nos encontramos con un esquema donde también aparece una explicación metalingüística del valor del lema. Esto se debe sin duda, en casos como (41)-(43), a la voluntad de Franciosini de distinguir los muchos valores que tiene la preposición *a*.

#### 4.4.2 Información complementaria

Otro aspecto relevante es el de las estrategias de presentación de la información lingüística complementaria en la entrada lexicográfica, que es espejo de las voluntades prácticas de la obra. Al mismo tiempo, su hallazgo y clasificación resulta relevante —como veremos— para la fase de etiquetado.

Los primeros dos casos que relatamos se relacionan con uno de los elementos que, en las primeras fases del proceso de transcripción, ha sido detectado como recurrente en el *Vocabolario*: el corchete, cuyo valor metalexicográfico reside, como hemos señalado, en su función de separador entre lema y equivalente(s). Más allá de esta regla general, en un análisis comparativo entre las dos partes del diccionario, ha sido posible detectar casos especiales relativos a su uso.

Como se ha anticipado en § 4.2.1, se ha constatado la aparición de un corchete de cierre, además del corchete de apertura, como forma de separación entre las lenguas: se trata de un mecanismo sistemático solo en el volumen italiano-español, mientras que en el volumen español-italiano ocurre de forma muy esporádica. El esquema que se sigue es: *lema it.* [ *lema es.* ] *Ejemplo it.* [ *Ejemplo es.*,

11 Lo presentamos como caso diferencial del primer volumen, ya que en el segundo se omite esta cantidad de ocurrencias de *a* en distintos contextos gramaticales: todo queda condensado en la entrada *A preposicion.* [ *a.*

fórmula que Franciosini usa cuando, en una misma entrada, se pasa más de una vez de una lengua a otra, como muestra el ejemplo a continuación:

(44) aspettare. [ aguardar. esperar. ] aspettatemi, che adesso vengo. [ aguardadme, que yà voy.

Por el hecho de tratarse de un elemento sistemático en el primer volumen en los casos de vaivén idiomático, se ha tomado la decisión de poner el corchete de cierre cuando ausente.

Un segundo caso reseñable tiene que ver con la introducción de frases hechas, modismos y significados metafóricos de algunos lemas y, más específicamente, con la lengua en la que se presentan. La obtención de un texto plano de cada letra ha facilitado el proceso de comparación entre los dos tomos: después de haber terminado la transcripción de la letra A, se ha hecho una búsqueda en ambos volúmenes, realizada a través de *metafor* y *modo d’/di* en italiano y *manera de* y *metafor* en español. El resultado de este cotejo muestra, en primera instancia, que no hay ninguna ocurrencia de *manera de* y *metafor*(-ico, -icamente) en contextos de introducción de informaciones de naturaleza metafórica o fraseológica. En consecuencia, resulta clara una diferencia en la presentación de tales informaciones: el modismo, la frase hecha, o el significado metafórico se presenta en italiano. Esto implica que en el primer tomo del diccionario esta información se presenta antes del corchete, es decir, en la lengua de entrada, mientras que en el segundo tomo se presenta después, esto es, en la lengua de llegada<sup>12</sup>.

#### Italiano-español

- (45) asino metaforicamente per rozzo, zotico, e di grossi costumi. [ asno.  
(46) attignere metaforicamente si intende per penetrare i concetti altrui in fatti, & in parole. [ sacar. colegir. entender. alcançar.  
(47) A farla male, modo di dire. [ A mal suceder.

#### Español-italiano

- (48) alto. [ metaforicamente si trasferisce all’animo, e significa grande, e profondo, nascosto, e segreto.  
(49) alumbrar à uno. [ metaforicamente significa guidare, e metter uno per la buona strada.  
(50) alla se lo aya con su herencia. [ modo di dire hironico, e vale tenghisi la sua redità che buon pro gli faccia.

Los últimos casos de información lingüística complementaria que presentamos son mecanismos más frecuentes en el volumen italiano-español y tienen que ver con el uso, en la lengua de entrada, de definiciones del lema (51) y de aclaraciones

12 En concreto, a través de esta búsqueda se han encontrado 12 ocurrencias de este tipo en la letra A de la primera parte del diccionario, y 13 casos en la segunda.

sobre el significado del lema o de un equivalente (52):

(51) abbagliare, l'effetto, che fa il Sole o il lume quando ferisce ne gli occhi. [ encandilar, o deslumar.

(52) baldacchino che si tiene nelle stanze de Principi, e gran personaggi in segno d'honore. [ dosèl.

En la fase de etiquetado distinguimos casos como (51) de casos como (52): aunque ambos aportan información lingüística, (51) contiene una definición, por lo tanto, esta se codifica a través del *tag* anidado *definition*. En cambio, (52) presenta una muestra de glosa (etiquetada como *gloss*) que, al servir para orientar un significado, se especifica con el atributo *semantic*.

Por último, señalamos que no siempre la información complementaria es de naturaleza estrictamente lingüística. De hecho, de vez en cuando, Franciosini interpola pequeños cuentos que guardan relación con el lema en cuestión:

(53) ammazzar uno per niente, cioè per leggier causa. [ matar a uno por quitame allà essas pajas. Un Gentil'huomo teneva in su casa un pazzo, & un Cameriere pigliandosi spasso con lui gli diceva. Io vi voglio ammazzare, guardatevi, che io voglio ammazzare. Il pazzo se n'andò alla volta del Padrone, egli disse. Il vostro Cameriere mi vuole ammazzare; il Padrone gli rispose se t'ammazza, io l'impiccherò. Replicò il pazzo; non voglio che tu l'impicchi se non va giorno innanzi, che ammazzi me. [ Un Cavallero tenía [...]

Como es evidente, estas glosas tienen una naturaleza distinta de las de índole semántica: para garantizar una codificación coherente, codificamos también estas muestras textuales como *gloss*, pero su atributo pasa a ser *cultural*.

Los casos expuestos son indicios de que la explicación de una palabra se presenta de forma más exhaustiva en italiano, la lengua materna de Franciosini, razón que se combina con la voluntad didáctica de la obra descrita en la portada por el mismo autor, la cual implica una muestra de la lengua y de su uso concreto: “[...] Con le Frasi, & alcuni Proverbi, che in ambe due le lingue giornalmente occorrono”.

Más allá de las características intrínsecas del diccionario, un análisis de este tipo lleva a reflexionar sobre la utilidad y las posibilidades de estudio que un texto plano ofrece, permitiendo búsquedas cruzadas tanto entre los dos tomos del diccionario de Franciosini —nuestro objeto de estudio—, como entre diccionarios distintos, oportunidad que posibilitará *TELEI*. Agregamos que, además, a partir de los hallazgos en el texto digital de la *editio princeps*, se vuelve más viable la comparación de aspectos como los aquí reseñados con ediciones sucesivas de



este mismo diccionario. Estas ventajas conllevan la consecuencia implícita de la aplicabilidad del proceso descrito en la presente contribución a otras obras lexicográficas, para su estudio y difusión.

## 5. Conclusiones

La transcripción armónica y coherente del conjunto del *Vocabolario* de Lorenzo Franciosini ha requerido de una labor sistemática, mantenida en el tiempo y reflexiva.

En nuestra contribución, por un lado, hemos presentado las fases y metodología de trabajo que han sustentado la transcripción de esta obra lexicográfica y, por el otro, hemos reseñado los principales desajustes y peculiaridades detectados a lo largo del —y gracias al— dilatado proceso que nos llevará, como último paso, a la codificación del diccionario en XML-TEI. Además, hemos expuesto y justificado los criterios de edición en los cuales se rigen los limitados casos de intervención en el texto: como hemos mostrado, se proponen unas normas de edición que retoquen mínimamente el cuerpo del texto y que conserven las características propias de la época y del estado de ambas lenguas involucradas. De hecho, los casos en que el texto se interviene se limitan a la corrección de erratas, la regularización de *u* frente a *v* y la armonización de los elementos metalexicográficos que, por su propia índole, consideramos oportuno homogeneizar. Estos últimos, por lo tanto, se reponen en casos de olvido y se sustituyen cuando su lugar en la estructura diccionarística lo ocupa otro signo ortográfico. Es lo que hemos visto con el punto como elemento para marcar el fin del lema o la separación entre equivalentes, o el corchete como elemento separador entre lema y microestructura. En esta misma línea, hemos cambiado ciertas mayúsculas iniciales por sus minúsculas correspondientes en el caso de una parte de los lemas de la letra A de ambos volúmenes, con el propósito de presentar de manera coherente la macroestructura del conjunto del *Vocabolario*.

Los criterios presentados no se han establecido *a priori*, sino que se han fijado gracias a la transcripción y a la acumulación de notas acerca de desajustes y características peculiares de la obra. Las dudas, en resumen, se han sorteado gracias al conocimiento empírico del conjunto del diccionario. Al respecto, es preciso considerar que nuestro trabajo de revisión y las notas aludidas redundarán en la realización de un aparato crítico que sirva para apuntalar de manera fiable la consulta de la edición digital del *Vocabolario* de Franciosini.

Por último, la inclusión de este insoslayable eslabón de la lexicografía bilingüe italoespañola en *TELEI* contribuirá a un conocimiento más profundo de los diccionarios considerados y de su tradición textual, gracias a las amplísimas posibilidades de búsqueda y la oportunidad de poner en relación los datos que garantizarán la plataforma diseñada para albergar el tesoro y la codificación en XML-TEI.

## Bibliografía citada

- Accademia della Crusca (1612, 1623, 1691, 1729-1738), *Lessicografia della Crusca in rete (Vocabolario degli Accademici della Crusca)* [09/03/2025], <<http://www.lessicografia.it>>.
- Cervantes Saavedra, Miguel de (1622-1625), *L'ingegnoso Cittadino Don Chisciotte della Mancia*, trad. it. Lorenzo Franciosini, Venetia, Andrea Baba, 2 vols.
- Casas, Cristóbal de las (1570), *Vocabulario de las dos lenguas toscana y castellana*, Sevilla, Alonso Escriuano.
- Castillo Peña, Carmen (2020), “Epigrama: Un portal para la edición digital de textos gramaticales”, *Anales de Lingüística*, 4: 201-17.
- Castro-Prieto, M.<sup>a</sup> Rosa (2022), “Recuperación del léxico especializado del *Diccionario Tecnológico Hispano Americano* (1930). Diseño de un proyecto XML-TEI”, *Cuadernos del Instituto de la Lengua*, 15: 209-50.
- Franciosini, Lorenzo (1620), *Vocabolario italiano e spagnolo*, Roma, Gio. Paolo Profilio, 2 vols.
- Franciosini, Lorenzo (1624), *Gramatica spagnola e italiana*, Venetia, Sarzina.
- Gallina, Anna Maria (1959), *Contributi alla storia della lessicografia italo-spagnola dei secoli XVI e XVII*, Firenze, Olschki.
- Landucci, Nicolao (1562), *Il dizionario spagnolo-italiano*, eds. Elena Liverani y Alessandro Parenti, Firenze, Olschki [Biblioteca Nacional de España, ms. 8431].
- Martínez Egido, José Joaquín (2002), *La obra lexicográfica de Lorenzo Franciosini: Vocabulario italiano-español, español-italiano (1620)*, Alicante, Biblioteca Virtual Miguel de Cervantes. [30/01/2025] <<https://www.cervantesvirtual.com/obra/la-obra-lexicografica-de-lorenzo-franciosini-vocabulario-italianoespanol-espanolitaliano-1620-0/>>.
- READ-COOP SCE (2024), *Transkribus* [03/02/2025] <<https://www.transkribus.org/>>.
- Real Academia Española; Asociación de Academias de la Lengua Española, *Diccionario de la lengua española*, 23.<sup>a</sup> ed. [04/02/2025] <<https://dle.rae.es>>.
- San Vicente, Félix (2010), “Diccionarios y didáctica en la tradición italoespañola (siglos XVI-XVII)”, *Diccionario y aprendizaje del español*, eds. Stefan Ruhstaller; María Dolores Gordón. Bern, Berlin, Bruxelles, Frankfurt am Main, New York, Oxford,

Wien, Peter Lang: 47-88.

TEI Consortium (2024), *Guidelines for Electronic Text Encoding and Interchange*, versión 4.8.0 del 02/09/2024 [03/02/2025] <<https://guidelines.tei-c.de/en/html/index.html>>.

Valencia Mirón, María Dolores; Peña, Victoriano (2003), “Teoría y práctica de la traducción hispano-italiana en el siglo XVII: reflexión gramatical y labor traductora de Lorenzo Franciosini”, *Seis estudios sobre la traducción en los siglos XVI y XVII*, coords. José Antonio Sabio Pinilla; María Dolores Valencia Mirón. Valencia, Granada, Comares: 119-67.

Wiegand, Herbert Ernst (2007), “Aspectos escogidos de una teoría de las estructuras textuales de los diccionarios”, *Reflexiones sobre el diccionario*, eds. Mar Campos Souto; Eugenia Conde Noguero; José Ignacio Pérez Pascual; José-Álvaro Porto Dapena. A Coruña, Universidade da Coruña: 35-63.

Zappulli, Andrea; Iorio, Sabrina (2019), “La digitalizzazione dell’Archivio Storico del Banco di Napoli”, *DigItalia*, 13(2): 46-51 [04/02/2025] <<https://digitalia.cultura.gov.it/article/view/2169>>.

**Amparo Alemany Martínez** es *assegnista di ricerca* en la Universidad de Verona. Sus intereses de investigación incluyen la semántica y la lexicología teóricas, la marcación discursiva, el análisis del discurso oral-coloquial. Participa en el PRIN 2022 “Un nuovo ambiente digitale per il recupero del patrimonio lessicografico: il *Tesoro digitale della lessicografia bilingue spagnolo-italiano*”.

**amparo.alemanymartinez@univr.it**

**Matteo De Beni** es catedrático de Lengua, Traducción y Lingüística Españolas en la Universidad de Verona. Sus principales intereses científicos son la lexicografía, la lexicología y la historia de la lengua española. Coordina el grupo de investigación *El léxico del español en su historia (LEHist)* y participa, entre otros proyectos, en el PRIN 2022 “Un nuovo ambiente digitale per il recupero del patrimonio lessicografico: il *Tesoro digitale della lessicografia bilingue spagnolo-italiano*”.

**matteo.debeni@univr.it**

**Alessandra La Manna** es doctoranda en Lenguas y Literaturas Extranjeras en la Universidad de Verona. Su principal interés académico es la lexicografía histórica. Su tesis doctoral se centra en el estudio comparado de distintas ediciones del *Vocabolario italiano e spagnolo* de Franciosini y en su digitización. Participa en el PRIN 2022 “Un nuovo ambiente digitale per il recupero del patrimonio lessicografico: il *Tesoro digitale della lessicografia bilingue spagnolo-italiano*”.

**alessandra.lamanna@univr.it**



**FLORENCIA FERRANTE**  
**CHIARA VALENTE**  
**ANA LOURDES DE HÉRIZ**  
**EL LEMARIO ITALIANO-ESPAÑOL**  
**DEL *DICCIONARIO DE FALTRIQUERA***  
**DE CORMON Y MANNI (1805):**  
**TRANSCRIPCIÓN AUTOMÁTICA Y**  
**CLAVES ECDÓTICAS**

Università degli Studi di Genova

**Resumen**

El *Diccionario de faltriquera* italiano-español y español-italiano de Cormon y Manni (1805) es una de las obras transcritas para su digitalización e inclusión en el *Tesoro TELEI*. El artículo expone las características tipográficas de la macroestructura y microestructura del lemmario italiano-español que han supuesto una dificultad en la transcripción automática, así como las soluciones que se han tomado tras el cotejo de fuentes primarias y obras publicadas posteriormente en las que este diccionario influyó.

palabras clave: lexicografía bilingüe, italiano-español, transcripción automática, Cormon y Manni

**Abstract**

***The Italian-Spanish section of Cormon and Manni's Dictionary (1805): automatic transcription and ecdotic clues***

*The Diccionario de faltriquera italiano-español y español-italiano by Cormon and Manni (1805) is one of the dictionaries transcribed for digitalisation and publication in the Tesoro TELEI. This paper describes the typographical characteristics of the macrostructure and microstructure of the Italian-Spanish section, which have represented a difficulty in the automatic transcription, as well as the solutions that have been taken, also after a comparison with the primary sources and later published dictionaries on which this dictionary had an influence.*

keywords: bilingual lexicography, Italian-Spanish, automatic transcription, Cormon & Manni

## 1. Introducción

Las páginas que siguen<sup>1</sup> presentan una investigación aplicada que se inscribe en el marco de un proyecto<sup>2</sup> cuyo objetivo general es la retrodigitalización del patrimonio lexicográfico bilingüe italiano-español (s. XVI-XX) para su encontrabilidad, reutilización, interoperabilidad y accesibilidad universales, según los principios FAIR (Wilkinson *et al.* 2016). El proyecto se sitúa en una tradición reciente de estudios que intentan combinar, en lo que se ha denominado un verdadero *paradigm shift*, la lexicografía, la ontología y las herramientas de la lingüística computacional en las Humanidades Digitales (Costa *et al.* 2021).

El término ‘retrodigitalización’ suele hacer referencia, por un lado, al proceso de transformación en documento digital legible por máquina (PDF, por ejemplo) de un objeto textual publicado en papel antes del advenimiento de las Humanidades Digitales; por el otro, ‘retrodigitalizar’ significa también crear, a partir de la fuente antigua en papel, un objeto digital en formato de archivo estructurado y codificado en XML, que permita la interoperabilidad con el usuario y la búsqueda y selección de datos y metadatos específicos de cada objeto a partir de una interfaz interactiva (Castillo Peña 2020, Nalesso 2024). Huelga decir que el tipo de retrodigitalización al que aspira el proyecto en curso es del segundo tipo, o sea el diseño y la creación de un tesoro de la lexicografía bilingüe italiano-español en el que los objetos textuales digitales permitan a cualquier tipo de usuario la realización de búsquedas parametrizadas complejas de información semántica y estructural, así como también la comparación y alineación de los contenidos seleccionados entre los distintos objetos digitales.

El primer paso de la investigación ha consistido, por tanto, en el desarrollo de un flujo de trabajo para experimentar un protocolo de retrodigitalización de un corpus reducido de obras lexicográficas bilingües, escogidas según cuatro criterios fundamentales descritos y justificados en Nalesso (2024: 139): cronológico, tipológico, historiográfico y metodológico.

Cada una de las universidades que participan en el proyecto TELEI (Padua, Génova, Bolonia, Pisa, Turín y Verona) constituyen sendos grupos de investiga-

1 Las tres autoras han colaborado a la par en el proyecto de este artículo y en la revisión final. No obstante, cabe precisar que las secciones §1 y 2 se deben a Florencia Ferrante, las secciones §3, 4, 5 y 6 a Chiara Valente y las secciones §7 y 8 a Ana Lourdes de Hériz.

2 El proyecto lleva como título *Un nuevo espacio digital para el patrimonio lexicográfico: el ‘Tesoro digital de la lexicografía bilingüe español-italiano’*, tesoro digital representado por la sigla TELEI. Esta investigación está siendo financiada por La Unión Europea-Next Generation EU, Missione 4 Componente 1 CUP C53D23004010006 / PRIN 2022 MUR 20229W73WR.

ción que se ocupan del proceso integral de retrodigitalización de una o más de las obras lexicográficas seleccionadas para el *Tesoro* digital. Así, cada grupo tiene a cargo objetos textuales distintos con características específicas. El grupo de la Universidad de Génova, cuyo protocolo de trabajo se presenta en estas páginas, tiene asignada la retrodigitalización de los dos volúmenes del *Diccionario de faltriquera italiano-español y español-italiano* de Cormon y Manni (primera edición en Lyon, 1805)<sup>3</sup>.

A la selección de los materiales, la asignación de estos a cada uno de los grupos y el estudio de las principales características filológicas y metalexicográficas del corpus, sigue el proceso de retrodigitalización propiamente dicho. Obviamente, estaba previsto que, antes de la fase de codificación y modelización de los textos para su etiquetación semántica en formato XML según las directrices TEI Lex-0<sup>4</sup>, se llevara a cabo la transcripción de los diccionarios, la mayoría de los cuales ya estaban disponibles en formato digital (.pdf o .jpg, por ejemplo), para su transformación en documentos legibles por máquina (en formatos como .txt o .doc, por ejemplo). Para ello, cuando se ha optado por una transcripción automática y no manual, se ha utilizado el programa de reconocimiento y transcripción de documentos antiguos *Transkribus*<sup>5</sup>, desarrollado por la Universidad de Innsbruck en el marco del proyecto READ<sup>6</sup> (Rabus 2019) y hoy ampliamente utilizado en los estudios filológico-literarios.

*Transkribus*, sin embargo, no se ha experimentado hasta ahora exhaustivamente para el reconocimiento y la transcripción de materiales lexicográficos. Estos

3 A este diccionario se le han dedicado extensos estudios por ser uno de los textos fundamentales de la lexicografía italoespañola del siglo XIX (San Vicente 2010; Martínez Egidio 2008, 2010; Alvar Ezquerro 2010). Remitimos a estas referencias bibliográficas a quien no conozca esta obra, que atravesó con muchas reimpresiones y ediciones revisadas todo el siglo XIX y que fue fuente primaria de otras obras lexicográficas.

4 TEI Lex-0 es una especificación técnica y también un conjunto de recomendaciones generales para codificar diccionarios legibles por máquina. Está basado en las directrices generales de la *Text Encoding Initiative* (TEI) y se ofrece como una personalización de sus directrices generales. El principal propósito del proyecto TEI Lex-0 es el de proporcionar un esquema de codificación simplificado y estandarizado para cualquier tipo de recurso lexicográfico antiguo o moderno, así como también para otros tipos de texto con características similares (enciclopedias, catálogos), que sea fácil de adoptar y utilizar por parte de los editores y que permita, además, la interoperabilidad y la reutilización de los contenidos.

5 Transkribus® READ-COOP SCE [22/03/2025] <<https://www.transkribus.org/>>.

6 *European Commission Horizon 2020 Research and Innovation Programme, Recognition and Enrichment of Archival Documents* (READ [22/03/2025] <<https://readcoop.org/>>).

tienen, como es sabido, una serie de características tipográficas, macro y microestructurales específicas, que no siempre son de fácil procesamiento para este tipo de *software*. Asimismo, el diccionario asignado a nuestro grupo de investigación, el *Diccionario de faltriquera italiano-español y español-italiano* de Cormon y Manini (Lyon 1805) ha supuesto, por una serie de motivos de los que daremos cuenta enseguida, un desafío aun mayor para la transcripción automática.

A continuación, tras un breve estado de la cuestión acerca del uso de *Transkribus* en proyectos de corte lexicográfico (§2), se dará cuenta de las dificultades para el reconocimiento y la transcripción del texto macro y microestructural de la obra y de las soluciones adoptadas, para las que ha sido a menudo necesario tomar en consideración la historia editorial y filológica de esta obra.

## 2. *Transkribus* y diccionarios

La creación, desarrollo y validación del protocolo de retrodigitalización de materiales lexicográficos bilingües de italiano-español es uno de los objetivos específicos del proyecto TELEI. Uno de los pasos esenciales para generar un flujo de trabajo eficiente es la elección de las herramientas informáticas adecuadas. Entre los recursos informáticos más difundidos hoy en día para la retrodigitalización de textos antiguos y modernos se encuentran los siguientes: el ya mencionado *software Transkribus*, especializado en el reconocimiento y transcripción automática de documentos históricos impresos y manuscritos, y GROBID<sup>7</sup>, un programa también configurado para el reconocimiento y estructuración automática de distintos tipos de documentos (PDF, por ejemplo) en un archivo en formato XML/TEI. Sobre este programa se ha desarrollado, en los últimos años, un módulo específico para la codificación en formato XML/TEI Lex-0 de cualquier tipo de documento lexicográfico en formato digital, conocido como GROBID *Dictionaries* (Khemakhem *et al.* 2017).

La elección de *Transkribus* para el reconocimiento y transcripción del corpus experimental de nuestro proyecto fue una decisión ponderada por todos los grupos de esta investigación, en la medida en que nos permite registrar y evaluar la utilidad, los problemas, las ventajas y desventajas de su aplicación para la digitalización de documentos lexicográficos y sentar un precedente para investigaciones y proyectos futuros. El proceso compartido de la experimentación, las dificultades

---

7 GROBID (GeneRation Of Bibliographic Data) [22/03/2025] <<https://grobid.readthedocs.io/en/latest/>>.



encontradas, así como las soluciones adoptadas, son parte, pues, de los resultados de la investigación en curso.

*Transkribus*, como ya se ha indicado más arriba, no ha sido utilizado masivamente para la transcripción y digitalización de textos lexicográficos. Entre los escasos empleos de este *software* en proyectos de digitalización de diccionarios o textos similares como bibliografías, se puede mencionar el estudio de Lindemann *et al.* (2018), quienes integran el uso de *Transkribus* con GROBID *Dictionaries*. En otro proyecto similar más reciente, Lindemann y Alonso (2021) presentan un flujo de trabajo para la digitalización de diccionarios históricos, pero optan por utilizar *Kraken*<sup>8</sup>, un programa con una función equivalente a la de *Transkribus* pero con características y especificaciones distintas, más adecuadas para resolver cierto tipo de requerimientos infraestructurales. Otros dos grandes proyectos de mención obligada de retrodigitalización de obras lexicográficas llevados a cabo en los últimos años son el *Nenúfar Project*<sup>9</sup> y el *MorDigital Project*<sup>10</sup>. El *Nenúfar Project*, cuyo objetivo es la retrodigitalización de distintas ediciones del diccionario monolingüe *Petit Larousse Illustré*, utiliza también el *software* GROBID *Dictionaries* sobre un corpus del que ya se habían obtenido archivos legibles por máquina (Bohbot *et al.* 2018). El *MorDigital Project*, con el fin de retrodigitalizar las primeras tres ediciones del *Diccionario da Lingua portuguesa* de Antonio de Moraes (publicado en 1789), utiliza también GROBID *Dictionaries* para generar directamente, a partir de los documentos en .pdf, un archivo XML estructurado en formato TEI Lex-0 (Costa *et al.* 2021). Como puede verse, los proyectos más recientes de retrodigitalización a gran escala del patrimonio lexicográfico monolingüe en lenguas románicas parecen haber optado por otras herramientas de transcripción, digitalización y codificación como *Kraken* o GROBID *Dictionaries* en lugar de *Transkribus*.

Las dificultades más frecuentes a las que se suele hacer referencia en los estudios recién citados, por lo que atañe a la transcripción automática para la posterior codificación de textos lexicográficos, son: i) la frecuente baja calidad de las imágenes del documento fuente (Khemakhem *et al.* 2019); ii) las características tipográficas del texto antiguo, dado que suelen ser fuente de problemas para el reconocimiento automático de caracteres; y iii) respecto a algunos casos específicos, la escasa correspondencia entre la macro y microestructura del diccionario y las etiquetas establecidas para la codificación (Lindemann, Alonso 2021).

8 [22/03/2025] <<https://kraken.re/main/index.html>>.

9 [22/03/2025] <<https://nenufar.huma-num.fr/presentation/>>.

10 [22/03/2025] <[mordigital.fcsh.unl.pt/en/homepage/](https://mordigital.fcsh.unl.pt/en/homepage/)>.

### 3. Características tipográficas de valor metalexicográfico

El leuario que ya ha sido transcrito completamente (tomo I, del italiano al español) es una de las dos partes de lo que el título de la 1ª edición de la obra presenta como un diccionario de bolsillo (“de faltriquera”). Este formato respondía a necesidades de tipo comercial y escolar (Cazorla Vivas 2010: 32-33) y obligaba a la maquetación de una macroestructura y una microestructura peculiares. Si se observa una página del leuario, se nota enseguida una disposición de los elementos de las entradas que prima el ahorro de espacio, con el objetivo de incluir el mayor número de palabras posible en una obra de solo 14 cm de largo.

El tomo I contiene un leuario de 422 páginas; a este se añaden los siguientes paratextos iniciales, que no hemos transcrito por ahora: el prólogo (páginas I-II), el apartado dedicado a las “conjugaciones de los verbos regulares e irregulares” (III-XXII), la lista del catálogo “d’alcuni libri italiani e spagnuoli che si trovano nella stessa libreria” y la lista “explicación de las señales y abreviaturas usadas en este tomo”; lista que se ha tomado en consideración cuando han surgido dudas de interpretación del texto producido por el escaneado automático.

Por lo que se refiere al formato de las páginas del leuario, el texto está repartido en tres columnas; encima de cada una de ellas aparecen las primeras tres iniciales-guía de la primera palabra lematizada en dicha columna. En lo alto de la página, a la derecha o a la izquierda (en las pares o impares, respectivamente), están las cifras arábigas de numeración de cada página. En el pie de página, a la izquierda, cada dieciséis páginas aparece la inscripción *Tomo I*. En la primera de cada pliego (16º), empieza una serie de signaturas tipográficas representadas en orden alfabético por una letra y un número arábigo (A, A2, A3, A4, B, B2, etc.).

Cada columna incluye un promedio de unas 30 entradas, es decir, alrededor de 100 entradas por página, en un volumen de dimensiones muy reducidas. Para aprovechar al máximo este espacio limitado, la microestructura presenta algunas estrategias peculiares que, aunque no fueran innovaciones en la lexicografía del tiempo –como se detallará más abajo–, tampoco estaban muy difundidas.

La microestructura básica del artículo lexicográfico más simple se desarrolla del siguiente modo: el lema está escrito en redondilla y solo con la inicial mayúscula; le sigue una coma que lo divide de las abreviaturas de la categoría gramatical, separadas cada una de ellas por un punto<sup>11</sup>. Tras estas sigue, en cursiva, el equivalente de traducción. Si se propone más de un equivalente de traducción en

11 Se ha observado un uso no siempre coherente de las abreviaturas; algunas variantes no aparecen en la lista incluida en el volumen del tomo I del diccionario.

las entradas monosémicas o con más de una acepción de significado, se separan con una coma (es lo que se observa más abajo en la imagen 1 de los artículos de *Faccenda* y *Faccendiere*).

A partir de esta estructura básica, se pueden encontrar diferentes variantes. Si el artículo comprende más de una acepción de significado con sus correspondientes equivalentes de traducción (véase *Laidezza*, imagen 2), se separan mediante el símbolo de la doble daga (‡, explicada en la lista inicial de “señales” del diccionario como “señal para separar las varias acepciones de una misma voz”). Este símbolo se emplea también para diferenciar dentro de un artículo las subentradas por categoría gramatical; en este caso, a la doble daga le sigue la abreviatura de la nueva categoría gramatical (v. imagen 3 de *Preciso*, adjetivo y adverbio en italiano).

<b>Faccenda, s. f. hacienda</b> <b>Faccendiere, s. m. fachenda, fachendon</b>	<b>Laidezza, s. f. fealdad</b> <b>‡ obscenidad ‡ suciedad</b>	<b>Preciso, sa, a. preciso,</b> <b>positivo ‡ preciso, conciso ‡ adv. precisamente</b>
Imagen 1	Imagen 2	Imagen 3

Además de este uso del símbolo de la doble daga, resulta peculiar el recurso al signo gráfico de la llave de cierre en forma de arco para recoger entradas de lemas que comparten el mismo equivalente de traducción, pero que tienen una estructura morfológica diferente (normalmente con variación sufijal). A la izquierda de la llave se listan verticalmente los lemas y, a la derecha, se indica la categoría gramatical seguida del equivalente o los equivalentes de traducción (v. imagen 4, lematización de *Importunezza* e *Importunità*). Si cada lema tiene una categoría gramatical diferente (o incluso solo una diferencia de género) se indica antes de la apertura de la llave tras cada forma de las entradas (v. imagen 5 de *Potagione* y *Potamento*).

<b>Importunezza, } sub. f.</b> <b>Importunità, } impor-</b> <b>tunidad</b>	<b>Potagione, s. f. } poda-</b> <b>Potamento, s. m. } dura,</b> <b>poda</b>
Imagen 4	Imagen 5

En todo el texto de las tres columnas se puede apreciar además el uso muy frecuente de guiones que sirven para cortar una palabra y completarla en la línea siguiente. Este recurso tiene el evidente objetivo de aprovechar el espacio de las

líneas hasta el final. Sin embargo, y es esta otra de las peculiaridades tipográficas de la obra, muchas veces las palabras equivalentes de traducción se completan en la línea de arriba, si el espacio sobrante del artículo anterior lo permite, separando esa fracción de la palabra cortada del texto de la línea superior mediante el símbolo de apertura de un corchete. Es una estrategia que permite aprovechar al máximo el espacio, pero que hace más difícil la lectura, como puede verse en la secuencia de 4 lemas *Implicato-Implicitamente* de la imagen 6.

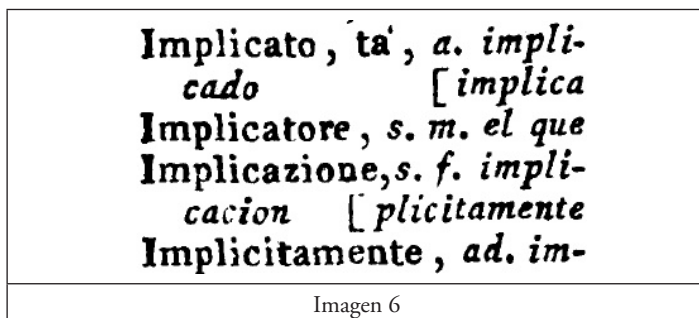


Imagen 6

Con todo ello no se debe pensar que consideramos el uso de estos símbolos y estrategias como algo extraño en la lexicografía de principios del s. XIX; sí, en cambio, que no era lo más frecuente. Por ejemplo, el uso del corchete de apertura para terminar una entrada en el espacio sobrante de la anterior ya se había visto en otros diccionarios compactos con espacio reducido como el de Gattel *Nuevo diccionario portatil español y francés*, de 1798 (Bruña Cuevas 2008: 54; Cazorla Vivas 2002: 311-28). El arco del símbolo de la llave que une varias formas flexivas de un lema se usaba también en diccionarios no portátiles como el *Dictionnaire italien, latin e françois* del Abad Antonini, publicado en París en 1743 (Lillo 2019: 152-54; Mormile 1993: 52-6). En el *Nuevo diccionario portatil español é ingles* de Gattel, publicado en 1803 (Lombardero Caparrós 2015: 149) –dos años antes que el Cormon y Manni de nuestro proyecto– coinciden ambas estrategias de ahorro de espacio, las llaves y los corchetes que encierran la continuación del artículo en la línea de arriba. Hemos dedicado un espacio a exponer estos recursos tipográficos por su valor metalexicográfico y por las complicaciones que han provocado en la fase de transcripción automática del texto, pues, como se verá en §5, *Transkribus* no los ha reconocido automáticamente.

## 4. El proceso de transcripción del tomo I del diccionario

Tal como se ha anticipado, la transcripción del diccionario seleccionado para el proyecto se ha llevado a cabo con el *software* en línea *Transkribus*. A continuación listaremos sucintamente los pasos que se han seguido para la transcripción, control, revisión y fijación definitiva del texto del tomo I, deteniéndonos luego en algunos detalles que consideramos que pueden ser interesantes.

### 4.1 *Pasos realizados*

1. Revisión de todos los ejemplares del tomo I a disposición en formato .pdf y elección del que se iba a transcribir.
2. Selección de las páginas del volumen que se iban a cargar en el *software* y subida de archivos PDF a *Transkribus*.
3. Transcripción con *Transkribus* en este orden: segmentación de la página en regiones (correspondientes a las tres columnas del leuario de cada página), escaneo, reconocimiento de las líneas y del texto mediante modelo público y generación de transcripción automática. Este proceso se ha aplicado inicialmente a las primeras 120 páginas del leuario (letras A, B, C, D y E), cuyo texto generado automáticamente se ha ido corrigiendo para educar al sistema de escaneo respecto a algunos errores de reconocimiento de letras y otras peculiaridades (como las presentadas en §5 y §6).
4. Creación de varios modelos privados de HTR (*Handwritten Text Recognition*) y aplicación al leuario segmentado en diferentes archivos PDF.

#### 4.1.1 Elección del ejemplar para la transcripción

Los ejemplares del tomo I de la edición de 1805 de este diccionario que teníamos a disposición en formato .pdf son de las siguientes bibliotecas<sup>12</sup>: Biblioteca de la Universidad Complutense de Madrid (a partir de ahora, ejemplar BUCM),

---

12 Nos basamos en los sellos que se pueden observar en las portadas u otras páginas de los ejemplares y en el hecho de que los números de registro que aparecen escritos en su interior coinciden con los que constan en los catálogos. Junto con la sigla de las bibliotecas mencionadas, indicamos la signatura del catálogo de estos ejemplares: BUCM (ID 50877); BNF (MAGL. 7.8.5 0001), BNE (U/2548). Sobre el ejemplar de la Biblioteca municipal de Forlì, véase la siguiente nota.

Biblioteca Nazionale di Firenze (ejemplar BNF), Biblioteca Nacional de España (BNE) y Biblioteca Municipale di Forlì (BMF)<sup>13</sup>.

Para la transcripción automática decidimos recurrir, tras varias pruebas, al ejemplar de la BUCM. Se tomó esta decisión por razones prácticas: ante todo, se descartaron los ejemplares de la BNE y de la BMF por contener ambos dos páginas del volumen en cada página del PDF (en formato apaisado). El hecho de presentar dos páginas del diccionario en una misma del PDF dificultaba aún más el reconocimiento de las regiones por parte del programa. De los otros dos ejemplares, se ha seleccionado el BUCM porque parecía que *Transkribus* lo escaneaba generando menos errores. Esta elección, sin embargo, no debe llevar a la conclusión de que la transcripción se basa únicamente en este ejemplar; para resolver todas las dudas que surgieron a lo largo de la revisión de la transcripción, hubo que cotejar los diferentes ejemplares en PDF.

#### 4.1.2 Selección de los apartados para cargar en el *software* y subida de estos a *Transkribus*

Tras elegir el ejemplar para transcribir, se dividió en archivos más pequeños, para agilizar el proceso de transcripción y trabajar con calas “letra por letra”, o sea archivos que contuvieran todos los lemas que iniciaban por la misma letra. El número de páginas de cada uno de estos archivos no coincide ya que depende de la cantidad de lemas que empiezan por cada letra. El primer archivo era el más extenso (de unas 120 páginas) y contenía la cala que va de la letra inicial A hasta la E (incluida)<sup>14</sup>.

---

13 La biblioteca en la que está guardado el ejemplar que llamamos BMF es la Aurelio Saffi de Forlì, actualmente cerrada por obras. El ejemplar forma parte del fondo antiguo, cuyo catálogo no ha sido digitalizado. El PDF de este tomo I del diccionario está disponible en el portal Contrastiva, it, de San Vicente (dir.).

14 En el momento de esa fase del trabajo se pensó que entrenar al *software* con un archivo muy extenso perfectamente revisado iba a facilitar que no se produjeran tantos errores al escanear los archivos más breves de las siguientes letras. La experiencia que se está llevando a cabo con la transcripción del tomo II nos ha llevado a entender que conviene educar al *software* al principio con archivos breves que se van convirtiendo en modelos cada vez más extensos a medida que se les añade el siguiente archivo revisado y aprobado.

#### 4.1.3 Transcripción con *Transkribus*

Esta fase del trabajo se puede sintetizar con los siguientes pasos realizados en cada una de las páginas de un archivo PDF: segmentación de la página en regiones, reconocimiento de líneas y del texto con un modelo público y correcciones de las imperfecciones y otras peculiaridades que queríamos enseñar al programa. Tras haber subido a la web de *Transkribus* los archivos correspondientes a las letras, se ha iniciado el reconocimiento de texto utilizando el modelo público llamado *Transkribus print M1*. Se ha elegido este modelo porque ya lo habían experimentado otros grupos de investigación del proyecto TELEI y habían llegado a la conclusión de que funciona bien con textos en español e italiano de los siglos XVIII y XIX. Al principio de la transcripción de las primeras páginas, la división en regiones se llevó a cabo de forma manual, utilizando el instrumento *Regions* de *Transkribus*; luego dejó de ser necesario.

#### 4.1.4 Creación de varios modelos privados de HTR y aplicación a los restantes archivos

A continuación, se creó el primer modelo privado de HTR, basado en el modelo público antes mencionado y en la transcripción revisada y aprobada de todos los artículos de las letras A-E. Dicho primer modelo de texto se aplicó a los archivos de los lemas con las letras iniciales F, G, H, I, J y L. El reconocimiento del texto y la transcripción automática con este primer modelo dio algunos resultados interesantes: el sistema reconocía el símbolo que separa en este diccionario las acepciones con la doble daga (‡) y lo transcribía con el símbolo que habíamos decidido adoptar (//)<sup>15</sup>. Asimismo, reconocía mejor que el modelo público la letra <ñ>. Sin embargo, con este primer modelo privado seguían produciéndose errores de reconocimiento de algunas letras –sobre todo mayúsculas– y de varios signos de puntuación.

15 Una breve nota para justificar esta equivalencia (//) en la transcripción de la doble daga. Este símbolo tipográfico, de antigua tradición y con múltiples valores según fuera el impresor que lo usaba, puede ser representado actualmente con programas de escritura de texto mediante el código U+2021, o sea, así: ‡. En el momento en que se tomaron algunas decisiones respecto a la revisión y aprobación del texto generado para poder ir creando modelos, se pensó que este símbolo podía ser difícilmente aceptado por el sistema de escaneo automático y por ello se adoptó la doble barra inclinada. Nada impide intervenir luego en el archivo global extraído de *Transkribus* en un archivo TXT con una operación busca/sustituye y cambiar todas las dobles barras inclinadas por una doble daga ‡.

A partir de ahí se creó un segundo modelo que se aplicó a los archivos de las letras iniciales M y N. Cabe decir que, en general, dicho modelo funcionaba mejor en el reconocimiento de las letras y, sobre todo, de las comas y los guiones, que antes se confundían o no se transcribían. El tercer modelo, creado a partir de la transcripción revisada de los lemas de las letras iniciales M y N, no funcionaba bien, probablemente por algún error humano cometido en la fase de creación del modelo. Por ello, hubo que preparar un cuarto modelo con una técnica diferente para mejorar el reconocimiento de las letras mayúsculas iniciales: para generar este modelo no nos hemos basado en archivos enteros del lemario con todos los lemas con varias letras iniciales como hasta ese momento, sino que tomamos como modelo las primeras tres páginas transcritas y revisadas de las letras iniciales O, P y Q, para que el sistema tuviera una muestra de las letras mayúsculas que había que reconocer y pudiera aplicar el nuevo conocimiento a las páginas por transcribir con estas mismas letras mayúsculas iniciales. Con este método se han obtenido buenos resultados. El nuevo modelo reconoce bien las letras mayúsculas, ha aprendido a no leer la parte de la línea que sigue a los corchetes, ya que en todas las correcciones manuales esta parte se había borrado y añadido manualmente en la línea de abajo; obviamente, esta última operación tuvo que seguir siendo manual durante todo el proceso.

El quinto modelo se ha creado con la misma técnica, con las primeras páginas de las letras R, S y T, y se ha aplicado luego al resto de las páginas de dichas letras. Lo mismo se ha hecho con el sexto y último modelo y las letras U, V, X y Z. Estos últimos dos modelos se han creado con el objetivo de mejorar cada vez más el reconocimiento del texto. Además, se ha observado un evidente progreso en la lectura de las palabras, aun cuando algunas letras eran casi ilegibles o estaban borradas, como si el sistema, conociendo la lengua de esa parte del texto, pudiera reconstruir la palabra a pesar de estar incompleta.

Por lo que se refiere al entrenamiento de *Transkribus* con modelos, cabe añadir que también creamos un modelo *layout*, con el objetivo de que leyera las tres regiones de las tres columnas de manera automática, sin tener que enmarcarlas manualmente antes de empezar el escaneo de cada página. Este modelo, a pesar de que ha sido útil en su aplicación, tiene tres inconvenientes: el primero es que para reconocer las regiones de forma automática se necesita el mismo tiempo de la transcripción de una página (doblando el tiempo total); además, el recurso a este modelo consume también créditos de uso del *software*; por último, no siempre funciona perfectamente y, en algunas páginas, se necesita igualmente intervenir manualmente, o escribir en el archivo TXT los símbolos de puntuación que quedan fuera de las regiones en el margen derecho.



## 5. Principales dificultades de transcripción y toma de decisiones

Tomando en consideración las características tipográficas y metalexicográficas de la disposición del texto en las páginas y de la macroestructura y microestructura antes descritas, se van a ilustrar aquí las principales dificultades y las decisiones que se han tomado para resolver algunos problemas surgidos con la transcripción.

### 5.1 *Por lo que concierne a la disposición del texto*

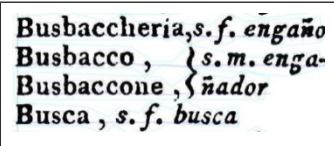
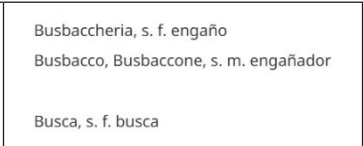
El primer obstáculo atañe a la división de la página en tres columnas. Como se ha descrito más arriba, fue necesaria la segmentación del texto en tres regiones, previa a la transcripción. Si se solicitaba al *software* directamente el reconocimiento del texto sin crear las regiones, el programa leía las líneas de las tres columnas seguidas, como si se tratara de una sola y de un único párrafo, de izquierda a derecha. Por esta razón, cuando no se aplicaba la transcripción automática con el modelo *layout*, antes de empezar la transcripción de cada página había que intervenir manualmente con la herramienta *Region* de *Transkribus* y crear las tres regiones. Con ello, se conseguía dejar fuera de las regiones los elementos que no se deseaba transcribir, como, por ejemplo, el número de las páginas, las tres letras-guía encima de cada columna o las signatures tipográficas de numeración del tomo o de los pliegos. Los resultados obtenidos ya en el primer archivo fueron satisfactorios: después de la división en regiones de forma manual, el programa leía separadamente las líneas de cada región, aunque a veces fallaba y era necesario segmentar las líneas a mano. Este fallo esporádico se ha mantenido hasta el final del trabajo.

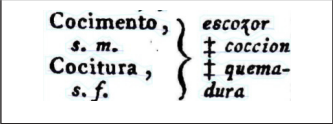
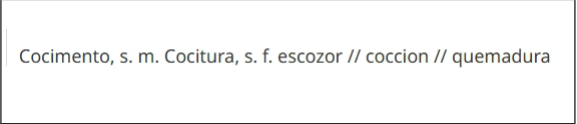
### 5.2 *Por lo que concierne a la microestructura*

La microestructura del leuario también proporcionó bastantes retos durante el reconocimiento del texto y la transcripción. Por lo que atañe al símbolo de la doble daga, el problema se resolvió rápida y fácilmente. Al principio, el programa leía este símbolo con resultados de transcripción diferentes (a veces con letras, como una < t > minúscula, por ejemplo, o bien con un corchete). Se decidió intervenir manualmente en las primeras páginas, convirtiendo la doble daga en dos barras inclinadas (//) y el programa aprendió muy rápidamente a sustituir este símbolo. En cambio, la llaves y los corchetes han representado problemas que no

se han conseguido solucionar.

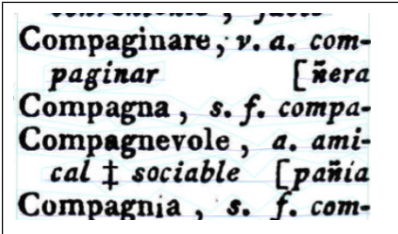
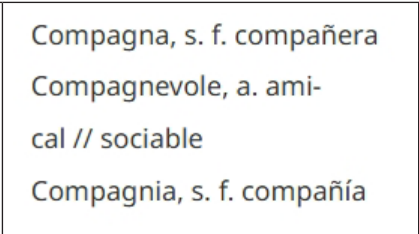
Como puede imaginarse, la llave no se puede leer en una sola línea (ya que recoge los lemas de al menos dos líneas, si no más). El programa detectaba la sinuosidad de la parte superior del símbolo de la llave y la transcribía como una C o una S; además, no repetía los equivalentes de traducción que comparten esos lemas. En general, las propuestas de reconocimiento del símbolo eran variadas y confusas y, por ello, hubo que tomar una decisión al respecto, interviniendo cada vez manualmente. Se ha adoptado la convención de transcribir, seguidas en serie, todas las entradas englobadas dentro de una llave con sendas iniciales mayúsculas y separadas por una coma (o por un punto si antes hay una abreviación de la categoría gramatical), transformándolas así en una única entrada con más lemas; siguiendo luego con la serie microestructural básica de las demás entradas (categoría gramatical y equivalente de traducción), como puede verse en las imágenes 7 (a, b) y 8 (a, b), en las que se muestra el texto original y la transcripción manual en *Transkribus* de dos ejemplos de uso de la llave.

	
Imagen 7a	Imagen 7b

	
Imagen 8a	Imagen 8b

Por lo que atañe al corchete, en cambio, la dificultad se ha podido solucionar parcialmente. El programa detectaba, exactamente como estaba escrito, el corchete de apertura en la línea de arriba, siguiendo al equivalente de traducción de la entrada anterior. Pensamos que la presencia de este corchete podía causar problemas a la hora de convertir el TXT en formato digital y dar instrucciones al programa de digitalización. Por esta razón, se ha preferido borrar totalmente el contenido del corchete y añadirlo manualmente donde correspondía, es decir, en la línea de abajo (véanse imágenes 9a y 9b). Después de varias operaciones de eliminación del contenido del corchete, el programa aprendió a no leer esta parte de línea.

Obviamente, la operación de completar la línea de abajo se ha tenido que realizar siempre manualmente.

	
Imagen 9a	Imagen 9b

Por lo que respecta, en cambio, al guion al final de la línea que indica que la palabra cortada sigue en la línea siguiente (véase *amical* en la imagen 9a y su transcripción en 9b), se ha tomado una primera decisión, la cual consiste en mantener dicha división para obtener un archivo de texto TXT lo más fiel posible al original –para posibles futuros usos del texto, como podría ser una edición filológica crítica–, ya que se está a tiempo de tomar una segunda decisión por la que en el archivo TXT se elimina el guion seguido de un salto de línea, juntando así las palabras cortadas.

## 6. Detección y corrección de erratas

El control de la transcripción automática del texto del tomo I ha identificado una serie de erratas y de errores<sup>16</sup>. Del análisis ecdótico de los errores (conceptuales) y de las decisiones que se han tenido que tomar al respecto se proporcionará algún ejemplo en §7. Nos concentraremos en este párrafo en las erratas, anticipando que no ha sido siempre posible reconocer e interpretar los fallos para intervenir con convicción. Como se ha especificado más arriba, para poder tomar una decisión ante una posible errata se han cotejado los cuatro ejemplares en PDF a disposición del leuario italiano-español del diccionario de Cormon y Manni de 1805: BUCM, BNF, BNE y BMF.

16 Adoptamos esta distinción terminológica, considerando *errata* los fallos de ejecución o errores materiales; es decir, una equivocación humana, del autor o del cajista. Recurriremos, en cambio, a la denominación de *error* para clasificar los fallos de tipo cognitivo, que derivan de la falta de conocimiento del autor, del impresor, tipógrafo, cajista, etc. La errata es material y el error, conceptual (Fernández-Quesada, Rodríguez-Rubio 2022).

## 6.1 Erratas de fallos mecánicos

Las erratas que hemos identificado en la obra son de diferentes categorías, siendo preponderantes los fallos mecánicos o de distracción que se listan a continuación. Se trata, por proporcionar solamente uno o pocos ejemplos de cada caso, de fallos como los siguientes:

- Inversión de tipos: algunas palabras presentan dos tipos posicionados de manera invertida con respecto a la palabra correcta, como sucede con “ciu dado”<sup>17</sup> en vez de “cuidado”.
- Sustitución de tipos: algunas letras se han sustituido erróneamente con otras, como en “palēbras” donde debía decir “palabras”, “antivedimengo” por “antivedimento”.
- Omisión de tipos: cuando algunas palabras carecen de una letra. Es lo que sucede con “cabeller\_” (por “cabellera”) o en “asolv\_miento” por “asolvimiento”, “cub\_ir de moho” en vez de “cubrir”.
- Colocación de un tipo al revés: se ha encontrado palabras que contienen un tipo girado (derecha/izquierda o bocabajo); el caso más repetido es el de las letras < u > y < n > (“Dnbbioso” en vez de “Dubbioso”).
- Letras mayúsculas fuera de lugar: aparecen letras mayúsculas en medio de algunas palabras y lemas, como en “BAldo”.
- Adición de sílabas: se trata de palabras que contienen una sílaba de más, como en “dañadado” (en vez de “dañado”) o que las tienen repetidas (“impepenitencia”).
- Inversión de sílabas: “peresozo” en vez de “perezoso”.
- Errores en las abreviaturas de las categorías gramaticales (que no se han hallado en diccionarios que pudieron ser fuente de este): por ejemplo, “s. f.” para “Ateo” en vez de “s. m.”, debido, probablemente, al hecho de que se ha unido con una llave al lema “Ateista”.

Estos fallos se han considerado totalmente mecánicos, de distracción o descuido. La primera decisión que se tomó al respecto –salvo en casos como el de los tipos colocados al revés– fue la de no intervenir, pues el primer objetivo de la investigación y del proceso de reconocimiento y transcripción automática era educar al *software* para que aprendiese a leer y transcribir el texto original lo más fielmente posible. De cada una de estas erratas y decisiones se tomó oportuna nota en un diario de trabajo para no perder la huella de los cambios realizados,

17 Los subrayados para resaltar el error son siempre nuestros.

poder localizarlas más adelante y corregirlas en el archivo TXT que será el texto base de la codificación.

## 6.2 Erratas en el uso de las tildes

No va a ser posible detenerse en este estudio sobre cuestiones ortográficas en sentido amplio; incluso remitir a estudios fundamentales de historiografía ortográfica ocuparía excesivo espacio<sup>18</sup>. Se presentará aquí un breve listado de ejemplos de erratas en el uso de las tildes que no nos parece que correspondan a la ausencia o caos normativo de aquella época. La variedad de imprecisiones encontradas en la acentuación es evidente y no ha resultado siempre fácil tomar una decisión. Entender si la adición u omisión de una tilde se debe a la distracción del tipógrafo, a un desconocimiento sobre la lengua extranjera con la que se estaba editando, a una diferente y efímera regla de acentuación que se estaba aplicando en ese momento o a una mancha de tinta ha sido un trabajo arduo. Este tipo de análisis –con continuos cotejos de los ejemplares a disposición, del *Nuevo tesoro lexicográfico de la lengua española*, de las cinco ediciones del *Vocabolario* de la Crusca en línea y del corpus CORDE– puede llevar a convertir una errata en un error de valor ecdótico, para el que hay que aplicar un tratamiento diferente. Siendo copiosas las erratas de acentuación, listaremos solo algún ejemplo de los diferentes tipos más difundidos.

- Tildes en vocales no tónicas: “anímal”, “provídencia”, “arbóles”, “claúsula”, “fástidioso”, “fuligínoso”, “márfil”, etc.
- Ausencia de tildes: “montaños”.
- Tildes en letras de abreviaturas: “Dolcione, á.” (de adjetivo).
- Tildes graves en palabras escritas en español: “jugar à”, “cèdula”, “de à dos pies”.
- Tildes agudas en palabras del italiano que no hemos conseguido documentar que se escribieran en aquel periodo también así: “giovedì”, “così”.
- Uso del acento circunflejo en palabras que no contienen la <x> con el sonido del grupo consonántico latino [ks] o falta de uniformidad en la aplicación de esta norma: “exâmen” frente a “exámen”, “divulgacìon”.

18 Es impensable poder indicar aquí todas nuestras referencias bibliográficas, estudios historiográficos sobre la ortografía, ni siquiera limitándonos al periodo del siglo XIX. Por lo que se refiere a la aplicación de las normas de ortografía en la lexicografía de principios del XIX, vamos a remitir solo a Terrón Vinagre (2022) y a Blanco Izquierdo (2021); por lo que se refiere a la ortografía del español en los bilingües de italiano-español, a De Hériz (2016, 2018).

## 7. El cotejo de fuentes y ediciones posteriores de la obra a favor de la transcripción y digitalización

Cuando se revisa la transcripción del texto original en formato .txt y se interviene en la corrección, se observa a menudo defectos de nitidez tipográfica que se pueden resolver, a veces, con el cotejo de los ejemplares de la misma edición a disposición; otras erratas como las ejemplificadas en §6 no se pueden interpretar sino con la deducción lógica o analógica. Por otro lado, también se van identificando erratas o errores que se repiten en todos los lemmas consultables y que por diferentes razones sorprenden más que las anteriores y obligan a tomar decisiones no siempre fáciles de adoptar. Si la transcripción de un original (automática o manual) tiene como único objetivo reproducir un texto antiguo en un formato de fácil lectura o bien en una edición crítica absolutamente fiel al original, no se suele intervenir sobre esos defectos considerados errores, sino que se remite a notas críticas que interpretan la intención del autor o señalan correcciones adoptadas en las siguientes ediciones o en textos de autores que tomaron como fuente el fragmento en cuestión.

Dado que los objetivos de nuestro proyecto no se limitan a generar protocolos de transcripción automática, a experimentar solamente programas informáticos de transcripción automática, sino que prevén el paso siguiente de la codificación del texto transcrito para hacerlo accesible junto con otros de la misma categoría lexicográfica bilingüe con el fin de constituir un tesoro lexicográfico digital consultable, en algunas ocasiones ha sido necesario plantearse cuál era la intención metalexicográfica del autor o autores del tomo I del Cormon y Manni de 1805 respecto a algunas peculiaridades de la macro y microestructura de esta versión. Claro está que estas extrañezas también pudieron ser fruto de los tipógrafos, quienes, ante una duda, consultaron quizás otros diccionarios para resolverla y no ralentizar el proceso de impresión de la obra. A continuación, se proporcionarán algunos ejemplos de resolución de problemas mediante el cotejo de obras lexicográficas anteriores y posteriores a la de Cormon y Manni italiano-español de 1805.

### *7.1 Cotejo de la primera (1805) y segunda (1821) edición del Cormon y Manni italiano-español*

Ante todo, cabe confirmar la convicción de Martínez Egidio (2010: 71) al respecto, el cual, tras haber cotejado una cala y los paratextos de la 2ª edición de 1821

del diccionario con la 1ª de 1805, aseguraba que el *Diccionario italiano-español y español-italiano* de 1821 de Cormon y Manni no contiene más diferencias que la ausencia en el título de la portada de la caracterización “de faltriquera” (tomo I) y “da tasca” (tomo II). Esto se puede corroborar ya que se han consultado varios ejemplares de 1821 para cada uno de los errores, caracteres poco nítidos, manchas extrañas, signaturas tipográficas, etc., presentes en los cuatro ejemplares de 1805 y no se ha observado la más mínima diferencia; solo la apenas mencionada del título de la obra.

## 7.2 Erratas, errores y curiosidades de la macroestructura

Tal como se ha descrito más arriba (§3), la macroestructura de este leuario se caracteriza por el uso de las llaves para englobar variantes morfológicas de una palabra-lemma que correspondan al mismo equivalente de traducción (o a más de uno y distintas acepciones de significado). A continuación presentamos dos ejemplos de lematización repetida (en *Apparato* y *Malaventura*). Propondremos como ejemplo de cotejo –aunque nos consten más– el C&M/ita-spa-1805<sup>19</sup> con otro diccionario bilingüe de 1802 de los mismos autores (Cormon y Manni) para la combinación italiano-francés, el *Dictionnaire portatif et de prononciation, italien-français et français-italien*, publicado también en Lyon por parte de los mismos libreros Cormon y Blanc (Lillo 2019: 219-223; Mormile 1993: 67-71). Alvar Ezquerro (2010: 48-9) ya supuso una probable filiación del bilingüe italiano-español de 1805 respecto al italiano-francés de 1802 de los mismos autores.

---

19 A partir de ahora y para abreviar al comentar cotejos entre diferentes obras, citaremos con la sigla C&M/ita-spa-1805 el tomo I del leuario italiano-español del *Diccionario de faltriquera* de Cormon y Manni de 1805 y con la sigla C&M/ita-fra-1802 el leuario italiano-francés del diccionario de Cormon y Manni de 1802.

<b>Apparato,</b> { <i>aparato, pre-</i> <i>s.m.</i> } <i>parativo</i> ‡ <b>Apparato,</b> { <i>adorno, de-</i> <i>s.m.</i> } <i>coracion</i>	<i>Apparato</i> , <i>ta</i> , adj. <i>Appris. + Pré-</i> <i>paré. etc. + Fourni ; pourvu.</i> <i>Apparato</i> , subst. m. ( <i>ap-pa-rá-to</i> ) <i>Ap-</i> <i>pareit; apprêt; préparatif. + Apparat.</i> <i>+ Apparato di chiesa. Ornement ;</i> <i>décoration ; pompe ; parure. + Appa-</i> <i>rato militare, apprêt, appareil militaire.</i>
<b>Malaventura,</b> { <i>s.f.mala-</i> <i>Malaventura,</i> } <i>ventura</i>	<b>Malaventura</b> , et <b>Mala ventura</b> , s. f. ( <i>-ven-toi-ra</i> ) <b>Malheur ; disgrâce ;</b> <b>malencontre.</b>
C&M-ita-spa-1805	C&M-ita-fra-1802

*Apparato* en C&M/ita-spa-1805 está etiquetado con una sola categoría gramatical (s. m.). En otros diccionarios anteriores se lematiza con dos entradas homónimas: una para el adjetivo derivado de *apparare* y la otra para el sustantivo masculino con varias acepciones<sup>20</sup>. Parece evidente que se puede considerar esta repetición del lema con la misma abreviatura gramatical como un error de distracción; por lo tanto, en la fase de digitalización de la entrada *apparato* podremos decidir que la 1ª forma se categorice como un adjetivo y la 2ª mantenga la abreviatura del original (s.m.). Conviene, sin embargo, ser conscientes de que una intervención de este tipo obliga a plantearse y a decidir si es necesario añadir a la forma del primer lema la desinencia femenina (*Apparato, ta*). Algo similar sucede con el lema compuesto *Malaventura* y su variante gráfica *Mala ventura*; se trata de un error que habrá que resolver durante la codificación para la digitalización del texto.

En otras ocasiones, algunas erratas debidas aparentemente a la confusión de tipos llevan a la más que probable identificación de la fuente, la cual contiene un verdadero error. Así hemos interpretado la forma del lema *Coregrafia*, no documentada en ninguna otra fuente lexicográfica monolingüe del español y del italiano<sup>21</sup> y sí, únicamente, en el bilingüe C&M-ita-fra-1802. Se trata probablemente de un galicismo del momento; el sonido del francés *Chorégraphie* se calcó en una

20 Ya en la 3ª edición del *Vocabolario degli accademici della Crusca* (1691). En estos cotejos daremos muestra de semejanzas o diferencias con diccionarios con fecha de edición más cercana a la del nuestro.

21 No se debe confundir este lema con el de *Coregrafia* (lematizado y traducido correctamente en los dos diccionarios que se están cotejando). Por otro lado, *Coregrafia* no está recogido en el *Vocabolario* de la Crusca hasta la 5ª edición (1863-1923) y, según el *NTLLE*, hasta 1825 en español en el *Diccionario* de Núñez de Taboada. Hemos encontrado en Gherardini (1840) un comentario con el que explica por qué recurre en un diccionario a la marca diatécnica *Coregrafia* y no a *Coreografia* (1840: 766).



inexistente forma italiana y se trasladó a otra española que tampoco encontramos documentada en el *NTLLE*.

<b>Coregrafia, s. f. coregrafia</b>	<b>Coregrafia , s. f. (-gra-fi-a) Choré-graphie, l'art de noter les pas et les figures d'une danse.</b>
C&M-ita-spa-1805	C&M-ita-fra-1802

En este caso, habrá que decidir si se digitaliza y codifica tal cual la entrada en el Tesoro TELEI, manteniendo presente en el diccionario en línea una “palabra fantasma” o de muy breve vida (Álvarez de Miranda 2000: 56-73)<sup>22</sup>.

No hemos hallado siempre la huella de las erratas de colocación de tipos en las palabras-lemma. Aun así, el cotejo con otros diccionarios anteriores al C&M-ita-spa-1805 –o posteriores, pero cercanos a la fecha de edición– han avalado la toma de una decisión respecto a una corrección en la transcripción y a la forma que se digitalizará. Véanse estos dos ejemplos:

<b>Insieme, ó beu'insieme, s. m. conjunto, total, reunion [ con</b>	<b>Insieme, et Ben' insieme, s. m. T. de Peint. et Sculp. L'ensemble, ce qui résulte de la réunion des parties d'un tout.</b>
C&M-ita-spa-1805	C&M-ita-fra-1802
<b>Mercecchè, conj. con motivo de .. porque</b>	<b>MÈRCÉ CUE, MERCECCHÈ, e MERCECCHÈ. V. MERC-E.</b>
C&M-ita-spa-1805	Vanzon 1841

Es lógico –pero podemos basarlo en modelos lexicográficos– que la segunda forma *Beu'insieme* de la entrada *Insieme* se debe corregir como *Ben' insieme*. Asimismo nos parece coherente que la intención de lematizar la conjunción *Mercecchè* con dos <cc> se respete con una forma correcta: o sea, *Mercecché* (en vez de *Merceceché*).

<sup>22</sup> El error perdura en el C&M italiano-español de 1821 y Blanc en su revisión de la obra de 1843 no lo elimina. Salvo que se consiga demostrar el uso de este término en ese periodo, este lema y su equivalente bien sirven de ejemplo de las palabras fantasma de la lexicografía.

7.3 Erratas, errores, imperfecciones y curiosidades de la microestructura

Los problemas o dudas surgidos durante la transcripción de los contenidos de los artículos también se han intentado resolver con el cotejo de otras obras lexicográficas. Como se ha explicado en §3 la imprenta en prensa con tipos y tinta producía muchas manchas o falta de nitidez en la impresión. Para completar algunas palabras en las que no se leen bien los caracteres, la comparación entre la 1ª edición (1805) y la 2ª (1821) no ha resuelto nunca las dudas, pues las galeras parecen ser exactas. En cambio, ha sido muy útil poder comparar dichos artículos con los de la edición de 1843 o 1848 del Cormon y Manni italiano-español, la revisada por S. H. Blanc (Alvar Ezquerro 2010: 52-5). Como muestra de muchas comparaciones realizadas presentamos el artículo del lema *Patto*, en el que dos palabras de la penúltima y última línea están incompletas:

<b>Patto</b> , s. m. <i>pacto, convenio, ajuste, condición</i> ‡ con patto <i>che...</i> con tal <i>que...</i> ‡ adv. per al un patto, <i>de ningun m. do</i>	<b>Patto</b> , sm. <i>pacto, convenio, ajuste, condición</i> ‡ con patto <i>che...</i> con tal <i>que...</i> ‡ ad. per al un patto, <i>de ningun modo</i> [convenir	<b>Patto</b> , sm. <i>pacto, convenio, ajuste, condición</i> ‡ con — <i>che...</i> con tal <i>que...</i> ad. <i>per alcun—</i> , <i>de ningun modo.</i>
C&M-ita-spa-1805	C&M rev. Blanc-ita-spa-1843	C&M-ita-spa-1848

En las tres muestras propuestas, la microestructura de las tres ediciones no es idéntica por lo que a símbolos metalexográficos se refiere, pero con este doble cotejo se completan con seguridad “alcun patto” y “de ningun modo” de la edición de 1805.

Cuando nos sorprende la ausencia de equivalentes en español dentro del artículo o la traducción de ejemplos, también consultamos diccionarios anteriores y el Cormon y Manni de italiano-francés de 1802. De esta manera solemos entender cuál era el proyecto del contenido de la entrada, si bien en ningún caso hemos añadido todo lo que probablemente falta. Es lo que sucede en el artículo del lema *Colà*, comparado aquí abajo, donde no se distinguen las dos funciones adverbiales y no se traduce el ejemplo *colà di maggio*.

<b>Colà</b> , adv. <i>alli</i> ‡ <i>cerca</i> <b>colà di maggio</b>	<b>Colà</b> , adv. de lieu. ( <i>ko-là</i> ) Là. ‡ <i>Di là a colà</i> , de l'endroit d'où l'on part. jusqu'où l'on veut se rendre. ‡ adv. de temps. <i>Euvron ; vers. Colà di maggio.</i>
C&M-ita-spa-1805	C&M-ita-fra-1802

También ha sucedido que nos ha extrañado la ausencia de algún símbolo de valor metalexicográfico, como la doble daga que antecede a las diferentes acepciones de significado. Presentamos aquí el ejemplo del artículo del lema *Manco*, en el que, al principio, nos llamó la atención la ausencia de una coma entre “menos” (español) y “venir manco” (en italiano). La edición del Cormon y Manni revisada por Blanc (1843) no resolvía la duda respecto al elemento que faltaba, mientras que el cotejo con el diccionario italiano-español del diccionario anónimo publicado en 1853 por la librería parisina de Rosa y Bouret (Castillo Peña 2010)<sup>23</sup> nos ha permitido entender que entre el equivalente “ménos” y la colocación “venir manco” falta en C&M de 1805 el símbolo de acepción (doble daga en C&M 1805), que podremos introducir en la digitalización del artículo.

<b>Manco</b> , <i>ad.</i> ménos venir mauco , faltar ; etc. ‡ senza manco, sin falta ‡ non ci ho manco	<del>manco</del> <b>miento</b> ‡ <i>ad.</i> <del>me-</del> nos venir mauco, faltar, etc. ‡ senza mauco, sin falta ‡ non ci ho mauco pensato, ni siquiera he pensado <i>á</i> ello [miento]	<b>MANGO</b> , <i>ad.</i> Menos.   Venir manco; faltar.   Senza mauco; sin falta.   Non et ho mauco pensato; ni siquiera he pensado en ello.
C&M-ita-spa-1805	C&M rev Blanc-ita-spa-1843	R&B-ita-spa-1853

Consideramos oportuno concluir esta sección con un ejemplo útil para demostrar que el Cormon y Manni italiano-francés (1802), aunque se pueda confirmar que fue la fuente principal del leuario italiano-español (1805) de los mismos autores, no constituyó la única base de inspiración del quehacer lexicográfico. El elemento microestructural que aportamos como botón de muestra para apoyar esta tesis es una glosa diatécnica añadida al final del artículo del lema *Aormare*: “es voz de caza”, información que no se encuentra en el leuario italiano-francés, como puede verse abajo. La Crusca lematiza *Aormare* a partir de la 3ª edición de su *Vocabolario* (1691) y marca este uso relacionado con la caza tanto en esa edición como en la 4ª (1729-1738), ambas anteriores a nuestro diccionario. Sugerimos aquí una comparación con el artículo registrado en el *Dictionaire* italiano, latín y francés de Antonini (ya mencionado en §3 a propósito del uso de las llaves para unir formas de lemas).

23 Carmen Castillo Peña (2010: 154-68) ha demostrado con un largo estudio de cotejos macro y microestructurales que el Cormon y Manni italiano-español/español-italiano de 1805 es una de las fuentes principales (no la única) del leuario italiano-español del anónimo *Nuevo diccionario italiano-español* de los editores parisinos Rosa y Bouret (1853) y no lo son, en cambio, las ediciones del Cormon y Manni de 1843 y 1848 revisadas por Blanc.

<b>Aormare, v. a. rastrear, buscar; es voz de caza</b>	<i>Aormare, v. a. (aor-má-re) Quêter, suivre à la piste.</i>	<b>AORMARE. Voce de' Cacciatori; e vale, Cercar la fiera, sieguendone l'orme. Chasser à la piste.</b>
C&M-ita-spa-1805	C&M-ita-fra-1802	Antonini-ita-fra-1743

## 8. Conclusiones

La transcripción automática de un diccionario bilingüe impreso a principios del siglo XIX, como la del leuario italiano-español del Cormon y Manni de 1805 que se ha descrito, puede parecer *a priori* un desafío más simple de afrontar que el que representarían obras de siglos anteriores. Se ha visto, en cambio, que cada una de estas obras conlleva diferentes dificultades por las condiciones materiales de los ejemplares, por la peculiaridad tipográfica de los textos de los leuarios y por las muchas erratas y errores que se van encontrando en la revisión de las 422 páginas del texto transcrito automáticamente.

Dado que la transcripción de dicho texto no ha sido el objetivo final del proyecto de investigación, sino el paso previo a la digitalización con la que se codificará cada uno de los elementos macro y microestructurales, cabe recordar que los ejemplos que se han aportado en esta publicación son un botón de muestra de todas las consideraciones que hubo que plantearse durante la revisión del texto transcrito; consideraciones intra y extratextuales.

## Bibliografía citada

- Alvar Ezquerro, Manuel (2010), “Un siglo de lexicografía bilingüe español-italiano”, *Diversidad lingüística y diccionario*, eds. Marta Concepción Ayala; Antonia María Medina. Málaga, Universidad de Málaga: 45-118.
- Álvarez de Miranda, Pedro (2007), “Palabras y acepciones fantasma en los diccionarios de la Academia”, Alicante, *Biblioteca virtual Miguel de Cervantes* [20/03/2025] <<https://www.cervantesvirtual.com/nd/ark:/59851/bmc2f836>>.
- Anónimo (1853), *Nuevo diccionario italiano-español*, París, Librería de Rosa y Bouret.
- Antonini (1743), *Dictionnaire italien, latin e françois*, París, Prault Fils.
- Blanco Izquierdo, M.<sup>a</sup> Ángeles (2021), “Más allá de las letras: la acentuación gráfica en el *DRAE* (1869, 1884 y 1899)”, *El diccionario académico en la segunda mitad del siglo XIX*:

- evolución y revolución*, eds. María Ángeles Blanco; Gloria Clavería. Berlin, Peter Lang: 91-125.
- Bohbot, Hervé; Frontini, Francesca; Luxardo, Giancarlo; Khemakhem, Mohamed; Romary, Laurent (2018), “Presenting the Nénufar Project: a Diachronic Digital Edition of the Petit Larousse Illustré”, *GLOBALEX 2018*, May 2018. Miyazaki, Japan: 1-6 [7/3/25] <<https://hal.science/hal-01728328v1>>.
- Bruña Cuevas, Manuel (2008), “La producción lexicográfica con el español y el francés durante los siglos XVI al XIX”, *Philologia Hispalensis*, 22: 37-111.
- Castillo Peña, Carmen (2010), “El *Nuevo diccionario italiano-español* (1853) de los editores Rosa y Bouret”, *Textos fundamentales de la lexicografía italoespañola: (1805-1916)*, dir. Félix San Vicente. Monza, Polimetrica: 146-92.
- Castillo Peña, Carmen (2020), “Epigrama: Un portal para la edición digital de textos gramaticales”, *Anales De Lingüística*, 4: 201-217 [20/03/2025] <<https://revistas.uncu.edu.ar/ojs3/index.php/analeslinguistica/article/view/4395>>.
- Cazorla Vivas, M.<sup>a</sup> del Carmen (2002), *Lexicografía bilingüe de los siglos XVIII y XIX con el español y el francés*, Tesis doctoral, Madrid, Universidad Complutense.
- Cazorla Vivas, M.<sup>a</sup> del Carmen (2010), “Panorama de la lexicografía bilingüe y plurilingüe del español a comienzos del siglo XIX”, *Textos fundamentales de la lexicografía italoespañola (1805-1916)*, dir. Félix San Vicente. Monza, Polimetrica: 27-56.
- Cormon, J. L. B.; Manni, Vincenzo (1802), *Dictionnaire portatif et de prononciation italien-français et français-italien*, Lyon, Libraires B. Cormon et Blanc.
- Cormon, J. L. B.; Manni, Vincenzo (1805). *Diccionario de faltriquera italiano-español y español-italiano*, Leon, Librería de B. Cormon y Blanc.
- Cormon, J. L. B.; Manni, Vincenzo (1821), *Diccionario italiano-español y español-italiano*, Leon, Librería de B. Cormon y Blanc.
- Cormon, J. L. B.; Manni, Vincenzo (1843), *Diccionario italiano-español y español-italiano*, Nueva edición revista y aumentada por S. H. Blanc, Leon, Librería Cormon y Blanc.
- Cormon, J. L. B.; Manni, Vincenzo (1848), *Diccionario italiano-español y español-italiano*. Nueva edición revista y aumentada por S. H. Blanc, Leon y Paris, Librería St. Hilaire Blanc y C<sup>ia</sup>.
- Costa, Rute; Salgado, Ana; Khan, Anas Fahad; Carvalho, Sara; Romary, Laurent *et al.* (2021) “MORDigital: The Advent of a New Lexicographical Portuguese Project”, *eLex 2021 - Seventh biennial conference on electronic lexicography*, Jul 2021, Brno, Czech Republic [20/03/2025] <<https://inria.hal.science/hal-03195362v2>>.
- De Hériz, Ana Lourdes (2016), “Normas ortográficas de la lengua española en la lexicografía ítalo-española del siglo XIX”, *Orillas*, 5 [30/03/2025] <<https://www.orillas.net/orillas/index.php/orillas/article/view/355/350>>.
- De Hériz, Ana Lourdes (2018), “Las reformas de la ortografía en los diccionarios bilingües de italiano-español de los siglos XIX, XX y XXI”, *Boletín de la Sociedad Española de la Historiografía Lingüística*, 12: 95-121 [30/03/2025] <[http://www.sehl.es/010\\_05\\_ana-lourdes\\_de-heacuteriz.html](http://www.sehl.es/010_05_ana-lourdes_de-heacuteriz.html)>.

- Fernández-Quesada, Nuria; Rodríguez-Rubio, Santiago (2022), “El tratamiento del error textual y de la errata en la era digital: elogio de la corrección”, *Detección y tratamiento de errores y erratas: un diagnóstico para el siglo XXI*, eds. Nuria Fernández-Quesada, Santiago Rodríguez-Rubio. Madrid, Dykinson: 13-28.
- Gallina, Anna Maria (1991). “La lexicographie bilingue espagnol-italien, italien-espagnol”. eds. Franz Josef J. Hausmann *et al.* *Wörterbücher Dictionaries Dictionnaires*. Berlin-New York, Walter de Gruyter: 2991-97.
- Gattel, Claude Marie (1798), *Nuevo diccionario portátil español y francés*, Paris, Casa de Bossange, Masson y Besson.
- Gattel, Claude Marie (1803), *Nuevo diccionario portátil español é inglés*, Valencia, P. J. Mallen y C.
- Gherardini, Giovanni (1840), *Voci e maniere di dire italiane additate a'futuri vocabolaristi*, Volume II, Milano, Gio. Bat. Bianchi di Giac.
- Khemakhem, Mohamed; Foppiano, Luca; Romary, Laurent (2017), “Automatic Extraction of TEI Structures in Digitized Lexical Resources using Conditional Random Fields”, *Electronic Lexicography*, eLex 2017, Sep 2017, Leiden, Netherlands [07/3/2025] <<https://hal.science/hal-01508868v2>>.
- Khemakhem, Mohamed; Galleron, Ioana; Williams, Geoffrey; Romary, Laurent; Ortiz Suárez, Pedro Javier (2019), “How OCR Performance can Impact on the Automatic Extraction of Dictionary Content Structures”, *19th annual Conference and Members' Meeting of the Text Encoding Initiative Consortium (TEI) -What is text, really? TEI and beyond*, Sep 2019, Graz, Austria [07/03/2025] <<https://hal.science/hal-02263276v1>>
- Lillo, Jacqueline, ed. (2019), *1583-2010: Quattro secoli e più di lessicografia italo-francese. Repertorio analitico di dizionari bilingue*, Vol I, 2ª edizione riveduta e ampliata, *Quaderni del Cirsil*, 14. Bologna, Università di Bologna.
- Lindemann, David; Alonso, Mikel (2021), “A workflow for historical dictionary digitization: Larramendi's Trilingual Dictionary”, *Proceedings of eLex*, 598-614 [07/03/2025] <[https://elex.link/elex2021/wp-content/uploads/2021/08/eLex\\_2021\\_39\\_pp598-614.pdf](https://elex.link/elex2021/wp-content/uploads/2021/08/eLex_2021_39_pp598-614.pdf)>.
- Lindemann, David; Khemakhem, Mohamed; Romary, Laurent (2018), “Retro-digitizing and Automatically Structuring a Large Bibliography Collection”, *European Association for Digital Humanities (EADH) Conference*, EADH, Dec 2018, Galway, Ireland [07/03/2025] <<https://hal.science/hal-01941534v1>>
- Lombardero Caparrós, Alberto (2015), *The Historiography of English Language Teaching in Spain: A Corpus of Grammars and Dictionaries (1769-1900)*, Tarragona, Universitat Rovira i Virgili.
- Martínez Egido, José Joaquín (2008), “Origen y desarrollo positivo de la lexicografía bilingüe español-italiano (siglos XVI-XIX)”, *Philologia hispalensis*, 22: 213-58.
- Martínez Egido, José Joaquín (2010), “El *Diccionario de faltriguera italiano-español y español-italiano* de J. L. B. Cormon y V. Manni (1805)”, *Textos fundamentales de la lexicografía italoespañola (1805-1916)*, dir. Félix San Vicente, Monza, Polimetrica: 57-92.

- Mormile, Mario (1993), *Storia dei dizionari bilingui italo-francesi. La lessicografia italo-francese dalle origini al 1900*, Fasano, Schena Editore.
- Nalesso, Giulia (2024), “Teoría y práctica de la retrodigitalización de diccionarios: el caso del *Vocabulario de las dos lenguas toscana y castellana*”, *Artifara*, 24.1: 135-53 [20/03/2025] <<https://ojs.unito.it/index.php/artifara/article/view/9988/8605>>
- Núñez de Taboada, Melchor Manuel (1825), *Diccionario de la lengua castellana*, Paris, Seguin.
- Rabus, Achim (2019), “Recognizing Handwritten Text in Slavic Manuscripts: a Neural Network Approach Using Transkribus”, *Scripta & e-Scripta*, 19: 9-32.
- San Vicente Santiago, Félix, dir. (2010), *Textos fundamentales de la lexicografía italoespañola: (1805-1916)*. Monza, Polimetrica.
- San Vicente, Félix, dir., *Contrastiva. Portal de gramática y de lingüística contrastiva español-italiano*. [20/03/2025] < <http://www.contrastiva.it>>.
- Terrón Vinagre, Natalia (2022), *Lexicografía y ortografía en el siglo XIX*, Berlin, Peter Lang.
- Vanzon, Carlo Antonio (1841), *Dizionario universale della lingua italiana*, Palermo, Tipografia Demetrio Barcellona.
- Wilkinson, Mark, *et al.* (2016), “The FAIR guiding principles for scientific data management and stewardship”, *Sci. Data*, 3:160018 [7/3/2025] <<https://www.nature.com/articles/sdata201618>>.

**Florencia Ferrante** es licenciada en Letras por la Universidad de Buenos Aires y en Italianística por la Universidad de Bolonia. En 2019 se doctoró en la Universidad de Módena y Reggio Emilia con una tesis sobre el pensamiento crítico de Juan Rodolfo Wilcock. Se ha ocupado principalmente de las relaciones literarias e intelectuales entre Italia e Hispanoamérica, especialmente en el ámbito de la teoría literaria y de la historia de la traducción. Es autora de una monografía (*Juan Rodolfo Wilcock crítico*, ETS, 2022) y de varios artículos publicados en revistas como *Strumenti Critici*, *Hispanérica* y *Meta*.  
**florencia.ferrante@unige.it**

**Chiara Valente** es licenciada en Lenguas y culturas modernas por la Universidad de Génova. De 2018 a 2021 formó parte del grupo de investigación GramHisGram de la Universidad de Salamanca, donde también defendió su tesis doctoral en 2021 con un estudio sincrónico, diatópico y comparativo español-italiano del uso de las formas verbales del PPS y PPC. Actualmente es profesora de lengua y literatura española en la *Scuola Secondaria di II grado* y profesora *a contratto* de la Universidad de Génova. Es autora de capítulos de libros en editoriales como Dykinson, De Gruyter y Arco/Libros y de varios artículos en revistas como *Moenia* y *Lingua e Stile*.  
**chiara.valente@edu.unige.it**

**Ana Lourdes de Hériz** es profesora titular de Lengua española en la Universidad de Génova. Se doctoró en la Universidad de Pisa con una tesis doctoral sobre la colaboración en la redacción de heterobiografías, estudio realizado con un enfoque pragmático. Durante los años en los que trabajó en la universidad como lectora investigó sobre el análisis del error, dedicándose luego a la historia de la traducción al español en España en el siglo XIX, a la gramática contrastiva (español-italiano) y, sobre todo, a la lexicografía en el siglo XIX y a la metalexicografía actual. Ha coordinado dos grupos de investigación PRIN (2017 y 2022) y publicado capítulos de libros en editoriales italianas y europeas.  
**ana.deheriz@unige.it**





MARIA FELISA BERMEJO CALLEJA  
GIULIA LANTERI  
MARÍA JOAQUINA VALERO GISBERT  
MARIA CARMELA ZACCONE  
LA TRASCRIPCIÓN DIGITAL  
DEL DICCIONARIO DE L. AMBRUZZI

Università degli Studi di Torino

Università degli Studi di Parma

#### Resumen

El *Nuovo dizionario* de Lucio Ambruzzi (DBA), publicado en 1948-1949, es una obra bilingüe con elementos enciclopédicos. Su digitalización enfrenta desafíos por su compleja microestructura y abundancia de abreviaturas. Se busca sistematizar su ordenación, clarificar abreviaturas y mejorar su consulta. El estudio aborda aspectos metalexicográficos y lingüísticos para optimizar su uso en formato digital.

palabras clave: lexicografía bilingüe, digitalización, abreviaturas

#### Abstract

##### *The digital transcription of the dictionary by L. Ambruzzi*

*Lucio Ambruzzi's bilingual dictionary (DBA), published between 1948 and 1949, is a work that features encyclopedic elements. Its digitization presents challenges due to its complex microstructure and numerous abbreviations. The goal is to systematize its organization, clarify the abbreviations, and enhance its usability. This study explores both metalexicographical and linguistic aspects to optimize the dictionary's function in digital format.*

*keywords: bilingual lexicography, digitization, abbreviations*

## 1. Introducción

El *Nuovo dizionario spagnolo-italiano e italiano-spagnolo* de Lucio Ambruzzi (*DBA*) es una obra bilingüe, pasiva y general con elementos enciclopédicos, cuya primera edición salió a la luz en 1948 y 1949, respectivamente, primer y segundo volumen (Bermejo 2008).

El *DBA* forma parte de los materiales lexicográficos bilingües español-italiano e italiano-español sometidos al proceso de retrodigitalización (Khemakhem 2018), cuyo objetivo es ponerlos a disposición en la red. Este proceso forma parte de un proyecto que busca conformar el primer *Tesoro digital de la lexicografía bilingüe español-italiano*. El *DBA*, a diferencia de otras obras lexicográficas cronológicamente anteriores, presenta una gran complejidad microestructural, lo cual obligó al autor, en busca de unas pautas coherentes de ordenación, a diseñar un sistema tipográfico que, aplicado a las abreviaturas, se tradujo en una multiplicidad y variedad de formas. Su esfuerzo, sin duda meritorio, no alcanza, sin embargo, la sistematicidad deseada, debido a varios factores entre los que destacaremos por un lado la sobreabundancia de información, así como su amplia diversificación (enciclopédica, didáctica, cultural, curiosidades, etc.) (Calvi 2006) y, por otro lado, la estrechez que presentan los límites de las páginas de un volumen impreso.

Todo ello supone un reto a la hora de realizar el proceso de transcripción y codificación del *DBA*. Dicho desarrollo se ha articulado en distintas fases: la primera se ha concentrado en el estudio de la macroestructura y en el análisis de la aplicación de los criterios declarados por el autor en la microestructura; la segunda se ha dedicado al establecimiento de criterios de sistematización; la tercera ha consistido en la transcripción mediante la herramienta *Transkribus*, seguida de la revisión y corrección de los errores generados por el programa; y, por último, una cuarta fase de uniformización.

Para ello, centrándonos en la segunda fase, se han determinado los criterios para lograr sistematicidad y eficacia en la ordenación de la microestructura. Entre ellos, y a partir de los elementos utilizados en el *DBA*, se han compensado las faltas por omisión, repetición o incongruencia en su función, a fin de lograr la homogeneidad deseada. Tras un profundo análisis de las numerosas abreviaturas utilizadas, se ha determinado cuáles de ellas son efectivamente marcas de uso (Hausman 1989, Valero Gisbert 2016, 2018, 2020) o funcionales para la relación interna de los elementos lexicográficos, y, por tanto, se han mantenido como abreviaturas. En cambio, las demás se han explicitado logrando de este modo

superar la dificultad de la lectura y comprensión del artículo lexicográfico, favoreciendo, con ello, su interpretación cabal y su eficacia a la hora de seleccionar los equivalentes. Lo que aquí se presenta corresponde al resultado de la transcripción de la letra A en ambos volúmenes del diccionario, por lo que las decisiones adoptadas se reflejan específicamente en la transcripción de dicha letra.

En definitiva, en este artículo abordamos las principales cuestiones metalexicográficas, lingüísticas y ecdóticas que caracterizan la problemática específica de la digitalización del diccionario de L. Ambruzzi.

## 2. El diccionario

Entre los principales actores de la lexicografía bilingüe, destaca la figura de Lucio Ambruzzi. Como apunta Bermejo (2008: 201), estamos ante “un gran estudio, lexicógrafo, gramático y filólogo, cuyos campos de interés se centraban en la lengua y se extendían a la literatura y a la cultura en general.” Esta obra ha sido objeto de un estudio pormenorizado por parte de Bermejo (2008), quien realizó un detallado análisis desde su génesis, pasando por los aspectos tanto macroestructurales como microestructurales que la definen.

Brevemente y en relación con nuestro estudio, recordamos algunos elementos especialmente interesantes para afrontar, en este estadio de la investigación, la transcripción de la letra A, dada la complejidad y la abundancia de datos.

La idea de crear esta obra no fue una iniciativa del propio autor, sino de la editorial que en distintas ocasiones lo invitó a realizarla. El diccionario se divide en dos volúmenes y si bien inicialmente aparece con el título referido a cada lemmario, posteriormente, en el año 1952, se unificó en “*Nuovo dizionario Spagnolo-Italiano e Italiano-Spagnolo*” (Bermejo 2008: 132).

Esta partición se debe al diferente receptor a quien se destinaba, ya que según Bermejo (2008: 151) “la división en dos volúmenes con los aparatos críticos escritos en lenguas diferentes, al igual que las marcas, demuestra la intención de dirigirse a usuarios de dos lenguas distintas.”

En cuanto a las fuentes del diccionario, Bermejo explica que el autor declara en la introducción que se basó en obras de prestigio de distintas editoriales. Sin embargo, esta circunstancia no significa que no se pueda considerar como una obra de autor, pues como apunta la estudiosa (2008: 130), “fue él, y no la editorial, quien decidió cuántos y qué lemas incluir [...]”.

Esta breve introducción a la obra y su autor nos permite contextualizar su

génesis y comprender con mayor profundidad el proyecto y las intenciones que lo sustentan en relación con las problemáticas planteadas. A continuación, exponemos las dificultades encontradas durante el proceso de transcripción, así como los criterios adoptados para garantizar una mayor claridad.

3. Abreviaturas

El diccionario cuenta con dos listados de abreviaturas, uno para cada volumen, que preceden al leuario.

La sección ‘Abbreviature’ del volumen italiano-español registra 438; mientras que la del volumen español-italiano recoge un número menor, en concreto 337.

El listado no es circular, no son siempre las mismas, ni coinciden en la forma. En algunas ocasiones, encontramos para el mismo significado dos abreviaturas, por ejemplo, en el I volumen (sección esp-ita) Ambruzzi presenta para ‘accusativo’ las formas: ‘acc.’ y ‘accusat.’, mientras que en el II volumen (sección ita-esp) podemos leer, para referirse a la albañilería, las formas ‘Alb.’ y ‘Albañ.’. Son marcas de distinto tipo, a veces se refieren a lenguajes de especialidad, o son gramaticales, o geográficas, en definitiva, con diferentes funciones.

Por otro lado, en la microestructura se observa el uso de abreviaturas que no están presentes en el listado de la macroestructura:

<i>Banc.</i> (anticipazione)	1
<i>Curt.</i> (abbancare)	1
<i>Meteor.</i> (annuvolamento)	1
<i>Telef.</i> (autocommutatore)	1
<i>Bols.</i> (arbitraggio)	1
<i>Cul.</i> (animella)	1

Fig. 1. Abreviaturas con el índice de frecuencia pertenecientes al volumen italiano-español (1949)

<i>Mur.</i> (abanto – alzar)	2
<i>Fis.</i>	18
<i>Farm.</i>	18
<i>Sport.</i> (andino – arquero)	2

masch. (algún)	1
patron. (Álvaro)	1

Fig. 2. Abreviaturas con el índice de frecuencia pertenecientes al volumen español-italiano (1948)

Las implicaciones que se derivan de la numerosidad de abreviaturas, junto con otras no presentes en la macroestructura e incluso el uso de diferentes formas para un mismo fenómeno, representan un aspecto complejo en la transcripción mediante el *software Transkribus*. Sirva como ejemplo la diversidad de formas para referirse a la fraseología que, como explica Bermejo (2008: 161), se encuentran: “[...] *fr.* (frase, locuzione), *loc.* (locuzione), *mod.* (modismo, modo di dire) [...]”.

Esta problemática nos ha llevado a plantearnos la necesidad de establecer una serie de criterios orientados a sistematizar dichos elementos para lograr una mayor claridad microestructural en la presentación de la información. Aunque es cierto, en palabras de Bermejo (2008: 158), que desde el punto de vista tipográfico, Ambruzzi “realiza un gran esfuerzo en crear y mostrar la diversificación de los distintos tipos de información que da en el artículo lexicográfico a través de características tipográficas y de sus posibilidades combinatorias”, este cuidado por diferenciar la información contenida en el leuario se ha traducido en realidad en una dificultad a la hora de mantener la sistematicidad necesaria en toda la obra y consecuentemente en la transcripción.

Por lo que se refiere a las marcas diatópicas, la profusión de indicaciones representa, sin duda, una riqueza, pero a su vez supone un reto en términos de sistematización, especialmente si se consideran los recursos de los que el autor dispuso en su época. Para una visión más profunda, remitimos a Valero Gisbert (2020: 169-170).

#### 4. Criterios adoptados para la transcripción del diccionario de Lucio Ambruzzi (1948-49)

En la segunda fase, hemos determinado una serie de pautas para lograr un orden coherente y uniforme de las abreviaturas.

Como criterio general y por lo que respecta a las abreviaturas presentes en la lista de Ambruzzi, si no presentan ambigüedades, se mantienen con la misma forma. Hemos constatado que, en ocasiones, utiliza indistintamente la abreviatura en español o italiano, por ejemplo, para la expresión ‘etcétera’ alterna ‘etc.’ y ‘ecc’.

La elección de una lengua u otra parece que se debe al contexto en que la usa. Es el caso de los ejemplos en italiano para los que el autor opta por la abreviatura en italiano y viceversa. No se ha intervenido en estos casos, ya que el dato puede ser de interés para investigaciones futuras.

Hemos decidido registrar algunas abreviaturas en su forma completa para una mayor claridad en los dos volúmenes del diccionario. Sospechamos que Ambruzzi las dejó abreviadas por economía espacial. Es el caso de Acad. → Academia; antig. → antiguo o antiguamente; confeder. comerc. → confederación comercial; Conj. c./Conj. como/ conj. como → conjúgase como; conson. → consonante (lemma a); corresp. → corresponde; determ. → determinados; dipt. → diptongo; n. de 7 emper. rom. ...incl. → nombre de 7 emperador romano ...incluido; esdrúj. → esdrújula; especialm. → especialmente, entre otras. En el volumen español-italiano contamos 97, mientras que son un total de 151 abreviaturas en el italiano-español. A modo de ejemplo, la forma abreviada ‘pal.’ se ha transcrito en su forma plena: ‘palabra’.

Por lo que se refiere a abreviaturas que se encuentran listadas pero que no usa en la microestructura, como por ejemplo ‘patr.’ y ‘patron.’, se ha optado por no intervenir y dejar que en futuros estudios se desambigüe el significado.

Para una mayor claridad, hemos seguido este mismo criterio con otras que, en cambio, sí estaban en la lista de Ambruzzi. En la microestructura hemos contabilizado un total de 9 en la sección ita-esp, se puede observar bajo el lema “arginale [...] adj. ref. al dique.”, mientras que son mucho más numerosas en la sección esp-ita, pues se cuentan 31 abreviaturas.

#### 4.1 *Abreviaturas con errores en la microestructura*

Se trata de formas incluidas en la lista de Ambruzzi que contienen errores en su utilización dentro de la microestructura. En estos casos el criterio adoptado ha sido el de la corrección, como se puede observar en el lema “avuto *pd* habido, tenido, poseído, recibido” donde ha sido corregido con la forma ‘pp.’. Por otro lado, anotamos el empleo de dos formas para señalar sinónimos, la abreviatura más usada es ‘Sin.’ con mayúscula o minúscula, en cambio Ambruzzi escribe ‘Sinon’ y ‘sinon’. En este caso se ha uniformado con la forma ‘Sin.’.

## 4.2 *Abreviaturas referidas a la marcación*

Se observa que en numerosas ocasiones la abreviatura aparece entre paréntesis, otras en cursiva o redonda sin una aparente explicación. No se ha modificado, ya que puede ser importante para futuros trabajos de investigación.

A propósito de las marcas diafrecuentes, se ha intervenido, principalmente, en las que resultaban poco comprensibles como ‘p. us.’, ya que, en el primer volumen, se refiere a ‘più usato’, mientras que en el segundo se refiere a ‘poco usado’. Por lo tanto, se han sustituido con ‘più us.’ (esp-ita) y ‘poco us.’ (ita-esp). La forma ‘men. us.’ se ha transcrito con ‘menos usado’ así como ‘n. us.’ por ‘nombre us.’, del mismo modo ‘más común’ se ha uniformado con ‘más us.’

En lo que respecta a las marcas cronológicas, la abreviatura ‘ant.’ presenta cierta ambigüedad, ya que puede interpretarse como ‘anticuado’, ‘antiguo’ o ‘antillano’. En este caso se ha optado por conservarla en su forma original, considerando que futuros estudios sobre este aspecto podrían arrojar datos de interés.

Se observa el caso de la marca gramatical correspondiente a ‘impersonal’ que en la lista aparece como ‘impers.’, mientras que Ambrozzi emplea en algunas ocasiones la forma abreviada ‘imp.’. En estos casos hemos cambiado esta última uniformándola a la marca que está en la lista, es decir, ‘impers.’ También la utiliza para aludir a ‘imperio’, en este caso la hemos desarrollado eliminando la ambigüedad presente en ‘imp.’.

Por lo que se refiere a las marcas diaevaluativas o pragmáticas se han mantenido tal como las emplea el autor.

En el volumen español-italiano, observamos que la marca diatécnica correspondiente al ámbito de la medicina presenta formas diferentes: se presenta de una manera en la lista de abreviaturas, y adopta una forma diferente en el cuerpo del diccionario. En la lista emplea ‘med.’ para referirse al ámbito médico, mientras que en la microestructura aparece ‘medic.’ En estos casos se ha optado por registrar la marca completa.

En el mismo volumen, las marcas diatópicas se han simplificado y uniformado de la manera siguiente: ‘spagn.’ → ‘sp.’; ‘ital.’ → ‘it.’

En el caso de la abreviatura ‘rom.’, se ha optado por conservar íntegramente las formas utilizadas por Ambrozzi, con la única excepción de esta marca, interpretada como ‘romano’, decisión tomada siempre en aras de una mayor claridad. Un ejemplo representativo se encuentra en la entrada “*Apuleio* Lit. Lucio Apuleyo, escritor rom.”

Por último, hemos encontrado alguna marca abreviada (ausente en la lista)

que hemos decidido mantener precisamente para no contaminar posibles estudios futuros, como la referida al dólar bajo el lema “*aquila* [...] moneda de a 10 doll. de Estados Unidos [...]”.

Para terminar, cabe señalar que se han eliminado restos de frases que probablemente por distracción se habían quedado en el texto.

## 5. Herramientas utilizadas para la transcripción mediante Transkribus

El proceso de retrodigitalización de una obra implica varias etapas, a veces largas y complejas, que son posibles gracias a las herramientas desarrolladas en el ámbito de las Humanidades Digitales (HD).

Aplicando la distinción de Bazzaco (2020), la primera fase es la digitalización, es decir, el escaneo de la obra en papel en imágenes procesables por un ordenador. A continuación, sigue la digitización: “la transformación de un texto analógico en un texto electrónico en formato mrf (*machine readable form*)” (Bazzaco 2020: 536).

Esta segunda etapa se simplifica gracias a los sistemas OCR (*Optical Character Recognition*) y HTR (*Handwritten Text Recognition*) que permiten el reconocimiento automático y semi-automático de las imágenes del texto y su transcripción a texto electrónico, sustituyendo en parte la transcripción manual (Bazzaco 2018). Estas herramientas, ya útiles desde las primeras fases de desarrollo, en las últimas décadas han mejorado cada vez más gracias a la integración de algoritmos de Inteligencia Artificial (IA) y aprendizaje automático (*machine learning*) que permiten al *software* aprender y mejorar a partir de los datos que proporciona el usuario (Bazzaco 2020), consiguiendo así márgenes de error cada vez menores.

Entre los diferentes recursos disponibles, se ha decidido utilizar la plataforma Transkribus<sup>1</sup>, un *software* en continua actualización que, gracias al uso de sistemas HTR e IA, reconoce tanto textos manuscritos como impresos y permite entrenar modelos personalizados que reconocen las peculiaridades del texto propio. La plataforma nace en 2015 en el seno del proyecto READ y ahora está dirigida por la cooperativa READ-COOP SCE cuyo objetivo es: “*to provide a comprehensive range of tools and services that empower researchers, institutions, and individuals to collaboratively discover and explore the rich tapestry of history*”<sup>2</sup>. Dicha cooperativa cuenta con más de 200 miembros en 30 países del mundo y más de 300 000

1 Cfr. <https://www.transkribus.org/it>

2 Cfr. <https://shorturl.at/6lmiv>



usuarios, lo que lo convierte en uno de los programas de transcripción automática más utilizados.

Para la transcripción de la letra A de los dos volúmenes de la primera edición del diccionario de Ambruzzi, se ha seguido un proceso metodológico similar al que se ha seguido en otros proyectos previos (e.j. Bazzaco 2018, 2020, Nalesso 2024). Tras el escaneo de los dos volúmenes, se procedió a trabajar en la plataforma *Transkribus*, comenzando con la carga de las imágenes. El siguiente paso fue el *Layout Analysis*, es decir, la segmentación de cada página en varias áreas para que el *software* leyera el texto en el orden apropiado. En el caso de nuestra obra, como es típico en los diccionarios bilingües, la página está dividida en dos columnas, por lo que fue necesario segmentarla en dos regiones para que se leyera todo correctamente. A continuación, habría sido necesario transcribir manualmente unas páginas para crear el *ground truth*, la base para el desarrollo de un modelo personalizado. Sin embargo, como sugiere la misma plataforma de *Transkribus*, las páginas de *ground truth* se transcribieron automáticamente con un modelo público disponible para todos los usuarios. De hecho, tanto para el primer volumen (esp-ita) como para el segundo (ita-esp), se utilizó inicialmente el modelo *TRANSKRIBUS PRINT MI* entrenado con 10 281 páginas y 14 idiomas diferentes, español e italiano incluidos. Las páginas transcritas con este modelo se corrigieron manualmente para lograr transcripciones perfectas y luego se utilizaron como entrenamiento para los dos primeros modelos personalizados.

A pesar de los diferentes problemas y procedimientos, el resultado final ha sido una transcripción correcta de la primera letra del diccionario, descargada de la plataforma en formato TXT.

A continuación, se analiza más detalladamente el proceso de transcripción y entrenamiento, los errores que surgieron en ese proceso y las correcciones que se hicieron manualmente para lograr una transcripción correcta de la letra A.

## 6. Primer volumen SPA-ITA

Este primer volumen de la primera edición del DBA se titula *Nuovo Dizionario Spagnolo-Italiano* y, aunque por cuestiones de *copyright* siempre aparece como publicado junto con el segundo volumen en 1949, en realidad apareció un año antes (Bermejo 2008).

En cuanto a la macroestructura, cuyas páginas están marcadas con números romanos, hay 4 partes: la ‘Presentazione’, la ‘Pronunzia dello spagnolo’ y las ‘Abbreviature’ están escritas en italiano mientras que la ‘Advertencia’ está en español porque el autor se dirige a los hispanohablantes explicando el sistema de la acentuación italiana y cómo les ayudará a entender la pronunciación de las palabras.

Por lo que respecta a la microestructura, cuenta con un total de 1 095 páginas y aproximadamente 65 000 entradas. Como ya se ha explicado, dada la extensión y la abundancia de información, para esta primera parte del trabajo nos hemos centrado únicamente en la digitación de la letra A, que consta de 147 páginas y 9 987 entradas.

### 6.1 *Transkribus: problemas y resultados*

A continuación, mostramos la metodología empleada para la transcripción. En primer lugar, se escaneó el diccionario, seguidamente se subieron las imágenes a *Transkribus* para empezar el proceso de transcripción automática. La parte transcrita con el modelo público *TRANSKRIBUS PRINT M1* fueron las primeras 14 páginas, es decir, el 10% del total. En el caso del primer volumen, el modelo entrenado con estas primeras páginas fue tan eficaz que no se consideró necesario entrenar nuevamente el *software*. Seguidamente, se presentan los principales errores antes y después del entrenamiento para mostrar su eficacia.

La transcripción con el modelo público contaba 1 358 errores, con un mínimo de 60 fallos en la página que contaba un número menor y un máximo de 133 equivocaciones en la página con mayor cantidad.

En cuanto a la tipología de erratas pueden identificarse 5 categorías diferentes:

#### 1. Ausencia de signos de puntuación

En todas las páginas faltan casi todas las barras verticales ( | ) que el autor utiliza para dividir las diferentes acepciones de un lema, y las barras dobles ( || ) que se emplean cuando una palabra cambia también de categoría gramatical (Bermejo 2008). De hecho, del total de los errores en estas 14 páginas, 583 son omisiones de barra vertical.

Además, a menudo omite los guiones para dividir palabras al final de las líneas y rayas (—) utilizadas como sustitutos del lema en los ejemplos. Por lo que se refiere a puntos, comas, paréntesis y signos de exclamación e interrogativos, normalmente el programa los transcribe.

## 2. Equivocación entre signos de puntuación

Como hemos adelantado, el guion (salvo cuando se omite) se transcribe con el símbolo ‘-’. Este es el error más común de este grupo y se repite en todas las páginas con un total de 155 veces. Además de confundir la coma con el punto en 11 páginas, solo se equivoca en otras ocasiones alterando signos parecidos como los exclamativos de apertura o cierre con barras o doble barras verticales o puntos y comas y dos puntos con signos de exclamación.

## 3. Ausencia de letras o abreviaturas

En esta primera transcripción sin entrenamiento, el programa no leía algunas letras de ciertas palabras, como por ejemplo la ‘i’, la ‘u’ o la ‘o’. Por otro lado, tampoco ha reconocido en la transcripción algunas marcas gramaticales, como la ‘r.’ para los verbos reflexivos, ‘s.’ para los sustantivos, ‘agg.’ para marcar los adjetivos y en cuatro casos no aparece la abreviatura ‘fig.’ que señala el sentido figurado de las palabras.

## 4. Equivocación entre letras

Este es el grupo donde aparece la mayor variedad de errores, aunque no se repiten con frecuencia. Salvo en los casos de confusión entre ‘l’ y ‘f’ en la marca gramatical ‘sf.’, así como entre las letras ‘l’ e ‘i’—errores que se registran con una frecuencia de 41 y 31 ocurrencias respectivamente—, el programa presenta numerosos errores en la transcripción de letras, sustituyéndolas por otras. No obstante, dichos errores no se repiten de manera sistemática.

En tres ocasiones acorta algunas abreviaturas escribiendo solo ‘f.’ o ‘g.’ en lugar de ‘fig.’ o ‘a.’ para ‘agg.’ y dos veces escribe la letra ‘j’ confundiéndola con los signos de exclamación de apertura y cierre. Altera algunas letras mayúsculas, como la ‘T’ en lugar de la ‘l’ o la ‘l’ en lugar de la ‘i’; sin embargo, en general, presenta dificultades en el reconocimiento de determinadas letras, que a menudo confunde entre sí. Por ejemplo:

- ‘t’ en lugar de ‘i, í, f o l’
- ‘e’ en lugar de ‘v, c, o’
- ‘o’ en lugar de ‘g, e, o, c’
- ‘m’ escrita como ‘n, ni, nt’

## 5. Desaciertos en la acentuación

Siendo un diccionario español-italiano donde el autor marca también la sílaba

tónica de algunas palabras italianas, hay muchísimos acentos que el *software* no siempre transcribe o reconoce bien. El error más común es la tilde en la letra ‘i’; este aparece, por lo menos una vez en todas las páginas, con un total de 57 veces.

Todos los errores en estas páginas se corrigieron manualmente y luego constituyeron la base para entrenar un nuevo modelo capaz de leer mejor las letras y los símbolos de este volumen.

En las 133 páginas siguientes se utilizó el nuevo modelo personalizado y de 1 358 errores en 14 páginas se obtuvo un resultado de 1 860 errores en 133. Por lo tanto, proporcionalmente, el sistema entrenado funcionó mejor.

En cuanto a la tipología, los errores son similares a los de antes del entrenamiento, pero el número ha disminuido notablemente. El más común sigue siendo la ausencia de barra vertical ‘|’, que ahora aparece en 128 páginas con un total de 475 ausencias. Si antes el segundo tipo de equivocación era la omisión de la doble barra, ahora son las rayas (101 veces) y los guiones para dividir la sílaba al final de las líneas que ya no se confunden con el símbolo ‘¬’. Sin embargo, se omiten en 61 páginas con un total de 105 veces. Las comas y los puntos resultan ausentes 42 veces y hay muy pocas omisiones de paréntesis, signos de exclamación, puntos y comas o dos puntos.

Por lo que respecta a los errores de confusión entre letras y signos de puntuación, se presentan también después del entrenamiento. En cuanto a la equivocación entre signos, como hay más páginas, emerge una mayor variedad de errores. El más repetido es la barra vertical en lugar de la doble barra, que aparece 25 veces, seguido de la confusión entre coma y punto. La barra vertical se transcribe algunas veces también en lugar de punto, raya, guion, coma, paréntesis o las letras ‘f’ o ‘i’, así como la coma confundida con paréntesis, guion, punto y coma o dos puntos. En las páginas de *ground truth* no aparecen comillas angulares («»), a veces utilizadas por el autor para las citas, y por eso, cuando se dan, la plataforma las transcribe de forma distinta: una vez con paréntesis ‘(’, y dos veces con ‘s’ y ‘v’; lo mismo con los corchetes ‘[ ]’, que dos veces se confunden con paréntesis de apertura y otras dos con la letra ‘l’, y con los números que se mezclan entre ellos. El número ‘4’ se transcribe una vez con la letra ‘l’, el ‘58’ con ‘ss’, el ‘7’ con ‘T’ y el número ordinal ‘2º’ que aparece una vez como ‘?’ y otra como ‘ze’. En cuanto a las letras, se confunde con otras similares como la ‘f’ en lugar de ‘l’ o la ‘l’ en lugar de ‘i’, los dos errores más comunes, mientras que otras letras aparecen ocasionalmente por error. Como antes, casi todas las erratas con los acentos son la ausencia de tilde o confusión entre acento agudo y grave, de hecho, la más común sigue siendo la omisión de tilde sobre la ‘i’. Sin embargo, en estas páginas

se encuentran dos acentos que no estaban en las primeras catorce: la diéresis en la ‘u’ y el circunflejo en la ‘i’. El primero casi siempre se transcribe como ‘ú’ o ‘u’ en lugar de ‘ü’, mientras que ‘í’ aparece algunas veces sin acento.

El único tipo de error que no aparece en la transcripción con el modelo público, y en cambio sí se encuentra después, es la inclusión de signos de puntuación y letras donde, en realidad, no las hay. A lo largo de las 133 páginas, el programa añade dos ‘i’ y una ‘c’, así como 29 barras verticales, 25 guiones, 15 rayas, un paréntesis y algunos puntos donde el original no los presenta.

En general, gracias a este breve resumen del número y el tipo de errores que produce *Transkribus* antes y después del entrenamiento con la IA, se puede ver cómo funciona este último y conduce a una disminución significativa de los errores, haciendo que el proceso de transcripción automática sea eficaz y útil durante la digitización.

## 6.2 *Detección y corrección de erratas en la versión en papel*

En el apartado anterior hemos hablado de los errores cometidos por el programa de transcripción automática. Ahora, en cambio, se presentan algunos desaciertos e incongruencias detectados en el volumen impreso, que se corrigieron y modificaron manteniendo el criterio de claridad y uniformidad. Durante la corrección de la transcripción se observó que la obra presenta erratas, a veces probablemente de impresión, y contradicciones con lo declarado en la macroestructura.

Entre los principales cambios realizados se encuentran los referidos a las abreviaturas y marcas de uso. En las entradas del diccionario se observa un considerable número de palabras que, si bien no corresponden a marcas de uso, se abreviaron seguramente por razones de espacio y formato de impresión. Sin embargo, dado que una edición digital no tiene tales limitaciones, se optó por desarrollar la mayoría de estas formas abreviadas con el objetivo de facilitar la comprensión del texto. Muchas de las palabras que se han extendido no figuran en la lista de ‘Abbreviature’ proporcionada por el autor. Excluyendo las repeticiones, se han identificado 97 términos abreviados que no están registrados en dicha lista. Algunos ejemplos son: ‘alfab.’ por ‘alfabeto’, ‘condiz’. por ‘condizione’, ‘esclam’. o ‘escl’. por ‘esclamazione’, y ‘mis.’ por ‘misura’, esta última con tres apariciones. Asimismo, se hallaron abreviaturas que sí están incluidas en el listado original, así como formas abreviadas que difieren de las registradas por el autor. En ambos casos, dado que no se trataba de marcas de uso, se ha decidido igualmente

expandir las palabras. En el primer caso—abreviaturas coincidentes con la lista— se contabilizaron 164 casos (incluyendo repeticiones), entre los cuales destacan: ‘dic.’ por ‘dicesi’ (la más frecuente, con 50 apariciones), ‘qc.’ y ‘qu.’ por ‘qualcosa’ y ‘qualcuno’, ‘c.’ y ‘ab.’ por ‘città’ y ‘abitanti’, o ‘sim.’ por ‘simile’. En el segundo caso—abreviaturas con formas diferentes a las del listado— se registraron 31 casos. Algunos ejemplos representativos son:

- ‘accusat.’ para ‘accusativo’ que es en la lista aparece como ‘acc.’ y ‘accusat.’
- ‘medic.’ para ‘medicinale’ que en la lista se abrevia ‘med.’
- ‘abit.’ para ‘abitante’ que, como acabamos de ver, en la lista aparece como ‘ab.’
- ‘tradiz.’ se abrevia ‘trad.’ para ‘tradizione’

En cuanto a las marcas de uso, casi todas están en el listado de Ambruzzi, sin embargo, figuran numerosas incoherencias en la forma. Como ya se ha explicado anteriormente, el sistema de abreviaturas creado por el autor es muy amplio y, ya en la macroestructura, presenta distintas formas para una misma información. Por otro lado, la situación se complica aún más por el hecho de que en la microestructura, además de las numerosas formas de la lista, se encuentran otras adicionales. A continuación, se presentan los tres escenarios encontrados y los cambios que se han realizado:

1. En ‘Abbreviature’ aparecen dos formas, pero en la microestructura se utiliza casi siempre solo una. En este caso se vio cuál era la marca más frecuente y se uniformó todo según esta. Aunque la tabla se refiere únicamente a la letra A, las decisiones se han tomado sobre la base de la frecuencia. Por ello, si bien la marca ‘gal.’ aparece solo cinco veces en la letra A, es en realidad la más frecuente a lo largo de todo el diccionario.

LISTA				
<i>ital. – it.</i>	it.	16	ital.	5
<i>spagn. – sp.</i>	sp.	16	spagn.	7
<i>gall. – gal.</i>	gal.	5	gall.	24
<i>Ven. – Venez.</i>	Venez.	70	Ven.	1
<i>Comm. – Com.</i>	Com.	41	Comm.	11
<i>Eq. – Equat.</i>	Equat.	69	Eq.	4
<i>Ed. – Edil.</i>	Edil.	7	Ed.	1

Fig. 3 Marcas presentes en la lista con dos formas

2. Marcas de uso que en la lista tienen una forma y en la microestructura otra e, incluso, en raras ocasiones, se usa otra diferente que no está presente.

LISTA	ENTRADAS	
Arg.	Argent.	1
am.; amer.	americ.	2
dif.	difett.	2
Fil.	Filipp.	2
Pp.; p.p	pp.	11
Ven.	Venat.	1
Par.	Parag.	1
lat.	l.	1
superl.	superlat.	1
	sup.	8
St.	Stor.	3
loc.	locuz.	2
	l.	1
Fot.	fotogr.	1
et.; etim.	etimol.	1
inv.	invar.	1
Gram.	Gramm.	4
acc.; accr.	accresc.	2
Am. M.	Amer. Merid.	5
per est.	per estens.	2
f.; fem.	femm.	5
Mess.	Mes.	1
Dom.; Domen.	Domin.	1
antill.	antil.	2
Ant.	Antil.	5
spr.	spreg.	1
con.; coniug.	si con.	3
Filos.	Fil.	1
Metal.	Met.	1
fr. prov.	fr. proverb.	1

Fig. 4 Marcas diferentes en las entradas

Aquí, son escasas las ocasiones en las que aparece la forma que no está en la lista, en estos casos se cambió por la de la lista que es también la más común.

3. Marcas que en el listado tienen una forma que apenas se utiliza porque es más frecuente otra. En estos casos se ha dejado la marca más utilizada, aunque no figure en la lista.

Estas son:

- ‘Met.’, marca diatécnica para ‘Metereologia’ que solo aparece 5 veces en todo el volumen, sustituida por la más usada ‘Meteor.’
- ‘C. Am.’, marca diatópica para ‘America Centrale’, se usa solamente 26 veces en todo el diccionario frente a las 368 ocasiones en las que aparece ‘C. Amer.’
- ‘Per.’, marca diatópica para ‘Perú’ aparece solo una vez, mientras que las otras veces se escribe por entero.

Las únicas dos excepciones que no entran en estas categorías son la abreviatura ‘Centroam.’ que no se cambia por la marca diatópica presente en la lista, sino con la marca más frecuente, es decir, ‘C. Amer.’, y las dos veces que figura ‘Cub.’, ausente en la lista, ya que como marca diatópica para referirse a Cuba se utiliza el nombre entero.

Además de los cambios con las abreviaturas y marcas de uso, también se corrigieron otros errores encontrados en la primera letra.

#### 1. Fallos de puntuación

En las entradas, el autor utiliza diferentes signos de puntuación que, a veces, son incorrectos. Muchas veces faltan ciertos signos como el punto al final de una entrada o después de una abreviatura, la coma para separar los equivalentes, el guion para dividir una palabra al final de la línea o los paréntesis de apertura o cierre.

En ese grupo se pueden incluir también la ausencia del guion para marcar la desinencia del femenino.

#### 2. Errores de acentuación

Con respecto a las palabras españolas, hay tantos lemas como palabras dentro de las entradas que llevan tilde donde no deberían como: ‘afluír’ → ‘afluir’, ‘Albanía’ → ‘Albania’ o ‘algortimía’ → ‘algortimia’, y otras que no la llevan (e.j. ‘actínico’ → ‘actínico’ o ‘aligator’ → ‘aligátor’). Por otra parte, en el diccionario de Ambruzzi, además de los acentos españoles, el autor marca también la sílaba tónica de algunas palabras italianas para ayudar a quienes no conocen bien el idioma a entender la pronunciación correcta. En la ‘Advertencia’ afirma que pon-



drá el acento agudo sobre las vocales cerradas (é, í, ó, ú) y el grave sobre las vocales abiertas (à, è, ò). Sin embargo, se presentan algunas excepciones en las que no utiliza el acento correcto con respecto a lo que declara en la macroestructura. Por ejemplo: ‘rullìo’ → ‘rullío’, ‘fucìle’ → ‘fucíle’ o ‘cùmulo’ → ‘cúmulo’. En total se cuentan 7 fallos en los acentos italianos.

### 3. Marcas y palabras en español

En el primer volumen esp-ita, los lemas están en español y las marcas y las entradas en italiano con excepciones de los ejemplos en español. A pesar de eso, en algunos casos se encontraron palabras escritas en español donde deberían haber estado en italiano. Ocurre dos veces con la palabra ‘Córdoba’ y una vez con ‘Castilla’ que deberían ser ‘Cordova’ y ‘Castiglia’ respectivamente, y con las marcas:

- ‘Chile’ para ‘Cile’ (8 veces)
- ‘Cile y Esp.’ para ‘Cile e Sp.’
- ‘desp.’ (despectivo) para ‘dispr.’ (dispregiativo)

El error más significativo que no se ha modificado es la presencia de la entrada “ascetismo” escrita toda en español.

### 4. Errores de imprenta

Se incluyen aquí todos los casos en los que una palabra o frase presenta errores ortográficos debidos a la omisión de una letra—como en ‘essicare’ por ‘essicare’ u ‘olfato’ por ‘olfatto’—o a la sustitución de una letra por otra, como en ‘doce’ por ‘dove’, ‘alealoideo’ por ‘alcaloideo’ o ‘agalologia’ por ‘agatologia’. Este tipo de errores se observa ocasionalmente también en las marcas, por ejemplo: ‘sgg’ en lugar de ‘agg’ o ‘Zoot’ en lugar de ‘Zool’.

Además, se señala que en el primer volumen se encontró un lema repetido dos veces y un orden alfabético incorrecto que, aunque en un tesoro lexicográfico no resulta un problema, en la corrección de la transcripción se ha corregido.

A pesar de los varios cambios que se hicieron para que el volumen resulte lo más uniforme y claro posible, hay algunos aspectos que no se han abordado en este trabajo, como, por ejemplo, los significados de la marca ‘ant.’ o el aspecto metalexicográfico relacionado con el uso de la puntuación, las mayúsculas, la cursiva y la negrita, que presentan varias incongruencias.

## 7. Segundo volumen ITA-SPA (1949)

El segundo volumen de la primera edición del DBA se titula *Nuovo Dizionario*

*Italiano-Spagnolo* publicado en el año 1949 (Bermejo 2008).

El volumen está compuesto por 1 337 páginas que comprenden una introducción en la que el autor presenta la historia de la lengua italiana con anotaciones sobre la fonética, la morfología y la sintaxis.

Además, dedica un apartado a la pronunciación de vocales y consonantes, una sección sobre las reglas de acentuación, subrayando también que la lengua italiana no tiene reglas precisas de acentuación gráfica como el español. A continuación, incluye otras notas sobre el uso de preposiciones, partículas relativas y nexos, así como sobre las fuentes utilizadas para elaborar el volumen y, por último, la lista de abreviaturas. Al final del volumen, añade un anexo con algunas erratas. Como ya se ha explicado, dada la extensión y la abundancia de información, en esta primera parte del trabajo —al igual que en el primer volumen— nos hemos centrado únicamente en la digitación de la letra A, que consta de 131 páginas y entradas.

### 7.1 Transkribus: *problemas y resultados*

Al igual que para la transcripción del primer volumen, para la del segundo se ha utilizado el programa *Transkribus*.

Inicialmente, se utilizó el modelo del programa, *TRANSKRIBUS PRINT M1*, a partir del cual se detectaron numerosos errores.

Los principales problemas se debieron principalmente a la escasa legibilidad del PDF, que no permitía distinguir letras aparentemente similares; a los signos añadidos posteriormente a lápiz o bolígrafo; a las líneas discontinuas especialmente en correspondencia del guion largo, utilizado para sustituir el lema italiano en los ejemplos, seguido de la pleca; y a la ausencia o inversión de acentos tanto en palabras italianas como españolas.

A diferencia del primer volumen, la transcripción del segundo exigió la creación de tres modelos más en los que, aunque mínimas, se observaron mejoras. Si se calcula el promedio de errores, los resultados obtenidos son los siguientes: promedio de página 1 a 13 (modelo *Transkribus*): 115 erratas; de página 14 a 30 (modelo I): 62 equívocos; de página 31 (modelo II): 88 faltas; de página 100 (modelo III): 55 fallos.

A continuación, la lista de los errores más frecuentes que se han corregido manualmente a lo largo de la letra A.

- —
- - (al final de renglón)
- ) → j
- / → f
- : → << >>
- ¬ → -
- l → i
- A → À
- a → á; à; e; g; o; c
- à → á; u
- á → à
- \* → a
- b → h
- c → e; v
- D → B
- é → á
- è → è
- e → o; ú
- f → i; t
- fg. → fig.
- í → á; f
- i → f; í; j; ú;
- í → i; s
- ifl → ¼
- il → XII
- j → f; p; h
- l → !; 3º; f; h; í; i, ñ; t; T
- L → í
- l 9y → 107
- ł → l
- ll → II
- m → n
- ñ → f
- n → u
- ó → è
- o → e
- P → V

$q \rightarrow g$   
 $r \rightarrow m; v$   
 $sfs \rightarrow 678$   
 $t \rightarrow f; i; í; l$   
 $u \rightarrow g; ú; a; n$   
 $ú \rightarrow u$   
 $v \rightarrow n$   
 $V \rightarrow y$

## 7.2 Detección y corrección de erratas en la versión en papel

Además de los errores derivados de la transcripción, se detectaron numerosas erratas. Concretamente, tras la transcripción, se exportó el archivo en TXT sobre el que se realizó una revisión minuciosa para corregir lemas, abreviaciones, marcas de uso, origen, equivalentes, ejemplos y fraseología. En vista del gran número de fallos encontrados en el segundo volumen (387 totales), se optó por dividirlos en 23 grupos y establecer una tipología.

Con el fin de ser más precisos, a continuación, se presentan los grupos y la tipología acompañados por el número total de las erratas encontradas y los ejemplos más explicativos. En el caso de los grupos más numerosos se ha indicado un ejemplo por tipo.

1. División silábica incorrecta al final de renglón (9 veces): (lema 'adito') recur- -rido  $\rightarrow$  recu- -rrido; (lema 'anillo') epi- -scopal  $\rightarrow$  epis- -copal.
2. Inversión de letras presentes en lemas, marcas, equivalentes, plurales, etc. (8 veces): (lema 'A.E.I.O.U.') Absburgo  $\rightarrow$  Asburgo; (lema 'affannare') hjio  $\rightarrow$  hijo; (lema 'ammodo') ocn  $\rightarrow$  con; (lema 'ausònico') -ic  $\rightarrow$  -ci; (lema 'Australia') Geogr.  $\rightarrow$  Geogr.; (lema 'assiepare') rt.  $\rightarrow$  tr.; (lema 'astuccio') caijtaà  $\rightarrow$  cajita; (lema 'atteggiare') acittud  $\rightarrow$  actitud.
3. Falta de orden alfabético en las entradas (5 veces): sobre los lemas: 'appillottare' desplazado antes de 'appinzare'; 'animetta' desplazado antes de 'animismo'; 'allotriofagia' desplazado antes de 'alotropía'; 'astensionista' desplazado antes de 'aster'; 'annaffiata' desplazado antes de 'annaffiato'.
4. Errores que quedan por descifrar: (lema 'año') corrtre.
5. Sustitución de letras que se pronuncian igual (2 veces): (lema 'agenzia') mensagería  $\rightarrow$  mensajería; (lema 'avventato') lijero  $\rightarrow$  ligero.
6. Mayúsculas (2 veces): (lema 'ascrittizio') arqueol.  $\rightarrow$  Arqueol.; (lema 'Attico') ático  $\rightarrow$  Ático, al tratarse de un nombre propio, se ha corregido con la mayúscula.
7. Uso de letras por otras (44 veces): (lema 'aggiustatezza') precición  $\rightarrow$  precisión; (lema

- ‘Alighieri’) Danre → Dante; (lema ‘automobile’) velocità → velocidades; (lema ‘arricciare’) francir → fruncir; (lema ‘arruffato’) anmarañado → enmarañado; (lema ‘aspettativa’) expectatica → expectativa; (lema ‘addestrare’) sorvirle → servirle; (lema ‘agitazione’) commoción → conmoción; (lema ‘ameno’) extrano → extraño; (lema ‘appiccatura’) añabidura → añadidura; (lema ‘Aristarco’) gramárico → gramático; (lema ‘appoggio’) Ginn. → Gimn.; (lema ‘àrnica’) Rot. → Bot.; (lema ‘ardore’) sn. → sm.
8. Empleo de palabras en italiano en frases españolas (8 veces): (lema ‘accrèscere’) famiglia → familia; (lema ‘agitare’) notizia → noticia; (lema ‘alto’) elevato → elevado; (lema ‘aggiogàbile’) che → que; (lema ‘applicazione’) ingegnería → ingeniería; (lema ‘abbominio’) e deriv. → y deriv.; (lema ‘asfissiante’) asfissiante → asfixiante; (lema ‘assuntivo’) gentilizio → gentilicio.
  9. Palabras que terminan en *-tor* en lugar de *-dor* (3 veces): (lema ‘aletta’) radiator → radiador; (lema ‘apòstolo’) predicator → predicador; (lema ‘aratro’) regulator → regulador.
  10. Términos españoles con acentos italianos (8 veces): (lema ‘annunciare’) età → está; (lema ‘acinoso’) glàndula → glándula; (lema ‘Aventino’) estòmagò → estómago; (lema ‘Àbila’) Àbila → Ábila; (lema ‘Àngelo’). Àngel → Ángel; (lema ‘Àtropo’) Àtropos → Átropos; (lema ‘Aristòfane’) comediògrafo → comediógrafo; (lema ‘ànitra’) ànade → ánade.
  11. Palabras que deben separarse (3 veces): (lema ‘accozzare’) cadauno → cada uno; (lema ‘alcìonico’) sunido → su nido; (lema ‘anagíride’) leñoh. → leño h.
  12. Repeticiones (5 veces): antinazionale (lemma) → antinazionale (equivalente); (lema ‘aiutante’) id. id. → id.; el lema ‘atomística’ se repite dos veces, uno con acento fónico y el otro sin acento, respetando las indicaciones de Ambruzzi, hemos decidido dejar el lema con acento fónico; repetición del lema ‘avvampamento’; repetición de la marca adj. (lema ‘antipùtrido’).
  13. Presencia de acentos en palabras que no los llevan (9 veces): (lema ‘addentallato’) denticulato → denticulado; (lema ‘accerito’) enrojécido → enrojecido; (lema ‘Aladino’) medío → medio; (lema ‘ametista’) amétista → ametista; (lema ‘ametria’) arritmía → arritmia; (lema ‘animaletto’) cojfo → cojijo; (lema ‘antirelligioso’) antirreligioso → antirreligioso; (lema ‘angelicamente’) angélicamente → angelicamente; (lema ‘accordare’) sentímiento → sentimiento.
  14. Letras o sílabas que se omiten o añaden (31 veces): (lemma ‘ammazzare’) a polar → apiolar; lema ‘adustar’ → adustare; (lema ‘azzardare’) quell’uomo (ejemplo) → quell’uomo; (lema ‘alessandrinismo’) alejandrinismo; (lema ‘arma’) estar bajo las, → estar bajo las a.; (lema ‘autenticità’) autencidad → autenticidad; (lema ‘asserire’) en las conjugaciones Pre. → Pres.
  15. Palabras divididas (3 veces): (lema ‘affezionare’) n os → nos; (lema ‘affocare’) pasio nes → pasiones.
  16. Lema en español: ‘aerificación’ → aerificazione.
  17. Letras dobles (5 veces): (lema ‘abbattimento’) postracción → postración; (lema

- ‘apostolato’) misión → misión; (lema ‘arcifanfano’) fanfarón → fanfarrón.
18. Letras que hay que eliminar (3 veces): (lema ‘avaccio’) prontitudo → prontitud; (lema ‘avanti’) ;marchen, mar! → marchen, ar!
  19. Acento en el diptongo *-ui* (24 veces): (lema ‘accomiatare’) fuímos → fuimos (ejemplo); (lema ‘accluso’) incluído → incluido (equivalente); (lema ‘agnosticismo’) descuído → descuido; (lema ‘altruismo’) altruísmo → altruismo; (lema ‘acuto’) buído/buída → buido/buida; (lema ‘assommare’) concluír → concluir; (lema ‘assentire’) anuír → anuir; (lema ‘avvilire’) ruín → ruin.
  20. Ausencia de acentos en palabras españolas (31 veces): (lema ‘appello’) tambien → también; (lema ‘appaltatore’) compaña → compañía; (lema ‘attingimento’) liquidos → líquidos; (lema ‘atleta’) campeon → campeón; (lema ‘attivístico’) activistico → activístico; (lema ‘avvertire’) cuidate → cuídate; (lema ‘acciò’) tu → tú; (lema ‘aria’) ventilacion → ventilación; (lema ‘alzata’) mas alto → más alto.
  21. Ausencia de *-s* final (verbos, artículos, sustantivos, etc.) (6 veces): (lema ‘antenora’) traidore → traidores; (lema ‘abitino’) terciario → terciarios; (lema ‘Adamástor’) la → las; (lema ‘affinché’) pueda → puedas; (lema ‘Ampolla’) lo → los.
  22. Errores de imprenta en las marcas (acentuación y varias) (96 veces): iron → irón; Quim → Quím; ár. → árabe o voz árabe; Amer. → Amér.; contracc. → contrac.; Sider. → Sid.; St → Hist; Méjco → Méj.; Geograf. → Geogr.; superlat. → superl.; infinit. → infin.; transit. → tr. - intransit. → intr.; s. → sf.; anglos. → ingl.; agg. → adj.; avv. → adv.; abr. → abrev.; antig. → ant.
  23. Errores de puntuación (80 veces): ausencia de paréntesis al final del artículo lexicográfico; ausencia del punto final; presencia del punto final dentro del paréntesis; sustitución de punto con la coma y viceversa; falta del guion en los plurales y al final del renglón; falta de la coma entre equivalentes; falta el punto después de las marcas; presencia de puntos, comas, paréntesis y plecas no necesarios; sustitución de la doble pleca con el punto.

## 8. Conclusiones

Tal como se ha expuesto a lo largo de este estudio, el diccionario Ambruzzi constituye una obra valiosa por la abundancia y diversidad de datos que proporciona al lector. Sin embargo, esta misma riqueza representa un desafío en el proceso de digitalización, ya que requiere la unificación de múltiples aspectos para alcanzar una coherencia interna. Esta coherencia, sin duda, no estaba al alcance del autor debido a las limitaciones de los medios manuales con los que trabajó. En este proyecto, nuestro principal objetivo ha sido resolver dichas inconsistencias, tal como se ha detallado en esta investigación.

## Bibliografía citada

- Bazzaco, Stefano (2018), “El Progetto Mambrino y las tecnologías OCR: estado de la cuestión”, *Historias Fingidas*, 6: 257–72 [03/04/2025] <<https://historiasfingidas.dlls.univr.it/article/view/89/135>>
- Bazzaco, Stefano (2020), “Siglo de Oro: creación de un modelo HTR basado en libros de caballerías del siglo XVI en la plataforma Transkribus”, *JANUS*, 9: 534–61 [03/04/2025] <<https://www.janusdigital.es/descargar.htm?id=160>>
- Bermejo Calleja, M. Felisa (2008), “El Nuovo dizionario spagnolo-italiano e italiano-spagnolo (1948-1949) de L. Ambruzzi”, *Textos fundamentales de la lexicografía italoespañola (1917-2007)*, ed. Félix San Vicente. Milán, Polimetrica, Vol. 1: 125-97.
- Calvi, Maria Vittoria (2006), “Il componente culturale del dizionario Ambruzzi di spagnolo e italiano”, *Lessicografia bilingue e traduzione: metodi, strumenti, approcci attuali*, ed. Félix San Vicente. Monza, Polimetrica: 83-10 [03/04/2025] <<https://www.yumpu.com/it/document/read/16097139/open-access-publications-eurac>>
- Hausmann, Franz Josef (1989), “Die Markierung in allgemeinen einsprachigen Wörterbuch: eine Übersicht”, 1989-1991, Wörterbücher: Ein internationales handbuch zur lexikographie, eds. Franz Josef Hausmann, Oskar Reichman, Herbert Ernst Wiegand, Ladislav Zgusta. Berlin-New York, Walter de Gruyter, 3 vols: 649-57.
- Khemakhem, M. *et al.* (2018), “Fueling Time Machine: Information Extraction from Retro-Digitised Address Directories. Leveraging Open Data”. Tokio, Japón.
- Nalesso, Giulia (2024), “Humanidades digitales y lexicografía bilingüe: recuperación y valorización del patrimonio lexicográfico español-italiano (REVALSI)”, *Lexicografía del español: internacionalización y contrastes*, eds. Alejandro Fajardo Aguirre; Dolores Torres Medina; Cristian Díaz Rodríguez. Frankfurt am Main, Peter Lang: 117-34.
- San Vicente, Félix ed. (2008-2010), *Textos Fundamentales de la lexicografía italoespañola (1917-2007) / (1805-1916)*, Monza, Polimetrica.
- Terras, Melissa (2010), “The Rise of Digitization: An Overview”, *Digital Libraries*, ed. Rico Rukowski. Países Bajos, Sense Publishers: 3–20.
- Valero Gisbert, María J. (2016), “Fraseología y marcas diacrónicas en la lexicografía bilingüe de los diccionarios más representativos de español/italiano del s. XXI”, *El diccionario en la encrucijada: de la sintaxis y la cultura al desafío digital*, eds. Ignacio, Sario López; Cecilio, Garriga Santander, AELEX: 395-411.
- Valero Gisbert, María J. (2018), “Las marcas cronológicas en la lexicografía bilingüe italoespañola del XIX y XX”, *Quaderns d'Italia*, 23: 131-44.
- Valero Gisbert, María J. (2020), “Fraseología y marcas diatópicas del español en la lexicografía bilingüe italoespañola”, *ELUA: Fraseología y variaciones (socio)lingüísticas y diatópicas*, eds. Pedro Mogorrón Huerta, Analía Cuadrado Rey, Anexo 7: 161-79.

**María Felisa Bermejo Calleja**, doctora en Filología Romanza e Linguistica Generale (Università di Perugia), fue profesora titular de Lingua, Traduzione e Linguistica Spagnola, Departamento de Lingue e Letterature straniere e culture moderne de la Università degli Studi di Torino. Entre sus áreas de investigación se pueden distinguir las siguientes: gramática de la lengua española; estudio contrastivo de la lengua española e italiana; metalexicografía bilingüe italiano-español.

**Giulia Lanteri**, tras una licenciatura en Ciencias de la Mediación Lingüística y un máster en Lenguas Extranjeras para la Comunicación Lingüística, prosigue sus estudios lingüísticos centrándose en la lexicografía española. Actualmente es estudiante de doctorado en Humanidades Digitales en las Universidades de Turín y Génova con un proyecto dedicado al estudio del lenguaje juvenil español en las redes sociales.

**giulia.lanteri@hotmail.com**

**María Joaquina Valero Gisbert** es doctora en Lingüística por la Universidad de Valencia (España). Profesora titular de Lingua, Traduzione e Linguistica Spagnola en la Universidad de Parma, Departamento de Discipline Umanistiche, Sociali e delle Imprese Culturali, unità di Lingue e Letterature Straniere. Sus áreas de investigación son la lexicografía, la fraseología hispánica, la traducción audiovisual y la audiodescripción. Dirige el máster oficial 'MTAV' en Traducción Audiovisual (<https://elly2022.postlaurea.unipr.it/course/index.php?categoryid=27>).

**mariajoaquina.valerogisbert@unipr.it**

**Maria Carmela Zaccone** es doctora en Digital Humanities en las Universidades de Génova y Turín. Actualmente es profesora contratada de Lengua Española en la Universidad de Milán y la de Turín y post-doc en lexicografía y fraseología en la Universidad de Turín. Su investigación se sitúa en el campo de la fraseología española e italiana y la fraseodidáctica digital. Su experiencia en el campo de la didáctica le ha permitido explorar las nuevas tecnologías utilizadas en la enseñanza ELE. Ha participado en varios congresos de carácter general y específico de su campo de investigación.

**mariacarmela.zaccone@unito.it**



# HUGO E. LOMBARDINI

## EL ESPACIO DE LO FEMENINO EN EL ÁREA LEXICOGRÁFICA DE *L'ITALIANO ISTRUITO...* DE F. MARÍN

Università di Bologna

### Resumen

Este artículo analiza el tratamiento de lo femenino en las partes lexicográficas de *L'italiano istruito nella cognizione della lingua spagnuola* de Francisco Marín. Para ello explora las cuestiones relacionadas con el género tanto en su dimensión lingüística como en la sociocultural. Se examinan así, por un lado, los mecanismos gráficos y tipográficos empleados para la representación del género gramatical y, por otro, las ideas expresadas directa o indirectamente sobre la mujer como figura social. Veremos, de esta manera, que las estrategias tipográficas utilizadas para la representación del femenino gramatical se revelarán en extremo variadas y, en cierto sentido, asistemáticas. Y, por parte de la dimensión sociocultural de la mujer, nos enfrentaremos a valores y roles tradicionales, a una visión de la mujer marcada por su papel en la familia y el matrimonio, a su identificación con la maternidad y con una posición subordinada en la esfera pública. En definitiva, observaremos cómo la obra de Marín no solo cumple una función didáctica en la enseñanza del español, sino que también actúa como testimonio de las concepciones ideológicas de su tiempo.

palabras clave: lexicografía, género gramatical, representación de la mujer, Francisco Marín, enseñanza del español

### Abstract

**The treatment of the feminine gender in the lexicographical sections of F. Marín's *L'italiano istruito...***

This article examines how the feminine gender is addressed in the lexicographical sections of Francisco Marín's *L'italiano istruito nella cognizione della lingua spagnuola*, with a focus on both linguistic and sociocultural perspectives. On a linguistic level, the graphic and typographic mechanisms used to represent grammatical gender will be analyzed. On a sociocultural level, both explicit and implicit ideas regarding womanhood will be explored. The graphic strategies used to represent grammatical gender will be shown to be varied and, in some respects, asymmetrical. However, from a sociocultural perspective, the main associations made are with traditional values and roles, representing womanhood primarily in relation to family, marriage, motherhood and as having a subordinate role in the public sphere. The overall analysis will demonstrate how Marín's work not only serves a pedagogical function in the teaching of Spanish, but also offers valuable insight into the ideological conceptions of its time.

keywords: lexicography, grammatical gender, representation of women, Francisco Marín, Spanish language teaching

## 1. Introducción

En este artículo, al referirnos a “lo femenino”, abordamos dos aspectos del género que, si bien distintos entre sí, son, en cierto sentido, colindantes. El primero está representado por el género lingüístico, habitualmente asociado a cuestiones morfológicas y, en menor medida, a aspectos gramaticales en sentido amplio. El segundo está estrechamente relacionado con la figura de la mujer, con su representación ideológica y su papel social.

Nuestra atención se centrará en cómo *L'italiano istruito nella cognizione della lingua spagnuola* de Francisco Marín articula estas dos perspectivas del género. Desde este punto de vista, examinaremos, por un lado, algunos componentes puramente formales (“gráficos” o “tipográficos”) de la obra y, por otro, consideraremos los contenidos socioculturales expresados directa o indirectamente por sus textos.

Cabe también recordar que, tratándose de una obra de enseñanza del español para hablantes de italiano, la perspectiva contrastiva del género gramatical estará constantemente presente y, por tanto, aunque nos refiramos principalmente a la lengua castellana, en algunos casos recurriremos también a la italiana.

En este estudio, restringiremos nuestro corpus a los sectores lexicográficos<sup>1</sup> que se encuentran en el *Supplemento* de *L'italiano istruito*..., dejando para investigaciones futuras los amplios sectores gramaticográficos que constituyen su cuerpo principal.

En síntesis, nuestro interés radicará tanto en la representación lexicográfica del femenino gramatical como en la concepción de la función social de la mujer transmitida por la obra. A través del primer enfoque, evaluaremos el grado de sistematicidad formal alcanzado por la gramaticografía de la época y, con el segundo, apuntaremos a una reconstrucción del ambiente ideológico del que la obra fue testigo y portavoz al mismo tiempo.

Tras esta necesaria introducción (§1.) se incluirá una descripción de la obra objeto de estudio (§2.), un análisis del género gramatical femenino en sus apartados lexicográficos (§3.), una evaluación de la figura de la mujer en dichos sectores (§4.), una breve incursión en sectores no estrictamente lexicográficos del *Supplemento* (§5.), unas conclusiones (§6.) y, finalmente, las referencias bibliográficas.

---

\* Este trabajo se ha realizado en el marco del proyecto “Un nuovo ambiente digitale per il recupero del patrimonio lessicografico: il *Tesoro digitale della lessicografia bilingue spagnolo-italiano*”, financiado por La Unión Europea- Next Generation EU, Missione 4 Componente 1 CUP C53D23004010006 / PRIN 2022 MUR 20229W73WR.

1 Esta delimitación se debe, principalmente, a que nos movemos —como se ha indicado— en el marco del Proyecto PRIN antes mencionado.

## 2. La obra: *L'italiano istruito nella cognizione della lingua spagnola*

### 2.1 *Autor y ediciones*

No son muchas ni demasiado seguras las noticias biográficas que poseemos de Francisco Marín, autor de *L'italiano istruito nella cognizione della lingua spagnola*. Sabemos que fue un religioso español, probablemente de origen vasco o navarro<sup>2</sup>, quien hacia 1809 se trasladó a Roma, presumiblemente forzado por el cierre de monasterios y conventos del norte de España impuesto por los ejércitos franceses tras su entrada en territorio español (1807). En la capital del Estado Pontificio, se estableció inicialmente en la Real Iglesia de Santiago de los Españoles situada en Piazza Navona<sup>3</sup>. Allí ejerció el cargo de capellán hasta 1818, año en que la iglesia nacional de España trasladó su sede a Santa María de Monserrat y Santiago de los Españoles (via Giulia, n. 151). En esta nueva sede, Marín se desempeñó como capellán (1818-1833), sacristán mayor (1833-1835) y, por último, rector de la Iglesia (desde 1835 probablemente hasta su fallecimiento el 7 de enero de 1843).

Además de su misión sacerdotal, Marín fue profesor de español durante al menos quince años y publicó una gramática con apéndices lexicográficos destinada a la didáctica del español. Con sus propias palabras:

Tali riflessi, uniti alla necessita, che noi abbiamo avuta, di una grammatica moderna, per insegnare nella sua purità la lingua patria a quei che si sono compiaciuti a noi dirigersi nello spazio di circa tre lustri, ci hanno determinati ad imprendere quest'opera così utile e necessaria. (Marín 1833: VII)

El texto de Marín fue objeto de, al menos, nueve ediciones distintas: una romana (1833 de Salviucci) y ocho milanesas (1837, 1853 y 1860 de Silvestri; 1862 de Recchioni; 1867, 1869, 1874 y ¿1885? de Carrara). A lo largo de sus múltiples ediciones, la obra acusó cuantiosos cambios, incluso sustanciales. Nosotros basaremos nuestro estudio en la edición romana (1833), la única que podemos considerar en todo adherente a la voluntad del autor<sup>4</sup>.

2 Ignoramos la fecha de su nacimiento, pero la hipótesis más atendible debería fijarla entre 1783 y 1793 (Lombardini 2014: 45-50).

3 El edificio hoy está ocupado por la Iglesia de Nuestra Señora del Sagrado Corazón.

4 La lejanía geográfica, el enfrentamiento político entre Roma y Milán en la tercera década del siglo XIX (ambas resultaban capitales de dos estados distintos y contrapuestos: el Estado pontificio y el Reino lombardo-veneto dependiente del Imperio austríaco), la falta de un prólogo puesto al día

## 2.2 Estructura de la obra

Además de unos preliminares iniciales<sup>5</sup> y de varios apéndices<sup>6</sup>, la obra consta de un cuerpo principal, subdividido entre una *Grammatica della lingua spagnuola* (pp. 1-259) y un *Supplemento* (pp. 261-376). En este último sector, que denominamos “lexicográfico”, pueden identificarse nueve grandes apartados:

1. Osservazioni sull’analogia delle parole italiane e spagnuole (pp. 261-264);
2. Degli aggettivi che vanno accompagnati dal verbo *ser* (pp. 265-268);
3. Degli aggettivi e participj passati, che devono essere accompagnati dal verbo *estar* (pp. 269-274);
4. Degli aggettivi che possono usarsi e col verbo *ser* e coll’*estar* (pp. 274-279);
5. Raccolta di alcuni nomi che sono femminini in italiano e mascolini in ispagnuolo (p. 280);
6. Raccolta di alcuni nomi che sono mascolini in italiano e femminini in ispagnuolo (p. 280);
7. Raccolta di nomi sostantivi più necessarj a sapersi (pp. 281-327);
8. Raccolta di verbi più necessarj a sapersi (pp. 328-342);
9. Frasi familiari per cominciare a parlare spagnuolo (pp. 343-376).

Todos estos apartados adoptan la forma de una *nomenclatura*: listados léxicos o fraseológicos organizados onomasiológicamente, es decir, agrupados por áreas temáticas.

---

por parte del autor y la inexistencia de correspondencia conocida entre el autor y el nuevo editor desaconsejarían considerar que la voluntad del primero esté reflejada también en la edición milanesa de 1837. Por supuesto, su voluntad tampoco pudo reflejarse —dado el año de fallecimiento de Marín (1843)— en las ediciones posteriores.

5 Un frontispicio (p. I), un epígrafe (p. II), una dedicatoria (pp. III-IV) y un prólogo (pp. V-VIII).

6 Una antología de textos tomados de las *Cartas Marruecas* de Cadalso (pp. 377-394), varios modelos de cartas (pp. 395-398), una antología de fábulas de Iriarte (pp. 399-400), unas indicaciones bibliográficas (pp. 401-404), un índice general (pp. 405-411) y un *nihil ostat e imprimatur* (pp. 412).

### 3. El género gramatical en los suplementos lexicográficos<sup>7</sup>

Este tercer apartado tiene como objetivo catalogar los recursos tipográficos y/o metalexicográficos empleados en la *Grammatica* de Francisco Marín para presentar el género gramatical<sup>8</sup>.

#### 3.1 Remisión a las reglas gramaticales ya expuestas en otras partes del texto

Este primer recurso para dar indicaciones sobre el género consiste en remitir en nota a lo dicho en el cuerpo gramatical de la obra. Así, en la primera entrada (*abile* > *hábil*) del listado “*Degli aggettivi che vanno accompagnati dal verbo ser*”, una nota nos advierte lo siguiente<sup>9</sup>:

A suo luogo si è detto che negli aggettivi in *o* per farli femminini basta cambiare l'*o* in *a*, e che quelli che non finiscono in *o* sono comuni ai due generi. (Marín 1833: 265)

El pasaje al que remite la nota se sitúa en el subcapítulo “Del genere dei nomi” (pp. 42-44). Allí se expresa lo siguiente:

I nomi aggettivi che sono mascolini se finiscono in *o*, divengono femminini [*sic*] cambiando l'*o* in *a*:

*Hermoso, corto, largo,  
buono, malo, ancho.*

*Bello, corto, lungo,  
buono, cattivo, largo*

*hermosa, corta, larga,  
buena, mala, ancha.*

*bella, corta, lunga,  
buona, cattiva, larga.*

Se l'aggettivo non finisce in *o*, qualunque sia la sua terminazione, è comune ai due generi:

7 Tal como se exponen en otras contribuciones de este monográfico, es importante notar aquí las dificultades que causan a la hora de codificar en xml-TEI LEX 0 todas las variaciones en el tratamiento del género vistas en el §3.

8 No consideraremos en este estudio los puntos 1 (*Osservazioni sull'analogia delle parole italiane e spagnuole*), 3 (*Degli aggettivi e participj passati, che devono essere accompagnati dal verbo estar*), 4 (*Degli aggettivi che possono usarsi e col verbo ser e coll'estar*), 8 (*Raccolta di verbi più necessari a sapersi*) y 9 (*Frase familiari per cominciare a parlare spagnuolo*) por no encontrarse en ellos ningún recurso específico vinculado con el género gramatical.

9 De aquí en adelante, las citas reproducirán fielmente el texto original, tanto por lo que se refiere a sus mayúsculas y minúsculas como a su acentuación y puntuación.

Hombre *alegre, fiel, ruin,*  
*capaz, regular, feliz.*

Mujer *alegre, fiel, ruin,*  
*capaz, regular, feliz.*

*Uomo* allegro, fedele, vile,  
capace, regolare, felice.

*Donna* allegra, fedele, vile,  
capace, ec.

(Marín 1833: 43)

### 3.2 Los listados contrastivos

Un segundo modo de Marín para indicar el género gramatical es proponer listados contrastivos con lemas femeninos italianos y equivalentes masculinos españoles (subcapítulo “*Raccolta di alcuni nomi che sono femminini in italiano e mascholini in ispagnuolo*”) o lemas masculinos italianos y equivalentes femeninos castellanos (subcapítulo “*Raccolta di alcuni nomi che sono mascholini in italiano e femminini in ispagnuolo*”).

l'aria	<i>el aire</i>
la carta	<i>el papel</i>
la cioccolata	<i>el chocolate</i>

(Marín 1833: 280)

l'affronto	la afrenta
il carcofo	la alcachofa
il cerino	la cerilla

(Marín 1833: 280)

### 3.3 El artículo como marca de género

En el recurso que acabamos de ilustrar (§3.2), no solo se contraponen lemas a equivalentes en virtud de sus distintos géneros<sup>10</sup>, sino que además se refuerza dicha contraposición con la inclusión de artículos femeninos o masculinos en lemas y equivalentes.

10 Lo que hubiera sido suficiente, visto los epígrafes que introducen las entradas.

Recurrir al artículo para marcar el género de un término es, muy probablemente, la estrategia metalexicográfica más frecuente de *L'italiano istruito...* Y son numerosas las variantes de este recurso que propone el autor en su texto:

1. Un lema italiano contrapuesto a un equivalente español, en ambos casos con sus respectivos artículos:

la corporatura	<i>el talle</i>
	(Marín 1833: 288)

i vormiglion	<i>las viruelas</i>
	(Marín 1833: 290)

2. Dos lemas italianos (con artículos) contrastados a un solo equivalente español (con artículo):

il giorno	}	<i>el día</i>
la giornata		
		(Marín 1833: 282)

3. Un lema italiano opuesto a varios equivalentes españoles, en ambos casos con artículo:

gli occhiali	{	<i>los anteojos</i>
		<i>las gafas</i>
		(Marín 1833: 291)

4. Variantes libres de las tres modalidades que se acaban de exponer.

Indicación del artículo en una sola de las lenguas, aunque en sentido estricto hubieran sido necesarios en ambas:

il rivolta	<i>solapa</i>
	(Marín 1833: 290)

Omisión del artículo tanto en el lema como en la equivalencia, si bien también en este caso hubieran sido imprescindibles:

cervice	<i>pescuezo</i>
	(Marín 1833: 286)

Oposición de lemas italianos con artículo a dos equivalentes españoles, de los cuales solo uno lo lleva:

l'armamento	<i>las járcias, ó aparejos</i>
	(Marín 1833: 310)

Indicación de artículos cuando no hubieran sido estrictamente necesarios, pues tanto lemas como equivalencias pertenecen a mismo género gramatical:

la seminagione	<i>la siembra</i>
la mietitura	<i>la siega</i>
la raccolta	<i>la cosecha</i>
la vendemia	<i>la vendimia</i>
	(Marín 1833: 282)

### 3.4 Las parejas de entradas sucesivas

Cuando la contraposición de géneros gramaticales coincide con la oposición de sexos, es decir, cuando los enfrentados son seres vivientes, Marín suele recurrir a la oposición de dos entradas sucesivas<sup>11</sup>. Véanse estas dos series de entradas subsiguientes<sup>12</sup>:

Uomo	<i>Hombre</i>
donna	<i>muger</i>
fanciullo	<i>niño</i>
fanciulla	<i>niña</i>
donzella	<i>doncella</i>

11 Y no como en los casos vistos hasta ahora, en los que siempre se contraponían el lema y la equivalencia de una misma entrada.

12 Nos referimos a las parejas de entradas *uomo>hombre* / *donna>muger*, *fanciullo>niño* / *fanciulla>niña*, *donzella>doncella* / *scapolo>soltero*, etc.



scapolo	<i>soltero</i>
ragazzo	<i>muchacho</i>
ragazza	<i>muchacha</i>

(Marín 1833: 389)

principe	<i>príncipe</i>
principessa	<i>princesa</i>
duca	<i>duque</i>
duchessa	<i>duquesa</i>
conte	<i>conde</i>
contessa	<i>condesa</i>
viceconte	<i>vizconde</i>
vicecontessa	<i>vizcondesa</i>
marchese	<i>marques</i>
marchesa	<i>marquesa</i>
barone	<i>baron</i>
baronessa	<i>baronesa</i>
cavaliere	<i>caballero</i>
dama	<i>dama</i>

(Marín 1833: 283)

Como se puede observar, suelen contraponerse términos con la misma raíz (*fanciullo*>*niño* contrapuesto a *fanciulla*>*niña*), aunque no faltan ejemplos de contraposiciones entre términos de base léxica diferentes (*uomo*>*hombre* contrapuesto a *donna*>*mujer*). Por otra parte, también suele iniciar la oposición la forma masculina (*ragazzo*>*niño* primero y *ragazza*>*niña* inmediatamente después, etc.) y solo en algún caso se constata la secuencia contraria (*donzella*>*doncella* primero y *scapolo*>*soltero* después).

### 3.5 Las entradas múltiples

Cuando la oposición entre distintos sexos se constata en el lema y/o en la equivalencia podría hablarse de *entradas múltiples*.

marito e moglie                      *marido y muger*  
 (Marín 1833: 285)

Y puede asimismo ocurrir que una equivalencia múltiple se encuentre contrapuesta a un lema/definición.

Chi va a sposare,  
 o ha sposato di                      {    *Novio, novia*  
 recente

(Marín 1833: 285)

### 3.6 *Las marcas lexicográficas*

El género también puede señalarse mediante marcas lexicográficas específicas, como *f.* para femenino y *m.* para masculino. Este recurso puede combinarse —y de hecho lo hace frecuentemente— con algún otro de los recursos hasta aquí mencionados.

il padrone                      *el amo*  
 la padrona                      *la ama f.*  
 (Marín 1833: 285)

rinoceronte                      {    *rinoceronte*  
    *la abada f.*  
 (Marín 1833: 322)

*tigre, m.*  
 tigre, f.                      *tigre, f.*  
 (Marín 1833: 322)

cinghiale, m.                      *jabalí*  
 cinghiale, f.                      *jabalina*  
 (Marín 1833: 322)

En ciertos casos, la marca de género adopta la forma de un término especificativo en función apositiva (*macho*).

mula	<i>mula</i>
mulo	<i>mulo macho</i>
(Marín 1833: 323)	

### 3.7 Las vocales (o sílabas) del género

La última forma de marcación encontrada incluye la indicación explícita de la vocal genérica (*prohibado, a*) o de su sílaba distintiva (*nieto, ta*, etc.).

figlio	<i>hijo</i>
figlia	<i>hija</i>
... per adozione	<i>prohibado, a</i>
(Marín 1833: 285)	

nipote in linea retta	<i>nieto, ta</i>
bisnepote	<i>biznieto, ta</i>
pronipote	<i>tataranieto, ta</i>
(Marín 1833: 286)	

## 4. La figura de la mujer en los suplementos lexicográficos

En este cuarto apartado abordaremos ideas relacionadas con la mujer y su papel en la sociedad tal como aparecen delineados en la *Grammatica* de Francisco Marín. Para este análisis, recurriremos principalmente al apartado *Frase familiari per cominciare a parlare spagnuolo* (pp. 343-376), pues el resto de los apartados lexicográficos no aporta ningún elemento en este sentido.

En las frases incluidas en el epígrafe *Per salutarsi e congedarsi* (42 entradas) destacan cuatro por su relación con el universo femenino.

Sea V. bienvenida. (p. 345)

Como está su parienta de V.? (p. 345)

Mucho tiempo há que no la he visto. (p. 345)

Espresiones á Doña Clara. (p. 346)

Son, en su mayoría, expresiones corteses y formales en las que se apela a la mujer con especial deferencia gracias al empleo de tratamientos como *usted* o *doña*. Tal tratamiento denota la posición socialmente prestigiosa alcanzada por las personas de sexo femenino en la (alta) burguesía<sup>13</sup> europea del siglo XIX. Además, cabe mencionar que, en el proceso de construcción de la identidad femenina, cumple un papel relevante su estado civil (*su parienta* ‘su esposa’). Por otra parte, el encomio de su función social parece no estar estrictamente vinculado a la presencia física de su persona en el trato social (*mucho tiempo há que no la he visto*).

En el quinto apartado (*Del levarsí*, 17 entradas) se recoge el siguiente diálogo:

– Á qué hora suele levantarse la señora?

– No sé decírselo a V., no hay regla. (p. 352)

Nos encontramos en el ámbito de la rutina doméstica. Una señora de la (alta) burguesía, dada su posición social, no tenía a su cargo los trabajos físicos que la economía de un hogar requería y, por tanto, la administración de su tiempo se revelaba más flexible e independiente que la de los hombres, quienes, en cambio, solían desarrollar sus obligaciones sociales y laborales fuera del hogar. Las rutinas de las mujeres se enmarcaban, así, en un ámbito privado y, en cierto sentido, oculto a la visión de los extraños (*no sé decírselo a V.*).

Entre las entradas del apartado *Per sapere, conoscere, udire, ascoltare* (28 entradas) se encuentra este pequeño diálogo:

– Quienes eran aquellas mugeres?

– No las he visto jamás. (p. 353)

De aquí podría inferirse la importancia de que la mujer sea reconocible en su entorno social: no ser identificada podría implicar la pertenencia a un grupo social ajeno o, peor aún, su irrelevancia social.

Otro diálogo con una temática similar aparece en el apartado *Della passeggiata* (42 entradas).

13 La enseñanza de lenguas extranjeras en la Italia del siglo XIX estaba dirigida, en buena parte, a un público femenino de la alta y mediana burguesía (cfr. Lombardini en prensa).

- Sabe V. quienes son estas señoras que vienen?
- Parece que buscan asiento.
- Dejémosles este banco. (p. 357)

Si bien este fragmento matiza el anterior con la idea de caballerosidad, en otras palabras, la presunción de que las mujeres merecen deferencia y/o protección por parte de los hombres.

Dos diálogos más se hallan en el apartado *Della commedia* (40 entradas), dedicado al mundo del espectáculo.

- Yo soy uno de los apasionados de la primera dama.
- Tiene V. razón, la naturaleza la ha dotado de grandes talentos. (p. 368)
- Perdone V., quien es esa cómica, que se ha asomado al bastidor?
- Es la segunda dama.
- Parece muy moza.
- No sabe V. que en las tablas gozan estas señoras de una primavera eterna?
- Si fuera el solo deseo de agradarnos, el que las moviese á adornarse tanto, yo les estimaría el trabajo.
- No piense V. tan mal.
- Sea el que fuere el motivo, ellas me gustan mucho. (p. 369)

En su brevedad, el primer diálogo refleja una visión tradicional de la cantante lírica: una figura admirada por su talento natural más que por su esfuerzo o formación; una imagen ligada más a la musa, a la diva o al objeto de deseo que a una verdadera artista.

El segundo diálogo, más extenso e interesante, nos sitúa en el teatro de prosa. En este contexto las actrices deben mantener una apariencia juvenil para seguir siendo valoradas en escena (*No sabe V. que en las tablas gozan estas señoras de una primavera eterna?*): la presión social las obliga a conservar su juventud y belleza, pues su valoración social se basa en las reacciones estéticas o emocionales que suscitan en los espectadores (*Sea el que fuere el motivo, ellas me gustan mucho.*). Por último, cabe aludir a las sospechas insinuadas –con malicia y moralismo– sobre las verdaderas razones por las que estas mujeres cuidan tanto su apariencia.

El apartado *Di varie cose che possono occorrere in una conversazione* –el más extenso de las *Frase familiari* (131 entradas)– contiene el mayor número de referencias a la mujer. Cuatro bloques de diálogos abordan dos temas centrales de la vida social de las mujeres: la maternidad y el matrimonio.

En cuanto a la maternidad se refiere, un primer bloque –constituido por una sola frase– refleja noticias introducidas en la conversación de un grupo de cono-

cidos tras la lectura de un periódico.

El artículo de Viena habla de una mujer que ha parido tres hijos y de una epidemia que hace mucho estrago. (p. 371)

Quien transmite al grupo las noticias supone de igual interés público un alumbramiento múltiple y una epidemia que produce “mucho estrago”. Se define a la mujer por su papel (y deber) social de madre. La maternidad, primera y más importante de las funciones sociales femeninas, al ser múltiple, adquiere inmensa relevancia.

El segundo bloque presenta un diálogo –mantenido, quizás, por el mismo grupo de personas– sobre el embarazo de una vecina:

- Pues sepan grandes y chicos, que la muger del boticario está embarazada.
- Hombre! Es posible!
- Tiene mas de cincuenta años.
- Es el parto milagroso de Sara.
- El marido está muy contento.
- Sentía mucho el ver que se acababa su apellido.
- Dice, que entre los boticarios no hay casa mas antigua. (p. 374)

También en este caso se define a la mujer por su capacidad reproductiva. La sorpresa ante un embarazo tardío (– *Hombre! Es posible! /– Tiene más de cincuenta años.*) denota la idea de que la fertilidad en una mujer mayor supone un caso excepcional, casi sobrenatural (*Es el parto milagroso de Sara.*). La pérdida de tal capacidad comporta uno de los acaecimientos más emblemáticos de la vida de una mujer.

Los dos últimos bloques están dedicados al matrimonio. El primero:

- Sepan que D.<sup>a</sup> N. se casa.
- Con quien?
- Con el señor de N.
- Hola! Es buen casamiento.
- Su hermana también se ha casado.
- Y quien es el novio?
- Un mercader muy rico.
- Que dote ha llevado?
- Cuatro cientos mil reales.
- Es un dineral.
- Lo merece el marido, que es hombre de bien.
- Pues la muchacha tambien está muy bien criada. Ella será muy muger de su casa.
- La madre no le consentia nada.

- Su hija se lo agradecerá.
- La hacienda mas segura que pueden dejarnos los padres, es una buena crianza. (p. 373)

En este diálogo se subraya que el matrimonio es el destino primordial de la mujer y que su vida y su futuro están determinados íntimamente por este acontecimiento. Un “buen casamiento” es el objetivo deseado (*Hola! Es buen casamiento.*), donde la conveniencia, el prestigio y la estabilidad económica son elementos clave de este evento. A la mujer, por otra parte, se la valora según la riqueza que aporte al enlace, aunque luego sea el marido quien disponga directamente de dicha dote. Por último, su idoneidad para desempeñar el papel de esposa y ama de casa (*Ella será muy muger de su casa.*) se vincula puntualmente a la buena educación (“crianza”) que haya recibido con anterioridad y a un intento de seducción por parte del hombre. El segundo bloque:

- Que? señora, se va V. ya?
- Me han de perdonar Vs., tengo que hacer en mi casa.
- Todavía es temprano.
- No, no, es cosa precisa. Me están aguardando ya.
- La acompañaré á V.
- No quiero que tan galan caballero se canse por una vieja como yo.
- V. se burla, señora.
- No, no, mire V. que ya tengo veinte y nueve años.
- Vd. no aparenta aun veinte.
- Quítese V. de ahi, que aparento mas de los que tengo. No falta quien me eche treinta y cinco años.
- No lo crea V. lo quiere V. decir así. Es V. casada señora?
- No señor, soy soltera, y V.?
- Soltero tambien, señora.
- Tambien V.? Me alegro.
- Si V. quisiera, en breve no pudiéramos decir otro tanto.
- Y como es eso?
- Discurro que V. me habrá entendido.
- Ya veo donde va V. á parar.
- Una señora de las circunstancias de V...
- Caballero, dejemos eso.
- Es menester pensarlo.
- Amigo, es para siempre.
- Ya lo sé, señora.
- Pues sigamos el refrán español, que dice *antes que te cases, mira lo que te haces.* (pp. 375-376)

Para la mujer mayor de edad, aun siendo soltera, las responsabilidades domésticas son prioritarias (*tengo que hacer en mi casa / Me están aguardando ya.*). Por otro lado, la preocupación por la edad que avanza es evidente (*que ya tengo veinte y nueve años*), pues la juventud femenina se asocia indefectiblemente a la belleza y a la posibilidad de contraer matrimonio. A pesar de todo esto, la “buena crianza” de la que se hablaba a raíz del bloque anterior recomienda siempre cautela (*antes que te cases, mira lo que te haces*), porque el matrimonio constituye una decisión crucial e irrevocable en la vida de la mujer.

## 5. La figura de la mujer en los suplementos no lexicográficos

A modo de anexo –por referirse a aquellos apartados del *Supplemento* que no tienen un corte específicamente lexicográfico–, en este punto quinto seguiremos tratando cuestiones vinculadas con la figura social de la mujer. Para ello nos focalizaremos en tres apartados: *Lettere estratte dalle Cartas Marruecas del Colonnello Cadahalso* (pp. 377-394), *Modelli di lettere secondo il costume presente degli spagnuoli* (pp. 395-398) y *Elenco de' classici spagnuoli* (pp. 399-402).

### 5.1 *Los textos de Cadalso*

Como hemos indicado en Lombardini (2014)

Son seis las cartas de Cadalso reproducidas en el texto de Marín y todas hacen referencia (salvo la primera, introductoria) a la descripción, historia de España o a la crítica de sus costumbres. Las cartas de Marín –numeradas de I a VI– corresponden, respectivamente, según la edición de Juan Tamayo (Cadalso 1979 [1935], [1789]) a la I («Da noticia Gazel a Ben-Beley de su detención en España, de su idea de viajar por ella, y de su amistad con Nuño. Le promete informarlo de cuanto observe y le pide que lo ayude con sus consejos» 460), la II («Se toma tiempo Gazel para informar a su maestro respecto a la diversidad que nota entre los europeos, y aun entre los mismos españoles.»), la XXVI («Diversidad de las Provincias de España.»), la III («Epítome de la Historia de España hasta el principio del siglo presente»), la V («Conquistas de las Américas») y la IX («Continuación de la carta V. Apología de Cortés. Retorsión de las declamaciones de los extranjereros.»). (Lombardini 2014: 281-282)



El mismo Marín explica el objetivo de incluir algunos textos de las *Cartas marruecas* de Cadalso:

Le seguenti sei lettere, che qui poniamo coll'intento di far esercitare gli studiosi nella lettura, contengono la relazione rimessa dal Moro Marrochino Gazal al suo Maestro in Africa Ben-Beley sopra le osservazioni da lui fatte ne' suoi viaggi per la Spagna nell'anno 1768. (Marín 1833: 376)

Se impone aquí una aclaración interpretativa. La red de voces responsables en este tipo de recurso presenta una articulación sumamente intrincada. En nuestro caso, por ejemplo, las apreciaciones sobre las mujeres expuestas en estas cartas deberían adjudicarse, en primer lugar, a la primera persona narradora o descriptora (el joven marroquí Gazel o el autor responsable de los textos transmitidos directa o indirectamente por Gazel en sus misivas), en segundo lugar, a Cadalso y, solo en tercer lugar, a Francisco Marín quien, al seleccionar y publicar sin comentarios críticos algunas cartas específicas, avala sus contenidos.

No son muchas las referencias dedicadas a las mujeres entre los textos de Cadalso publicadas por Marín y, sobre todo, se contraponen entre sí, se enfrentan, inducen a interpretaciones contradictorias. Es así que, aunque se describa con un tono aparentemente halagüeño a las andaluzas, se aconseja (entre líneas) mantenerse lo más posible alejado de ellas.

La viveza, astucia y atractivo de las Andaluzas las hace incomparables. Le aseguro á V. que una de ellas sería bastante para llenar de confusión el Imperio de Marruecos, de modo, que todos nos matásemos unos á otros. (Marín 1833: 382)

Pero se alaban sin reserva las capacidades de gobierno de la Reina Isabel de Castilla:

Incorporáronse tantas Provincias, y tan diferentes en dos coronas, la de Castilla y la de Aragon; y ambas en el matrimonio de D. Fernando y Doña Ysabel, Príncipes que serán inmortales entre cuantos sepan lo que es gobierno. (Marín 1833: 386)

En su afán de defender la actuación de Hernán Cortés durante la conquista de América, Gazel subraya la hipocresía de los europeos no españoles que, por un lado, critican con acritud las acciones españolas y, por otro, negocian y se enriquecen con la trata de esclavos africanos:<sup>14</sup>

14 Claramente, la perspectiva propuesta forma parte de la compleja cuestión de la “leyenda negra” de España cuya extensa tradición de estudios da cuenta de las posiciones encontradas que, especialmente durante los siglos XVIII y XIX, caracterizaron el debate y la disputa por el “nuevo mundo”. Por otra parte, consideramos que dicha discusión en nuestro caso es marginal, pues pro-

Acabo de leer algo de lo escrito por los Europeos que no son Españoles, acerca de la conquista de la América. Si del lado de los Españoles no se oye sino religion, heroismo, vasallage, y otras voces dignas de respeto, del lado de los estrangeros no suenan sino codicia, tiranía, perfidia y otras no ménos espantosas. No pude ménos de comunicarselo á mi amigo Nuño, quien me dijo, que era asunto dignísimo de un fino discernimiento, juiciosa crítica y madura reflexion; pero que entre tanto, y reservandome el derecho de formar el concepto que mas justo me pareciese en adelante, reflexionase por ahora, que los Pueblos que tanto vocean la crueldad de los Españoles en América son precisamente los mismos que van á las costas de Africa, compran animales racionales de ambos sexos á sus padres, hermanos, amigos y guerreros victoriosos, sin mas derecho que ser los compradores blancos y los comprados negros; los embarcan como brutos; los llevan millares de leguas desnudos, hambrientos y sedientos; los desembarcan en América; los venden en público mercado como jumentos, á mas precio los mozos sanos y robustos, y á mucho mas las infelices mugeres que se hallan con otro fruto de miseria dentro de sí mismas; toman el dinero; se lo llevan á sus humanísimos paisajes; y con el producto de esta venta imprimen libros llenos de elegantes invectivas, retóricos insultos y elocuentes injurias contra Hernan Cortés por lo que hizo; ¿y qué hizo? (Marín 1833: 388-389)

En ese contexto, Gazel demuestra el buen tino de recordar las infinitas desgracias sufridas indefectiblemente por miles de esclavas africanas (*los venden en público mercado como jumentos, á mas precio los mozos sanos y robustos, y á mucho mas las infelices mugeres que se hallan con otro fruto de miseria dentro de sí mismas*).

Siempre en el ámbito de la defensa de Cortés, Gazel recuerda la utilidad de su amante Malinche quien, a pesar de ser mujer, jugó un papel positivo en el ejército de los españoles:

De la misma flaqueza humana sabe Cortés sacar fruto para su intento. Una Yndia noble, á quien se habia aficionado apasionadamente, le sirve de segundo intérprete, y es de suma utilidad en la expedicion. Primera muger que no ha perjudicado en un ejército, y notable egemplo de lo útil que puede ser el bello sexo, siempre que dirija su sutileza natural á fines loables y grandes. (Marín 1833: 390)

La mujer, con su fuerza de voluntad y gran empeño, es capaz de revocar la suerte aciaga a la que su condición femenina la destinaba (*lo útil que puede ser el bello sexo, siempre que dirija su sutileza natural á fines loables y grandes*).

---

porciona solo el marco expresivo para la figura de la mujer y, por tanto, no nos ocuparemos de ella sino solo en esta nota.

## 5.2 Los modelos epistolares

Todos los emisores previstos en los modelos de cartas propuestos por Marín son hombres. Entre los destinatarios se prevé un único caso en que el receptor puede ser una mujer: cuando se trate de la reina de España. En esas ocasiones, el encabezamiento de la misiva podría ser *Á la Reina Nuestra Señora*, en el caso que el emisor sea un súbdito español, o *Á la Sagrada R. Magestad / De la Señora D.a Cristina de Borbon / Reina de España e Indias*, si este fuera un extranjero (Marín 1833: 395). El resto de los destinatarios –ministros, secretarios de estado y otros grandes funcionarios, encargados comerciales, familiares y amistades– se presuponen siempre del género masculino. Solo en el caso de las cartas a amigos se ofrece la posibilidad de incluir en la despedida saludos para una mujer (*Dé V. muchas expresiones á mi Sra D.a Mariquita*, Marín 1833: 398).

## 5.3 Las obras recomendadas

Marín finaliza su texto con un listado de obras que el usuario de la *Grammatica* debería considerar de especial interés para el aprendizaje de la lengua y la cultura españolas. Se trata de obras recomendadas agrupadas según su género (poesía, prosa o comedia) y pertenecientes a unos sesenta autores, pero no de una bibliografía con indicación de publicaciones específicas<sup>15</sup>. Entre estas obras, se mencionan dos cuyas autoras son mujeres: la *Cartas* de Santa Teresa de Jesús (1515-1582) y la *Novelas* de María de Zayas y Sotomayor (1590-1661)<sup>16</sup>. Aunque sean solo dos obras –frente a unas setenta obras de autores masculinos– su presencia reviste un interés particular en el presente trabajo<sup>17</sup> por la condición de precursoras del feminismo que la crítica filológica ha adjudicado a sus autoras.<sup>18</sup> Santa Teresa,

15 Solo en un par de casos las referencias mencionan publicaciones concretas: Fernández (1785-1797) y López de Sedano (1768-1778).

16 Es bastante probable que con estas indicaciones Marín haya estado pensando en Teresa de Jesús (1771 [1658]) o (1778 [1658]) y en De Zayas (1805 [1637]) o (1750 [1647]).

17 No es baladí recordar que, en la edición de la *Gramática* de 1853, se sustituye el listado al que estamos haciendo referencia por un nuevo apartado (*Elenco di classici, e libri spagnuoli, e di traduzioni dallo spagnuolo*, pp. 345-351). De este texto, aun aumentado el número de obras mencionadas, desaparece toda mención a obras creadas por mujeres.

18 Cfr., por ejemplo, Pérez Cuenca *et al.* (2015) para Santa Teresa y Gorgas Berges (2018) para María de Zayas.

figura de gran inteligencia y perspicacia, habilísima estratega y diplomática nata, plasma en sus *Cartas* una profunda conciencia de las limitaciones impuestas a las mujeres en su época y fomenta determinación necesaria para superarlas. María de Zayas, por su parte, expresa en sus novelas fuertes críticas sociales mediante la representación de la vida de las mujeres; en palabras más modernas, denuncia la desigualdad de género y el sufrimiento de las mujeres en una sociedad patriarcal.

## 6. Conclusiones

Desde una perspectiva lingüística, los mecanismos de representación del género gramatical en *L'italiano istruito nella cognizione della lingua spagnuola* de Francisco Marín revelan una escasa sistematicidad y cierta inmadurez lexicográfica. Dichos mecanismos resultan sobreabundantes<sup>19</sup>, aplicados con rigor insuficiente y, en parte, superpuestos entre sí. Esto sugiere la ausencia de un plan previamente meditado y adoptado de manera consciente y sistemática. Si bien la dificultad añadida por la contrastividad del par español-italiano aminore, en cierta medida, la inconsistencia lexicográfica de los resultados obtenidos, la representación social de la mujer en los distintos sectores lexicográficos y no lexicográficos del *Supplemento* de la obra<sup>20</sup> revela una visión alineada con los valores tradicionalmente asociados a la mujer de la (alta) burguesía en la sociedad occidental; en otras palabras: (i) su papel central dentro de la familia; (ii) el matrimonio —conveniente, prestigioso y económicamente estable— como acontecimiento crucial en sus vidas, con la dote como criterio para establecer su valor; (iii) la necesidad de una buena educación acorde con su estatus social; (iv) la maternidad como su función primordial en la unión matrimonial; (v) la preocupación por el paso del tiempo y sus efectos en la concertación de un matrimonio ventajoso; (vi) el espacio doméstico como ámbito casi exclusivo de su actividad; (vii) su presencia en el espacio público generalmente mediada por la figura del marido o del jefe de familia; (viii) un trato social caracterizado por la deferencia de los otros, y (ix) el reconocimiento necesario del resto de los integrantes de su grupo social.

19 Se indica el género con listados contrastivos, artículos, parejas de entradas contiguas, entradas múltiples, marcas abreviadas, marcas con términos especificativos y vocales o sílabas finales.

20 Lexicográfico: *Frase familiari per cominciare a parlare spagnuolo*. No lexicográfico: *Lettere estratte dalle Cartas Marruecas del Colonnello Cadahalso, Modelli di lettere secondo il costume presente degli spagnuoli y Elenco de' classici spagnuoli*.

Más allá del ámbito de la mujer alto-burguesa, el texto menciona a las monarcas, cuyas figuras en nada difieren de la de sus pares masculinos. En el extremo opuesto, se alude a las mujeres del pueblo, quienes, pese a su condición femenina, pueden alcanzar logros notables con esfuerzo y determinación; y a las esclavas, retratadas con sincera conmiseración ante las múltiples desgracias de las que son objeto.

## Bibliografía citada

- Cadalso, José (1935 y 1979 [1789]), *Cartas marruecas*, ed., prólogo y notas de Juan Tamayo y Rubio, Madrid, Espasa-Calpe.
- Fernández, Ramón (1785-1797), *Colección de poetas españoles*, 20 vols., Madrid, Imprenta real.
- Gorgas Berges, Ana I. (2018) Mujeres transgresoras en el siglo de oro español: la voz de María de Zayas y Sotomayor. En Martín Clavijo, Milagro *et al.*, coords. *Mujeres dentro y fuera de la Academia*, Salamanca, Ediciones Universidad de Salamanca, 139-151.
- Lombardini, Hugo E. (2014), *Gramáticas para la enseñanza del español en la Italia del siglo XIX: el caso de Francesco Marín*, Salamanca, Ediciones Universidad de Salamanca.
- Lombardini, Hugo E. (en prensa), An Ideological Mirror: Lexicographic Insights into a 19th-Century Spanish Grammar book for Italian Speakers. En Vişan, Ruxandra (Ed.), *Lexicography, Discourse, and Power: Dictionary-Making in History and the Construction of Hegemonic Discourses*, “Lexicography, Lexicology, and Society” series, Berlin/Leiden: De Gruyter Brill.
- López de Sedano, José (1768-1778), *Parnaso español*, Madrid, Ibarra.
- Marín, Francesco (1833), *L'italiano istruito nella cognizione della lingua spagnuola*, Roma, Salviucci.
- Marín, Francesco (1837), *Gramatica della lingua spagnuola, o sia l'italiano istruito nella cognizione di questa lingua*, Milano, Giovanni Silvestri.
- Marín, Francesco (1853), *Gramatica della lingua spagnuola, o sia l'italiano istruito nella cognizione di questa lingua*, Milano, Giovanni Silvestri.
- Marín, Francesco (1860), *Grammatica della lingua spagnuola, ossia l'italiano istruito nella cognizione di questa lingua*, Milano, Giovanni Silvestri.
- Marín, Francesco (1862), *Grammatica della lingua spagnuola, ossia l'italiano istruito nella cognizione di questa lingua*, Milano, Antica casa libraria Giov. Silvestri ora di L. Recchioni.
- Marín, Francesco (1869), *Grammatica della lingua spagnuola, ossia l'italiano istruito nella cognizione di questa lingua*, Milano, Paolo Carrara.
- Marín, Francesco (1874), *Grammatica della lingua spagnuola, ossia l'italiano istruito nella*

- cognizione di questa lingua*, Milano, Paolo Carrara.
- Marín, Francesco (1885?), *Grammatica della lingua spagnuola, ossia l'italiano istruito nella cognizione di questa lingua*, Milano, Paolo Carrara.
- Pérez Cuenca, Isabel *et al.*, coords (2015), *Congreso Interuniversitario "Santa Teresa de Jesús, Maestra de Vida"*, Ávila: Universidad Católica de Ávila.
- Teresa de Jesús, Santa (1658) *Cartas de la gloriosa Madre Santa Teresa de Jesús*, con notas de Juan de Palafox, Zaragoza: Diego Dómer.
- Teresa de Jesús, Santa (1771 [1658]), *Cartas de Santa Teresa de Jesús fundadora del Carmen*, con notas del R. P. Fr. Antonio de San Joseph, Madrid, Joseph Doblado.
- Teresa de Jesús, Santa (1778 [1658]), *Cartas de Santa Teresa de Jesús*, con notas de Juan de Palafox, Madrid, Joseph Doblado.
- Zayas, María de (1637), *Novelas amorosas y ejemplares*. Madrid, Sebastián de Soubrier.
- Zayas, María de (1647), *Desengaño amoroso*. Madrid, Francisco Martínez.
- Zayas, María de (1750 [1677]), *Desengaño amoroso*. Madrid, Francisco del Hierro.
- Zayas, María de (1803 [1637]), *Novelas amorosas y ejemplares*. Madrid, Herederos de Sancha.

**Hugo E. Lombardini** es doctor en Filología Española (Universidad de Salamanca) y catedrático de Lengua Española en la Universidad de Bolonia. Se ha desempeñado como docente de Literatura Argentina en la Universidad Católica Argentina y como *lettore di madrelingua* en la Universidad Ca' Foscari. Es miembro del Consejo de Administración de la *Fundación Alma Mater Studiorum - Università di Bologna* con sede en Buenos Aires y director del *CIRSIL (Centro Interuniversitario di Ricerca sulla Storia degli Insegnamenti Linguistici)*. Ha publicado nueve volúmenes colectivos como editor y otras tantas monografías como autor; además de ser responsable de unos cincuenta estudios aparecidos en numerosas revistas científicas y obras colectivas.

**hugo.lombardini@unibo.it**

# NATALIA PEÑÍN FERNÁNDEZ

## ESTUDIOS PRELIMINARES PARA LA RETRODIGITALIZACIÓN DEL *VOCABULARIO ESPANYOL, ITALIANO* *Y TUDESCO* DE IGNACIO DE BORIA

Università di Bologna

### Resumen

El estudio del *Vocabulario Espanyol, Italiano y Tudesco* (1719) de Ignacio de Boria se inscribe en el proyecto PRIN *Tesoro digital de la lexicografía bilingüe español-italiano* (TELEI), dirigido por Carmen Castillo Peña. El objetivo principal del proyecto es la preservación del patrimonio lexicográfico bilingüe mediante ediciones críticas y facsimilares digitales. En este artículo se presenta este repertorio, parte de la tradición lexicográfica alemana y se describen los aspectos relevantes de los análisis preliminares necesarios para su edición crítica y digitalización.

palabras clave: lexicografía bilingüe, digitalización, bilingüismo, edición crítica, patrimonio digital

### Abstract

*Preliminary study for the retrodigitalisation of Ignacio de Boria's Vocabulario Espanyol, Italiano y Tudesco*

*The study of the Vocabulario Espanyol, Italiano y Tudesco (1719) by Ignacio de Boria is part of the PRIN project Tesoro digital de la lexicografía bilingüe español-italiano (TELEI), directed by Carmen Castillo Peña. The main objective of the project is the preservation of the bilingual lexicographical heritage through critical editions and digital facsimiles. This article presents this repertoire, based on the German lexicographical tradition, and describes the relevant aspects of the preliminary analyses necessary for its critical edition and digitization.*

keywords: bilingual lexicography, digitization, bilingualism, critical editing, digital heritage

## 1. Introducción

La digitalización del manuscrito titulado *Vocabulario Espanyol, Italiano y Tudesco*, escrito en Viena, por Ignacio de Boria en 1719, se enmarca dentro del proyecto *Un nuevo espacio digital para el patrimonio lexicográfico: el “Tesoro digital de la lexicografía bilingüe español-italiano (TELEI)”*<sup>1</sup>, dirigido por C. Castillo Peña (Universidad de Padua). Este proyecto surge del diálogo entre las ciencias, la tecnología y las humanidades, con el propósito de digitalizar el patrimonio lexicográfico bilingüe italiano-español.

Uno de los objetivos que se propone el equipo de investigación a cargo del proyecto es la creación de una biblioteca impresa con ediciones críticas de textos antiguos pertenecientes a la tradición gramatical del español en Italia. Esto resulta esencial para reconstruir la historiografía gramatical de esta combinación de lenguas entre los siglos XVI y XX, considerando obras de distinta naturaleza con finalidad didáctica.

Además, en la línea de iniciativas previas como el *Tesoro Lexicográfico* de Gili Gaya (1957), el *Nuevo Tesoro Lexicográfico de la Lengua Española* (NTLLE) de la Real Academia Española, el *Nuevo Tesoro Lexicográfico del Español* (S. XIV-1726) de Nieto Jiménez y Alvar Ezquerro (2007), el *Tesoro Lexicográfico del Español en América* (TLEAM) dirigido por Corbella Díaz y la *Opera del Vocabolario Italiano* (OVI) de la Accademia della Crusca (Nalesso, 2024:138) se está desarrollando *TELEI*, el *Tesoro digital de la lexicografía bilingüe español-italiano*. Concebido como un “diccionario de diccionarios”, compuesto por ediciones facsimilares digitales de textos clave en la lexicografía bilingüe ítaloespañola, gracias a la aplicación de las Humanidades Digitales, se logrará reunir un vasto acervo de datos lexicográficos que ninguna biblioteca física podría custodiar en su totalidad. El corpus en proceso de digitalización abarca diversos géneros lexicográficos desde el siglo XVI hasta el siglo XX que conforman la historia de la lexicografía ítaloespa-

---

\* El autor de la nomenclatura objeto de estudio aparece en los diferentes archivos consultados con diferentes nombres: Ignacio de Boria y de Gualba, Inacio de Boria e incluso Ignacio de Borja, su nombre castellano. Se decide utilizar a partir de ahora el nombre Ignacio de Boria, versión que el autor usa para presentarse en su obra.

<sup>1</sup> Este trabajo se inscribe en el marco del proyecto PRIN financiado por el MUR, Ministero dell’Università e della Ricerca italiano, “Un nuovo ambiente digitale per il recupero del patrimonio lessicografico: il “Tesoro digitale della lessicografia bilingue spagnolo-italiano”, financiado por La Unión Europea- Next Generation EU, Missione 4 Componente 1 CUP C53D23004010006 / PRIN 2022 MUR, cuya investigadora principal es Carmen Castillo Peña (Universidad de Padua).



ñola, entre las que se encuentra el vocabulario objeto de este estudio<sup>2</sup>.

En este artículo se abordará, en primer lugar, el análisis de la nomenclatura, situada dentro de la tradición lexicográfica alemana. A continuación, se examinarán aspectos relevantes de los análisis preliminares realizados para la edición crítica y la digitalización de la sección español-italiano del texto. En este marco, se analizan la macroestructura de la obra, determinada por la organización temática de los lemas, y su microestructura, considerando elementos como la presentación de variantes, irregularidades ortográficas, marcación de la flexión femenina, errores de transcripción y excepciones en la disposición de las entradas, aspectos que requieren la búsqueda de soluciones metodológicas adecuadas.

## 2. Datos generales sobre el autor y la publicación

La nomenclatura de Ignacio de Boria fue localizada durante la investigación realizada en el marco del proyecto historiográfico *LITIAS3* y la posterior publicación de la autora del *Catálogo analítico-crítico de la lexicografía italo-española*. El *Vocabulario Espanyol, Italiano y Tudesco* del marqués se conserva actualmente en el Fondo Antiguo de la Biblioteca Universitaria de Barcelona, procedente de la Biblioteca del Convento de San José. Hasta ese momento, las únicas referencias a esta obra se hallaban en la *Bibliografía de la Lexicografía Española del Siglo XVIII* de San Vicente (1995), donde se catalogan textos de autores españoles y extranje-

2 El corpus del proyecto está constituido por las siguientes obras:

- diccionarios bilingües generales: el *Vocabulario Español e Italiano* y *Vocabulario Italiano e Spagnolo* (1620) de Lorenzo Franciosini; el *Diccionario de faltriquera italiano español español italiano* (1805) de Cormon y Manni; el *Nuovo dizionario spagnolo-italiano italiano-spagnolo* (1948-1949) di Lucio Ambruzzi
- diccionarios especializados: el *Diccionario de música italiano español o el intérprete de las palabras empleadas en la música* (1841) de Moreali; el *Diccionario marítimo italiano-español* (1831) de Timothy O'Scanlan
- nomenclaturas: el *Quinque linguarum utilissimus Uocalista* (1526) de Francesco Garrone, la nomenclatura incluida en la gramática *L'italiano istruito nella cognizione della lingua spagnola* (1833) de Francisco Marín y el *Vocabulario Espanyol, Italiano y Tudesco* (1719) de Ignacio de Boria

3 Proyecto PRIN 2017, coordinado por Félix San Vicente: *La lingua italiana in territori ispanofoni, da lingua della cultura e della traduzione a lingua dell'educazione e del commercio: analisi storiografica attraverso il reperimento dei diversi testi e materiali metalinguistici con finalità descrittiva, contrastiva e didattica destinati a ispanofoni, dalle origini fino alla fine del ventesimo secolo* (ref: 2017J7H322\_004). Cfr. el portal LITIAS <<http://www.litias.it>>.

ros publicados entre 1700 y 1808 —siglo en el que de Boria concluyó su manuscrito—, así como en las bibliografías de Mauricio Fabbri: *A Bibliography of Hispanic Dictionaries. Catalan, Galician, Spanish. Spanish in Latin America and the Philippines* (1979, 2002). Sin embargo, este repertorio ha pasado prácticamente desapercibido para los estudiosos, pues no se menciona en los principales recopilatorios de referencia de historiografía lingüística, como las *Nomenclaturas del español* de Alvar Ezquerro, la *Bibliografía cronológica de la lingüística, la gramática y la lexicografía del español* (BICRES) o la *Biblioteca Virtual de Filología Española*.

La información disponible sobre Ignacio de Boria es, por lo tanto, escasa. En el *Vocabulario Espanyol, Italiano y Tudesco*, el autor firma con este nombre, aunque existen referencias a su identidad bajo la forma catalana *Ignasi de Bòria i Gualba*. Según se desprende de la breve introducción de su obra, Boria ostentaba el título de “señor del castillo y del término de Benviure”.

Gracias a la publicación del *Nobiliari General Català* de Armand de Fluvià, es posible acceder a algunos datos biográficos adicionales: se estima que nació alrededor de 1655; posteriormente, el 5 de febrero de 1690, contrajo matrimonio en Barcelona con María de Orlau y de Pastor.

Además, la participación de Boria en los asuntos políticos y militares de la Cataluña de su tiempo resulta particularmente relevante y se relacionan directamente con la Guerra de Sucesión Española (1701-1714), conflicto que desencadenó la muerte sin descendencia de Carlos II de España entre los principales pretendientes que aspiraban al trono: Felipe de Anjou, nieto de Luis XIV de Francia, y el archiduque Carlos de Austria. La designación de Felipe como heredero provocó la formación de una coalición internacional en su contra, integrada por potencias como Inglaterra, Holanda y Austria, que respaldaban al archiduque Carlos. En el ámbito interno, la guerra adquirió características de conflicto civil. Mientras que la mayor parte de Castilla apoyaba a Felipe V, los territorios de la Corona de Aragón, incluyendo Cataluña, se inclinaron mayoritariamente por el archiduque Carlos. En 1705, Cataluña reconoció al archiduque como rey con el título de Carlos III, estableciendo su corte en Barcelona. La resistencia catalana culminó en el asedio de Barcelona, que se prolongó desde el 25 de julio de 1713 hasta el 11 de septiembre de 1714. La caída de la ciudad marcó el fin de la resistencia austracista en la península y consolidó a Felipe V en el trono español.

Durante este periodo de resistencia, en 1709 nuestro autor fue designado consejero ciudadano del Brazo Militar en la *Junta general de brazos* (Martí, 2005), instancia en la que se tomó la decisión de defender el Principado por la vía armada, con el propósito de mantenerlo bajo la soberanía del Archiduque Carlos de

Austria y salvaguardar “la conservació de constitucions, privilegis, usos, costums e immunitats del Principat de Catalunya” (Eva Serra i Puig, 2008: 113). Su implicación en la resistencia continuó durante los años siguientes: el 30 de junio de 1713, participó en la *Junta de armas* de la Diputación del General junto con su hijo, José de Bòria y Orlau y en 1714, como Capitán de *La Coronela*, tomó parte en la defensa de Barcelona contra las tropas borbónicas.

Este episodio condujo al exilio forzoso de un gran número de partidarios austracistas. Es posible que Boria se encontrara entre ellos, ya que en 1714 emigró a Viena tras la confiscación de sus bienes y la anulación de sus títulos nobiliarios. Este dato coincide con la fecha de elaboración del *Vocabulario Espanyol, Italiano y Tudesco*, en 1718 en la capital austríaca. No obstante, esta hipótesis, basada en la correlación entre los hechos históricos y la trayectoria del autor, requiere ser confirmada mediante futuras investigaciones.

Posteriormente, el 5 de mayo de 1719, el Archiduque Carlos de Austria le concedió el título de Marqués. Sin embargo, este reconocimiento nobiliario nunca fue oficializado ni confirmado.

### 3. Descripción de la obra

El *Vocabulario Espanyol, Italiano y Tudesco* es una nomenclatura trilingüe, en español, italiano y alemán. Recordamos que, en este género lexicográfico, las nomenclaturas se incluyen tradicionalmente entre los repertorios temáticos que se presentan en capítulos configurados en torno a un vocabulario básico y necesario para el desempeño de las diferentes actividades que gobernaban la vida cotidiana de los hombres de su época. Son repertorios de carácter práctico y didáctico cuya clasificación no respeta un orden alfabético, sino temático, organizando los lemas según una forma lógica, una visión del mundo en parcelas conceptuales (Alvar Ezquerra, 2013). Según afirma Castillo Peña (2005:2), más allá de su estructura onomasiológica,

resulta difícil establecer otros rasgos distintivos que permitan caracterizarlas como grupo homogéneo: hay obras multilingües, o solo bilingües; algunas de ellas son vocabularios de discreta extensión que forman parte de tratados gramaticales, o de diálogos, y otras se publican de forma independiente; en algunas la clasificación metódica se acompaña de una clasificación alfabética, y otras carecen de un sistema auxiliar que permita encontrar un determinado vocablo; finalmente, junto a repertorios que se presentan en forma de tablas con la indicación escueta de las palabras, hay nomenclaturas que aportan una cierta información metalingüística.

Teniendo en cuenta lo anterior y pudiendo afirmar que la obra comparte el carácter onomasiológico de sus compañeras de clasificación, conviene realizar una detallada descripción de la estructura de la obra.

### 3.1 Descripción de la macroestructura

La primera característica que la diferencia de otras nomenclaturas ‘gemelas’ es que carece de prólogo u otros paratextos que nos permitan conocer la intención del autor al redactarla, a pesar de que algunas de sus particularidades, como veremos más adelante, nos inclinan a conceder al autor una intención didáctica.

La nomenclatura está dividida en 5 apartados. El primero [pp. 7-79], sin título, está precedido por el título general de la obra:

*Vocabulario Espanyol Italiano y tudesco trabajado por mi Dn Ignacio de Boria, Señor del Castillo y término de Benbiure con jurisdiccion civil y criminal en el Principado de Cataluña, y Comienço el trabajo (por huyr al enemigo de la ociosidad, tan pernicioso a la mocedad) en la imperial ciudad de Viena, Corte de los emperadores de Alemania, a los 20 de [diciem]bre 1718.*

Este capítulo se divide en 44 epígrafes, con un total de 2016 voces que no respetan un orden alfabético, eminentemente sustantivos, con rarísimas excepciones como las unidades pluriverbales que se refieren a colocaciones frecuentes que siguen: *a medio dia, tirar al bersallo, giocare a cartelai dadi/alla tavola/al suonol/agli escacos/alla dama/alle pedine/al gallo/al pisello/all'escacomo/alla morra/alle bongas, ir a cavallo, ir bien a caballo, dar bueltas, montar a cavallo, desmontar de cavallo, trotar el cavallo, a mano derecha, a mano izquierda, fare la comunione, confessare*. La conformación espacial se caracteriza por una división en tres columnas separadas por guiones como se puede observar en la figura 1. La primera equivale al español, la segunda al italiano, la tercera al alemán.

Vocabulario Espanyol, Italiano, y Tudesco, trabajado por mi D<sup>no</sup> Ignacio de Boria Senor del Castillo y Camarero de Benavise Con Jurisdiccion Civil y Criminal En el Principado de Catalunya, y Comienso el trabajo (por suya el tiempo de la vida de suena Corte de los Imperadores de Alemania año 20 de mayo 1718

Espanyol	Italiano	Tudesco
Dios	Dio	Gott
El Cielo	Il Cielo	Der himel
El Angel	l'Angelo	Der Enguel
El Paraiso	Il Paradijo	Der Paradijs
La Tierra	La Terra	Die Erde
El Purgatorio	Il Purgatorio	Der Purgator
El Limbo	Il Limbo	die Forkell
El Inferno	l'Inferno	Die Hölle
El tiempo	Il tempo	Die Zeit
El Dia	Il Giorno	Der Tag
a medio dia	a mezzo di	Der mitt Tag
La noche	la notte	Die nacht
a media noche	a mezza notte	Der mitt nacht
El Sol	Il Sole	Die Sonen
La Estrella	la Stella	Der Stern
El Elemento	l'Elemento	Der Element
La Luna	la Luna	Der mont
El mundo	Il mundo	Die welt
El hombre	l'omo	Der mensch
La muger	la Donna	Der weib
El alma	l'anima	Der Seel
El Demonio	Il diavolo	Der Teuffel
El Agua	l'Acia	Der Waff
El Fuego	Il Fuoco	Der Falger

Figura 1: Vocabulario Espanyol Italiano y Tudesco de Ignacio de Boria, 1719. Distribución en columnas

El segundo apartado (pp. 81-92) lleva el título *Algunas palabras que sirven para un passagero por pedir lo que ha menester*. A partir de su contenido y organización, se deduce que constituye una continuación de la sección anterior. Se compone de 12 capítulos que agrupan un total de 308 voces, predominantemente sustantivos. No obstante, la clasificación gramatical presenta menor uniformidad, incluyendo también expresiones verbales y léxico de uso general.

(1)

Dadme de beber	Datemi da berebere
Beber a la salud	Bere alla sanità
Le doy las gracias	La ringrassio
Por pedir	Per dimandare
Querría tener	Vorrey habere

verbos y léxico de uso general. Su estructura sugiere un propósito eminentemente conversacional. Hasta este punto, el contenido constituye la nomenclatura propiamente dicha, que actualmente se encuentra en proceso de digitalización dentro del marco del proyecto.

El tercer apartado (pp. 95-107) lleva por título *Vocabulario de algunos verbos más usados y necesarios* y consiste en una lista de 600 verbos en infinitivo en español y alemán, ordenados alfabéticamente; el italiano desaparece. El análisis comparativo con nomenclaturas previas de tradición italiana y alemana (§ 3.2) sugiere que esta sección guarda una correspondencia absoluta con una parte del *Vocabulario de los nombres y verbos más usados y necesarios*, incluido por Ernesto Joseph Eder en su *Florilegio español y alemán y gramática de la lengua alemana para los españoles, y española para alemanes*, publicado en Viena en 1714.

Al final de esta sección (p. 107), el autor anota la inscripción “FINIS, acompañada de la fecha a los 20 de febrero de 1719 se acabó este trabajo. Jesús María Joseph”. No obstante, la obra continúa con dos apartados adicionales.

El cuarto apartado (pp. 109-110), titulado *Declinación de algunos verbos más usados*, presenta un conjunto de 100 términos organizados en dos columnas. En la columna izquierda se expone la conjugación de ciertos tiempos verbales del verbo *amar* (tiempos presente, imperfecto, pretérito perfecto, futuro, imperativo e infinitivo), junto con su traducción al alemán. La columna derecha, en cambio, ofrece la traducción al alemán de diversos verbos y expresiones aisladas, tales como *tengo hambre, sed, sueño o hace frío, calor*, así como la construcción *puede ser*.



El quinto apartado (pp. 113-120), denominado *Algunas palabras de las más usadas*, recoge una lista de 450 adjetivos organizados en cuatro columnas, donde se presentan términos en español junto con sus equivalencias en alemán. Además, se incluyen otros vocablos, mayoritariamente adverbios, junto con expresiones de uso frecuente. La sección finaliza con una lista adicional de verbos traducidos. La disposición del contenido carece de un criterio organizativo claro y no sigue un orden alfabético, salvo en el caso de los verbos finales, que constituyen la única parte con una estructuración sistemática.

Como hemos adelantado, el léxico que recoge la nomenclatura en tres lenguas se reparte en 57 epígrafes en lengua española que constituyen las secciones I y II. Entre paréntesis se indican las entradas incluidas en cada uno de los epígrafes:

(Primera parte)

1. «De las Cosas de Dios» (55)
2. «De las Cosas de la Iglesia» (53)
3. «De los Estudios» (32)
4. «De la Corte y sus Oficios» (53)
5. «De la Mesa» (40)
6. «De las Viandas» (83)
7. «De los Pescados» (36)
8. «De las Aves» (41)
9. «Del Jardín» (58)
10. «De la Cavalleriza» (56)
11. «Diferentes Nombres de Cavallos» (43)
12. «De la Esgrima» (38)
13. «Del Juego» (17)
14. «Del Tiempo y de las Fiestas» (54)
15. «De las Partes del hombre» (68)
16. «De la Consanguinidad» (41)
17. «De las Ciudades» (42)
18. «De las Artes o Oficios Mecánicos» (48)
19. «De los Nombres Propios» (21)
20. «De la Familia y demás de Casa» (67)
21. «Del Trabajo de Mercante» (52)
22. «De la Moneda» (29)
23. «De la Materia de Vestidos» (24)
24. «Del Vestido» (50)
25. «De los Colores» (20)
26. «De los Instrumentos musicales» (27)
27. «De las Fortificaciones» (23)
28. «De la Guerra» (82)

29. «De los Oficios Militares» (42)
30. «Del Viaje» (34)
31. «De las Piedras o Joyas» (32)
32. «Del Magistrado» (82)
33. «Nombres de Provincias y Reynos» (56)
34. «De las Ciudades mas Principales» (67)
35. «De los Rios y las Montañas» (46)
36. «De los Signos Celestes» (33)
37. «De la Casa y de la Pesca» (39)
38. «De los Animales cuadrupedos» (44)
39. «De la Villa y agricultura» (42)
40. «De la cocina» (42)
41. «De los Arboles» (31)
42. «De los Frutos» (55)
43. «De los Sacramentos de la Iglesia» (32)
44. «De la Medicina» (71) Segunda parte

*Algunas palabras que sirven para un passagero por pedir lo que ha menester*

1. «Por saludar» (59)
2. «De las Aves» (12)
3. «De los Peces» (18)
4. «De los Frutos» (14)
5. «Del Beber» (16)
6. «De la Mesa» (27)
7. «Del Pasto» (12)
8. «Del pagamento» (14)
9. «Del Dormir» (16)
10. «Del Viaje» (49)
11. «Del Comprar» (7)
12. «De los Numeros» (44)
13. «Del Valor de la Moneda» (16)

### 3.2 Estudio de las fuentes

La sección de la nomenclatura en los tres idiomas consta de 2305 entradas, una cifra considerablemente reducida en comparación con la amplitud léxica que suelen abarcar no solo los diccionarios monolingües y bilingües, sino también las nomenclaturas.



Se han revisado antecedentes directos de nomenclaturas bilingües o trilingües con combinaciones lingüísticas similares (español con italiano o alemán), todas registradas en la *Biblioteca Virtual de la Filología Española* y descritas en *Las Nomenclaturas del Español* de Alvar Ezquerro (2013).

El análisis comparativo de los epígrafes de estas obras con el texto objeto de estudio revela importantes diferencias, más allá de algunas similitudes formales propias del género de las nomenclaturas. Estas discrepancias sugieren que el manuscrito podría ser un trabajo original. Esta posibilidad explicaría los numerosos errores de transcripción presentes en el documento, siempre que se confirme que efectivamente se trata de errores y no de variaciones intencionales.

Determinar la fuente o las fuentes utilizadas por el autor resulta complejo, ya que el repertorio no es particularmente extenso y las agrupaciones que presenta son comunes a muchas nomenclaturas precedentes. En algunos casos, recuerda a obras como la de Zumarán (1621) y Eder (1714)<sup>4</sup>. La condición militar del autor, evidenciada por la cantidad de términos relacionados con las artes militares y las armas, sugiere que existe un proceso de reelaboración y adaptación de los materiales a sus necesidades, tanto por las voces que recoge como por la distribución de los centros de interés, entre los cuales no se encuentran muchos de los habituales en este tipo de obras. El análisis comparativo con las posibles fuentes aún se encuentra en fase de investigación. Se prevé que, en futuras publicaciones, se lleve a cabo un estudio más exhaustivo sobre esta cuestión.

### 3.3 Características de la microestructura

Las equivalencias interlingüísticas se presentan en columnas separadas por guiones como ya se presentaban en obras equivalentes, en Franciosini (1626) o Zumarán (1626), por ejemplo. Las características de la microestructura derivan de la naturaleza de la obra; la información que se recoge en las entradas es coherente a lo largo de los dos primeros apartados por lo que también lo es la forma de la microestructura (voz española---equivalente italiano---equivalente alemán). Se

4 En este sentido, conviene recordar la facilidad con la que los autores reproducían el trabajo de sus predecesores. En relación con la obra de Eder, Alvar Ezquerro (2013:273) señala que, si bien la gramática de Zumarán es considerablemente más extensa que la de Eder, la nomenclatura de este último presenta claras similitudes con la del primero. Asimismo, destaca que la estructura de ambas obras no introduce innovaciones significativas, sino que sigue los modelos derivados de las gramáticas de César Oudin, a lo que añade que “la nomenclatura de Eder, es una copia de la de Juan Ángel de Zumarán, pues prácticamente, los capítulos en que se divide el léxico son los mismos”.

trata de entradas constituidas solo por sucintas equivalencias interlingüísticas, que generalmente (es casi una constante) presentan, como única información metalingüística de naturaleza gramatical, el artículo que ayuda a identificar el género de los sustantivos.

(2)

La Iglesia	La Chiesa
La Capilla	La Cappella
El Nuestro Señor	Il Nostro Signore

Podemos, sin embargo, señalar algunas excepciones:

a) Presentación de variantes. Van unidas por la conjunción *o* u otros elementos gráficos separadores: coma, espacio, etc., se encuentran tanto en la columna del español como en la de italiano

(3)

La Merla ó Merlo	Il Merlo
El Higo ó La Higa	Il Fico, La fica

b) Entradas con series sinonímicas. Se hallan tanto en la columna del español como en la del italiano con diferentes elementos separadores como la conjunción *o* u otros elementos gráficos separadores: coma, espacio, etc.

(4)

El Lirio o Azucena	Il Giglio
La Campaña o llanura	La Campagna Pianura
La Selva, El Bosque	La Selva, Bosco

c) Entradas con definiciones que sustituyen a las voces. Esta excepción se observa predominantemente en la columna española lo que indica que la direccionalidad de la nomenclatura puede variar a lo largo de la obra:

(5)

Comer â medio dia	Il Pranzo
La Yerba Rapa	Le Rappa
Un Pescado Grande	La Sfogliata
Ave que llama a las demás	Il Simbello

Desde un punto de vista metalingüístico, podríamos precisar que no existen marcas gramaticales. La indicación del género gramatical en el caso de de Boria la proporcionan los artículos que, con sorprendente uniformidad, preceden a casi todas las palabras principales, la mayoría con la forma determinada, a veces con la indeterminada:

(6)

El Pergamino	La Pergamena
El Quaderno De Papel	Il Quaderno di Carta
Una rasma de Papel	Una risma di Carta
Una hoja de Papel	Un foglio di Carta

En las entradas constituidas por verbos no proporciona información gramatical:

(7)

El Pergamino	La Pergamena
Tirar al bersallo	Tirar al bersaglio
Tirar ala mira	Tirar ala mira
Jugar	Giocare
A las cartas	A le carte
A los dados	Ai dadi

Por otra parte, notamos que la forma lematizada más frecuente es el masculino genérico sin especificar la flexión femenina en los sustantivos con variación de género; sin embargo, observamos algunas excepciones al respecto:

a) Lematización de formas femeninas tras el sustantivo masculino

(8)

El Hombre	Il Huomo
La Muger	La Donna
El Gentil Hombre	Il Gentil uomo
La Gentil dona	La Gentil Donna
El Padre	Il padre
La Madre	La madre

b) Lematización de la flexión femenina en los sustantivos con variación de género

(9)

El Cocinero	Il Cuoco
La Cocinera	La Cuoca

El padrino    Il Compadre  
 La padrina    La Comadre

c) Ejemplos donde el equivalente italiano presenta también la variante femenina (10)

El Moço        Il Giovane Ragazzo  
 La Moça        La Zitella Ragazza

d) La flexión femenina acompañando a la entrada del sustantivo masculino (11)

El Primo La Prima    Il Cugino a  
 El Esposo La Esposa    Lo Sposo a  
 El Viudo La Viuda    Il Vedovo La Vedova

e) Indicación de la desinencia femenina –a en los sustantivos de desinencia variable

(12)  
 El forastero la F<sup>ra</sup>        Il Forastero la Foras<sup>ra</sup>

Por otro lado, las nomenclaturas no suelen llevar marcas restrictivas y, en el caso de la nuestra, se respeta esta tradición. Los discriminadores semánticos a efectos de retrodigitalización suponen una cuestión compleja, aún en estudio; podemos entender, por ejemplo, los títulos y subtítulos en los que se distribuye el léxico como una forma de discriminación por ámbito de uso. Por ejemplo, el término «dama» lo encontramos en la sección:

(13)  
 4-Nº 4 DE LA CORTE Y SUS OFICIOS  
 La Dama        La Dama

(14)  
 38-Nº 38 LOS ANIMALES QUADRUPEDOS  
 La Dama        La Damma  
 así como en la expresión ‘a las damas’

(15)  
 13-Nº 13 EL JUEGO  
 (Jugar) A las Damas    Alle Dame

#### 4. Dificultades y desafíos de la fase de transcripción

A continuación se enumeran los desafíos ortotipográficos identificados durante el proceso de transcripción, así como las reflexiones que estos han suscitado. Dichos problemas pueden agruparse en tres categorías principales derivadas de:

a) Variantes ortográficas. Para el estudio de los elementos ortográficos de la nomenclatura, nos centramos en la ortografía de:

- *g, j, k*: observamos que la pronunciación de la fricativa velar /x/ está representada por las grafías *g, j, x*, y observamos que *jazmín, jubón, relojes, trabajo, rojo, ejercito* se transcriben con *j*, mientras que *muger, salvage, mensagero* con *g*, y también observamos algunos términos en los que la *x* sigue representando la pronunciación ‘fricativa velar sorda’ de algunas palabras latinas (ejemplo, *ejercicio*). También hay irregularidades dentro de la misma familia; *relojes, relojero*.
- *c, z, ç*: para indicar el sonido de la fricativa interdental sorda /z/, se utiliza la grafía *ç* ante *-a* y *-o* como en *pitança/ mançanal/ lançal/ moço* aunque también encontramos *ceniza, maza, zizania, cavallerizo, rizo*. Ante las vocales *-e, -i*, las grafías alternan con *z*: *hazer, mercanzia, vezino*, y con *c*: *cerradura, cerrojo, centinela, cocina, cilla*. También se observan irregularidades dentro de la misma familia: *cocina, cozinero/cozinera*.
- *b, v*: en los casos de la *b* y la *v* no existe coherencia con la grafía del DRAE, siendo algunas incoherencias: *escriviente, beber, cuervo, cerbeza* y *caballero, vino, ave, conserva*.
- *y* e *i*: un estudio de las grafías *y* e *i* como representación del sonido /i/ en el *Vocabulario*, nos permite confirmar la prevalencia de la grafía *-y-* como vocal tras *-a, -e*: *ayre, reyno, rey, buey, faysan, paysano, baylarino, bayle, maryscal, prisionero, azeyte, veynte* y la de la *-i* cuando se antepone a otra vocal: *Evangelio, oración, escritorio, frio, viento, demonio*, así como consonante: *ayuno, mayordomo yerba*.
- *cua* y *qua*: las secuencias /kwa/, /kwe/, /kwi/, /kwo/ se transcriben *qua, qüe, qüi, quo* en numerosos términos: *quaderno, quaresma, quartel, esquadra, esquadron, quanto* (*quíeres, cuesta...*).
- *h*: en cuanto a la *-h-* intervocálica, encontramos que de Boria sigue escribiendo *traher* a *traer* (recordemos que incluso hasta la edición del DRAE de 1739 se sigue encontrando la forma *traher*) y *parahiso*, pero encontramos que utiliza la grafía moderna para términos como *toalla* o *almohada*.
- La *h-* inicial se encuentra en la mayoría de los sustantivos que la usan según

la ortografía moderna: *honor, harina, hígado, hinojo, hoja, herida* excepto en *ospital*.

b) Errores de transcripción. Otras dificultades surgidas durante la transcripción y el análisis de la nomenclatura están relacionadas con la presencia de errores, relativamente frecuentes, derivados de la naturaleza manuscrita del texto, los cuales en ocasiones dificultan su correcta interpretación.

Cabe recordar que el autor era un militar en el exilio en contacto directo con las lenguas abordadas en la obra. En algunos casos, este contexto multilingüe ha dado lugar a una evidente influencia e interferencia entre ellas.

(16)

El mundo	Il mundo
La espiga	Lo spigo, la Espiga
La Missa	La Messa

c) Por último, no menos problemática es la dificultad de reconocer las variantes terminológicas obsoletas (en ocasiones diatópicas) que no están presentes en la lengua y los diccionarios modernos.

(17)

La Capa	Il Ferrayuolo
El Fenis	Il Bezzo
El Cedron ó Limera	Citrone

## 5. Conclusiones

Dado que el manuscrito no contiene prólogo, prefacio u otro tipo de paratexto, la intención del autor solo puede inferirse a partir de las breves líneas introductorias en las que señala haberlo redactado la intención de “[...] *huir al enemigo de la ociosidad, tan perniciosa a la mocedad*”. Esta declaración sugiere una posible finalidad didáctica, que también se refleja en el título de la segunda parte de la nomenclatura, “*Algunas palabras que sirven para un passagero por pedir lo que ha menester*”.

A pesar de su imperfección técnica y de las dudas con las que se afrontan algunas voces, es evidente que la obra trasciende el mero propósito lexicográfico y responde a una intención pedagógica y práctica más amplia, como se observa en la inclusión de un apartado dedicado a los verbos y otro con indicaciones de

carácter conversacional, lo que la aproxima a la estructura de un manual. Además, la presencia de numerosos errores tipográficos sugiere que el texto podría encontrarse en una fase de elaboración. Es plausible que el autor se haya basado en una nomenclatura preexistente, aunque no para copiarla literalmente, sino para reelaborarla y adaptarla a sus necesidades.

Desde una perspectiva microestructural y en el contexto del proceso de digitalización, se identifican diversas cuestiones relativas a la coherencia en la aplicación de ciertos criterios, como el uso de artículos y la disposición del contenido, con algunas excepciones. Asimismo, se observan variaciones en el tratamiento del género femenino y dudas en la transcripción gráfica de ciertos fonemas, como *j/g/x*, *c/z/ç*, *cu/qu* y, por último, *y/i*.

Por otro lado, desde una perspectiva macroestructural un análisis comparativo sistemático con nomenclaturas anteriores podría aportar información relevante sobre la selección y exclusión de determinados términos, así como sobre las preferencias del autor en relación con distintos ámbitos léxicos. Finalmente, si bien este estudio se ha centrado en las lenguas española e italiana, la relación del texto con la lengua alemana aún no ha sido explorada en profundidad. En este sentido, consideramos que este aspecto merece ser objeto de futuras investigaciones.

## Bibliografía citada

- Alva Ezquerro, Manuel (2013), *Las Nomenclaturas del español. Siglos XV-XIX*, Liceus.
- Bru de Sala, Lluís; Fluvià, Armand (1998), *Nobiliari del reial cos de la noblesa de Catalunya*, Barcelona, Editorial Joventut.
- Castillo Peña, Carmen (2005), “La *Nomenclatura italiana, francesa y española* de Guillaume Alexandre de Noviliers Clavel. La cuestión de las fuentes”, *Quaderni del CIRSIL* – 4, 1-14.
- Corvo Sánchez, M.<sup>a</sup> José (2008), “Breve historia de la lexicografía bilingüe española y alemana hasta el siglo XIX”, *Philologia Hispalensis*, 22: 113-139.
- Fabbri, Maurizio (2002), *A Bibliography of Hispanic Dictionaries. Basque, Catalan, Galician, Spanish, Spanish in Latin America and the Philippines*, Rimini, Panozzo Editore.
- Martí Fraga, Eduard (2005), “La Conferència dels Tres Comuns sota el govern de l'arxiduc Carles III. 1705-1713”, *L'Avenç*, 3-12.
- Nalesso, Giulia (2024), “Teoría y práctica de la retrodigitalización de diccionarios: el caso del *Vocabulario de las dos lenguas toscana y castellana*”, *Artifara*, 24.1, 135-153.
- San Vicente, Félix (1995), *Bibliografía de la lexicografía española del siglo XVIII*, Abano

Terme, Piovan.

Serra i Puig, Eva *et al.*, eds. (2008), *La junta General de Braços de 1713*, L'ambaixada Dalmases i altra documentació (1713-1714), Barcelona, Generalitat de Catalunya.

## Obras antiguas

- A. F. K. (1712), *El parte secondo, ò Nomenclator / Die Andere Abtheilung, Oder Einfache Wörter*, incluída en *Teutsch-Spanischer Richtiger Und Regulmäßiger Sprach-Zeiger*, Núremberg, J. L. Buggel.
- Eder, Ernesto Joseph (1714), *Vocabulario de los nombres y verbos más usados y necesarios*, incluída en *Florilegio español y alemán y gramática de la lengua alemana para los españoles, y española para alemanes*, Viena, Juan Jorge Schlegel, Impressor de la Universidad.
- Franciosini, Lorenzo (1626), *Nomenclator, o Registro de algunas cosas curiosas, y necessarias de saberse à los estudiosos de la lengua Española*, incluída en *Diálogos apazibles compuestos en castellano y traduzidos en toscano*, Venecia, Giacomo Sarzina.
- Kramer, Matthias (1670), *Nomenclatura toscana, tudesca e spagnuola di quei vocaboli che occorrono giornalmente. Nomenclatura toscana, tudesca, y spagnola de aquellas palabras o vocablos que se usan cada día. Italianisch, teutsch und spanisch Namen-Buchlein der jenigen Wörter, so täglich fürfallen*, Núremberg, Wolf Eberhard Felsecker.
- Noviliers Clavel, Guillaume Alexandre de (1629), *Nomenclatura italiana, francesa e spagnuola*, Venecia, Barezzo Barezzi.
- Oudin, César (1622), *Nomenclator o registro de algunas cosas curiosas y necessarias de saberse, a los estudiosos de la lengua española. Nomenclator ou mémoire de quelques choses curieuses & necessaires à sçavoir, aux studieux & amateurs de la langue espagnole, placé en appendice de: Diálogos muy apazibles escritos en lengua española y traduzidos en francés. Dialogues fort plaisans écrits en langue espagnolle et traduicts en françois, avec des annotations françoises [...]. Le tout fort utile à ceux qui désirent entendre ladite langue. Plus est adiousté un Nomenclator de quelques particularitez qui se présentent à tout propos*, Pierre Billaine, Paris.
- Perles y Campos, José Faustino (1689), *Nomenclatura español-italiano*, incluída en *Gramática española, o modo de entender, leer y escribir español*, Nápoles, Parrino & Mutti.
- Schweitzeren, Joannes (1665), *Tesaurus quinque linguarum aliarumque memorabilium, et cuicunque semper et ubique utilium annotationum*, Viena, Susanna Rickesin.
- Zumarán, Juan Ángel de (1621), *Nomenclatura en quatro lenguas, de las cosas más necessarias y comunes*, incluída en *Libro muy provechoso para aprender las lenguas*, Múnich. [Lenguas: francés-italiano y español-alemán].
- Zumarán, Juan Ángel de (1634), *La Grammatica y Pronunciación alemana y española*, Múnich. [Lenguas: español-alemán].



**Natalia Peñín Fernández** es investigadora Tenure Track en la Universidad de Bolonia, donde imparte cursos de Lengua y Lingüística Española. Doctora en 2019 por el DIT, ha trabajado en la enseñanza y la investigación en lexicografía bilingüe y didáctica de la traducción. Ha participado en los proyectos LITIAS y TELEI, y es coautora de *Lo Spagnolo Junior* (2016). Su labor académica se centra en la contrastividad italiano-español y la historiografía lexicográfica.

**natalia.penin2@unibo.it**



# CRISTIAN DÍAZ RODRÍGUEZ

## EL *DICCIONARIO DE VOCES AMERICANAS* DE AYALA (1777) EN EL *TLEAM*: SACRIFICAR EN ARAS DE HOMOGENEIZAR

Universidad de La Laguna

### Resumen

El *Tesoro lexicográfico del español en América (TLEAM)* aspira a reunir, en una misma plataforma, todas las obras lexicográficas que den cuenta del español hablado en América, lo que exige la homogeneización, tanto a nivel micro como macroestructural, de las obras repertoriadas. A partir del *Diccionario de voces americanas* de Ayala (1777), describiremos cómo se ha operado para conseguir un equilibrio entre el respeto a la fidelidad de los originales y la armonización exigida por la editorialización.

palabras clave: lexicografía, *TLEAM*, lematización, editorialización, armonización

### Abstract

*The Diccionario de voces americanas by Ayala (1777) in the TLEAM: to sacrifice in the name of homogenization*

*The Tesoro lexicográfico del español en América (TLEAM) aims to bring together, on a single platform, all lexicographic works that account for the Spanish spoken in America, which requires the homogenization, both at the micro and macrostructural levels, of the listed works. Starting from Ayala's Diccionario de voces americanas (1777), we will describe how efforts have been made to achieve a balance between respecting the fidelity of the originals and the harmonization required by editorialization.*

keywords: lexicography, *TLEAM*, lemmatization, editorialization, harmonisation

## 0. Introducción

En 2018, los lexicógrafos hispanistas Dolores Corbella, Alejandro Fajardo y Jutta Langenbacher-Liebgott editaron un libro que aspiraba a reunir la descripción de los proyectos más punteros en el ámbito de las Humanidades digitales pilotados por «labrantes de palabras»: *Historia del léxico español y Humanidades digitales*. Además de señalar las ventajas que presentaba la conjunción Informática-Filología, insistían ya en el prólogo de la obra en la necesidad de estar atentos a los desafíos que este maridaje iría revelando con el tiempo:

Estos avances, que han traído consigo un indiscutible progreso, no están exentos de algunos problemas que se derivan de su concepción y de su funcionamiento. La dispersión de los recursos, la falta de interconexión, la carencia de estándares, la rápida obsolescencia, etc., dificultan el acceso a la información que contienen. Su diseño, por otra parte, no siempre resulta adecuado para determinadas investigaciones. Todo esto hace que sea necesaria una crítica especializada, desde el punto de vista del usuario filólogo, que contribuya a señalar sus puntos débiles para poder mejorar las próximas versiones y los nuevos productos (2018: 5-6)

Ciertamente, uno de los mayores retos a los que nos enfrentamos a la hora de crear herramientas o recursos digitales es el de la homogeneización exigida por la *editorialización*, entendida en su sentido más estricto<sup>1</sup>. En efecto, las enormes posibilidades que nos ofrece el manejo informatizado de cantidades ingentes de datos conllevan, por desgracia, una pérdida tributaria a la especificidad de cada fuente original, un pequeño precio que hay que pagar en aras de garantizar el tratamiento eficaz de la información manejada, pero que tiene una impronta mayor desde una perspectiva patrimonial.

Évidemment, cela a aussi un impact sur les contenus eux-mêmes : le concept d'éditorialisation souligne comment la technologie façonne les contenus. Selon cette définition, on serait tenté d'assimiler l'éditorialisation à la curation des contenus (*digital curation*) – qui désigne le processus d'organisation des contenus dans un environnement numérique déterminé. Mais il y a une distinction fondamentale entre les deux processus : le concept d'éditorialisation implique une dimension culturelle qui n'est pas présente dans l'idée de curation. Cette dernière renvoie plutôt aux pratiques visant à rassembler, organiser et afficher des contenus dans un environnement donné – ce qui met d'abord l'accent sur l'ensemble des compétences nécessaires pour réaliser une

---

1 [L']éditorialisation est une instance de mise en forme et de structuration d'un contenu dans un environnement numérique. On pourrait dire, en ce sens, que l'éditorialisation est ce que devient l'édition sous l'influence des technologies numériques. (Vitali Rosati 2020 § 31)

bonne curation. En revanche, l'éditorialisation fait référence à la façon qu'ont les outils, les pratiques émergentes et les structures déterminées par les outils d'engendrer une relation différente avec les contenus eux-mêmes. On pourrait dire que la curation désigne l'action d'un individu spécifique ou d'un groupe d'individus défini, tandis que l'éditorialisation met l'accent sur la façon dont cette action est structurée par les caractéristiques de l'environnement numérique. Il faut enfin souligner que ces caractéristiques ne sont pas seulement techniques, mais aussi culturelles. (Vitali Rosati 2020 § 32)

En este artículo nos centraremos en el análisis de cómo se ha tenido que proceder, en el caso de la elaboración del *Tesoro lexicográfico del español en América*, al que nos referiremos por sus siglas *TLEAM*<sup>2</sup>, para conciliar la homogeneización de los materiales reunidos con la salvaguarda de la especificidad metalexicográfica de cada una de las fuentes almacenadas. En efecto, el ejercicio de digitalización homogeneizadora, indispensable para garantizar la eficiencia de las eventuales consultas, exige una reflexión profunda en torno a diversos aspectos relacionados tanto con la variedad identificada en el nivel macroestructural como con aquella propia de la microestructura de las obras lexicográficas: órdenes macroestructurales no necesariamente alfabéticos, ortografías vacilantes, lematización geminada, mezcla de categorías gramaticales en la lematización, falta de sistematización de las lematizaciones anidadas... Describir cómo se ha procedido en el *TLEAM* para dar respuesta a cada uno de estos desafíos es el objetivo principal de este artículo. En concreto, nos valdremos de la forma en la que ha sido adaptado el *Diccionario de voces americanas* de Ayala (1777) para su introducción en la susodicha plataforma tesórica.

Tras hacer una pequeña descripción del proyecto *TLEAM*, así como de la obra que hemos seleccionado para ilustrar el proceso de homogeneización, nos centraremos en la presentación y discusión de las diferentes decisiones metodológicas que se han tenido que implementar para poder garantizar su inclusión en la plataforma de hospedaje y aumentar la eficiencia de su consulta a través de los procedimientos de búsqueda previstos.

---

2 Esta publicación forma parte de los proyectos de I+D+i PID2020-117659GB-I00 y PID2023-149847NB-I00, financiados por el MCIN/AEI/10.13039/501100011033.

## 1. El *TLEAM*: una estructura para salvaguardar el léxico del español de América

El *Tesoro léxico del español en América-TLEAM* (cfr. Corbella Díaz; Díaz Rodríguez 2023; Corbella Díaz, Fajardo Aguirre, Díaz Rodríguez 2024) es un proyecto que, siguiendo la estela de la tradición tesorística hispánica, inaugurada por Gili Gaya y su *Tesoro Lexicográfico* (1492-1726), aspira a convertirse en un repositorio lexicográfico del español en América y, por consiguiente, en un depositario de la memoria patrimonial léxica de las regiones hispanófonas del Nuevo Mundo<sup>3</sup>. El proyecto se articula en tres fases de volcado de materiales lexicográficos a la plataforma: la primera, ya concluida, incluía las obras publicadas hasta 1884, fecha en la que la Real Academia Española sistematizó, por primera vez, la marcación diatópica, introduciendo, además, el concepto de *americanismo*; la segunda, actualmente en marcha, incluye las obras datadas entre 1884 y 1925, año marcado por la publicación de la decimoquinta edición del *Diccionario de la Lengua Española* (conocido tradicionalmente como *DRAE* hasta su edición de 2014); la tercera, por su parte, tomará como fecha de cierre la creación de la Asociación de Academias de la Lengua Española (ASALE) en 1951<sup>4</sup>.

En esta infraestructura se almacenan, en formato digital, los testimonios diferenciales que han visto la luz en la América hispana, con independencia de que puedan catalogarse exactamente como obras lexicográficas *stricto sensu*: desde diccionarios generales o vocabularios a glosas a listados con disposición lexicográfica. En efecto, el *TLEAM* aspira a ser una herramienta de la que se puedan valer tanto el usuario profano como el especializado, lo que exige un alto grado de versatilidad en el diseño de sus prestaciones. Por ello, además de contar con un sistema de búsqueda por lema al uso, o guiado por índice, se ha adoptado un rico sistema de marcación textual, basada en estándares de codificación TEI basado en meta-lenguaje XML, implementado a través de la plataforma TEITOK (Arrabal Rodríguez 2022), que aumente, por un lado, las posibilidades de interrogación del corpus (etimología, información gramatical, lemas, forma normalizada, variantes,

3 En este sentido, bien es cierto que se cuenta ya con algunos proyectos —finalizados o *in fieri*— que tienen por objetivo el crear «diccionarios de diccionarios» que den buena cuenta del español de ciertas zonas geolectales, con una perspectiva diacrónica, p. ej.: *Tesoro Lexicográfico del español de Puerto Rico* (*Tesoro. pr*) o el *TLEX-Tesoro Lexicográfico de Cuba* (Camacho Barreiro 2010) pero ninguno aspira a cubrir toda el área americana de habla hispana.

4 Aunque sería deseable completar la horquilla temporal cubierta con obras posteriores a 1951, lo cierto es que, esto implicaría el tener que depender enormemente de los permisos de propiedad intelectual o explotación de cada una de las obras recogidas, lo que nos hace desistir ante tal improbable desafío.

ejemplos, acepciones, información fraseológica, marcas de variación, categoría ontológica, zona geolingüística...); y, por otro, la personalización del corpus de búsqueda en función del objetivo perseguido. En otras palabras, al posibilitar la parametrización de los subcorpus de trabajo, así como las posibilidades de búsqueda, para adaptarlas a los objetivos individuales de cada usuario, el *TLEAM* funcionaría como un auténtico *hub* multicorpus y multipropósito. Es justamente esta faceta de corpus consultable y explotable —concebido en términos de Rastier (2004)<sup>5</sup>—, la que exige una homogeneización o, si se prefiere, una armonización, a varios niveles, del material almacenado, de manera que se permita una búsqueda y una explotación eficaz, al mismo tiempo que se asegure una visualización clara y ergonómica de la información volcada, que salve las especificidades tributarias a la diversidad introducida por la disparidad de documentos reunidos, cuya materialización puede afectar a distintos niveles: diferencias diamésicas, multiplicidad de criterios para organizar la macroestructura, falta de sistematicidad en la lematización, descarte de los topónimos y antropónimos, variación ortográfica en diacronía, discrepancias geolingüísticas, etc.

Ahora bien, a pesar de que cualquier ejercicio de homogeneización exige un sacrificio en detrimento de la fidelidad absoluta de la reproducción de los originales, no es menos cierto que:

En cualquiera de los casos, esta renuncia procedimental no es óbice para que el investigador interesado en estos aspectos pueda acceder a la disposición primigenia, puesto que el *TLEAM* ofrece la posibilidad de acceder a la ficha documental [...] de cada una de las obras que acoge, en la que [...] se permite el acceso tanto a la información paratextual [...] como a la copia facsímil o escaneada del documento original, cuando esta obra esté digitalizada. (Corbella Díaz, Fajardo Aguirre, Díaz Rodríguez 2024: 281)

### 1.1 *El Diccionario de voces americanas de Manuel José de Ayala (1777)*

Antes de proceder a escrutar algunas de las intervenciones realizadas durante el proceso de editorialización del *TLEAM*, haremos una pequeña descripción de la

5 “Un corpus est un regroupement structuré de textes intégraux, documentés, éventuellement enrichis par des étiquetages et rassemblés : (i) de manière théorique réflexive en tenant compte des discours et des genres, et (ii) de manière pratique en vue d’une gamme d’applications”.

6 La versatilidad a la que se aspira es tal que también se pone a disposición del usuario la consulta del material liminar paralexográfico, como pueden ser las introducciones, las notas al lector, los prólogos o los prefacios.

obra de la que nos valdremos para ejemplificarlas: el *Diccionario de voces americanas* de Manuel José de Ayala (1777). Este es el título que se le asignó a un manuscrito, compuesto por dos tomos —II-884 y II-855, de 455 folios numerados y 141 folios sin numerar respectivamente— conservado en la Biblioteca del Palacio Real de Madrid dentro del catálogo de *Manuscritos de América* con el título de *Diccionario de América* y que fue editado por Miguel Ángel Quesada Pacheco para Arco/Libros en 1995. Aunque se desconoce con certeza el autor de esta obra, se suele atribuir a Manuel José de Ayala<sup>7</sup>, ya que, además de haber estado estrechamente ligado a la Corte<sup>8</sup> —lo que habría posibilitado que su obra se conservara en la Biblioteca del Palacio Real—, existen similitudes gráficas y estilísticas, amén de una interreferencialidad notable entre esta obra y su *Cedulario Índico* y su *Miscelánea*. No en vano, este diccionario figura en la relación de méritos y servicios de Ayala.

A pesar de que no llegó a publicarse en el momento de su redacción, a saber, el siglo XVIII, se puede presuponer que se elaboró con esta finalidad pues, como nos recuerda en el prólogo de la edición Quesada Pacheco, pese a que hay folios completamente en blanco: “[e]l libro fue concebido de antemano, ya que casi todos los folios contienen letras de ordenamiento, donde se irían a poner los artículos de acuerdo con el criterio metodológico establecido por el que concibió la obra” (1995: x), lo que implica que los folios se iban llenando progresivamente a medida que se fuese recopilando el material. De hecho, “a lo largo del manuscrito hay diferentes fechas que consignan el día en que se introdujo la palabra, lo cual fortalece la idea de su composición gradual y paulatina” (Quesada Pacheco 1995: xi). Justamente, gracias a la consignación de estas fechas, ha sido posible establecer la fecha, o, mejor dicho, la horquilla temporal, de elaboración de este proyecto lexicográfico: entre 1751 (s.v. *bejuquillo*) y 1777 (s.v. *holpatán*).

---

7 Otras hipótesis consideran a Alcedo como potencial autor de esta obra fundamental para la lexicografía hispanoamericana, pues se arguyen coincidencias parciales entre esta y otra cuya autoría sí que está comprobada, a saber: *Diccionario geográfico histórico de las Indias Occidentales* 1786-1789. Sin embargo, las discrepancias estructurales y estilísticas parecen descartar esta autoría.

8 Manuel José de Ayala fue ilustre recopilador de las leyes de las Indias, así como archivero de la Secretaría de Estado del Despacho Universal de Indias. Para más inri, en el artículo dedicado a la voz *anaquine* se hace referencia explícita a la profesión del supuesto artífice de este diccionario: “Esta noticia comprobé en los papeles del Archivo de mi cargo, en carta de aquel govern[ad]or d[o]n Jossep L Antonio Gil de 26 de sep[tiem]bre de 1763, escrita al ex[celentísi]mo S[en]or Bo. F. d[o]n Julián de Arraiga remitiendo un cajoncito con 14 beas de ella que tuve en mi mano”. (Tomo I, f. 49 v. *apud* Quesada Pacheco 1995: 13).



## 2. Retos de la homogeneización exigida por la *editorialización*

Sin que podamos abordar de manera exhaustiva todas las adaptaciones que, para la elaboración del *MLEAM*, hemos tenido que operar en los diferentes textos-meta de carácter lexicográfico con la intención de poder garantizar un equilibrio entre fidelidad / autenticidad del documento original, por una parte, y la eficacia de búsqueda / la organicidad de visionado de la información, por otra; presentaremos, a continuación, a título de ejemplo representativo, una serie de adaptaciones que han tenido que operarse en el texto transcrito a partir de la obra original de Manuel José de Ayala, *Diccionario de voces americanas*, para poder garantizar su integración efectiva en la plataforma de custodia y consulta *MLEAM*.

### 2.1 Sobre el orden determinado por la macroestructura

Una de las principales características del *Diccionario de voces americanas de Ayala* es la forma en la que se disponen los lemas en la macroestructura, pues no siguen el orden alfabético al uso, sino según una aplicación parcial del mismo: se ordenan las entradas según la disposición gráfica de las dos primeras letras que componen cada vocablo —aa, ab, ac, ad...—, sin importar si conforman o no una única sílaba; pero, a partir de ahí, no se sigue un orden marcado por el abecedario<sup>9</sup>, lo que produce que nos encontremos con secuencias como: AL {alcribís, aloes, alhuehuetes, algarrobas, alabastro}; JA {jagüelles, jaboncillos, jalsontles, jayo, Jaén, Janta, Jardín del rey, Java Mayor y Menor, Juana / Cuba / Fernandina, Jayna, Joujuas, Janos, Jalapa, Jabalí, Jayba}. Consciente de esta dificultad, en su edición, Quesada Pacheco, completa la reproducción del diccionario con un índice alfabético de palabras:

Para facilitar la búsqueda de un término, se ha hecho un índice alfabético de todas las entradas del manuscrito, al final del libro. En él se han incluido también los sinó-

9 Al optar por esta particular disposición del lemario, Ayala dotaba de flexibilidad a su obra, pues proponía un sistema de repertoriado a caballo entre el diccionario y un archivo vertical o base de datos *avant la lettre*, idóneo para una alimentación paulatina, no necesariamente secuencial. Su contrapartida, como hemos señalado, la dificultad de su consulta dada la imprevisibilidad de la secuenciación de lemas. En cualquiera de los casos, si tenemos en cuenta otros aspectos que también apuntan a la poca sistematización de la obra —p. ej.: diversidad de caligrafías, presencia de rasgos ortotipográficos vacilantes, variedad de tintas, uso caprichoso de cursiva y letra de imprenta, disparidad en el esmero caligráfico...—, todo parece apuntar a que esta no sería la disposición final prevista por el autor, sino tan solo una forma de facilitar su trabajo de recopilación.

nimos y otras palabras de interés (nombres propios, americanismos, indigenismos) que figuran dentro de un artículo, pero que no aparecen como entrada, y las entradas que presentan variaciones gráficas en la actualidad. En vista de que se ha respetado la escritura y el orden de entradas del manuscrito, se recomienda consultar el índice antes de pasar al diccionario [...]. Con esto se facilita la búsqueda sin correr el riesgo de pasar por alto algún artículo. (1995: XLIV)

Desde el punto de vista de la editorialización, esta peculiaridad no presenta ningún problema a la hora de integrar la información en el *TLEAM*, pues el almacenamiento de los diferentes lemas puede llevarse a cabo siguiendo la disposición propuesta por Ayala, lo que permite garantizar cierta fidelidad al consultar el leuario de manera íntegra, como se muestra en la Imagen 1.

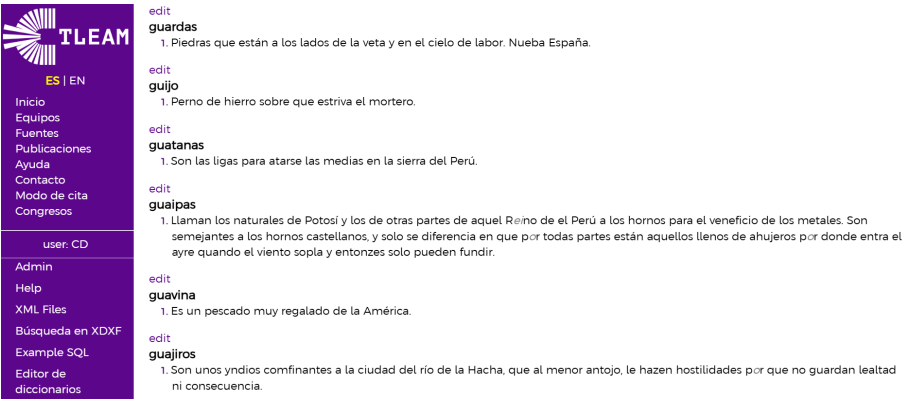


Imagen 1: ejemplo de secuenciación de lemas que no siguen el orden alfabético.

Ahora bien, en principio, el común de los usuarios nunca accederá al contenido de las palabras de manera secuencial (tampoco alfabética) sino a través de un motor de búsqueda que permita mostrar en pantalla la información solicitada. La única salvedad la podríamos encontrar a la hora de realizar una búsqueda a través de una eventual rueda de deslizamiento (ver imagen 2) en la que aparecieran todas las palabras contenidas en nuestra plataforma; posibilidad, a nuestro juicio, poco operante, dado que este leuario general forzosamente tendría que seguir un orden alfabético estricto, que iría en contra de la disposición seguida por Ayala.

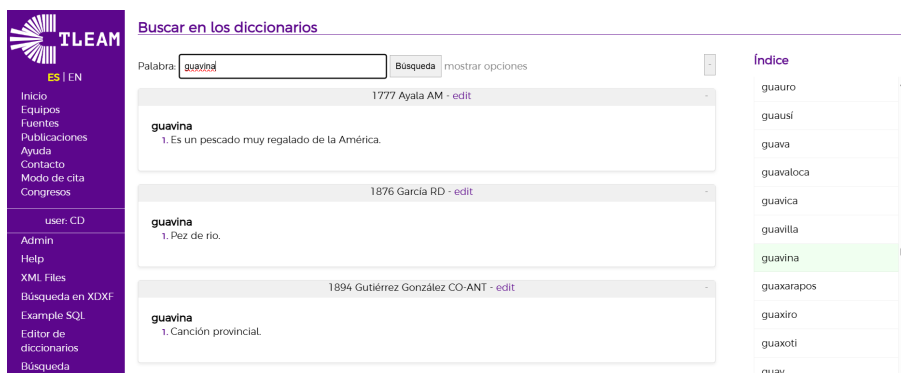


Imagen 2: interfaz de búsqueda a través de un motor de búsqueda y de un leuario deslizable global ordenado alfabéticamente.

## 2.2 Sobre la ortografía vacilante

En general, uno de los mayores retos a los que nos enfrentamos durante un proceso de editorialización que abarque una horquilla temporal tan amplia como la propuesta por el *TLEAM* es el de intentar repertoriar, y brindar un trato similar, a todas las variantes gráficas de una palabra recogidas en diacronía, en muchos casos atribuibles a la existencia de una norma poco marcada o, incluso, totalmente ausente. En el caso del *Diccionario de voces americanas*, las principales razones de la abundante variación gráfica se deben a la falta de definición clara de los valores y usos de ciertos grafemas, así como a una vacilante correlación entre fonemas y grafemas:

- Confusión entre “i” e “y”, pues no se reserva el primero de estos grafemas para los valores vocálicos y el segundo, para los consonánticos. Así, por ejemplo, si observamos simplemente el caso de las palabras que comienzan gráficamente por “i” o “y”, podemos constatar la cohabitación de lemas que efectivamente respetan la repartición funcional preconizada en el siglo XVIII —valor vocálico de “i”: *ixistineses*, *iztaes*, *icotea*, *inca* o *inga*, *incienso*, *indio timaba*; y valor consonántico de “y”: *yanaconas*, *yareta*, *yamas*, *yato*, *yacaré*, *yamacuá*, *yalamacuá*, *yapayaes*, *tabaqué*, *Yabucoa*, *yaquino*, *yaqui*, *yare*, *yayrepa*, *yantayanta*, *yavar chunca*, *yeguana*, *yerba*— con numerosos ejemplos que contravienen esta distinción y que, por tanto, reflejan un valor consonántico de “i” —p. ej.: *iantem*— y un valor vocálico de “y”:

*Ybagui, ybagué, Ybueras* (o *Higüeras*), *ycho, ychuris, ycacos, yca, ycho, yllo, ylbol, ymagua, ymperial, yncorporadero, yndehé, ynguarán, ynfeles, yngenos, indios*. En algunas ocasiones, hay que añadir un elemento suplementario que también interviene en el paradigma, ora la “j” —*icotea* vs. *jicotea*—, ora la “ll”: *yamas, iantén*.

- Uso vacilante de la “h”, tanto inicial como interno-etimológica: *hamana, icotea, Carthago, uipeles, higuana*.
- Confusión de la “b” y la “v”: *bálsamo perubiano, báquiras, Córdova, Varco* (Santiago del Estero).
- Inestabilidad en el sistema de sibilantes: *senzontle, passote, Epazote*.
- Falta de sistematicidad en el uso de la “e” protética: *Spíritu Santo*.
- Hesitación para representar gráficamente el fonema /j/: *Xeniquén / He-nequén, jurón / hurón, Xaca, Xagua, Xaguay, Xalisco, Xamayca, xengibre, Xihuitl (Jigüite)*.

Esta situación se resuelve, en la mayoría de casos, consignando una de las variantes gráficas —generalmente la aceptada sincrónicamente por los diccionarios de referencia, en nuestro caso particular el *DLE* o el *DAmer*— como forma normalizada (<k reg>), por ejemplo *pava*, y enlazar todas las variantes registradas a esta, en este caso (cfr. tercera línea de código en la Imagen 3) *pabas*, de manera que, al lanzar una búsqueda, ese lema sea recuperable tanto si se introduce esa variante en el motor de búsqueda como si se introduce la forma normalizada.

```
</deftext></def></ar><ar id="ar-999"><k reg="payagu&#xE1;n">payaguanes</k><def id="ar-999.1" n="1"><
</deftext></def></ar><ar id="ar-1005"><k reg="mora">paniquiris</k> o <k>moras</k><def id="ar-1005.1"
</deftext></def></ar><ar id="ar-1007"><k reg="pava">pabas silvestres </k><def id="ar-1007.1" n="1"><
</deftext></def></ar><ar id="ar-1010"><k reg="palometa">paguara</k> o <k>palometa</k> o <k>payara</k>
</deftext></def></ar><ar id="ar-1011"><k reg="payaguarima">payaguarima</k><def id="ar-1011.1" n="1">
</deftext></def></ar><ar id="ar-1012"><k reg="palo">palo amarillo</k><def id="ar-1012.1" n="1"><gr v
```

Imagen 3: código XML crudo para vincular las variantes de un lema con una forma normalizada común.

### 2.3 Sobre la homogeneización de la lematización

El asociar cada una de las variantes de los lemas a una forma normalizada también es la solución que se ha implementado para dar respuesta a la falta de sistemati-

cidad de la lematización. En efecto, contrariamente a lo que se preconiza actualmente en la práctica lexicográfica, según la cual el lema debe aparecer, si es un verbo, en su forma infinitiva; o en su forma masculina singular, eventualmente acompañada de los morfemas de género y, rara vez, de número, cuando se trata de adjetivos o sustantivos, se comprueba que, en el proceso de editorialización del *TLEAM*, se ha apostado por mantener la variante original del lema que sirve de entrada para cada artículo. Así, por ejemplo, como se observa en la Imagen 3, se ha conservado como lema *payaguanes*, en plural, pero se ha asociado a la forma normalizada *payaguán*, en singular, de manera que la interfaz de volcado de resultados mostrará los mismos resultados con independencia de que el usuario introduzca *payaguán* o *payaguanes* en el motor de búsqueda. Esto no se cumple en el sentido inverso, es decir, mientras que una búsqueda a partir de la forma normalizada *payaguán*<sup>10</sup> proporcionará al usuario todas las entradas encabezadas tanto por *payaguán* como por *payaguanes* (cfr. Imagen 4), una búsqueda a través de la forma *payaguanes* solo volcará la información relacionada con las entradas cuyo lema también aparezca en plural, obviando, pues, las que solo aparezcan en singular.

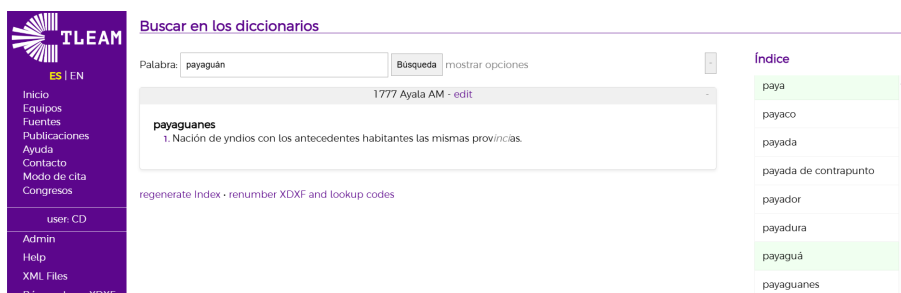


Imagen 4: volcado de la información del lema *payaguanes* al que se ha accedido a través de la forma normalizada *payaguán*.

10 Sin embargo, como se puede apreciar en la Imagen 4, la variante *payaguán*, al no haber sido reconocida como lema de entrada en ninguna de las obras registradas hasta ahora, no aparece en el leuario deslizante global, por lo que, el acceso a la información a través de este procedimiento —clicar en el leuario deslizante— estaría bloqueado para las formas normalizadas que no cuentan con una variante de un lema isomórfico registrado.

## 2.4 Sobre la lematización geminada

La asociación de las distintas variantes con una forma normalizada, que, en los casos anteriores, ha resultado ser tan satisfactoria, en el caso de la lematización geminada —p. ej.: *acmulí* / *amole*; *almorajifazgo* / *diezmos*; *antena* / *entena*—, en cambio, presenta limitaciones. Este escollo es especialmente significativo dada la alta frecuencia de entradas con lematización geminada a las que recurre Ayala, adoptando, para ello, distintas disposiciones:

- Enumeración yuxtapuesta: *Juana, Cuba, Fernandina*.
- Enumeración copulativa: *Java Mayor y Menor*.
- Enumeración disyuntiva total: *Arbinas, o tembladeras; Bejuquillo, hebecaquana, o raicilla; Pacae o pacay*.
- Enumeración disyuntiva parcial: *bálsamo negro, o [bálsamo] peruviano*.
- Reenvío hiperonímico o paronímico: *Páquira / báquiras*: vide *Jabalí*
- Amalgama de distintas categorías gramaticales: *demora o demorados*

Si observamos, nuevamente, la Imagen 3, en concreto la segunda y la cuarta líneas de código, se puede comprobar cómo las entradas vinculadas respectivamente con las formas normalizadas *mora* y *palometa*, corresponden a las entradas con lematización geminada: *paniquiris* / *moras* y *paguara* / *palometa* / *payara*, cada una marcada como co-lemas a través de la etiqueta “k” entre las que, por cuestiones de legibilidad, se ha introducido la conjunción disyuntiva “o”. La limitación a la que acabamos de aludir consistiría en la imposibilidad de asociar cada uno de los co-lemas a una forma normalizada casi-isomorfa, teniendo que seleccionar una única *k* reg.

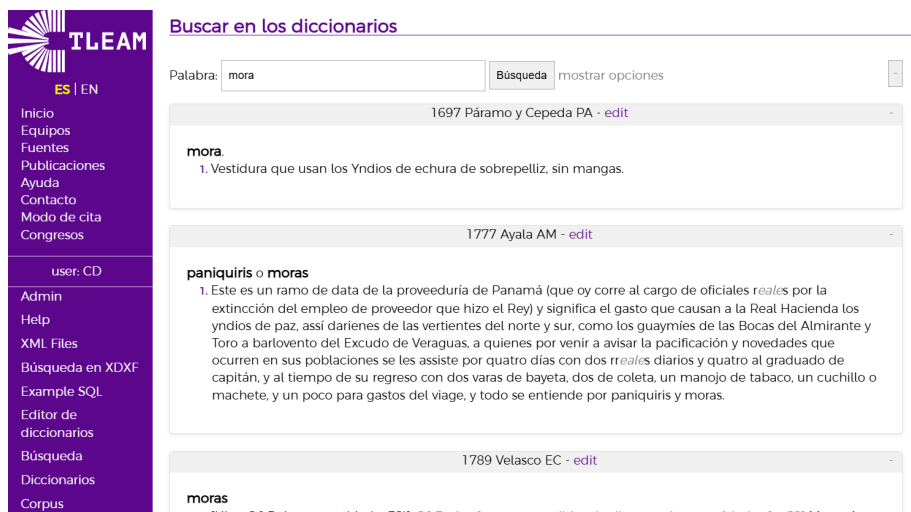


Imagen 5: interfaz de salida de búsqueda en el caso de buscar a través de una palabra normalizada ligada a un lema geminado.

La consecuencia de esta limitación es especialmente significativa en el momento de volcar los datos resultantes de la búsqueda. Así, por ejemplo, la información contenida en el artículo encabezado por el lema geminado *paniquiris* o *moras*, se mostrará cuando el usuario introduzca en el motor de búsqueda los valores “paniquiris” y “moras” —coincidentes con las variantes originales de los lemas, o el valor “mora”— carente de formas flexivas, que ha sido identificada como forma normalizada. Sin embargo, si el usuario interroga el motor de búsqueda a partir del valor “paniquiri”, forma, en principio, esperable según las preconizaciones de la lematización contemporáneas, no obtendríamos ningún resultado, puesto que no hay nada que lo vincule con nuestro artículo.

## 2.5 Sobre la duplicidad de lemas

La naturaleza *in fieri* del proyecto lexicográfico de Ayala bastaría para justificar la inclusión de repeticiones, que, presumiblemente, habrían desaparecido una vez se hubiera procedido a corregir sistemáticamente la obra con la finalidad de dotarla de su forma definitiva. En algunos casos, no hay duda de que se trata de duplicados no deseados por el autor, ya que se reproducen de manera exacta, o práctica-

mente sin alteraciones, las definiciones vinculadas a ciertos lemas. En otros casos, en cambio, la repetición es totalmente voluntaria, pues es la forma adoptada por el autor para enfrentarse tanto a los fenómenos de homonimia como a los de polisemia. Este es el caso, por ejemplo, de la voz *víbora*, que se encuentra definida tanto en el folio 100v como en el folio 101r, con sendas acepciones: “isletas junto a Xamayca, donde se han perdido por su peligro muchos navíos” y “en todos los países fríos y aun en los cálidos y templados de Yndias las hay en mucha abundancia, y de tan mortífera picadura como las de España; pero al mismo t[iem]po abundan aquellos parages de antídotos contra su ponzoña.”; o de la voz *patí*, cuyas acepciones aparecen definidas hasta en tres páginas diferentes, cada una constituyendo un nuevo duplicado lematizado: “es un pez de las veinte y dos especies de que se haze relación en la antecedente voz.” (f. 25v, s. v.); “río caudaloso, que pasa por la ciudad de Santa Fee en el Nuevo Reyno de Granada” (f. 30r, s. v.) y “es un lienzo de algodón que fabrican los yndios de la provincia de Yucatán; cuyas piezas tienen dos varas de largo y tres quartas de ancho, y su valor regular son doce reales. De estas dan dos anualmente todos los indios a sus curas por vía de limosna que han introducido estos.” (f. 32r, s. v.).

Para intentar paliar esta situación, en el *TLEAM* se ha optado por mantener la multiplicidad de entradas con independencia de que su separación responda o no a un caso de homonimia, lo que, atendiendo a la lexicografía moderna, sí que implicaría la duplicidad de las entradas. Si bien esta opción, en el caso del *Diccionario de voces americanas* de Ayala, parece responder a la necesidad de respetar fielmente la obra original, esta decisión no deja de tener consecuencias bajo una percepción global del proyecto, ya que, como resultado, en el *TLEAM*, podemos encontrar duplicidad de entradas que reflejan etimologías diferentes, duplicidad de entradas que nada tienen que ver con la etimología, como en Ayala, sino que dan respuesta a un tratamiento organizado por acepciones y, por último, entradas que se articulan en diferentes acepciones. Desde el punto de vista de la consulta de la información, esta situación no ofrece mayor dificultad, pues la solución pasa por considerar la acepción y no el artículo lexicográfico en sí, como unidad de muestra. Desde el punto de vista de la búsqueda simple, al no exigirse ninguna restricción, si se introduce una palabra, se mostrarán todas las acepciones ligadas a un lema coincidente con el valor de entrada, pero, desde el punto de vista de la búsqueda compleja, sí que se permite parametrizar las restricciones, con lo que, al estructurar la búsqueda a nivel de las acepciones y no de los lemas, se garantiza una mayor precisión y exactitud en la selección de resultados que se han de mostrar en la interfaz de volcado de resultados. Como era de esperar, esta decisión



metodológica también conlleva algún efecto no deseable: el falseo del leuario, especialmente desde el punto de vista cuantitativo, aunque también, desde el punto de vista cualitativo, si el usuario decidiera optar por realizar una búsqueda a través del leuario deslizante, opción, en principio, como hemos apuntado, poco operante.

## 2.6 Sobre la anidación de elementos dependientes de un lema ordenante

En relación con el punto anterior, encontramos la imposibilidad de anidar, en el caso de la obra de Ayala, elementos dependientes, especialmente compuestos y locuciones, bajo un lema ordenante. Como sabemos, la lexicografía contemporánea, en particular aquella que se edita en papel, dispone de ciertas reglas estables para incluir los elementos poliléxicos dentro de artículos encabezados por lemas que funcionan como palabras ordenantes. Así, por ejemplo, tanto el compuesto *bola criolla* como la locución *dar bola* [a alguien] se incluirían dentro del artículo encabezado por *bola*, el nombre que funciona como núcleo o palabra ordenante. Esta estructuración, desde una perspectiva digital, es irrelevante, pues la información será recuperable siempre que se introduzca alguno de los elementos que integran la polilexía, con independencia del lugar en el que se almacene, pero como la filosofía de la editorialización que subyace al TLEAM aspira a encontrar un equilibrio entre la obra original, la tradición lexicográfica y la eficiencia en la explotación de la herramienta, se ha optado por consignar, como en la obra original, una entrada independiente a los términos poliléxicos, pero, al mismo tiempo, se han vinculado, mediante una falsa forma normalizada, al elemento que según la tradición lexicográfica tendría que funcionar como palabra ordenante. Esto se puede comprobar si se observa, una vez más, la Imagen 3, en la que se puede apreciar cómo los compuestos *pabas silvestre* o *palo amarillo* se vinculan respectivamente a *pava* y *palo*. Como se señalaba en el apartado anterior, esto produce una multiplicación en la casuística repertoriada, ya que, en las entrañas del TLEAM, conviven elementos poliléxicos que encabezan entradas independientes y otros que se encuentran anidados dentro de los artículos lexicográficos encabezados por los lemas que funcionan como palabras ordenantes, en función de la opción elegida para la elaboración de cada una de las obras originales aquí editorializadas. Sin embargo, nuevamente, esta decisión metodológica parece estar justificada, pues el único efecto pernicioso que produce es de tipo cuantitativo, ya que, como señalábamos en el apartado anterior, se falsean los números referidos al leuario

global, falsedad que bien vale la enorme eficiencia y exactitud en la búsqueda de información que se consigue a cambio.

### 3. Conclusiones

La editorialización de cualquier obra implica priorizar ciertos elementos frente a otros, pero, cuando el proyecto al que se aspira es tan ambicioso como el *TLEAM*, en el que coincide la diversidad de materiales reunidos con la dificultad inherente al abarcamiento de una horquilla temporal tan amplia, puesta en correlación con una tradición lexicográfica en constante evolución, la priorización puede implicar directamente el renunciar a plasmar ciertos aspectos, generalmente relacionados con el respeto de la autenticidad y la fidelidad de la obra original, con el propósito de garantizar una gran eficiencia en el almacenaje, la búsqueda y la forma en la que se muestra la información al usuario. A través del escrutinio de la obra *Diccionario de voces americanas*, atribuida a Ayala, hemos podido comentar los mayores obstáculos a los que nos hemos enfrentado durante la implementación de la estructura en la que se sostiene el *TLEAM*, así como las soluciones que se han propuesto. Evidentemente, la enorme casuística que, casi de manera inherente, caracteriza este proyecto, hace que sea imposible anticipar todas las situaciones posibles, por lo que, en más de una ocasión nos vemos obligados a renunciar a los objetivos y motivaciones que sustentan este proyecto, pero, al mismo tiempo, creemos que una buena planificación, basada en el estudio detallado de los materiales y la anticipación ante eventuales adversidades, de las que en este artículo solo hemos presentado una pequeña muestra, dejando de lado otras cuestiones igualmente medulares en la concepción de cualquier estructura de este tipo (inclusión de los elementos onomásticos, etimologías, desarrollo de abreviaturas, comentarios internos de autores, tratamiento de ejemplos, etc.) son la base que garantizan el éxito de un proyecto de editorialización de tal envergadura.

### Bibliografía citada

Arrabal Rodríguez, Pilar (2022), «TEITOK, a visual solution for XML/TEI encoding: editing, annotating and hosting linguistic corpora», *RIDE – A review journal for digital editions and resources* 15, doi: 10.18716/ride.a.15.5.

- Ayala, José Manuel de (1777), *Diccionario de voces americanas*. Edición de Miguel Ángel Quesada Pacheco (1995). Madrid, Arco Libros.
- Camacho Barreiro, Aurora (2010), «Tesoro lexicográfico de Cuba: un recorrido a través de la historia de las palabras (siglos XIX-XXI)», *Revista digital Librinsula* 1-3 [https://librinsula.bnjm.cu/secciones/300/nombrar/300\\_nombrar\\_2.html](https://librinsula.bnjm.cu/secciones/300/nombrar/300_nombrar_2.html) [15/12/2024].
- Corbella Díaz, Dolores; Díaz Rodríguez, Cristian (2023), “El Tesoro lexicográfico del Español en América: entre repositorio lexicográfico y memoria patrimonial”, *Perspectives de recherche en linguistique et philologie romanes*, eds. Dolores Corbella; Josefa Dorta; Rafael Padrón, Strasbourg, ELiPhi: 689-702.
- Corbella Díaz, Dolores; Fajardo Aguirre, Alejandro; Díaz Rodríguez, Cristian (2024), “TLEAM: Tesoro lexicográfico del español en américa o cómo atesorar el patrimonio léxico en la era digital”, *Lexicografía del español: panhispanismo e internacionalización*, eds. Alejandro Fajardo Aguirre; Dolores Torres Medina; Cristian Díaz Rodríguez, Berlin, Peter Lang: 269-286.
- Corbella Díaz, Dolores; Fajardo Aguirre, Alejandro; Langenbacher-Liebgott, Jutta, eds. (2018), *Historia del léxico español y Humanidades digitales*, Berlin, Peter Lang, Studia Romanica et Lingüística.
- DAMer = Asociación de Academias de la Lengua Española, (2010). *Diccionario de americanismos*. <https://www.asale.org/damer/>
- DLE = Real Academia Española, [versión 23.8 en línea]. <https://dle.rae.es>.
- Rastier, François (2004), «Enjeux épistémologiques de la linguistique de corpus», *Textol*, sección Dits et inédits. [www.revue-texto.net/Inedits/Rastier/Rastier\\_Enjeux.html](http://www.revue-texto.net/Inedits/Rastier/Rastier_Enjeux.html) [01/12/2024].
- Tesoro.pr = Academia Puertorriqueña de La Lengua Española, (2016), *Tesoro lexicográfico del español de Puerto Rico*. <https://tesoro.pr> [24/12/2024]. Es la versión electrónica del texto de Vaquero, María / Morales, Amparo, 2005, *Tesoro lexicográfico del español de Puerto Rico*, San Juan: Academia Puertorriqueña de la Lengua Española.
- TLEAM = Corbella, Dolores, coord. (2021-), *Tesoro lexicográfico del español en América* [en línea]. Disponible en [www.ull.es/tleam](http://www.ull.es/tleam). [23/01/2025].
- Vitali-Rosati, Marcelo (2020), «Pour une théorie de l'éditorialisation», *Humanités numériques* 1. DOI : <https://doi.org/10.4000/revuehn.371>

**Cristian Díaz Rodríguez** es doctor en Estudios Filológicos y Ciencias del Lenguaje. Actualmente es profesor ayudante doctor en el Departamento de Filología Clásica, Francesa, Árabe y Románica de la Universidad de La Laguna. Sus investigaciones se orientan, principalmente, hacia los ámbitos de la fraseología y la paremiología francesa y española, así como de la lexicografía bilingüe. Ha sido coeditor de las obras *Lexicografía del español: intercomunicación y diálogos* y *Lexicografía del español: panhispanismo e internacionalización*.

**cdiarodr@ull.edu.es**



# FRANCESCA DALLE PEZZE ELISA SARTOR EL *VOCABOLARIO ITALIANO E SPAGNOLO* (1620) DE LORENZO FRANCIOSINI: REPRESENTACIÓN DEL DISCURSO REPETIDO PARA SU DIGITALIZACIÓN

Università degli Studi di Verona

## Resumen

La investigación examina los retos que plantea el etiquetado del discurso repetido en el *Vocabolario italiano e spagnolo* de Franciosini (1620) según el estándar TEI Lex0 empleado en el proyecto TELEI (*Tesoro Lexicográfico Español-Italiano*). Se trata de un estudio piloto que permite prever problemas de etiquetado, extensibles posteriormente al conjunto de la obra, que atesora un caudal de expresiones fijas que no es fácil sistematizar, debido a la incoherencia de su representación lexicográfica.

palabras clave: Franciosini, lexicografía bilingüe italiano-español, discurso repetido, codificación XML-TEI, siglo XVII

## Abstract

*Vocabolario italiano e spagnolo (1620) by Lorenzo Franciosini: representation of phrases for its digitization*

*This research examines the challenges posed by the tagging of phrases in Franciosini's Vocabolario italiano e spagnolo (1620) using the TEI Lex0 standard according to the TELEI (Tesoro Lexicográfico Español-Italiano) project protocol. This is a pilot study that allows the detections of tagging problems that could later be applied to the whole dictionary, which contains a wealth of fixed expressions that are not easy to systematize due to the inconsistency of their lexicographic representation.*

*keywords: Franciosini, Spanish-Italian bilingual lexicography, phrases, XML-TEI encoding, 17th Century*

## 1. Introducción

El objetivo de este artículo es ofrecer una contribución al estudio del discurso repetido en el *Vocabolario italiano e spagnolo* de Lorenzo Franciosini, cuya primera edición se publicó en Roma en 1620<sup>1</sup>, enfocada hacia las reflexiones y los retos que su volcado en formato XML-TEI supone desde un punto de vista teórico-metodológico.

En efecto, este trabajo se enmarca en el proyecto PRIN 2022 “Un nuovo ambiente digitale per il recupero del patrimonio lessicografico: il *Tesoro digitale della lessicografia bilingue spagnolo-italiano*”, (TELEI), cuyo objetivo es el de poner a disposición en una plataforma de acceso abierto el patrimonio lexicográfico bilingüe italiano-español en formato digital e hipertextual.

En el presente estudio vamos a analizar los retos que supone el tratamiento informático del *discurso repetido*. Elegimos este hiperónimo, de raíz coseriana, por dos razones principalmente: en primer lugar, porque, ante la asistematicidad de Franciosini, preferimos una etiqueta más abarcadora que diferenciadora; en segundo lugar, porque el lexicógrafo no se limitó a los subtipos que para el siglo XXI son “clásicos”, como refrán, locución, colocación, sino que lematizó otras estructuras, más periféricas, como unidades poliléxicas de la lengua común o de dominios especializados.

El enfoque de este estudio ha sido el de cómo tratar la fraseología en una plataforma digital, acotando los problemas metodológicos que surgen a la hora de volcar un contenido lexicográfico antiguo en un medio digital moderno. Puesto que se trata de un proceso *in fieri*, creemos que el análisis realizado sobre un porcentaje del archivo txt, fruto de la transcripción completa de la obra, no perjudica la validez de los resultados, ya que nuestro estudio piloto nos permite prever precisamente eventuales problemas de etiquetado, extensibles en un segundo momento a todo el conjunto de la obra. Dentro de las diferentes fases de trabajo este estudio piloto representa la ocasión para operar una reflexión global e idear soluciones operativas.

---

\* Este trabajo se ha realizado en el marco del proyecto “Un nuovo ambiente digitale per il recupero del patrimonio lessicografico: il *Tesoro digitale della lessicografia bilingue spagnolo-italiano*”, financiado por la Unión Europea-Next Generation EU, Missione 4 Componente 1 CUP C53D23004010006 / PRIN 2022 MUR 20229W73WR. Aunque el presente ensayo se ha elaborado de forma conjunta, a efectos académicos Francesca Dalle Pezze es responsable de los apartados 1, 2 y 6 y Elisa Sartor de los apartados 3, 4 y 5.

1 El estudio se ha llevado a cabo sobre esta edición, que es la de referencia para nuestra digitalización.

## 2. Estado de la cuestión

A principios del siglo XXI y en una coyuntura de general interés por la labor de Lorenzo Franciosini, varios estudios coincidieron en tratar, entre otros, un aspecto específico —si bien sobresaliente— de su *Vocabolario*: la lematización de las expresiones pluriverbales. Así pues, Martínez Egido y Ruiz Gurillo (2004: 1055) destacaron la sorpresa que nos reserva no solamente descubrir su alto número sino también apreciar su variada riqueza. Ya anteriormente Scandola (2001) había resaltado la especial inclinación del autor hacia el mundo fraseológico, que se refleja en el caudal de expresiones recogidas en sus *Diálogos apazibles*. En esta obra —traducción de los *Pleasant and Delightfull Dialogues in Spanish and English* del lexicógrafo John Minsheu—, Franciosini muestra su comportamiento de hispanista no nativo hacia todos aquellos coloquialismos y construcciones que ni una gramática ni un diccionario podían recoger exhaustivamente, la primera por su planteamiento mismo, el segundo por falta de espacio (Scandola 2001: 82). Asimismo, la riqueza del material fraseológico del *Vocabolario* abrió el camino a la investigación sobre su relación con el universo conversacional contenido en los *Diálogos apazibles*, puesto que ambos tipos de obra persiguen una finalidad parecida, aunque no coincidente. Martínez Egido (2008), entonces, analizó los elementos fraseológicos repetidos en ambas obras y también los añadidos en los *Diálogos*, que se publicaron seis años después del repertorio.

Sin embargo, la falta de sistematicidad del autor en la inserción de unidades pluriverbales fue algo resaltado por todos los estudiosos (Alvar 2002: 59; Acero Durántez 2004: 202; Martínez Egido, Ruiz Gurillo 2004: 1057), quienes señalaron su dependencia de una práctica lexicográfica todavía en desarrollo.

Los criterios de lematización de la fraseología cuentan con una bibliografía extensa, siendo uno de los problemas más llamativos de la técnica lexicográfica (Martínez Marín 1996; Ruiz Gurillo 2000; Porto Dapena 2002: 148-170). Se puede afirmar que el interés por el *Vocabolario* de principios del siglo XXI coincidió con la simultánea propensión de una parte de los estudios lingüísticos por la compleja cuestión de esos criterios. El estudio diacrónico del quehacer fraseográfico se convirtió en uno de los numerosos bancos de pruebas para la teoría general, que se interrogaba sobre cuánta fraseología insertar en un diccionario modélico, qué lugar reservarle (en la macroestructura o en la microestructura) y qué elemento del sintagma escoger para registrar la unidad (Acero Durántez 2004: 199-200).

La cuestión del tratamiento reservado por parte de Franciosini a las expresiones pluriverbales revistió gran importancia porque, empezando por un enfoque simplemente cuantitativo, sirvió para realizar un cálculo global del tipo de entradas que recoge el autor, labor llevada a cabo por Martínez Egido (2002a: 183). Por lo tanto, en la primera parte el estudioso detectó el 9,5 % de sintagmas, como *palla di legno*, y el 6,67 % de oraciones, como *raffrenar le voglie* (2002a: 184), mientras que en la segunda comprobó que el 13,59 % de las entradas está representado por sintagmas, como *llaga reciente*, y el 11,97 % corresponde a oraciones, como *labrar la tierra* (2002a: 206). En el estudio, asimismo, se destacaba el acrecentamiento de estos dos tipos de entrada compleja en el paso de la primera a la segunda parte (2002a: 206), que puede cuantificarse en un 10 %. Descripciones exhaustivas como estas demostraron claramente la diferencia de composición de los dos volúmenes, independientes más que complementarios y con finalidades diferentes. Como señaló Alvar (2002: 53), la primera parte (italiano-español) se asemeja a un diccionario monolingüe italiano con la adición de equivalentes en español. Esta peculiar combinación de lo monolingüe con lo bilingüe demuestra la prontitud del autor en aprovechar los avances culturales contemporáneos, como la reciente publicación del *Tesoro* de Covarrubias (2002: 57).

En Acero Durántez (2004) se estableció una minuciosa taxonomía de todos los tipos de unidades pluriverbales, basada en la clasificación de Corpas Pastor (1996). Según las calas de la autora, el lexicógrafo recogió profusamente unidades pertenecientes a las tres esferas, a saber, colocaciones, locuciones y enunciados fraseológicos. Se detectaron también algunas tendencias en la general falta de sistematicidad: las colocaciones son más frecuentemente entradas autónomas en la macroestructura (Acero Durántez 2004: 204), mientras que los refranes y las fórmulas rutinarias forman parte habitualmente de la microestructura (2004: 207-8). Además, en cuanto a las locuciones, algunas se recogen incluso bajo un lema que no forma parte de la unidad pluriverbal, como *a horcajadas*, tras *horadador* (2004: 205).

Fragapane (2020) ha hecho hincapié en la presencia de elementos extrafraseológicos que constituyen el contorno sintáctico, que se convierte, sin embargo, en información muy valiosa para el usuario a la hora de reconstruir la forma citativa de la unidad misma: se trata de la explicitación de los actantes, mediante pronombres como *uno o algo* (*convenir a uno en juicio*) (170).

Otro asunto importante fue la detección de las (si bien escasas) notaciones metalingüísticas que Franciosini emplea para indicarle al usuario que se encuentra frente a un elemento del discurso repetido. Así pues, Martínez Egido y Ruiz



Gurillo (2004: 1058) señalaron *modo di dire*, que efectivamente se asocia en numerosas ocasiones a locuciones verbales —que son la categoría más nutrida de las recopiladas por el lexicógrafo, según el recuento de Acero Durántez (2004: 205)—, pero también a locuciones adverbiales, a enunciados fraseológicos e incluso a interjecciones. Por lo tanto, el tratamiento de las unidades fraseológicas en la codificación XML del diccionario de Franciosini requiere una definición más bien ‘operativa’ de UF, lo que se pretende hacer en el apartado 3. Finalmente, desde la vertiente de la lexicografía del portugués, llega una fructuosa aportación al estudio de la sistematización de la fraseología en la codificación de un diccionario en XML (Tasovac, Salgado, Costa 2020), que orienta también nuestras reflexiones en el apartado 5.

### 3. Metodología

La conversión en formato txt de los dos volúmenes del diccionario nos ha brindado la oportunidad de realizar búsquedas en el texto por palabras clave. Esto nos ha permitido encontrar no solo las expresiones pluriverbales lematizadas como lemas o sublemas, sino también aquellas que se hallan en la microestructura, así como las marcas lexicográficas que las acompañan. A este respecto, cabe aclarar que, pese a que la transcripción de ambos tomos está completa, la investigación se ha llevado a cabo empleando solo las páginas revisadas de forma definitiva a la fecha de redacción del presente estudio (enero de 2025), con el objetivo de obtener datos más fiables desde el punto de vista textual. De ahí que se haya analizado el 46,7 % de las entradas que conforman el primer volumen y el 35,3 % de las del segundo, con una composición variable: algunas letras se han considerado parcialmente, otras en su totalidad (E, F, H, L, N, O, Q, U, V, X, Z del primer tomo y G, L, N, Ñ, O, Q, X, Z del segundo); finalmente, algunas han quedado excluidas por no estar terminado el proceso de revisión textual, concretamente, D, G, T de la parte italiano-español y F, S, T, U, V de la parte español-italiano.

Un rastreo efectuado a través de las palabras clave “modo di dire” y “proverbio” ha permitido extraer aquellos fraseologismos que el autor consideraba útil señalar a los usuarios de su diccionario; en concreto se han encontrado 80 unidades pluriverbales marcadas con la etiqueta “modo di dire” y 8 con “proverbio”. Citamos como ejemplos, respectivamente,

(1) moliente, y corriente. [ è un modo di dire che significa che uno è apparecchiato, e lesto a far ogni cosa.

(2) quando vieres la barba de tu vezino pelar, echa la tuya in remojo. [ quando la casa del tuo vicino abbrucia, porta l'acqua a casa tua. Inferisce questo proverbio, che quando vediamo intervenir qualche disgrazia a persona della nostra condizione, e sorte, haviamo a pensar che un'altro giorno toccherà anche a noi.

Otras pesquisas realizadas con la palabra clave “suol” han puesto de relieve 18 unidades pluriverbales que muestran una fuerte solidaridad semántica entre sus componentes, encajando con la actual definición de colocación, por ejemplo,

(3) legione, o gran numero, come si suol tal volta dire de'demoni. [ legion de demonios.

(4) manàr, agua o otra cosa liquida. [scaturire, o surgere acqua, o altra cosa liquida, e si suol dire comunemente delle fonti, che scaturiscono acqua.

Desde luego, este primer sondeo ha dejado claro que Franciosini no emplea las notaciones metalingüísticas en modo sistemático, como ya destacaron Martínez Egido y Ruiz Gurillo (2004). Por un lado, parece atribuir las etiquetas “modo di dire” y “proverbio” a las unidades que actualmente consideramos locuciones y enunciados fraseológicos<sup>2</sup>, respectivamente, mientras que la fórmula “suol” por lo general introduce sintagmas caracterizados por solidaridad semántica, que adscribiríamos al ámbito de las colocaciones. Por otro lado, este patrón no se aplica de forma rigurosa —véase, por ejemplo, el enunciado

(5) buen jubon me tengo en Francia. [ è un modo di dire che si usa quando uno ci promette una cosa, che al presente non l'ha, ne ce la puol dare.

y, además, cuantiosos fraseologismos no llevan ningún tipo de marcación que los describa, como

(6) perro ladrador, poco mordedor. [ can che abbaia non morde.

Por lo tanto, se ha decidido construir un corpus piloto<sup>3</sup> para comprobar en qué

2 Nos apoyamos en la clasificación tradicional de Corpas Pastor (1996) por ser bien asentada y aceptada en el contexto de la fraseología actual del español, además de suficientemente abarcadora en relación con las cuestiones metodológicas que nos atañen.

3 En estudios anteriores dedicados al tratamiento de la fraseología en Franciosini, se han analizado corpus compilados a partir de criterios distintos. Por ejemplo, Martínez Egido (2002b) selecciona una muestra de entradas (263) procedentes de la primera parte del diccionario (italiano-español), concretamente de las letras B, L y T. El corpus del trabajo de Ruiz Gurillo y Martínez Egido (2006),

medida Franciosini recoge unidades fraseológicas en la nomenclatura de su diccionario y las emplea en la microestructura de las entradas, con vistas a examinar el tratamiento lexicográfico de las expresiones pluriverbales, incluso las que él no marca como tales. En este trabajo se ha decidido emplear una definición muy amplia de *expresión pluriverbal* para poder abarcar distintas tipologías de discurso repetido, hasta las colocaciones, que representan una categoría resbaladiza, debido a la variabilidad de la solidaridad semántica de los elementos que las componen, especialmente en perspectiva diacrónica. Además, estimamos útil para este estudio tener en cuenta también las unidades poliléxicas nominales con sentido estable y significado tanto (semi)composicional como no composicional, que Montoro del Arco (2008) denomina *formaciones nominales pluriverbales*: esto nos permite considerar los compuestos sintagmáticos de la lengua común así como los que pertenecen al lenguaje especializado (por ejemplo, al dominio de la botánica), que de otra forma quedarían excluidos de nuestro análisis —y posterior tratamiento informático— por no ser fraseologismos *strictu sensu*.

El corpus se ha construido a partir de una selección de somatismos, puesto que, como ya se ha demostrado en numerosos estudios (Navarro 2008, Mellado Blanco 2023, entre otros), estos son especialmente productivos; concretamente, en esta ocasión la búsqueda se ha realizado empleando lexemas relacionados con la cabeza y la cara: *barba*, *ojo*, *diente*, *oreja*, *nariz* y *boca*<sup>4</sup>, sus equivalentes en italiano y las formas del plural en ambos idiomas. La búsqueda ha devuelto 17 expresiones pluriverbales que contienen la palabra *barba(s)/barb\**, 58 con *ojo(s)/occhi(o)*, 15 con *diente(s)/dent\**, 11 con *oreja(s)/orecch\**, 16 con *nariz(es)/nas\** y 31 con *boca(s)/bocc\**, que se estudiarán desde el punto de vista cualitativo en el apartado siguiente.

Para ello, se han consultado dos obras lexicográficas contemporáneas que Franciosini tuvo muy presentes a la hora de componer su diccionario bilingüe, a saber, el *Vocabolario degli Accademici della Crusca* (1612) y el *Tesoro de la lengua castellana o española* de Covarrubias (1611), que fueron una fuente directa para la elaboración de la microestructura del tomo italiano-español y español-italiano, respectivamente (Martínez Egido 2002b y 2003).

---

en cambio, se ha construido a partir de cuatro verbos (*dar*, *echar*, *saber* y *sacar*), con el resultado de estar formado únicamente por locuciones verbales.

4 La creación de un corpus piloto a partir de somatismos derivados de partes móviles del cuerpo humano se ha llevado a cabo recientemente en otro estudio fraseográfico basado en un diccionario del s. XX, a saber, el *Diccionario de uso del español* de María Moliner. Para profundizar en la metodología empleada y los resultados del estudio véase Dal Maso y Sartor (2025).

#### 4. Resultados

El análisis cualitativo del corpus de estudio nos ha permitido mirar desde cerca los procedimientos lexicográficos adoptados por Franciosini, comprobando algunos datos señalados en estudios anteriores y que, sin embargo, no exploraremos por exceder el propósito de este trabajo. Baste con citar la marcación lexicográfica, aspecto ya estudiado por Acero Durántez (2002), de la que se encuentra un ejemplo múltiple en la entrada

(7) hallar horma de su çapato. [ per parlar chièremente, e secondo l'uso, e a proprietà di questo, si dice correntemente, se bene con poca honesta, trovar naso da suo'culo, cioè trovar persona che ti sappia rispondere, e ti mostri il viso.

Franciosini, como se puede ver, proporciona información sobre la frecuencia de uso (*si dice correntemente*) y a la vez advierte sobre la malsonancia del fraseologismo italiano (*con poca honesta*).

Desde el punto de vista léxico, merece la pena observar las estrategias empleadas para establecer equivalencias semánticas entre los dos sistemas lingüísticos lo más adecuadas posible, básicamente a través de dos procedimientos, a saber: una expresión pluriverbal equivalente en la lengua meta, aunque formalmente distinta, o bien una explicación del significado de la unidad. Estas dos técnicas a veces se emplean de forma híbrida, proporcionando un equivalente y también una glosa, especialmente cuando la lengua de redacción es el italiano, tanto si es la lengua fuente como si es la lengua meta. En los ejemplos siguientes se muestran casos que caben en la primera tipología (8-12), en la segunda (13-14), o que presentan una mezcla de las dos (15-18):

(8) ojos que no ven, coraçon no quebrantan. [ quello che gl'occhi non veggano, il cuor non crede. (*s.v. ojo*)

(9) batir los dientes. [ battere i denti. (*s.v. batir los dientes*)

(10) boca de leon. [ leonina herba. (*s.v. leon*)

(11) nuve del ojo. [ maglia dell'occhio. (*s.v. nuve del ojo*)

(12) lienço de narizes. [ fazzoletto. pezzuola. (*s.v. lienço*)

(13) a poca barba, poca verguença. [ inferisce che i giovani con la sua poca prudenza soglion esser talvolta molto sfacciati, e persuntuosi. (*s.v. barba ES*)

(14) muelas de gallo. [ si dice a colui che non hà nissun dente. (*s.v. gallo ES*)

(15) levarsi le mosche d'intorno al naso, vale non si lasciar burlare ne far ingiuria. [ no

sufrir cosquillas. (*s.v. mosca* IT)

(16) a bulto. [ a occhio, senza pesare o misurare. (*s.v. a bulto*)

(17) diente de ajo. [ spicchio d'aglio, cioè una di quelle particelle in che si divide. (*s.v. cabeça de ajo*)

(18) morselletto, boccone di pasta di varie sorte d'ingredienti, che si fanno da speciali. [ tableta, o pastilla de boca. (*s.v. morselletto*).

Desde un punto de vista taxonómico actual, las unidades pluriverbales citadas pertenecen a distintas categorías, a pesar de que, en ninguno de los ejemplos mencionados, Franciosini emplee etiquetas que lo señalen: se podrían considerar enunciados fraseológicos los ejemplos (8) y (13), locuciones (14), (15) y (16) —respectivamente nominal, verbal y adverbial—, colocaciones (9) y (17), compuestos *lienço de narizes* y *pastilla de boca* y, finalmente, formaciones poliléxicas propias del lenguaje de especialidad *boca de leon* (del dominio de la botánica) y *nuve del ojo* (término de la medicina).

Otra cuestión que merece la pena destacar es la posición de la unidad pluriverbal en la entrada o subentrada. Se han encontrado casos en los que la expresión pluriverbal es el lema mismo (véase *infra*, 19-20) o constituye un sinónimo del lema (*a occhio e croce*, *orejas de Abad*, 21 y 22, respectivamente); en otros casos se halla en la microestructura y puede ser el equivalente del lema (23 y 24, donde se ofrecen hasta tres variantes fraseológicas: *guardar con mal occhio uno*, *non gli far buon viso*, [*non gli far*] *buona cera*) o bien un sinónimo del equivalente (*digrignare i denti* y *guarda la gamba*, 25 y 26, respectivamente):

(19) vale un ojo de la cara. [ modo d'esaggerare il prezzo d'una cosa, vale un'occhio. (*s.v. ojo*)

(20) dexelo, que no le hiede la boca. [ modo di dire che usano le bottegaie quando non gli si dà per la lor mercanzia quanto domandano. (*s.v. heder*)

(21) a vista, avverbialmente giudicare a vista, cioè con la semplice vista, senza venire ad altro cimento, che diremmo anche giudicare a occhio, e croce. [ a ojo. a bulto. (*s.v. a vista* IT)

(22) lasañas, o orejas de Abad. [ certa pasta frita nella padella in forma di frittelle, ò crespelli. (*s.v. lasañas*)

(23) mangiar a mal in corpo, cioè mal volentieri, come tal volta fanno gli ammalati. [ comer a regaña dientes. (*s.v. mangiar a mal in corpo*)

(24) estrañarse. [ si dice il guardar con mal occhio uno, e non gli far buon viso, o buona cera. (*s.v. estrañarse*)

(25) chillar. [ stridere, e far rumore come quando si frigge, o si strofinano i denti, l'uno con l'altro, che diciamo propriamente sdirugginare i denti, o digrignare i denti. (*s.v. chillar*)

(26) abre el ojo. [ maniera di dire per avvertire. guarda la gamba. (*s.v. ojo*).

Es oportuno señalar, sin embargo, que se dan casos ambiguos en los que resulta muy difícil averiguar en diacronía si un determinado sintagma constituía una forma de discurso repetido en la época de Franciosini, como en el ejemplo siguiente:

(27) A, in vece d'inguisa, ò come. Hà i denti fatti à bischeri. [ Tiene los dientes como clavijas, o a manera, o a semejança de clavijas. (*s.v. a IT*).

Concretamente, *denti fatti a bischeri* se recoge como ejemplo en una entrada del diccionario de los académicos de la Crusca (1612). Se trata de una cita de una *novella* de Boccaccio, que representó uno de los modelos, junto con Petrarca, para la codificación de la lengua italiana culta en las etapas humanista y renacentista: “E mi par vederti morderle, con cotesti tuoi denti fatti a bíscheri, quella sua bocca vermigliuzza, ec.” (*s.v. bischero*). Esta misma cita se mantiene hasta la cuarta edición del *Vocabolario* (1729-1738) y, en forma abreviada, se emplea también para ejemplificar el lema *mordere* de la primera a la cuarta edición. Debido a la falta de documentación de este sintagma en sincronía, podemos suponer que se usó como ejemplo en las obras lexicográficas del s. XVII y posteriores<sup>5</sup> por el prestigio literario de su autor, sin pasar a convertirse en un elemento de discurso repetido.

Otro caso límite es la ocurrencia, generalmente en la microestructura, de enunciados que tienen una estructura que trae a la memoria los dichos proverbiales o las sentencias y que, sin embargo, no se documentan en otras fuentes consultadas<sup>6</sup>:

(28) A uno, che parlava molto, e dava poco, disse una gentil donna il miglior huomo del mondo saresti, se le serrature che havete nella borsa, l'havessi nella bocca. [ A uno, che hablava mucho, y dava poco, le dixo una Señora. El mejor hombre del mundo seriades, si los cerraderos que teneis en la bolsa, tuviessedes en la boca. (*s.v. borsa*).

5 Señalamos que la misma cita, en una versión más extensa y con la marca de uso *per similitudine*, se recoge también en el *Grande dizionario della lingua italiana* de Salvatore Battaglia (1961-2004), *s.v. bischero*.

6 Como es sabido (Alvar 2002: 43), Franciosini conocía y se sirvió de las breves historietas registradas en la *Floresta española* de Melchor de Santa Cruz (1.<sup>a</sup> edición, 1574), como la siguiente: “A Vno que tenia vna cuchillada por la cara, y hablaua demasiadamente, dixo vno, Pues teneys rienda en la cara, tened freno en la boca” (Santa Cruz, 1576: 238v).

Desde luego, no se considera que sean formas de discurso repetido sino más bien glosas o breves anécdotas con las que Franciosini pretende ilustrar ejemplos de uso en las (sub)entradas.

A este propósito, finalmente, Fragapane (2020: 175) subraya el hecho de que la finalidad didáctica, que prima en la obra sobre las demás, acaba convirtiendo el límite entre fraseología y ejemplos de uso bastante borroso, ya que varios casos se ubican en una vía intermedia entre lo fraseológico y lo no fraseológico. Para ilustrar este fenómeno la autora menciona los ejemplos de *esto limpio es, bien se puede comer sin asco* y *yo no puedo comer de asco*, bloques que, más que verdadero discurso repetido (ya fijado e institucionalizado), son fórmulas que el usuario podría proyectar en su propio acto comunicativo si la situación lo permitiera.

## 5. Etiquetado con XLM-TEI Lex0

La pregunta de investigación que nos planteamos, no solo teórica sino práctica también por las implicaciones a la hora de volcar las entradas y subentradas en la plataforma XLM-TEI, atañe al tipo de información sobre las unidades pluriverbales que el usuario querrá o podrá obtener cuando consulte el diccionario de Franciosini en formato digital. El estándar Lex0 empleado en el proyecto TELEI (*Tesoro Lexicográfico Español-Italiano*) nos brinda la oportunidad de etiquetar las unidades pluriverbales con *multiword expression* (*mwe*), lo que va a permitir recuperar la información a la hora de realizar consultas y de emplear la etiqueta *mwe* como filtro.

En el lenguaje de codificación empleado en el proyecto TELEI, las líneas de código que sirven para identificar una unidad pluriverbal en la macroestructura seguirán el modelo siguiente, basado en el ejemplo *nuve del ojo*:

```
<entry type="mainEntry" xml:id="Francio1620.nuve del ojo." xml:lang="es">
  <form type="lemma">
    <orth>nuve del ojo</orth></form>
    <gramGrp><gram type="mwe"/>
```

En concreto, la primera línea sirve para atribuir un *id* (identificativo) al lema y clasificar la expresión pluriverbal como perteneciente a la primera edición del diccionario de Franciosini, especificando la lengua del lema, en este caso el español<sup>7</sup>;

7 No hay que olvidar que la plataforma TELEI almacenará en formato XML todas las obras

la segunda línea caracteriza la unidad como lema del diccionario, mientras que la etiqueta <orth> es la que permite ingresar manualmente la expresión pluriverbal (en nuestro ejemplo, *nuve del ojo*); finalmente, la última línea permite la caracterización gramatical del lema como *multiword expression*.

Ahora bien, a fin de que el usuario pueda aprovechar plenamente los avances en la consulta que la digitalización supone, es necesario que todas las unidades pluriverbales sean correctamente etiquetadas para ser luego recuperadas a través del filtro *mwe*. Desde el punto de vista operativo, esto conlleva marcar con la etiqueta *mwe* todos los fraseologismos que el mismo Franciosini distinguió a través de las etiquetas “proverbio”, “modo di dire”, “si suol dire”, etc., tanto los que siguen vigentes (por ejemplo, *hasta los ojos*) como los que están desusados en la época actual (*no me dixo negro has el ojo*). Además, es preciso etiquetar también los demás elementos que se reconozcan como formas de discurso repetido en diacronía y sincronía (por ejemplo, la variante ortográfica *a regaña dientes*, hoy en desuso), incluso aquellos que hayan quedado obsoletos (*hombre de barba*), considerando también las unidades pluriverbales de uso general (*ojo del aguja*) o propias del lenguaje de especialidad (*oreja de raton*) que caben bajo el marbete de expresión pluriverbal, esto es, *multiword expression*. Obviamente, los ejemplos (27) y (28), discutidos en el apartado anterior, quedarían excluidos por no ser formas de discurso repetido; concretamente, el caso (28) se etiquetaría como *gloss*, o sea, glosa.

Coherentemente con los criterios compartidos que rigen la creación de la plataforma TELEI, las expresiones pluriverbales que se clasifiquen como *multiword expression* se etiquetarán como *entry* (lema) si aparecen en la macroestructura; si conforman una subentrada, en cambio, se etiquetarán como una *entry* anidada en la etiqueta <sense>, que corresponde al sentido o acepción. Las ocurrencias en la microestructura, a su vez, se clasificarán como *multiword expression* dentro del *sense* de la entrada correspondiente. Esta jerarquización de la información se debe, por un lado, a las características lexicográficas del diccionario mismo, que registra expresiones pluriverbales tanto en la macro- como en la microestructura, según se ha comentado en el apartado anterior, y, por otro, a la estructura del estándar Lex0.

De acuerdo con Tasovac, Salgado y Costa (2020: 51), “TEI Lex0 is fully capable of consistently marking up polylexical units as constituent parts of the

lexicográficas digitalizadas en el marco del proyecto “Un nuovo ambiente digitale per il recupero del patrimonio lessicografico: il *Tesoro digitale della lessicografia bilingue spagnolo-italiano*”, por lo tanto, es imprescindible que cada elemento esté vinculado en forma unívoca con el diccionario correspondiente.



dictionary macro- and microstructure, regardless of whether they appear as head-words in independent entries, or in nested entry-like structures inside entries for monolexical units”. Desde luego, a diferencia del estudio de caso presentado por estos autores, que trabajan en la codificación de un diccionario actual de la lengua portuguesa, para la codificación de Franciosini se ha decidido desestimar algunas de las funciones de Lex0 que permite, entre otras cosas, marcar el núcleo de la expresión poliléxica o señalar los compuestos con guion (*compounds*), atendiendo a los rasgos propios de una obra lexicográfica del s. XVII.

## 6. Conclusiones

El proyecto PRIN 2022 “Un nuovo ambiente digitale per il recupero del patrimonio lessicografico: il *Tesoro digitale della lessicografia bilingue spagnolo-italiano*” (TELEI) y, más específicamente, la digitalización del *Vocabolario italiano e spagnolo* de Lorenzo Franciosini —obra cumbre de la producción de diccionarios español-italiano durante el siglo XVII— nos pone ante el reto de un adecuado tratamiento del discurso repetido.

El *Vocabolario italiano e spagnolo* de Lorenzo Franciosini sigue brindando hoy en día ocasiones de interés, no solamente cultural (por el rico y abigarrado caudal de informaciones que ofrece), sino también para la lexicografía informatizada y la filología actuales. El rastreo de los elementos que forman parte del discurso repetido, tanto si es prototípico como periférico, nos depara numerosos desafíos a la hora de posibilitar una búsqueda provechosa por parte del usuario de la plataforma digital e hipertextual.

Como hemos intentado demostrar, la reflexión teórico-metodológica que hemos llevado a cabo tendrá un impacto en la codificación; sin embargo, al ser este un trabajo en progreso, los posibles problemas y desajustes deberán ser tomados en consideración como retroalimentación para mejorar el proceso. Por lo tanto, sigue abierta una vía para futuras investigaciones y mejoras.

## Bibliografía citada

- Accademici della Crusca (1612), *Vocabolario degli Accademici della Crusca*, Venezia, Giovanni Alberti [15/01/2025] <<http://www.lessicografia.it/>>.
- Acero Durántez, Isabel (2004), “El tratamiento de la fraseología española e italiana en un diccionario bilingüe del siglo XVII”, *Nuevas aportaciones a la historiografía lingüística. Actas del IV Congreso Internacional de la SEHL. La Laguna (Tenerife), 22 al 25 de octubre de 2003*, eds. Cristóbal José Corrales Zumbado; Josefa Dorta Luis; Antonia Nelsi Torres González; Dolores Corbella Díaz; Francisca del Mar Plaza Picón. Madrid, Arco/Libros: I, 199-210.
- Alvar Ezquerro, Manuel (2002), “El *Vocabolario italiano e spagnolo* de Lorenzo Franciosini”, *Lengua y diccionarios. Estudios ofrecidos a Manuel Seco*, eds. Pedro Álvarez de Miranda; José Polo Polo. Madrid, Arco/Libros: 37-61.
- Battaglia, Salvatore (1961-2004), *Grande dizionario della lingua italiana*, Torino, UTET [15/01/2025] <http://www.gdli.it>.
- Covarrubias, Sebastián de (1611), *Tesoro de la lengua castellana o española*, Madrid, Luis Sánchez [15/01/2025] <https://apps2.rae.es/ntlle/SrvltGUILoginNtlle>.
- Dal Maso, Elena; Sartor, Elisa (2025), “Fraseologismos somáticos en el DUE de María Moliner: un estudio desde la perspectiva de los aprendices de ELE”, *Revista Internacional de Lenguas Extranjeras*, 24: 1-23 [13/06/2025] <https://www.revistes.urv.cat/index.php/rile/article/view/3973/4716>.
- Fragapane, Federica (2020), “El tratamiento de las unidades fraseológicas en el diccionario de Franciosini”, *Phrasis. Rivista di studi fraseologici e paremiologici*: 167-178 [24/01/2025] <https://www.phrasis.it/rivista/index.php/rp/article/view/64>.
- Franciosini, Lorenzo (1620), *Vocabolario italiano e spagnolo*, 1.<sup>a</sup> edición, Roma: Gio. Paolo Profilio, 2 vols.
- Martínez Egido, José Joaquín (2002a), *La obra lexicográfica de Lorenzo Franciosini: Vocabulario italiano-español, español-italiano (1620)*, Alicante, Biblioteca Virtual Miguel de Cervantes [21/01/2025] <https://www.cervantesvirtual.com/obra/la-obra-lexicografica-de-lorenzo-franciosini-vocabulario-italianoespanol-espanolitaliano-1620-0/>.
- Martínez Egido, José Joaquín (2002b), “Las fuentes de la microestructura en la primera parte del *Vocabulario italiano-español, español-italiano* (1620) de Lorenzo Franciosini”, *Estudios de lingüística*, 16: 5-35.
- Martínez Egido, José Joaquín (2003), “La utilización de fuentes en la realización de diccionarios: ¿Covarrubias (1611), fuente directa en la microestructura de Franciosini (1620)?”, *ELUA*, 17: 413-26.
- Martínez Egido, José Joaquín; Ruiz Gurillo, Leonor (2004), “La integración de la fraseología en los primeros diccionarios del español: el *Vocabulario* (1620) de Lorenzo Franciosini”, *Nuevas aportaciones a la historiografía lingüística. Actas del IV Congreso Internacional de la SEHL. La Laguna (Tenerife), 22 al 25 de octubre de 2003*, eds.

- Cristóbal José Corrales Zumbado; Josefa Dorta Luis; Antonia Nelsi Torres González; Dolores Corbella Díaz; Francisca del Mar Plaza Picón. Madrid, Arco/Libros: II, 1055-1066.
- Martínez Egidio, José Joaquín; Ruiz Gurillo, Leonor (2006), “Las unidades fraseológicas en el inicio de la lexicografía del español (1495-1620)”, *Actas del VI Congreso Internacional de Historia de la Lengua*, eds. José Luis Girón Alconchel; José Jesús de Bustos Tovar. Madrid, Arco/ Libros: 1-13 [21/01/2025] [https://rua.ua.es/dspace/bitstream/10045/24048/1/2006\\_Leonor\\_Ruiz\\_Congreso\\_Historia\\_Lengua.pdf](https://rua.ua.es/dspace/bitstream/10045/24048/1/2006_Leonor_Ruiz_Congreso_Historia_Lengua.pdf).
- Martínez Egidio, José Joaquín (2008), “El español conversacional y su reflejo en los diccionarios de Oudin y Franciosini”, *El diccionario como puente entre las lenguas y culturas del mundo. Actas del II Congreso Internacional de Lexicografía Hispánica*, (eds.) Dolores Azorín Fernández; Belén Alvarado Ortega; Jaume Climent de Benito; M. Isabel Guardiola i Savall; Ruth M. Lavale Ortiz; Carmen Marimón Llorca; José Joaquín Martínez Egidio; Xose A. Padilla García; Herminia Provencio Garrigós; Isabel Santamaría-Pérez; Larissa Timofeeva Timofeev; Elena Toro Lillo. Alicante, Biblioteca Virtual Miguel de Cervantes: 144-149 [21/01/2025] <https://www.cervantesvirtual.com/obra/el-espaol-conversacional-y-su-reflejo-en-los-diccionarios-de-oudin-y-franciosini-0/>.
- Martínez Marín, Juan (1996), *Estudios de fraseología española*, Málaga, Librería Ágora.
- Mellado Blanco, Carmen (2023), “Los somatismos y su transparencia semántica: los rasgos tipológicos”, *EntreLíneas*, 6: 1-23 [17/01/2025] [https://entrelineasrevista.wordpress.com/wp-content/uploads/2023/03/monografico\\_entrelineas.pdf](https://entrelineasrevista.wordpress.com/wp-content/uploads/2023/03/monografico_entrelineas.pdf).
- Montoro del Arco, Esteban T. (2008), “Relaciones entre Morfología y Fraseología: las formaciones nominales pluriverbales”, *Neologismo y morfología*, eds. Ramón Almela Pérez; Esteban Tomás Montoro del Arco. Murcia, Universidad de Murcia: 121-146.
- Navarro, Carmen (2008), *Aspectos de fraseología contrastiva*, Verona, Fiorini Editore.
- Porto Dapena, José-Álvaro (2002), *Manual de técnica lexicográfica*, Madrid, Arco/Libros.
- Ruiz Gurillo, Leonor (2000), “Cómo integrar la fraseología en los diccionarios monolingües”, *Las lenguas de Europa: estudios de fraseología, fraseografía y traducción*, (coord.) Gloria Corpas Pastor. Granada, Comares: 261-274.
- Santa Cruz, Melchor de (1576), *Floresta española de apothegmas o Sentencias*, Alcalá, Juan Íñiguez de Lequerica.
- Scandola, Viviana (2001), “Lorenzo Franciosini: un interés declarado por la fraseología”, *Paremia*, 10: 79-84 [16/01/2025] [https://cvc.cervantes.es/lengua/paremia/pdf/010/009\\_scandola.pdf](https://cvc.cervantes.es/lengua/paremia/pdf/010/009_scandola.pdf).
- Tasovac, Toma; Salgado, Ana; Costa, Rute (2020), “Encoding polylexical units with TEI Lex-0: A case study”, *Slovenščina 2.0*, 8(2): 28-57.
- TEI (Text Encoding Initiative) [22/01/2025] <https://tei-c.org>.
- TEI Lex-0 [22/01/2025] <https://dariah-eric.github.io/lexicalresources/pages/TEILex0/TEILex0.html>.

**Francesca Dalle Pezze** es profesora titular de Lengua, Traducción y Lingüística Españolas en la Universidad de Verona, donde participa en el grupo de investigación *El léxico del español en su historia (LEHist)* y en el PRIN 2022 “Un nuovo ambiente digitale per il recupero del patrimonio lessicografico: il *Tesoro digitale della lessicografia bilingue spagnolo-italiano*”. Sus principales líneas de investigación se centran en la lexicografía, la lexicología y la historia de la lengua española.

**francesca.dallepezze@univr.it**

**Elisa Sartor** es profesora titular de Lengua, Traducción y Lingüística Españolas en la Universidad de Verona, donde participa en el grupo de investigación *El léxico del español en su historia (LEHist)* y en el PRIN 2022 “Un nuovo ambiente digitale per il recupero del patrimonio lessicografico: il *Tesoro digitale della lessicografia bilingue spagnolo-italiano*”, entre otros proyectos. Sus líneas de investigación se centran en el estudio histórico del léxico y la lingüística de corpus.

**elisa.sartor@univr.it**

# ROSA MARÍA GARCÍA JIMÉNEZ MARÍA ENRIQUETA PÉREZ VÁZQUEZ LA RETRODIGITALIZACIÓN DEL *VOCABULARIO ITALIANO-ESPAÑOL* DEL *DICCIONARIO MARÍTIMO* *ESPAÑOL* DE O'SCANLAN

Università di Pisa

## Resumen

Este trabajo ofrece la descripción y el análisis inicial de uno de los textos del proyecto TELEI: el *Vocabulario italiano-español* del *Diccionario marítimo español* (1831) de T. O'Scanlan. Presentamos las principales características del *Diccionario* y del *Vocabulario*, como estudio previo para su estructuración en formato interoperable dentro de la base de datos del proyecto, y mostramos los criterios adoptados para la anotación del repertorio mediante TEI Lex-O.

palabras clave: lexicografía bilingüe, patrimonio digital, retrodigitalización, lexicografía naval, lenguaje especializado

## Abstract

**The retrodigitalization of the *Vocabulario italiano-español* of O'Scanlan's *Diccionario marítimo español***  
*This paper offers the description and initial analysis of one of the texts of the TELEI project: the *Vocabulario italiano-español* of the *Diccionario marítimo español* (1831) by T. O'Scanlan. We present the main characteristics of the *Diccionario* and the *Vocabulario*, as a preliminary study for its structuring in interoperable format within the project's database, and describe the criteria adopted for the annotation of the repertoire by TEI Lex-O.*

*keywords: bilingual lexicography, digital heritage, retrodigitization, nautical lexicography, specialised language*

## 0. Introducción

Este trabajo se ha realizado en el marco del proyecto “Un nuovo ambiente digitale per il recupero del patrimonio lessicografico: il “Tesoro digitale della lessicografia bilingue spagnolo-italiano” (TELEI), financiado por La Unión Europea- Next Generation EU, Missione 4 Componente 1 CUP C53D23004010006 / PRIN 2022 MUR 20229W73WR, dirigido por C. Castillo Peña. La finalidad principal del proyecto es la preservación del patrimonio lexicográfico bilingüe mediante ediciones críticas y facsimilares digitales<sup>1</sup>. Siguiendo este objetivo, el equipo de investigación de la Universidad de Pisa se ha dedicado a la transcripción, el estudio y la aplicación del estándar TEI Lex-0 a obras lexicográficas bilingües de ámbito especializado.

El propósito del presente estudio es mostrar el proceso que hemos llevado a cabo para la retrodigitalización del *Vocabulario marítimo italiano-español*, uno de los tres repertorios bilingües que ofrece el *Diccionario marítimo español*<sup>2</sup> (Madrid, Imprenta Real, 1831), realizado por Timoteo O'Scanlan (1774-1831), por encargo del historiador naval Martín Fernández de Navarrete (1765-1844).

La organización de este estudio es la siguiente: en el primer apartado presentamos el contexto histórico y los orígenes del *Diccionario marítimo español*; a continuación, se ilustran las características de las entradas del *Vocabulario* en italiano y sus equivalentes en español. En el tercer apartado, se exponen las diferentes fases del proceso de codificación del repertorio en un formato digital estructurado. El trabajo se concluye con unas primeras reflexiones en §4.

## 1. El *Diccionario marítimo español*. Orígenes

El *Diccionario marítimo español* puede considerarse el primero de este tipo en lengua española “realizado por un equipo según principios lexicográficos modernos y, asimismo, el primero que combina las fuentes escritas y el conocimiento

---

\* La redacción del presente artículo y la investigación en que se basa, en todas sus fases y partes, son fruto de la estrecha colaboración entre ambas autoras. Solo a efectos académicos, la autoría se disgrega como sigue: García Jiménez es responsable de 2.2 y 3, y Pérez Vázquez de 1 y 2.1. La introducción y las conclusiones son comunes.

1 El grupo de investigación de la Universidad de Pisa, que ha trabajado conjuntamente en el proceso de retrodigitalización, está formado por la profesora Elena Carpi y las autoras de este artículo.

2 En el original, sin tilde.

empírico del uso léxico de manera metódica” (Carriazo Ruiz, 2008). La moderidad de su estructura y de su planteamiento teórico se manifiesta en el empleo de marcas diatécnicas (arquitectura, naval, astronomía, navegación, “pilotage”, táctica naval, etc.), diacrónicas (anticuado), gramaticales (adjetivo, femenino, imperativo, infinitivo, irregular, etc.), etimológicas (arábigo, inglés, italiano, etc.), diafásicas (familiar, inusitado, poesía, etc.) y semánticas (figurado, metafórico, etc.). Y, por otra parte, (por influjo del *Diccionario de Autoridades*, 1726-1739 de la Real Academia) incluye citas de autoridades como aval de las definiciones y para ejemplificar el lema<sup>3</sup>, lo que supone que está confeccionado a partir de un corpus. El “alto sentido filológico del autor” se manifiesta además en la reflexión sobre las etimologías de los términos y sobre la variedad regional de muchos de ellos (Ahumada 2008).

El principio oficial de la historia del diccionario puede situarse cuando Fernando VII, por Real Orden del 16 de mayo de 1827, encarga a la empresa de realizar un diccionario marítimo a Martín Fernández de Navarrete<sup>4</sup>, entonces director del Depósito Hidrográfico, consejero de guerra jubilado. Este parece que, en un primer momento, pensó para su redacción en el teniente de navío Miguel Roldán, que ya había recopilado algunos textos, pero, meses más tarde, muy oportunamente, el 23 de junio, el Capitán de fragata Timoteo O’Scanlan<sup>5</sup> manda una

3 El listado de las mismas se encuentra en las dos últimas páginas del prólogo («autoridades que se citan en el mismo diccionario ú obras de donde se han tomado algunas noticias»). La inclusión de vocablos arcaicos en los diccionarios decimonónicos del español responde a una larga tradición que arranca, al menos, de la fundación de la Academia, cuando se sentaron las bases del diccionario académico y se publicó su planta, con una cuidadosa consideración por lo histórico (Jiménez, 2001: 21-46).

4 Martín Fernández de Navarrete y Ximénez de Tejada (1765-1844) ingresó en la Marina en 1780, estudió astronomía, navegación e ingeniería naval. En uno de sus destinos como marino, le fue encomendada la tarea de crear una biblioteca en la Real Isla de León y de elaborar un catálogo de obras para la biblioteca de los Reales Estudios de San Isidro (Madrid). Adquirió, entonces, una gran destreza como historiador y documentalista. Ocupó varios cargos políticos. Fue miembro de numerosas instituciones culturales, en las cuales llevó a cabo una intensa labor y se convirtió en un referente cultural en la vida intelectual española. Fue académico de la RAE, de la Real Academia de Bellas Artes de San Fernando, además de secretario y académico y director de la Real Academia de la Historia. Perteneció, también, a diferentes sociedades nacionales e internacionales. Como miembro de la Real Academia Española colaboró en la elaboración de la *Gramática*, en la *Ortografía de la Lengua Castellana*, en el *Silabario teórico-práctico*, y en las diferentes revisiones del *Diccionario*, estuvo destinado a la Biblioteca del Monasterio del Escorial con el encargo de copiar a mano varios diccionarios etimológicos antiguos que se conservaban allí. Véase Palacio, Cervera, Seco, Higuera y Anes (1995).

5 Timoteo O’Scanlan de Lacy nació en El Ferrol en 1774, en una conocida familia de origen irlandés, pues su padre era un conocido médico militar. Al acabar sus estudios, primero en Tolosa

carta en la que pide permiso al rey para publicar un *Vocabulario trilingüe. Pronuario de los términos de Marina en los tres idiomas Inglés, Francés e Italiano, con sus correspondientes en castellano, para el uso de los Marinos españoles, con un apéndice en Portugués*, fruto del trabajo de varios años y de la recopilación de numerosa bibliografía (hoy en el Museo Naval de Madrid). Se encomienda entonces oficialmente a O'Scanlan el perfeccionamiento y ampliación de su obra para lo que se le llama a Madrid, bajo la dirección de Navarrete, poniendo a su disposición un equipo de colaboradores y toda una recopilación de material metalexicográfico<sup>6</sup>.

En 1830 había ya varias pruebas de imprenta y parte del diccionario impreso (se conserva la correspondencia de Navarrete con recomendaciones y consejos sobre las primeras impresiones de la obra), lo que explica que ya en la primera edición existan unas páginas de la obra dedicadas a adiciones y rectificaciones del *Diccionario* y de los vocabularios bilingües<sup>7</sup>. La confección del diccionario, a partir del primer borrador de O'Scanlan, incluyó varias revisiones, entre ellas la del Capitán de fragata don Manuel del Castillo, encargado de «añadir otras [voces] antiguas que, siendo las más castizas y puras, son igualmente la raíz y fundamento de muchas actuales, y que usaron nuestros españoles cuando reducían a sistema el arte de navegar escribiendo los primeros tratados que d'él se conocieron», como explica el mismo Fernández de Navarrete (1831: 25-26) en el prólogo de la obra.

### 1.1 *La macroestructura*

Los preliminares o paratextos del *Diccionario marítimo español* constan de un prólogo de 43 páginas, escrito por su ideador y director, Martín Fernández de Navarrete, una página con las abreviaturas y una lista de las autoridades que se citan, con numeración romana. A continuación, se encuentra el leuario o nomenclatura monolingüe en español y, en algunas de sus entradas, tras la parte informativa del artículo, se incluye una cita y la palabra equivalente en francés,

---

y luego en Madrid regresó a El Ferrol, donde ingresó en el regimiento irlandés Hibernia, ayudante ingeniero de Marina, después alférez de navío, habiendo alcanzado el grado de capitán de fragata en el puerto de Salou, se traslada a Madrid para redactar unas nuevas Ordenanzas de Arsenales, además de recibir en 1827, por parte de Martín Fernández de Navarrete, el encargo de la redacción del *Diccionario Marítimo* que se publica pocas semanas antes de su muerte en 1831. Véase Guillén (1967).

6 En la Biblioteca de la RAE se conserva el volumen con los vocabularios y glosarios, con fecha de 1991 (Nieto 2002: 25).

7 Parte de la correspondencia entre los dos se puede consultar en Guillén (1967).



inglés e italiano, lo que ocupa 568 páginas, con un suplemento que sigue la misma numeración. A ello se añaden tres vocabularios bilingües, uno en francés, otro en inglés y, el último, que es el que nos ocupa en este estudio, en italiano. Finaliza la obra con 18 páginas de paratextos de rectificaciones y adiciones.

### 1.1.2 El prólogo de Martín Fernández de Navarrete

El prólogo de esta obra, de Martín Fernández de Navarrete, escrito en tercera persona y sin firmar, fue publicado de forma independiente el mismo año que el *Diccionario marítimo español*, en 1831, con el título *Discurso que sobre la utilidad de los diccionarios facultativos, con un examen de los que se han escrito de marina, y con las advertencias convenientes para formarlos y corregirlos en lo sucesivo escribió don Martín Fernández Navarrete para servir de prólogo al Diccionario marítimo español publicado en 1831*. A lo largo de las 47 páginas que ocupa el prólogo, Navarrete demuestra una visión muy moderna de la lengua, que relaciona con el nivel cultural de una nación, para lo que defiende la necesidad de buenos diccionarios, importantes para reflejar el léxico de los desarrollos tecnológicos a través de neologismos, que defiende solo cuando no existe el correspondiente castellano, demostrando así una actitud prescriptiva y purista. Navarrete justifica la inclusión de las palabras desusadas en un vocabulario técnico porque pueden ser útiles para la historia de la Marina o «de curiosidad para los eruditos anticuarios o para entender a Ercilla, Lope de Vega y otros clásicos escritores castellanos que usaron también muchas voces náuticas para describir una navegación y para entender la doctrina de nuestros antiguos maestros o apreciar las relaciones de nuestros intrépidos navegantes y primitivos descubridores» y para frenar neologismos innecesarios. Se lamenta de que todavía no existiera un buen diccionario de marina donde apareciesen todos los campos semánticos relacionados con lo marino, lo naval y la navegación e incluso con términos especializados de la agricultura, como los relacionados con el cultivo del lino para fabricar las velas, por ejemplo. En su opinión, no existía hasta la fecha un buen diccionario de marina porque todos los que hasta entonces habían redactado en forma manuscrita o publicado algún tipo de obra lexicográfica no habían unido sus fuerzas, comparando a España con Holanda, Inglaterra o Francia, que en cambio sí contaban con buenos diccionarios.

Navarrete realiza además una reseña que se extiende durante 17 páginas de los diccionarios marítimos en los que se ha basado el diccionario<sup>8</sup>. Desde el texto

8 Para la caracterización y filogenia de estos vocabularios, véase Nieto (2001).

más antiguo, de 1587, hasta el del padre Terreros de 1786. Dedicar unas cuantas líneas a cada uno de estos diccionarios en los que critica algunos aspectos de cada uno y celebra otros. En esta labor, aprovecha para criticar el atraso de España que achaca a la decadencia de los últimos reyes de la dinastía Habsburgo. Reflexiona sobre el hecho de que debe conocerse la náutica, pero también el arte de confeccionar diccionarios y en esta labor critica los plagios. La finalidad de Navarrete es demostrar que está bien documentado, que ha escogido solo lo mejor de cada obra para justificar la creación del diccionario como una necesidad.

Realiza una reflexión metalexicográfica sobre la formación y mejoras sucesivas de los diccionarios, obras largas, que requieren años y continuas revisiones (aquí se refleja la experiencia de Navarrete como académico de la RAE) y define el diccionario como “árbitro y juez” con palabras muy parecidas a las que se usan en el Prólogo del *DRAE* de 1832. En la última parte del prólogo, compuesta por 13 puntos numerados “da la idea del plan y método seguido en la redacción de este diccionario”. Nada añade, por desgracia, sobre los vocabularios bilingües inglés-español, francés-español, ni sobre el italiano-español, que aquí nos ocupa.

El diccionario venía a cubrir un gran vacío, según cuenta Navarrete en el prólogo, y parece que «la edición del tan suspirado *Diccionario Marítimo Español*, que así se denominó, debidamente distribuida despertó una profunda reacción léxica y menudearon las propuestas de correcciones y adiciones» (Guillén 1967). Veinte años más tarde de esta primera y única edición, en 1851, el jefe de dirección de Hidrografía don Francisco Chacón sugiere la reimpresión con las debidas adendas, mejoras, enmiendas y correcciones y se le concede el encargo en 1859, pero la obra no llegó nunca a ver la luz. Hasta 1867 la obra de O'Scanlan fue el diccionario de referencia de la Marina española, hasta que en 1867 se adopta por Real Orden otro más moderno, que incluye términos de las novedades aparecidas en aquellos años: el de Lorenzo, Murga, y Ferreiro de 1864: *Diccionario marítimo español que además de las voces de navegación y maniobra en los buques de vela, contiene las equivalencias en francés, inglés e italiano, así como términos utilizados en los buques de vapor*. El de O'Scanlan no se reeditó nunca, el Museo Naval ha realizado dos ediciones facsímiles de esta obra, una en 1974 y otra en 2003.

## 2. El Vocabulario italiano-español

El *Vocabulario italiano-español* consta de 33 páginas, con un total de 2.956 voces. La palabra lematizada aparece en carácter redondo separado por un punto de su

equivalente o definición, que se muestra en cursiva.

Las entradas del *Diccionario* monolingüe se encuentran en orden alfabético, considerando los dígrafos *ch* y *ll* como letras, separados de la *c* y de la *l*, en cambio en los vocabularios bilingües se sigue el orden alfabético de las otras lenguas, es decir, dentro de estas dos letras, como era y es corriente en francés, inglés e italiano.

Como se indica en el extenso prólogo, se le facilitaron a O'Scanlan nueve diccionarios extranjeros, los más destacados fueron el de Neuman (1800) y el de Rödning (1789), cuyos lemarios son, según hemos podido comprobar, muy similares al leuario del *Vocabulario italiano-español*.

## 2.1 Características del leuario italiano-español

En cuanto a la ortografía, cabe reseñar el uso de la grafía <j> como fonema vocálico, por ejemplo, en las voces de *Occhj* (por *Occhi*) o *Parasarzj* (por *Parasarzi*), grafía latinizante<sup>9</sup> que desaparece a principios del siglo XX. Hallamos algunos términos del italiano con acento gráfico, como *Marinería* o *Piratería*. Como veremos más adelante, en la retrodigitalización del diccionario estas formas aparecen “normalizadas”.

### 2.1.1 Marcas lexicográficas

Como los otros dos vocabularios bilingües presentes en el *Diccionario Marítimo*, el italiano también carece de marcas de categorización gramatical o morfológica (categoría léxica, género y número) o explícitas diatópicas, diafásicas, diastráticas o diatécnicas, aunque se recoge algún tipo de información para el lector por otros medios, como mostramos a continuación.

Los adjetivos están generalmente lematizados con un solo género, por lo general en masculino, si bien depende del término al que modifica, pues la mayoría forma parte de una colocación, así por ejemplo, hallamos “*Abbrumato. Comido de la broma (el bajel que se halla en tal estado)*” o también “*Arcatto. Quebrantado*

9 Esta grafía es frecuente en textos humanistas y renacentistas para imitar la grafía del latín clásico y medieval y en diccionarios bilingües o técnicos para marcar gráficamente el plural masculino en -i, evitando ambigüedades gráficas. Con las primeras normalizaciones ortográficas del italiano y el perfeccionamiento de la imprenta, la -j cayó en desuso.

(*el bajel que se halla en esta disposición*)”. Encontramos, sin embargo, un término con flexión explícita en la misma cabecera del artículo: “Levatojo, ja. *Levadizo, za*”. La variación morfológica de número aparece con la palabra en singular y plural separada por comas en “Asta, aste. *Asta. ||Botalon*”. Aparece la variante plural como entrada independiente. Es decir, como lemas diferentes: “Bitta. *Bita*”, y como entrada independiente se halla “Bitte. *Las bitas*”; otro ejemplo podría ser “Braccj. *Brazas*” y más abajo, como entrada independiente “Braccio. *Braza. ||Brazo*”.

Se encuentran también numerosos términos lematizados en plural, como “Giovani. *Grumetes*”, o “Mozzi. *Grumetes*”, lo que evidentemente responde al uso práctico que se le da a un vocabulario bilingüe.

Las entradas carecen de marcas de naturaleza sintáctica, pero hallamos indicaciones explícitas de este tipo en “Fossa (In). *Sobre cubierta, á plan. Úsase con el verbo amainare, amainar, arriar*”. En el siguiente ejemplo se encuentra una indicación sobre la preposición que selecciona el verbo en el mismo enunciado de la entrada: «Cacciare sul’ancora ó sopra l’ancora».

No se hallan tampoco marcas diatópicas ni diatécnicas, pero, como veremos más adelante, en la parte informativa de algunas entradas, se encuentra información de este tipo.

### 2.1.2 Características semántico-categoriales del leuario

La categoría léxica mayormente representada en el leuario es la de los sustantivos, seguida por la de los verbos, los adverbios y las interjecciones. Estas últimas usadas como enunciados formularios en situaciones comunicativas muy codificadas. Algunas aparecen solo con el término equivalente, como “Alto! *Forte*”, “Una! *A una!*”, “Sinistro! *A babor!*”, “Top! *Top! Forte!*”, y otras además con definición: “Allevoggie! *Vira al cabrestante! Forma de mando*”, “Loff! *Orza! Voz de mando al timonel*”, “Manesco (A)! *Mano entre mano!*”, “Dritto! *Derecha la caña! Voz de mando al timonel*”.

Los criterios de elección de las entradas del *Vocabulario* son, evidentemente, de tipo extralingüístico, siendo un diccionario de especialidad, se hallan sobre todo palabras pertenecientes al léxico marítimo. El criterio de la frecuencia de uso en la comunicación en la nave justifica palabras del caudal común como “Cucire. *Co-ser motones*”, “Cuoco, *Cocinero del equipage*”, “Giornale. *Diario*”, “Lanciamiento. *Lanzamiento*”, “Lanciare. V. *Slanciare*”, “Letto. *Cama*”, “Logorare. *Luir, rozar*”,

“Magazzino. *Almacen. || Pañol*”, “Giovani. *Grumetes*”, “Mozzo di camera. *Page de la cámara*”, “Fondo bruto. *Fondo sucio*” o “Quarantina. *Cuarentena*”, por ejemplo.

Por último, encontramos diferentes culturemas, términos enciclopédicos que, por resultar difícilmente traducibles, aparecen con una definición, sin un equivalente en español, como “Ganzaruolo. *Embarcación antigua veneciana*”, “Bojera. *Barco holandés á que dan este nombre*”, “Botto. *Nave holandesa á que dan este nombre*”, “Cature. *Buque de guerra del Rey de Bamtam*”, “Chiappi. *Estacas ó palos en que se amarran los barcos en los canales de Venecia*”, “Cocche. *Nombre de un buque Veneciano*” o “Lecht. *Medida marítima del Norte*”.

Se lematizan colocaciones o combinaciones sintagmáticas, como dos entradas diferentes, es decir en dos encabezamientos diferentes: uno para el lema simple y otro para la colocación<sup>10</sup>. Y así, hallamos “Bussola (osservare per la)” e inmediatamente después, como entrada independiente, se halla “Bussola”. Son también entradas independientes “Bordo”, “Bordo á bordo”, “Bordo á dritta”, “Bordo á sinistra”, “Bordo (Nave col) dentro”, “Bordo (Reversare il) ó virare per avanti” y “Bordo (Voltare il) ó virare di bordo”.

Aparecen también dos entradas independientes según un uso pronominal o no del verbo: “Incapellare” e “Incapellarsi” y también “Allargare” y “Allargarsi”. Del mismo modo, se encuentran algunas variantes morfo-fonéticas como “Barcajuolo” y debajo “Barcaruolo”.

### 2.1.3 Enunciados polimórficos y pluriverbales en el leuario

En el leuario se hallan enunciados complejos de dos tipos: los que presentan variantes de una misma palabra y los que presentan unidades pluriverbales (locuciones o colocaciones).

Las variantes de un mismo lema de nuestro vocabulario pueden ser morfo-fonéticas, como “Bertocci, bertucci, bertogli, bertolli” o “Ancoretta ó ancorotto”, por ejemplo, o dos términos sinónimos en italiano “Impalomare, relingare”, “Filo, fi, spagno fino” o “Bernarda ó magazzino da camera”.

La técnica metalexicográfica no ha sido del todo coherente para los enunciados polimórficos y así, al introducir estas variantes, encontramos tres casos: apa-

<sup>10</sup> Esta manera de proceder no se encuentra en línea con el DRAE de 1871 que “identifica las diversas acepciones de una palabra en distintas entradillas en las que se repite el lema; también se introducen como sublemas diferenciados dentro de una entrada las formas complejas. La distinción entre el lema y los sublemas descansa en la diferente tipografía empleada: mayúscula para el lema, versalita en el caso de los sublemas” (Freixas, Clavería 2018).

recen separados por una coma, otras veces se da con una conjunción disyuntiva y, a veces, con una remisión (V.):

- (i) separados por comas en el enunciado: “Attrazzi, attrezzi”.
- (ii) separados con una “ó” disyuntiva y acentuada: «Bernarda ó magazzino da camera», “Immorsiare ó immorsare”, “Bring ó loneta”.
- (iii) con una remisión (V.= vid., véase): “Banchi. V. Poggiapiedi” o “Lanche. V. Anche”.

Por otra parte, se hallan numerosas combinaciones sintagmáticas pluriverbales que se introducen como entradas independientes, y no al final del artículo del lema principal o base, como en “Batteria annegata”, “Batteria di coperta” y “Batteria di corridore”, que no se encuentran dentro del artículo del lema “Batteria”, sino bajo él, todas ellas como voces diferentes del *Vocabulario*.

Estas unidades pluriverbales presentan diferentes formatos lexicográficos:

- (i) Con el lexema genérico o lexema base entre paréntesis: “Acqua (Fare)”, “Bighe (Far le)”, “Bussola (Observare per la)”, “Picco (Andare á)”, “Pible (Albero á)” o “Pesci (Bastonare i)”.
- (ii) Como un sintagma (infinitivo con su complemento, sustantivo con complemento preposicional o con adjetivo restrictivo): “Affornellare i remi”, “Acqua pianna”, “Affrancare la nave”, “Bancassa dell’arganello”, “Bastingare una nave” o “Bocca di cane”. “Bancazze colle bigotte é lande”, “Conj per varare una nave”.
- (iii) Una frase: “Emendare la gomena quando prende sopra”.

En el *Vocabulario marítimo* de O’Scanlan no parece existir un criterio de lematización claro o coherente porque “en los diccionarios académicos anteriores al de 1871, la ordenación de las expresiones pluriverbales se realiza a partir de aquella voz que se considera más relevante, que habitualmente era un verbo” (Buenafuentes 2018: 119)<sup>11</sup>.

## 2.2 Características de los equivalentes y las definiciones

La microestructura de un diccionario puede definirse como “la ordenación de los distintos elementos que componen el artículo lexicográfico” (Rodríguez Barcia

11 Véase Buenafuentes (2018: 120-21) para los criterios de lematización de la Academia. Sobre la historia del tratamiento de las unidades pluriverbales en los diccionarios, véase Romero Aguilera (2014).

2016: 212). En el caso del *Vocabulario italiano-español*, esta estructura se caracteriza por una presentación sintética y orientada a la comunicación práctica, lo que influye en la disposición y el tipo de información contenida en cada entrada. Desde el punto de vista gráfico, el repertorio emplea la cursiva para señalar los equivalentes de traducción, los cuales aparecen separados por doble pleca (||), comas o puntos. La organización de estos elementos responde a la función primordial del vocabulario, que era servir de herramienta de referencia rápida en el ámbito marítimo.

Dado que esta obra se concibe como una síntesis del *Diccionario Marítimo Español*, su orientación estrictamente práctica se traduce en una estructura microtextual simplificada. En este sentido, la información que se proporciona queda reducida a lo que Seco (2003: 25) denomina el “segundo enunciado del artículo lexicográfico”, es decir, la definición. Esta particularidad responde a la finalidad del vocabulario, que no es la descripción exhaustiva de los términos sino la facilitación del acceso a las equivalencias terminológicas en español. Como señala Porto Dapena (2002: 268), el concepto de “definición” puede entenderse en un sentido amplio como cualquier tipo de relación de equivalencia entre la entrada y la expresión explicativa de la misma. En los diccionarios bilingües, esta relación se manifiesta en la identificación de términos equivalentes en la lengua meta, sin recurrir necesariamente a una definición explícita. Sin embargo, en aquellos casos en los que no existe un equivalente léxico directo, la estrategia lexicográfica recurre a procedimientos definitorios o parafrásticos que permitan explicar el significado del lema, como se observa en la obra que aquí analizamos. En el caso del *Vocabulario italiano-español*, la formulación de definiciones propiamente dicha obedece a un enfoque eminentemente funcional, concebido para optimizar la comunicación en entornos multilingües como los barcos y los puertos. La estructura microtextual del repertorio, por tanto, refleja una organización orientada a la eficiencia, en la que predominan las equivalencias directas sobre las definiciones explicativas.

El análisis del material lexicográfico ha permitido identificar y sistematizar cuatro tipos principales de definiciones, en sentido amplio, presentes en la obra. No obstante, es necesario señalar que los límites entre estas categorías no son siempre nítidos, ya que algunas definiciones presentan características híbridas que dificultan una clasificación estricta. A continuación, se presentan los distintos tipos de equivalencias documentadas en el *Vocabulario*, agrupados según su estructura.

### 2.2.1 Entrada + uno o varios equivalentes

En este grupo se encuentran aquellos artículos en los que la entrada se define exclusivamente a través de uno o más equivalentes léxicos en la lengua meta, sin proporcionar información adicional sobre su uso o significado. Ejemplos de este tipo de estructura son: “Calcagna. *Zapata*”, “Arrembare. *Arramblar, abordar*”, “Sponda. *Orilla, playa*” o “Braccio della campana. *Cigüeña de la campana*”. Cuando los equivalentes son sinónimos, se separan mediante comas, mientras que la doble pleca (||) suele emplearse para separar elementos que no tienen una relación de sinonimia exacta como en “Bugi. *Groeras. || Troneras ó aspilleras*”, donde el primer término, “groeras”, funciona como hiperónimo de los dos siguientes, que son sinónimos entre sí. En otras ocasiones, la doble pleca sirve para indicar referentes distintos: “Apostoli. *Astas de proa. || Ligazones de escoben*”. No obstante, en buena parte de los artículos, los significados distintos se separan simplemente mediante comas y, en algunos casos, se incorpora además un comentario metalexicográfico aclaratorio. Un ejemplo de ello es “Anima. *Eje, alma, corazón, madre, ánima, dado, chapeleta: tiene todos estos significados, según de lo que se trate*”, donde la enumeración de acepciones se complementa con una observación metalingüística que proporciona mayor precisión interpretativa.

Por otra parte, al igual que ocurre en algunos lemas, la conjunción “ó” se emplea para separar variantes que pueden usarse indistintamente en la lengua meta, como en “Battume del piano. *Betún ó sebo de los fondos*”. En cambio, la conjunción “y” se utiliza para diferenciar acepciones con significados distintos, como en “Vasi. *Anguilas y basos*”, donde se señalan términos que no son sinónimos, sino que designan conceptos diferenciados dentro del campo semántico correspondiente.

### 2.2.2 Entrada + definición

Dentro de esta categoría se incluyen aquellas entradas en las que no se proporciona un equivalente en la lengua meta, sino que el significado del término se explica mediante una definición propiamente dicha. Este procedimiento se emplea en el *Vocabulario* en casos como «Borgognone. *Pedazo de hielo suelto y flotante en el mar*» o «Madia. *La armazón en que se ponen a escurrir los cabos alquitranados*», o bien para términos que designan conceptos estrechamente ligados a la cultura o la tradición naval de una determinada zona geográfica y que no tienen un corre-



lato exacto en la lengua de destino. Ejemplos de este tipo son «Cature. *Buque de guerra del Rey de Bamtam*», «Chiappi. *Estacas ó palos en que se amarran los barcos en los canales de Venecia*», «Botto. *Nave holandesa á que dan este nombre*», «Cocche. *Nombre de un buque veneciano*» o «Lecht. *Medida marítima del Norte*». En estos casos, la definición no solo permite describir el significado del término, sino que también aporta información de carácter enciclopédico, contribuyendo a la comprensión contextual de la unidad léxica.

### 2.2.3 Entrada + uno o varios equivalentes + descripción

Este grupo está compuesto por aquellos artículos en los que, además de proporcionar uno o varios equivalentes léxicos en la lengua meta, se añade una breve descripción o explicación adicional que permite precisar su significado. Es el caso, por ejemplo, de “Vetta. *Beta, tira de aparejo*” o de “Legname d’arborare. *Arboladura: el conjunto de palos, masteleros y vergas*”. En otros casos, el orden se invierte, presentando primero la descripción y, posteriormente, el equivalente léxico, como ocurre en “Buonbordo. *El forro del casco del buque, buena bordada*”.

### 2.2.4 Entrada + uno o varios equivalentes + especificación

Dentro de esta categoría se incluyen aquellos artículos en los que, además de proporcionar uno o varios equivalentes léxicos en la lengua meta, se incorpora una especificación adicional que delimita su ámbito de uso o restringe su aplicación a un contexto determinado. Este tipo de estructuración es especialmente relevante en repertorios terminológicos especializados, donde la equivalencia entre lenguas no siempre es directa y requiere matices que permitan su correcta interpretación. Ejemplos representativos de este tipo de configuración lexicográfica pueden observarse en entradas como “Lora. *Eslora del buque*”, donde el equivalente se restringe explícitamente a un tipo específico de embarcación. Un caso similar se encuentra en “Caichieri. *Los marineros que reman en un bote*” o en “Calcese. *Calcés ó delgados del buque*”.

En este grupo podemos incluir los artículos que contienen datos que suplen a la marca lexicográfica. Esto se observa, por ejemplo, en la información diatópica presente en «Borra. *Viento de tierra. En Trieste, del esnordeste*» o en «Provenza. *En el mar Adriático, el viento oes-sudoeste*». Asimismo, se encuentran ejemplos de

información de uso relacionados con la función estilística de un término, como en el caso de «Vasto. *El Océano (en poesía)*», donde se señala explícitamente su empleo en un registro poético.

### 3. Primeros pasos en la retrodigitalización

La conversión del *Vocabulario italiano-español* en un formato digital estructurado conforme a los estándares de la *Text Encoding Initiative* (TEI) y, en particular, a la variante TEI Lex-0 (2024), responde a la necesidad de garantizar la interoperabilidad y la reutilización de los datos lexicográficos en entornos digitales<sup>12</sup>. Dado el carácter especializado del *Vocabulario*, la adaptación al entorno digital ha requerido un trabajo meticuloso en lo relativo a la codificación de las equivalencias terminológicas y la organización de los artículos lexicográficos.

El proceso de digitalización y codificación ha seguido una secuencia bien definida. Inicialmente, se recurrió al Reconocimiento Óptico de Caracteres (OCR) para la extracción automática del texto. Sin embargo, como es habitual en la digitalización de materiales históricos, se generaron errores sistemáticos que requirieron una revisión manual exhaustiva. La reconstrucción de signos diacríticos ausentes, la corrección de caracteres mal reconocidos y la comparación detallada con el documento original fueron tareas esenciales para garantizar la fidelidad textual.

La heterogeneidad de los criterios lexicográficos empleados en la obra original ha sido uno de los principales retos durante la codificación del repertorio en TEI Lex-0. La naturaleza práctica del vocabulario presentaba incongruencias en la lematización, en la presentación de equivalencias y en la organización interna de los artículos. Sin embargo, al tener como objetivo el respeto del original, no hemos realizado ningún tipo de intervención de tipo formal. Lo único que se ha añadido, para uniformar el *Vocabulario* con los otros textos lexicográficos del proyecto, es la información específica sobre categoría gramatical, género y número, mediante la etiqueta <gramGrp> (grupo gramatical).

#### 3.1 La codificación de los lemas

Las entradas polimórficas y polilexemáticas han sido etiquetadas bajo la categoría <variant>, asegurando una representación estructurada de las distintas formas

12 Véanse Bowers y Romary (2017), Tasovac (2020) y TEI Consortium (2024).

léxicas. Para los elementos que separan las variantes, como las conjunciones o las comas, indicados mediante la etiqueta <metamark>, se ha definido el atributo “variantLemmaDelimiter”. Este atributo permite una segmentación más precisa dentro de la codificación, facilitando la identificación y la organización de los lemas. Su implementación puede observarse en la figura 1.

```
<!-- Acque morte ó felle di acque. Aguas muertas. -->
<entry xml:id="VMarittimo_IS.Acque_morte" xml:lang="it" type="mainEntry">
  <form type="lemma">
    <orth norm="acque morte"> Acque morte</orth>
  </form>
  <metamark function="variantLemmaDelimiter"> ó </metamark>
  <form type="variant">
    <orth norm="felle di acque">felle di acque</orth>
  </form>
  <gramGrp>
    <gram type="mwe"/>
  </gramGrp>
  <metamark function="lemmaDelimiter">.</metamark>
  <sense xml:id="VMarittimo_IS.Acque_morte.s.1.">
    <cit type="translationEquivalent" xml:lang="es">
      <form>
        <orth>Aguas muertas</orth>
      </form>
      <gramGrp>
        <gram type="mwe"/>
      </gramGrp>
    </cit>
  </sense>
</entry>
```

Fig.1. Entrada polilexemática

Los lemas pluriverbales han sido anotados como “mwe” (multiword expressions) y así, también, se han registrado los enunciados fraseológicos que corresponden a fórmulas rutinarias y expresiones establecidas, utilizadas en situaciones comunicativas altamente codificadas. Estas expresiones ocurren en contextos específicos donde su uso es convencional y esperado. Se trata de voces como «Una!», «Sinistro!», «Top!», «Allevoggie!», «Loff!», «Manesco (A)!» y «Dritto!», para las cuales se ha incorporado información adicional mediante el atributo “forma exclamativa” bajo la etiqueta <norm> con el objetivo de facilitar su correcta interpretación. (Fig. 2)

```

<!--Allevoggie! Vira al cabrestante!-->
<entry xml:id="VMarittimo_IS.Allevoggie" type="mainEntry" xml:lang="it">
  <form type="lemma">
    <orth>Allevoggie!</orth>
  </form>
  <gramGrp>
    <gram type="mwe" norm="forma esclamativa"/>
  </gramGrp>
  <sense xml:id="VMarittimo_IS.Allevoggie.1">
    <cit type="translationEquivalent" xml:lang="es">
      <form>
        <orth>Vira al cabrestante!</orth>
      </form>
      <gramGrp>
        <gram type="mwe" norm="forma esclamativa"/>
      </gramGrp>
    </cit>
  </sense>
</entry>

```

Fig. 2. Enunciado fraseológico

### 3.2 La codificación de las equivalencias

Desde el punto de vista de la codificación, las entradas que presentan uno o varios equivalentes constituyen las estructuras más simples dentro del sistema de representación lexicográfica. En la figura 3, se puede observar la aplicación del atributo “translationDelimiter”, utilizado para marcar la separación entre el lema y el equivalente mediante el uso del punto. La etiqueta <gramGrp> permite estructurar la información gramatical asociada a la entrada y al equivalente.

```

<!-- Ancha. Anca. -->
<entry xml:id="VMaritimo_IS.Ancha" type="mainEntry" xml:lang="it">
  <form type="lemma">
    <orth>Ancha</orth>
  </form>
  <gramGrp>
    <gram type="pos" norm="sustantivo"/>
    <gram type="gender" norm="femenino"/>
  </gramGrp>
  <metamark function="translationDelimiter">.</metamark>
  <sense xml:id="VMaritimo_IS.Ancha.1">
    <cit type="translationEquivalent" xml:lang="es">
      <form>
        <orth>Anca</orth>
      </form>
      <gramGrp>
        <gram type="pos" norm="sustantivo"/>
        <gram type="gender" norm="femenino"/>
      </gramGrp>
    </cit>
  </sense>
</entry>

```

Fig. 3. Uso de la etiqueta &lt;translationDelimiter&gt;

En nuestra codificación, la información adicional, ya sea en forma de definición o de especificación, ha sido anotada con la etiqueta <gloss>, como puede observarse en las figuras 4 y 5. Esta etiqueta desempeña un papel fundamental al garantizar una clara diferenciación de los datos dentro de la estructura del vocabulario digitalizado. Su uso es especialmente relevante en la anotación de acepciones complejas, donde resulta esencial preservar la relación entre el término y su explicación contextual.

```

<!--Legname d'arborare. Arboladura: el conjunto de palos, masteleros y vergas.-->
<entry xml:id="VMaritimo_IS.Legname_arborare" type="mainEntry" xml:lang="it">
  <form type="lemma">
    <orth>Legname d<pc>'</pc>arborare</orth>
  </form>
  <gramGrp>
    <gram type="mwe" />
  </gramGrp>
  <metamark function="translationDelimiter">.</metamark>
  <sense xml:id="VMarittimo_IS.Legname_arborare.1">
    <cit type="translationEquivalent" xml:lang="es">
      <form>
        <orth>Arboladura</orth>
      </form>
      <gramGrp>
        <gram type="pos" norm="sostantivo" />
        <gram type="gender" norm="feminile" />
      </gramGrp>
    </cit>
    <pc>.</pc>
    <gloss xml:lang="es">el conjunto de palos, masteleros y vergas</gloss>
  </sense>
</entry>

```

Fig. 4. Uso de la etiqueta &lt;gloss&gt;

```

<!-- Calcese. Calcés, en los palos de vela latina. -->
<entry xml:id="VMaritimo-IS.Calcese" type="mainEntry" xml:lang="it">
  <form type="lemma">
    <orth>Calcese</orth>
  </form>
  <gramGrp>
    <gram type="pos" norm="sostantivo" />
    <gram type="gender" norm="maschile" />
  </gramGrp>
  <metamark function="translationDelimiter">.</metamark>
  <sense xml:id="VMaritimo-IS.Calcese.1">
    <cit type="translationEquivalent" xml:lang="es">
      <form>
        <orth>Calcés</orth>
      </form>
      <gramGrp>
        <gram type="pos" norm="sostantivo" />
        <gram type="gender" norm="maschile" />
      </gramGrp>
    </cit>
    <pc>,</pc>
    <gloss xml:lang="es">en los palos de vela latina</gloss>
  </sense>
</entry>

```

Fig. 5. Otro ejemplo de uso de la etiqueta &lt;gloss&gt;

Para concluir este apartado, es importante señalar que, aunque el objetivo principal era respetar íntegramente el texto original, hemos considerado pertinente anotar las erratas evidentes bajo la etiqueta <norm>. De este modo, se garantiza que los errores claramente identificables en el texto sean señalados sin alterar el documento original. Un ejemplo de este tipo de intervención puede observarse en la figura 6, donde el verbo “afferare” ha sido registrado con su forma correcta “afferrare”.

```

<!--Afferare una nave. Afferrare, amarrar un buque.-->
<entry xml:id="VMaritimo-IS.Afferare_nave" type="mainEntry" xml:lang="it">
  <form type="lemma">
    <orth norm="Afferrare una nave">Afferare una nave</orth>
  </form>
  <gramGrp>
    <gram type="mve"/>
  </gramGrp>
  <metamark function="lemmaDelimiter">.</metamark>
  <sense xml:id="VMaritimo-IS.Afferare_nave.s.1">
    <cit type="translationEquivalent" xml:lang="es">
      <form>
        <orth>Afferrar</orth>
      </form>
    </cit>
    <metamark function="translationDelimiter">.</metamark>
    <cit type="translationEquivalent" xml:lang="es">
      <form>
        <orth>amarrar un buque</orth>
      </form>
      <gramGrp>
        <gram type="mve"/>
      </gramGrp>
    </cit>
  </sense>
</entry>

```

Fig. 6. Uso de la etiqueta &lt;norm&gt;

#### 4. Conclusiones

Resulta sorprendente que, no obstante el gran interés para la lexicografía de especialidad que los pocos estudiosos que se han acercado a la obra de O'Scanlan le atribuyen, la bibliografía sea, de momento, escasa.

El análisis de las características lexicográficas del *Vocabulario italiano-español* nos ha permitido categorizar los distintos tipos de lemas y definiciones, así como constatar la heterogeneidad de su microestructura. Hemos comprobado la existencia de diferentes estrategias de ordenación del material lexicográfico en la obra original, algunas de las cuales han requerido una codificación específica para su correcta representación digital.

La integración del *Vocabulario* en la plataforma TELEI facilitará su interacción con los otros diccionarios presentes en la base de datos, así como su consulta y estudio para futuros trabajos de investigación en lexicografía histórica bilingüe y terminología especializada. La metodología empleada en este proyecto podría proporcionar en el futuro un marco replicable para otras iniciativas de digitalización de repertorios bilingües.

#### Bibliografía citada

- Ahumada, Ignacio (2008), "Anonimia desvelada de tres diccionarios terminológicos del siglo XIX español", *Limes. Lexicografía y lexicología de las lenguas de especialidad*, eds. Hugo E. Lombardini; Maria Carreras i Goicoechea, Monza/Italy, Polimetrica International Scientific Publisher: 29-45.
- Bowers, Jack; Romary, Laurent (2017), "Deep Encoding of Etymological Information in TEI", *Journal of the Text Encoding Initiative*, 10 [06/03/2025] <<https://doi.org/10.4000/jtei.1692>>
- Buenaufuentes, Cristina (2018), "Los criterios de lematización: las formas complejas", coords. Margarita Freixas; Gloria Clavería, *El diccionario de la Academia en el siglo XIX: la 5ª edición (1817) al microscopio*, Madrid, Arco Libros.
- Carriazo Ruiz, José Ramón (2008), "La marca anticuado en el *Diccionario Marítimo Español* de 1831", *Revista de Investigación Lingüística*, 11: 65-78.
- Fernández Navarrete, Martín (1831), *Discurso que sobre la utilidad de los diccionarios facultativos, con un examen de los que se han escrito de marina, y con las advertencias convenientes para formarlos y corregirlos en lo sucesivo escribió don Martín Fernández Navarrete para servir de prólogo al Diccionario marítimo español publicado en 1831*, Madrid, Imprenta Real.

- Freixas, Margarita; Clavería, Gloria (2018), “Los criterios de lematización: los lemas simples y los lemas múltiples”, coords. Margarita Freixas; Gloria Clavería, *El diccionario de la Academia en el siglo XIX: la 5ª edición (1817) al microscopio*, Madrid, Arco Libros.
- Guillén, Julio F. (1967), “El *Diccionario Marítimo Español* de 1831”, *Boletín de la Real Academia Española*, XLVII: 103-114.
- Jiménez Ríos, Enrique (2001), *Variación léxica y diccionario. Los arcaísmos en el diccionario de la Academia*, Madrid-Frankfurt, Iberoamericana-Vervuert.
- Lorenzo, José de; Murga, Gonzalo de; Ferreiro, Martín (1864), *Diccionario marítimo español que además de las voces de navegación y maniobra en los buques de vela, contiene las equivalencias en francés, inglés e italiano, así como términos utilizados en los buques de vapor*, Madrid, Establecimiento tipográfico de T. Fortanet.
- Neuman, Henry (1799), *A Marine Pocket-Dictionary: Of the Italian, Spanish, Portuguese, and German Languages, with an English-French, and French-English Index*, Londres, Impreso para el autor y vendido por Vernor y Hood.
- Nieto, Lidio (2001), “Inventario y filogenia de los repertorios de términos marineros españoles anteriores a 1726”, *Actas del II Congreso Internacional de la Sociedad Española de Historiografía Lingüística*, Madrid, Arco Libros: 93-124.
- Nieto, Lidio (2002), *Tesoro lexicográfico del español mariner*, Madrid, Arco Libros.
- O'Scanlan, Timoteo (1829), *Cartilla practica de construccion naval, dispuesta en forma de vocabulario con algunos apéndices, y las nomenclaturas francesa, inglesa é italiana, con su correspondencia en castellano; para uso de los principiantes en este ramo tan esencial de la marina, redactada por d. Timoteo O-Scanlan, caballero de la real y militar orden de San Hermenegildo, Capitan de fragata de la armada*, Madrid, De la Imprenta Nacional.
- O'Scanlan, Timoteo (1831), *Diccionario Marítimo Español*, Madrid, Imprenta Real.
- Real Academia Española, *DRAE*, 2001 [01/03/2025] <<https://www.rae.es/drae2001>>
- Real Academia Española, *Diccionario de Autoridades 1726-1739* [02/03/2025] <<https://apps2.rae.es>>
- Röding, Johann Hinrich (1793-1798), *Allgemeines Wörterbuch der Marine, in allen europäischen Seesprachen nebst vollständigen Erklärungen*, Hamburgo y Leipzig, Conrad Müller para Licentiat Nemnich y Adam Friedrich Böhme.
- Rodríguez Barcia, Susana (2016), *Introducción a la lexicografía*, Madrid, Síntesis.
- Romero Aguilera, Laura (2014), “Unidades pluriverbales y diccionario: el tratamiento de las colocaciones en la historia de la lexicografía española”, *Dándole cuerda al reloj: ampliando perspectivas en lingüística histórica de la lengua española*, coord. Vicente Álvarez Vives; Elena Díez del Corral; Natacha Reynaud Oudot. Valencia, Tirant lo Blanch: 223-240.
- Seco Reymundo, Manuel (2003), *Estudios de lexicografía española*, Madrid, Gredos.
- Tasovac, Tom (2020), *TEI Lex-0: A Target Format for TEI-Encoded Dictionaries. Journal of the Text Encoding Initiative*, 13 [29/01/2025] <<https://doi.org/10.4000/jtei.3006>>
- TEI Consortium (2024), *TEI Guidelines P5: Guidelines for Electronic Text Encoding and*



*Interchange* [12/02/2025] <<https://www.tei-c.org/release/doc/tei-p5-doc/en/html>>  
 Palacio, Vicente; Cervera, José; Seco, Carlos; Higuera, Dolores; Anes, Gonzalo (1995),  
*Martín Fernández de Navarrete, el marino historiador (1765-1844)*, Cuadernos  
 monográficos del Instituto de historia y cultura naval, 24, Madrid, Instituto de  
 Historia y Cultura Naval (IHCN).

**Rosa María García Jiménez**, doctora en Lingüística Hispánica, es investigadora y enseña Lengua Española en la Universidad de Pisa. Sus principales líneas de investigación son la lingüística aplicada a la enseñanza de ELE, el análisis del discurso, la pragmática y la lexicografía, campos en los que ha participado en diversos proyectos nacionales e internacionales.

**rosa.garcia@unipi.it**

**María Enriqueta Pérez Vázquez** es licenciada en Filología Hispánica por la Universidad Complutense, doctora en Lingüística General por la Scuola Normale Superiore de Pisa. Profesora titular en la Universidad de Pisa. Ha publicado una monografía, numerosas ediciones críticas y artículos en el campo de la lingüística contrastiva, en el ámbito de la sintaxis, la semántica, la lexicología y la lexicografía.

**maria.perez@unipi.it**



# CARMEN CASTILLO PEÑA

## LA DIGITALIZACIÓN DE NOMENCLATURAS BILINGÜES ESPAÑOL-ITALIANO

Università degli Studi di Padova

### Resumen

A partir de una breve presentación del proyecto de retrodigitalización de diccionarios TELEI, el artículo ofrece un estudio de las nomenclaturas bilingües hispano-italianas. Se analiza la macroestructura de estos diccionarios, así como la organización interna de sus entradas. La investigación aborda aspectos relacionados con su codificación en XML TEI Lex-O, con el objetivo de facilitar la retrodigitalización de las nomenclaturas.

palabras clave: tesoro lexicográfico bilingüe, nomenclaturas bilingües, retrodigitalización de diccionarios, TEI Lex-O

### Abstract

#### *Digitising Spanish-Italian bilingual Thesauri*

*Starting with a brief overview of the TELEI retrodigitised dictionary project, the article presents a study of Spanish Italian bilingual Thesaurus. It analyses the macrostructure of these dictionaries as well as the internal organization of their entries. The research addresses issues related to their encoding in XML TEI Lex-O, with the aim of retrodigitising these Thesaurus.*

*keywords: lexicographical corpus of bilingual dictionaries, bilingual thesaurus, retrodigitised dictionary, TEI Lex-O*

## 1. Introducción

En este estudio<sup>1</sup> presentamos las decisiones tomadas para la digitalización de nomenclaturas multilingües español-italiano utilizando el lenguaje de marcado XML TEI Lex-0<sup>2</sup> (Tasovac *et al.* 2001). Este proceso se realizó con el propósito de incorporar dichas nomenclaturas al *Tesoro digital de la lexicografía bilingüe español-italiano* (TELEI), cuyo objetivo principal es la creación en formato digital de un “Tesoro lexicográfico”, denominación habitual en la lexicografía hispánica para referirse a un *corpus glosariorum* (Nieto Jiménez, Alvar Ezquerro 2007: IX) o a un *diccionario de diccionarios* (Alvar Ezquerro 2008). En esencia, se trata de un diccionario digital que abarca todo el léxico inventariado en los repertorios lexicográficos bilingües o multilingües español-italiano.

Las nomenclaturas son obras lexicográficas bilingües o multilingües de carácter eminentemente didáctico, cuyo principal rasgo distintivo consiste en la organización temática<sup>3</sup> del léxico. A diferencia de los diccionarios alfabéticos, las nomenclaturas presentan las palabras agrupadas por temas en capítulos o secciones (por ejemplo “la casa”, “los oficios”, “las tareas del campo”). Cada entrada léxica se acompaña de sus equivalentes en una o varias lenguas meta, facilitando así el aprendizaje de lenguas extranjeras y la comparación interlingüística. Su estructura refleja una concepción de la organización del mundo (García Aranda 2011:

---

1 La investigación forma parte del proyecto “Un nuovo ambiente digitale per il recupero del patrimonio lessicografico: il “Tesoro digitale della lessicografia bilingue spagnolo-italiano”, financiado por La Unión Europea- Next Generation EU, Missione 4 Componente 1 CUP C53D23004010006 / PRIN 2022 MUR 20229W73WR.

2 TEI-Lex0 es una personalización del esquema TEI para la *retrodigitalización* de diccionarios, esto es, para la digitalización de diccionarios impresos.

3 Además de *temático*, otros términos para referirse a este tipo de ordenación del leuario son *conceptual* y *onomasiológico*. Hemos elegido el primero porque la clasificación del léxico que reflejan las nomenclaturas se basa en la realidad, por lo que es prevalentemente extralingüística, mientras que la ordenación de los diccionarios conceptuales u onomasiológicos tienen una base fundamentalmente lingüística (García Aranda 2003: 53-55). En esta diferencia se basa García Aranda para distinguir las nomenclaturas de los otros diccionarios no alfabéticos, como los ideológicos, analógicos, de ideas afines, de sinónimos y antónimos, etc. Vázquez y Fernández-Montraveta (2023: 554) por su parte, identifican en el bilingüismo el rasgo diferencial: las nomenclaturas son diccionarios temáticos que incluyen más de una lengua; sin embargo definen los diccionarios temáticos como aquellos que presentan el léxico organizado por temas, aunque en cada subdivisión temática el orden sea alfabético. Con todo, no es posible aplicar a las nomenclaturas una distinción neta entre ordenación temática y ordenación conceptual, por lo que en ocasiones utilizaremos este último término como variante estilística del primero.

293; Vázquez, Fernández-Montraveta 2023: 554), lo que las convierte en valiosos testimonios no solo lingüísticos, sino también culturales e históricos (Marello 1980).

La importancia de las nomenclaturas en la historia de la lexicografía bilingüe ha sido ampliamente reconocida por numerosos trabajos historiográficos (Gallina 1959; Quemada 1967; Alvar Ezquerro 1993; Ayala Castro 1992; Pablo Núñez 2008; García Aranda 2004, 2007), destacando especialmente García Aranda (2003) y Alvar Ezquerro (2013). Su finalidad didáctica, peculiaridad estructural y enorme difusión en los siglos XV-XVIII han sido factores clave en este reconocimiento.

La incorporación de nomenclaturas en un Tesoro lexicográfico digital plantea desafíos nuevos, ya que el etiquetado propuesto por TEI Lex-0 no contempla la codificación de diccionarios temáticos, conceptuales u onomasiológicos. Estas innovaciones exigen una reflexión crítica previa sobre la macroestructura y la microestructura de esta particular clase de diccionarios.

En las secciones siguientes abordaremos estas y otras cuestiones relacionadas con la edición de nomenclaturas en forma de Tesoro, proponiendo un sistema de codificación específico para las mismas. Con esta finalidad, se analizan los rasgos característicos de las nomenclaturas (§ 2), se presenta sucintamente el corpus de nomenclaturas con el español y el italiano de los siglos XVI-XIX, (§ 3), y, por último, se indican las principales soluciones adoptadas para la codificación de las entradas de una nomenclatura (§ 4).

## 2. Caracterización de las nomenclaturas como tipos de diccionarios. Rasgos fundamentales

Para la realización de TELEI, hemos adoptado la codificación propuesta por TEI Lex-0, la cual asume los fundamentos metodológicos de la lexicografía teórica, establecidos a partir de un prototipo de diccionario que es semasiológico y monolingüe. La codificación en XML consiste en reflejar a través de las etiquetas<sup>4</sup> los componentes básicos de una obra lexicográfica y su estructuración, ya que los datos lingüísticos del enunciado lexicográfico no constituyen una simple cadena ordenada (a, b, c, d, e, f, g), sino un conjunto estructurado ( $A = \{a, b, c, D\}$ ,  $D$

<sup>4</sup> Recordamos que un lenguaje de marcado es una forma de codificar un documento añadiendo al texto etiquetas o marcas que contienen información adicional sobre la estructura del texto para un análisis posterior o su presentación en forma digital.

= {E, f}, E = {g}). Conceptos básicos de la metalexicografía como *entrada*, *entrada pluriverbal*, *definición*, *acepción*, *marca gramatical*, *marca de uso*, *ejemplo*, etc. y su marcación en XML corresponden plenamente a la macroestructura y a la microestructura del diccionario alfabético monolingüe y se ajustan satisfactoriamente a las del alfabético bilingüe, pero no podemos decir lo mismo en lo que se refiere a las nomenclaturas y los diccionarios onomasiológicos en general:

Los investigadores en lexicología teórica han reflexionado sobre la macroestructura y la microestructura del diccionario monolingüe y, en todo caso, del bilingüe contemporáneo, nunca sobre los repertorios no alfabéticos, bilingües o plurilingües y de los siglos XVI, XVII o XVIII (García Aranda 2003: 67).

Como el propósito del proyecto TELEI consiste en digitalizar tanto los diccionarios alfabéticos como los no alfabéticos para la constitución de un solo corpus que abarque ambos tipos de diccionarios, es esencial, por tanto, establecer en primera instancia cuáles son las unidades constitutivas de una nomenclatura, cómo están estructuradas y el modo en que se pueden codificar con un etiquetado idéntico al utilizado para los diccionarios alfabéticos.

## 2.1 *La macroestructura*

La aplicación del concepto de macroestructura, ampliamente utilizado y aceptado en la lexicografía, al análisis de las nomenclaturas presenta desafíos particulares, a causa de las controversias que genera su pertinencia en relación con dicho género lexicográfico.

Según García Aranda (2003: 70), el concepto de macroestructura, tal como se entiende convencionalmente, no es adecuado para las nomenclaturas. Su razonamiento se fundamenta en dos aspectos clave: en primer lugar porque la ordenación de los materiales es de naturaleza extralingüística, organizando “ideas, conceptos o realidades” en lugar de “palabras, voces o segmentos de la lengua”; en segundo lugar, porque a menudo las nomenclaturas presentan la característica de la multidireccionalidad, a causa de la cual no es explícito el sistema que establece cuál es la lengua de partida y cuál la de llegada, como es habitual en los diccionarios bilingües alfabéticos. Ante estos argumentos, García Aranda propone el término “macrodisposición de los materiales” como alternativa preferible para describir la organización de las nomenclaturas.

No obstante, otros investigadores han optado por amoldar el concepto de

macroestructura al contexto específico de las nomenclaturas. Por ejemplo, Castillo Peña (2004, 2005) emplea el término para referirse al conjunto de capítulos que constituyen el repertorio y al orden con el que estos se presentan. Por su parte, también Carranza Torrejón (2017) aplica *macroestructura* a la ordenación temática en sí misma, reconociendo la heterogeneidad macroestructural característica de este tipo de repertorios. Esta autora señala que “el propósito docente de las nomenclaturas también se manifiesta en su macroestructura, es decir, en la distribución del vocabulario”, destacando la complejidad que puede alcanzar esta ordenación al combinar criterios temáticos, gramaticales e incluso alfabéticos.

Es interesante notar que, en contraste con estas posturas, en su exhaustivo estudio sobre las nomenclaturas españolas, Alvar Ezquerro (2013) opta por evitar el término *macroestructura* (y cualquier otro), prefiriendo expresiones más genéricas como “división” (la fórmula más repetida es “la nomenclatura está dividida en x capítulos”). Esta elección terminológica confirma la cautela con la que conviene abordar el uso de conceptos lexicográficos tradicionales cuando se describen o analizan las nomenclaturas.

En este estudio se parte de una hipótesis de trabajo según la cual las entradas principales de las nomenclaturas son los títulos de los capítulos o epígrafes (y no las unidades léxicas agrupadas bajo cada uno de esos capítulos), de ahí que mantengamos el término “macroestructura” para referirnos, precisamente, al conjunto de capítulos o epígrafes y al modo en que estos se presentan y suceden. Efectivamente, a diferencia de los diccionarios alfabéticos, en las nomenclaturas no es el conjunto de unidades léxicas inventariadas lo que determina su estructura, sino la serie de capítulos o epígrafes bajo los que se agrupa una serie de lemas, ordenados a su vez con criterios que pueden ir variando (extralingüísticos y lingüísticos, como la hiponimia, meronimia, etc.). Cada uno de estos capítulos suele tener una denominación (*las partes del cuerpo, los enseres de la cocina, los oficios*, etc.), de ahí que podamos aceptar en principio que tales denominaciones son las entradas (principales) de una nomenclatura y que, por ello, constituyen la macroestructura de la obra.

La cuestión de la direccionalidad en las nomenclaturas bilingües o plurilingües añade una capa adicional de complejidad a esta cuestión, ya que no hay un leuario para cada lengua. En las nomenclaturas que presentan sus materiales en columnas, se asume que la lengua de partida es la de la columna izquierda (por ejemplo, Alvar Ezquerro 2013: 80-81), mientras que en aquellas que los presentan en forma de párrafo, la primera palabra suele ser de la lengua de entrada. En los siguientes ejemplos de Franciosi (1626) —en columnas— y de Noviliers

Clavel (1629) —en párrafos— se aprecia que la lengua de partida es el español para el primero y el italiano para el segundo:

(Franciosini 1626: 236)

retrete, o scritorio.	scrittoio, o studio; il luogo dove l'huomo si ritira a studiare, o a scrivere
recamara.	la guardaroba.
alcova, o alcoba.	un camerino fatto nella sala; architettura, molto usata in Spagna.

Noviliers Clavel (1629: XVI)

64. conficcare un chiodo. *ficher un clou*. clavar.

97. un bastone. *un baston*. un palo.

Por otra parte, la presencia de sinónimos y de fórmulas definitorias o explicativas son factores que pueden alterar la direccionalidad aparente de una nomenclatura (Castillo Peña 2005: 8-10). Es lo que se observa en los siguientes casos de Franciosini (1626), en los que, a pesar de estar en la columna de la izquierda, el español es la lengua de llegada<sup>5</sup>:

Manteca de vaca. *burro o butiro*.

Manteca de puerco. *strutto*. (246)

O en el siguiente de Noviliers Clavel (1629), en donde se deduce que la lengua de partida es el francés por la ausencia de equivalentes en italiano y español:

169. quelli che portano a vendere gl'uccelli di rapina. *les cagiers, ceux qui en des cages, ou sur des perches portent à vendre des oiseaux de proie*. los que en jaulas llevan a vender aves de rapiña o halcones (XXXV).

Los ejemplos extraídos de obras como las de Franciosini (1626) y Noviliers Clavel (1629) ilustran tanto la norma general como las excepciones en la determinación de la lengua de partida, subrayando la complejidad inherente a estos repertorios lexicográficos.

Uno de los rasgos distintivos de la macroestructura en las nomenclaturas es el orden en que aparecen agrupadas las formas léxicas. Este orden no es arbitrario,

5 Alvar Ezquerro (2013) ofrece numerosos ejemplos de otras nomenclaturas, entre otros el del *Thesaurus puerilis* de Bernabé Soler, en donde a pesar de ser sistemática la posición del español en la columna de la izquierda (los bancos que se sientan los remeros, *transtrum*, i, n.) es evidente que “la nomenclatura está pensada desde el latín, pues si lo estuviesen desde el español, carecerían de sentido esas entradas que son la definición o descripción de la voz latina” (117).



sino que refleja relaciones léxicas significativas como la hiperonimia, la sinonimia y la meronimia. Además, se pueden observar agrupaciones basadas en familias léxicas o redes semánticas más complejas, como las acciones propias de un concepto, sus cualidades o los agentes relacionados. Se trata de una organización conceptual interna a cada capítulo y fundamental para comprender la verdadera naturaleza y el valor de las nomenclaturas. Como acertadamente señala García Aranda (2003: 71),

la nomenclatura contiene material léxico que no puede ser entendido, ni analizado, ni interpretado correctamente si no se tiene en cuenta el capítulo en el que se encuentra y las voces adyacentes. El estudio atomístico que permite un diccionario alfabético, en el que las palabras pueden ser aisladas y diferenciadas individualmente, no es posible en el caso de las nomenclaturas.

Esta característica distingue claramente las nomenclaturas de los diccionarios alfabéticos y requiere para las mismas un enfoque analítico diferente. En conclusión, aunque el concepto de macroestructura de una nomenclatura sea controvertido, se puede adoptar en la medida en que permite reconocer la estructura temática como base de la organización del material léxico inventariado, así como dar cuenta de las particularidades de cada capítulo en cuanto a direccionalidad y forma de presentar los materiales. Esta aproximación permite mantener la terminología lexicográfica tradicional, adaptándola a las características específicas de las nomenclaturas, y proporcionar un marco conceptual útil para el análisis y la digitalización de estos valiosos recursos lexicográficos.

## 2.2 *La microestructura*

También la aplicación de la noción de microestructura a las nomenclaturas requiere una serie de consideraciones. La microestructura “está constituida por la información ofrecida sobre un lema y por la secuencia de formas y contenidos que el lexicógrafo le impone” (Lombardini 2023: 252). Se trata de un concepto próximo al de “artículo lexicográfico”, que es, en palabras de Garriga (2003) la unidad “autónoma” formada por la unidad léxica y la información proporcionada sobre esa unidad. La información que aparece en la microestructura de un diccionario bilingüe comprende los equivalentes interlingüísticos, la pronunciación, las remisiones, las informaciones gramaticales, los sinónimos y antónimos, las marcas de uso, ejemplos y varios tipos de información cultural o enciclopédica. Todos estos

datos se refieren normalmente a la lengua de llegada, pero también a la lengua de partida (Lombardini 2023).

Como se ha mencionado, en las nomenclaturas la información sobre el lema se reduce habitualmente a la mera equivalencia interlingüística. Se trata del único elemento en común con los diccionarios alfabéticos bilingües, si bien, a diferencia de los mismos, se manifiesta de formas diversas y no siempre sistemáticas ya que es muy alto el grado la heterogeneidad estructural de los artículos lexicográficos en una misma nomenclatura.

En los ejemplos anteriores podemos observar entradas que se limitan a una simple equivalencia interlingüística (como en el caso de “bastón”), junto a aquellas que incorporan sinónimos (“retrete”), glosas explicativas (“alcoba”), o incluso definiciones completas (“cagiers”). García Aranda (2003: 72) ofrece una taxonomía detallada de los tipos de equivalentes más frecuentes en las nomenclaturas, que va desde los equivalentes lingüísticos conceptuales sinonímicos simples hasta los equivalentes enciclopédicos. Esta clasificación nos permite apreciar la riqueza y complejidad de la información lexicográfica contenida en estos repertorios, a pesar de la aparente simplicidad de su estructura.

Otro aspecto que hay que tener en cuenta a la hora de la codificación en XML TEI Lex-0 radica en el hecho de que la ausencia de marcas gramaticales o de uso explícitas no implica que la microestructura de estos diccionarios carezca de información gramatical o pragmática, ya que tales datos se suelen presentar de manera implícita y contextual. Por ejemplo, los sustantivos suelen aparecer acompañados de artículos (a modo de unidad ostensiva del género) y muchos adjetivos muestran su moción. Además, hay glosas ocasionales con información sobre el contexto de uso que incluye aspectos diastráticos, diafásicos o diatópicos. El caso siguiente ilustra la información diatópica en una nomenclatura de modesta extensión.

faltriquera, o faldriquera.                      tasca, ed a Roma saccoccia. (Franciosini 1626: 227)

Se trata, en todo caso, de información no codificada y asistemática, de ahí nuestra insistencia en la prudencia con la que conviene aplicar estas categorías propias de la lexicografía alfabética a las nomenclaturas bilingües antiguas.

En conclusión, el etiquetado de las nomenclaturas para su digitalización debe ser capaz de capturar toda esta riqueza y complejidad. Es esencial marcar no solo las entradas principales y sus subentradas, el orden de aparición, sino también las lenguas involucradas, la información gramatical implícita, las indicaciones de uso y las glosas culturales o ideológicas. Solo así podemos preservar y hacer accesible

la información contenida en estos repertorios lexicográficos.

### 3. Nomenclaturas que incluyen el español y el italiano

Para la constitución del corpus de TELEI se ha elaborado una lista, que aspira a ser exhaustiva, de los repertorios no alfabéticos publicados entre los siglos XVI y XIX en los que es posible identificar equivalencias interlingüísticas entre el español y el italiano. La lista ha sido realizada a partir de los estudios catalográficos llevados a cabo por Gallina (1959), Pablo Núñez (2008), Alvar Ezquerra (2013) y Peñín Fernández (2023)<sup>6</sup>. Abarca un amplio espectro temporal y tipológico, incluye obras plurilingües y bilingües, y comprende tanto textos independientes como nomenclaturas publicadas en el interior de obras didácticas. La tabla que se presenta a continuación ofrece una visión sinóptica de estos diccionarios<sup>7</sup>, organizados cronológicamente. Para cada obra, se especifica si es plurilingüe o bilingüe español-italiano y si está incluida en una gramática o libro de diálogos. Esta clasificación nos permite observar tanto la consistencia del corpus como las tendencias y cambios en la producción lexicográfica a lo largo de cuatro siglos, proporcionando una buena base para el análisis historiográfico previo a la digitalización.

	Plurilingüe	Bilingüe español-italiano	Incluida en una gramática o en libro de diálogos
<i>Quinque linguarum</i> (1526)	x		
<i>Quinque Lingvarum</i> (1534) <sup>8</sup>	x		
Junius (1567)	x		
Decimatur (1596)	x		x
Zumarán (1621)	x		x
Franciosini (1626)		x	x

6 Varias de estas obras (por ejemplo los *Quinque linguarum*, *Decimatur*, *Iunius*, *Zumarán*, *Perles*, *Marín*) han sido objeto de trabajos puntuales y de Tesis Doctorales, que no se consignan en bibliografía por problemas de espacio y porque el análisis historiográfico del estado de la cuestión queda fuera de los objetivos de este artículo.

7 Indicamos las referencias en la sección de la bibliografía dedicada a las fuentes primarias.

8 Las investigaciones de Rossebastiano Bart (1984), Redondo Rodríguez (2007) y Alvar Ezquerra (2013) permiten afirmar que, a pesar de ser una reedición del *Quinque* (1526), se puede considerar una nomenclatura nueva, por sus muchas variantes y adiciones.

	Plurilingüe	Bilingüe español-italiano	Incluida en una gramática o en libro de diálogos
Habrecht (1629)	x		x
Noviliers (1629)	x		
Barnabé (1657)	x		x
Giulliani (1659)	x		x
Howell (1659)	x		
Comenius (1661)	x		
Kramer (1670)	x		
Perles (1689)		x	x
Borja (1718)	x		
Andree (1725)	x		
Chirschmair (1734)		x	x
Rosterre/Terreros (1771)		x	x
Tomasi (1779)		x	x
Hervás y Panduro (1797)		x	x
Frutos, José de (1804)	x		
Boardman (1811)	x		
Vergani (1826)		x	x
Marín (1833)		x	x
Ainsa Royo (1837)	x		x
Ochoa (1842)		x	x
Costanzo (1848)		x	x
López de Morelle (1851)		x	x
Corona Bustamante (1870?)		x	x
<i>Vocabolario</i> (s.f.) <sup>9</sup>		x	x
<i>Vocabulaire europeen</i> (1878)	x		
Foulques (1898)		x	x

El corpus comprende treinta y un textos únicos, considerando que el *Vocabolario* (s.f.) es una reedición fiel de Corona Bustamante (1870). Aunque no se ha especificado el número de ediciones de cada texto, es importante señalar que la mayoría

<sup>9</sup> Alvar Ezquerro (2013: 681) afirma que no es sino una edición italiana de Corona Bustamante (s.f.), con la sola diferencia de que el italiano aparece como lengua de partida en la columna de la izquierda y en los escuetos paratextos.

ha sido reeditada en múltiples ocasiones. Ambos factores —la cantidad de textos y sus numerosas reediciones— subrayan la relevancia cuantitativa de este tipo de obras en la historia de la lexicografía.

El análisis de la tabla revela patrones interesantes en la evolución de este sector de la lexicografía hispano-italiana, que son, por otra parte, comunes a los de la historia general de las nomenclaturas del español (Alvar Ezquerro 2013).

Respecto a la distribución temporal, observamos que cuatro diccionarios pertenecen al siglo XVI, once al XVII, siete al XVIII y once al XIX. Esta distribución muestra una mayor producción en los siglos XVII y XIX, que juntos representan el 65 % del total de obras.

Respecto al número de lenguas, quince diccionarios (44%) son plurilingües, mientras que diecinueve (56%) son bilingües español-italiano. De manera similar a lo que sucede con todas las nomenclaturas que incluyen el español, la proporción de las bilingües aumenta significativamente con el pasar del tiempo. De hecho, todos los diccionarios del siglo XVI son plurilingües, mientras que, en el siglo XIX, nueve de once son bilingües.

En cuanto a la inclusión en gramáticas o libros de diálogos, veinte diccionarios (59%) están integrados en este tipo de obras, frente a catorce (41%) que se conciben y publican como repertorios independientes. Esta característica también muestra una evolución temporal: en el siglo XVI, solo una de las cuatro nomenclaturas está incluida en una gramática o libro de diálogos, mientras que, en el siglo XIX, diez de once lo están. El cruce de datos permite un análisis más detallado: de los quince diccionarios plurilingües, solo seis (40%) están incluidos en gramáticas o libros de diálogos; en contraste, de los diecinueve diccionarios bilingües, catorce (74%) están incluidos en este tipo de obras; en el siglo XVII, siete de once son diccionarios plurilingües, pero solo cuatro de estos siete están incluidos en gramáticas o libros de diálogos; en el siglo XIX, nueve de once diccionarios son bilingües, y ocho de estos nueve están incluidos en gramáticas o libros de diálogos.

Estas cifras confirman una tendencia hacia la especialización bilingüe y la integración en obras didácticas a lo largo del tiempo, tendencia que las nomenclaturas hispano-italianas comparten con el resto de las nomenclaturas que incluyen el español. El punto de inflexión parece situarse en el siglo XVIII, donde la proporción de diccionarios bilingües (cuatro de siete) supera por primera vez a los plurilingües.

Este sucinto análisis cuantitativo muestra la evolución desde diccionarios predominantemente plurilingües y autónomos en los siglos XVI y XVII, hacia obras

bilingües español-italiano integradas en gramáticas o libros de diálogos en los siglos XVIII y XIX. Esta transformación es gradual pero constante, con el siglo XVIII como período de transición.

#### 4. La digitalización de las nomenclaturas. El problema del orden conceptual

Tras haber examinado las características formales de las nomenclaturas (§ 2) y la relación entre el plurilingüismo, la independencia textual y el contexto cronológico (§ 3), nos adentramos ahora en la propuesta de etiquetado para su digitalización. Dicho proceso requiere una cuidadosa adaptación del sistema de codificación de diccionarios TEI Lex-0 a las necesidades específicas del proyecto TELEI. Nuestra propuesta se fundamenta en un principio esencial: preservar en lo posible la idiosincrasia estructural de las nomenclaturas al integrarlas en un Tesoro junto a repertorios alfabéticos<sup>10</sup>. Este enfoque busca evitar la dispersión de la información temática, inherente a las nomenclaturas, a causa de la arbitrariedad del orden alfabético. Por el contrario, aspiramos a mantener las relaciones temático-conceptuales que cada lema establece con los demás dentro de su grupo.

Para lograr este objetivo, hemos priorizado dos aspectos fundamentales: primero, conservar la vinculación de cada lema con el título del capítulo en el que aparece; segundo, desarrollar un sistema de visualización de resultados que permita recuperar todos los lemas agrupados en un capítulo, respetando su orden original. Esta metodología no solo facilita la digitalización, sino que también enriquece las posibilidades de consulta y análisis del material lexicográfico, manteniendo la esencia estructural de las nomenclaturas en un entorno digital.

##### 4.1 TEI Lex-0

TEI Lex-0 es un subconjunto de las directrices de la Iniciativa de Codificación de Textos (TEI, por sus siglas en inglés), diseñado específicamente para codifi-

---

10 El *Nuevo Tesoro Lexicográfico del español* (Nieto Jiménez, Alvar Ezquerro 2007) y el *Tesoro lexicográfico* (Gili Gaya 1960), ambos en ediciones impresas, incluyeron numerosas nomenclaturas bilingües en sus corpus. Sin embargo, en estos trabajos monumentales se recogen solo las equivalencias interlingüísticas, omitiendo las referencias a los títulos y subtítulos que encabezan las series léxicas. Por otra parte, las introducciones de estos Tesoros no abordan el problema crítico que supone la edición en orden alfabético de la información léxica contenida en un diccionario conceptual.

car recursos lexicográficos. Fue desarrollado con la finalidad de proporcionar un formato estandarizado para codificar diccionarios y otros recursos léxicos en un formato digital y legible por máquinas, garantizando y promoviendo la interoperabilidad entre diferentes proyectos y herramientas lexicográficas. Al ser un sistema de la TEI, utiliza un lenguaje XML (eXtensible Markup Language) para estructurar los datos lexicográficos, permitiendo una representación detallada y flexible de las entradas del diccionario. Al igual que el módulo para diccionarios de TEI, incluye elementos y atributos específicos para codificar las distintas partes de una entrada de diccionario, como lemas, definiciones, etimologías, ejemplos de uso e información gramatical. Proporciona un marco estándar y, a la vez, permite la personalización necesaria para satisfacer las necesidades específicas de diferentes proyectos de digitalización de diccionarios (Tasovac *et al.* 2001; Salgado *et al.* 2024).

En general, se trata por una parte de una reducción de etiquetas y atributos TEI y, por otra, de un sistema distinto de anidamiento que favorece estructuras más planas y con menos niveles jerárquicos.

#### 4.2 TEI Lex-0 para la codificación de nomenclaturas

A tenor de lo que se expuso en §2, el fundamento de la codificación que proponemos consiste en lo siguiente:

- la marcación de los títulos de capítulos como entradas principales (<entry type="mainEntry">),
- la marcación de los lemas que aparecen en cada capítulo como entradas secundarias relacionadas con la entrada principal (<entry type="relatedEntry">),
- la marcación del orden en que se consignan los lemas como una característica de la entrada secundaria, esto es, como un atributo de la etiqueta <entry>.

Veamos un ejemplo concreto, procedente del *Quinque linguarum* (1534):

El ix cap. De los servidores  
peon  
moço

El ix capitolo de li servitori  
il fante a pie  
il servo

servidor	famiglio
moça	la serva
servir	servire
mançebo	il garzone
mançebos	li garzoni
machacho	il fanciullo
mochachos	li fanciulli
varón	il maschio
mochacha	la puta

El elemento raíz <entry> contiene el segmento textual correspondiente al título del capítulo en las dos lenguas:

```
<entry xml:id="QuinLin_34_De_los_servidores" xml:lang="es" type="mainEntry">
```

La etiqueta <entry> del elemento raíz se caracteriza por tres atributos, siendo los dos primeros obligatorios en TEI Lex-0: @xml:id="QuinLin\_34\_De\_los\_servidores", @xml:lang="es" y @type="mainEntry". El primer atributo identifica unívocamente la entrada de la que se trata; en nuestro caso contiene la sigla correspondiente a la nomenclatura que estamos digitalizando y la identificación de la entrada; el segundo atributo define la lengua de la entrada y, por último, el tercer atributo especifica que se trata de una entrada principal, que en el caso de las nomenclaturas es un capítulo.

Además, <entry> agrupa una serie de componentes hijos: <form>, <sense> y varias <entry> anidadas.

El componente hijo <form>, contiene obligatoriamente un <orth>, que es el que marca el segmento textual del título del capítulo (*El ix cap. De los servidores*).

```
<entry xml:id="QuinLin_34_De_los_servidores" xml:lang="es" type="mainEntry">
  <form>
    <orth>El ix cap. De los servidores</orth>
  </form>
```

El segundo componente hijo es <sense>, del cual depende <cit>, etiqueta que en los diccionarios bilingües sirve para la marcación del equivalente traductivo (@type="translationEquivalent" y @xml:Lang="it"). En nuestro caso codifica el título en italiano (*El ix capitolo de li servitori*). Es un requisito de TEI Lex-0 que bajo <sense> haya un <cit>, que a su vez incluye un <form> con un <orth>. Es



este último el que etiqueta el segmento textual:

```
<entry xml:id="QuinLin_34_De_los_servidores" xml:lang="es" type="mainEntry">
  <!-- etc. -->
  <sense xml:id="QuinLin_34_De_los_servidores.s.1.">
    <cit type="translationEquivalent" xml:id="it">
      <form>
        <orth>El ix capitolo de li servitori</orth>
      </form>
    </cit>
  </sense>
```

En realidad, hasta ahora nos hemos limitado a etiquetar las denominaciones de los capítulos del mismo modo en que se codifican las entradas de un diccionario bilingüe, con la salvedad de que no consta la información gramatical. Hemos convenido arbitrariamente en que el español es la lengua de partida por estar en la columna de la izquierda con respecto al italiano, que es la lengua de llegada. Aun siendo una convención arbitraria, somos conscientes de que no se trata de una solución plenamente satisfactoria pero no tiene implicaciones importantes ni en la representación estructural del diccionario ni en su visualización.

El siguiente esquema proporcionado por el software *Oxygen* representa jerárquicamente la serie de nudos:

▼entry	@xml:id	QuinLin_34_De_los_servidores		
	@xml:lang	es		
	@type	mainEntry		
	▼form	orth	El ix cap. De los servidores	
	▼sense	@xml:id	QuinLin_34_De_los_servidores.s.1.	
	▼cit	@type	translationEquivalent	
		@xml:id	it	
		▼form	orth	El ix capitolo de li servitori

Imagen 1

El segundo paso en nuestro proceso de codificación implica la estructuración de las entradas secundarias. Para lograr esto, anidamos una nueva etiqueta <entry> dentro de la <entry> principal. Esta entrada anidada se distingue de la principal mediante el atributo @type="relatedEntry". Recordamos que TEI Lex-0 recomienda el uso de este atributo en diccionarios alfabéticos tradicionales para anidar entradas pluriverbales bajo un lema univerval. Aunque nuestro caso difiere

conceptualmente del uso para el que está pensado en los diccionarios alfabéticos, la funcionalidad del atributo es similar: anidar una entrada dentro de otra. Por esta razón, hemos decidido adoptar esta convención en nuestra codificación de nomenclaturas, adaptándola a nuestras necesidades específicas.

Además, a cada una de las `<entry type="relatedEntry">` le añadimos un atributo nuevo con respecto a los tres que tienen las `<entry>` principales; este nuevo atributo (`@n=""`) tiene como función identificar el lugar que ocupa cada lema en la serie:

```
<entry xml:id="QuinLin_34_De_los_servidores_peon" xml:lang="es" type="relatedEntry" n="1">
```

Lógicamente, cada `<entry>` anidada tiene sus propios elementos hijos: `<form>`, `<gramGrp>` y `<sense>`:

```
<entry xml:id="QuinLin_34_De_los_servidores" xml:lang="es" type="mainEntry">
<!-- etc. -->
  <entry xml:id="QuinLin_34_De_los_servidores_peon" xml:lang="es" type="relatedEntry" n="1">
    <form>
      <orth norm="peón">peon</orth>
    </form>
    <sense xml:id="QuinLin_34_De_los_servidores_peon.s.1.">
      <cit type="translationEquivalent" xml:id="it">
        <form>
          <gramGrp>
            <gram type="pos" norm="sustantivo"/>
            <gram type="gender" norm="masculino">il</gram>
          </gramGrp>
          <orth>fante à pie</orth>
        </form>
      </cit>
    </sense>
  </entry>
</entry>
```

entry	@xml:id	QuinLin_34_De_los_servidores						
	@xml:lang	es						
	@type	mainEntry						
	form	orth	El ix cap. De los servidores					
	sense	@xml:id	QuinLin_34_De_los_servidores.s.1.					
		cit						
	entry	@xml:id	QuinLin_34_De_los_servidores_peon					
		@xml:lang	es					
		@type	relatedEntry					
		gn	1					
		form	gramGrp					
			orth	#norm	peón			
				#text	peon			
		sense	@xml:id	QuinLin_34_De_los_servidores_peon.s.1.				
			cit	@type	translationEquivalent			
			@xml:id	it				
		form	gramGrp	gram	@type	#norm	#text	
				(2 rows)	1 pos	sustantivo		
					2 gender	masculino	il	
			orth	fante a pie				

Imagen 2

El atributo `@xml:id` desempeña un papel crucial. Su función principal es proporcionar un identificador único para cada elemento al que se aplica. Esto es especialmente importante por las siguientes razones:

- **Identificación única:** Cada `@xml:id` asignado a un elemento `<entry>` o `<sense>` es único dentro del documento, lo que permite referenciar específicamente ese elemento sin ambigüedades.
- **Relaciones entre elementos:** Los `@xml:id` permiten establecer relaciones entre diferentes partes del documento. Por ejemplo, otros elementos podrían hacer referencia a estas entradas utilizando estos identificadores únicos.
- **Trazabilidad:** Facilita el seguimiento y la vinculación de información relacionada a lo largo del documento o incluso entre diferentes documentos.
- **Procesamiento de datos:** Permite un procesamiento más eficiente de los datos, ya que cada elemento puede ser localizado y manipulado fácilmente mediante su identificador único.
- **Creación de enlaces:** En un contexto más amplio, estos identificadores pueden utilizarse para crear hipervínculos o referencias cruzadas en la versión digital del diccionario.

En el anterior fragmento, podemos ver cómo los `@xml:id` se utilizan tanto para la entrada principal (“QuinLin\_34\_De\_los\_servidores”) como para la entrada relacionada (“QuinLin\_34\_De\_los\_servidores\_peon”), así como para los elementos `<sense>` dentro de estas entradas. Esto permite establecer una clara jerarquía y relación entre los diferentes componentes del diccionario, facilitando su organi-

zación, recuperación y presentación.

Reviste también especial importancia para la codificación de las nomenclaturas el atributo `n=""`, ya que permite numerar los lemas que componen la serie. Si *peon* ocupa el número 1, *mochacho* ocupa el 8, por lo que la codificación será:

```
<entry xml:id="QuinLin_34_De_los_servidores_machacho" xml:lang="es"
type="relatedEntry" n="8">
```

Esta entrada da pie a una cuestión filológica de suma importancia en la edición de nomenclaturas antiguas y, en general de todo texto metalingüístico. Es sabido que especialmente este tipo de lexicografía “menor” multilingüe suele presentar numerosos errores. TELEI normaliza internamente todas las formas léxicas de manera que la base de datos las relacione en el seno del corpus lexicográfico. Para ello utilizamos el atributo `norm=""`:

```
<form>
  <orth norm="muchacho">mochacho</orth>
</form>
```

Notemos, por último, que hemos etiquetado el artículo como una marca gramatical, a la misma sazón de las marcas *m.* (masculino) o *adj.* (adjetivo):

```
<form>
  <gramGrp>
    <gram type="pos" norm="sustantivo"/>
    <gram type="gender" norm="masculino">il</gram>
  </gramGrp>
  <orth>fanciullo</orth>
</form>
```

## 5. Conclusión

La digitalización de nomenclaturas multilingües español-italiano utilizando TEI Lex-0 puede representar un avance significativo en la preservación y accesibilidad de estos recursos lexicográficos. La adaptación del esquema de codificación TEI Lex-0 a las particularidades estructurales de las nomenclaturas, especialmente en lo que respecta a la organización temática y la relación jerárquica entre entradas principales y relacionadas, permite mantener la integridad conceptual de estas

obras al integrarlas en un Tesoro lexicográfico. La utilización de atributos como @xml:id, @type, y @n facilita no solo la identificación unívoca de cada elemento, sino también la salvaguardia del orden original y de las relaciones entre los lemas. Además, la normalización de formas léxicas mediante el atributo @norm permite abordar las variaciones ortográficas y los errores comunes en estos textos históricos, mejorando su interoperabilidad y utilidad para la investigación lingüística y lexicográfica. Este enfoque de codificación, que equilibra la fidelidad al texto original con las necesidades de estandarización y recuperación digital, ofrece nuevas posibilidades para la exploración y análisis de estos importantes testimonios de la historia lingüística y cultural.

## Bibliografía citada

### Fuentes primarias

- Ainsa Royo, Manuel (1837), *Nuevo vocabulario de los idiomas modernos*, Barcelona, Imprenta de Valentín Torras. [Se encuentra en *Gramatica y vocabulario en castellano, inglés, francés é italiano*].
- Andree, John (1725), *A Vocabulary in Six Languages*, London, Vaillant and Meadows.
- Barnabé, Stephan (1657), *Unterweisung Der Spanischen Sprach*, Wien, Matthäus Rickhes.
- Boardman, James (1811?), *A Vocabulary in the English, Latin, German, French, Italian, Spanish, Portuguese*, London, White and Lewis.
- Borja, Ignacio de (1718), *Vocabulario Español, Italiano y Tudesco*, Wien. [Manuscrito].
- Chirchmair, Mattia (1734), *Nomenclatura spagnuola e italiana*, Firenze, Bernardo Paperini. [Se encuentra en *Grammatica spagnuola, e italiana*].
- Comenius, Joannes Amos (1661), *Janua linguarum reserata quinque linguis*, Amsterdam, Ludovicus et Danielis Ezelviros.
- Corona Bustamante, Francisco (s.f.), *Vocabulario*, Paris, Garnier. [Se encuentra en *Manual de Conversación y del estilo epistolar para el uso de los viajeros y de la juventud de las escuelas*].
- Costanzo, Salvador (1848), *Vocabulario italiano español*, Madrid, J. Llorente. [Se encuentra en *Gramatica italiana y española*].
- Decimator, Heinrich (1596), *Tertia pars Sylvae vocabulorum et phrasium, sive Nomenclator. Continens permultas omnium rerum, quae in probatis omnium dictionarum auctoribus inveniuntur*, Leipzig, Excudebat Michael Lantzenbergerus, Sumptibus et Impensis

Heningi Grosii.

- Foulques, Alessandro Edoardo (1898), *Nomenclatura per la conversazione italiana-spagnuola*, Milán, Bietti. [Se encuentra en *Nomenclatura e dialoghi per la conversazione italiana-spagnuola*].
- Franciosini, Lorenzo (1626), *Nomenclator*, Venezia, Giacomo Sarzina. [Se encuentra en *Dialogos apazibles, compuestos en Castellano y traduzidos en Toscano*].
- Frutos, José de (1804), *Diccionario manual de voces en las cinco lenguas española italiana francesa inglesa y latina*, Madrid, Gómez Fuentenebro y compañía.
- Giulliani, Blaise-Christophe (1659), *Nomenclature du Sieur Iulliani qui contient les mots exacts et choisis pour apprendre les langues Françaises, Italiennes et Espagnoles. Ensemble les dialogues familiers*, Paris, Jean Baptiste Loyson.
- Habrecht, Isaac (1629), *Ianua linguarum Silinguis*, Strasbourg, Eberhardus Zetznerus.
- Hervás y Panduro, Lorenzo (1797), *Vocabulario italiano español*, s.l. [Original manuscrito, se encuentra en *Gramatica de la lengua italiana*].
- Howell, James (1659), *A particular Vocabulary or Nomenclature In English, Italian, French, and Spanish*, London, Tomas Leech. [Se encuentra en *Lexicon Tetraglotton, an English-French-Italian-Spanish Dictionary*].
- Junius, Hadrianus (1567), *Nomenclator, omnium rerum propria nomina variis linguis explicata indicans*, Antwerpen, Christophori Plantini.
- Kramer, Matthias (1670), *Nomenclatura Toscana, Tudesca y Espagnole de aquellas palabras o vocablos que se usan cada día*, Nuremberg, Wolf Eberhard Felsecker.
- López de Morelle, José (1851), *Nomenclatura del Arte completo de la lengua italiana*, Madrid, Imprenta y oficina literaria de D. José López de Morelle. [Se encuentra en *Arte completo de la lengua italiana. Nuevo método para aprenderla sin necesidad de más libros y aunque sea sin maestro*].
- Marín, Francisco (1833), *Racolta di nomi sostantivi più necessari a sapersj, Racolta di verbi più necessari a sapersj*, Roma, Tipografia Salviucci. [Se encuentra en *L'italiano istruito nella cognizione della lingua spagnuola*].
- Noviliers Clavel, Guillaume Alexandre (1629), *Nomenclatura italiana, francese e spagnola con i termini propri di ciascun capitolo*, Venezia, Barezzi.
- Ochoa, Eugenio de (1842), *Vocabulario español italiano*, Paris, Madrid, Carlos Hingray, Casa de C. Monier. [Se encuentra en *Guía de la conversación español-italiano al uso de los Viageros y de los Estudiantes*].
- Perles y Campos, José Faustino (1689), *Nomenclatura*, Napoli, Perrino y Mutii. [Se encuentra en *Gramática española o modo de entender, leer y escribir español*].
- Pomey, François Antoine (1705), *Universo abreviado, adonde están contenidos en diversas listas casi todos los nombres de las obras de la Naturaleza, de todas las Ciencias, y de todos los Artes, con sus principales términos*, Lyon, Pierre Valfrey.
- Quinque linguarum utilissimus vocabulista. Latine. Tusche. Galliche. Hispane et Alemanice* (1526), Venezia, Francesco Garone.
- Quinque Linguarum, Latinae, Teuthonicae, Gallicae, Hispanicae, Italicae* (1534), Augsburg,

Stelsius.

- Rosterre, Esteban de (1771), *Colección o resumen de los nombres y verbos más necesarios para la locución y conversaciones ordinarias*, Forlì, Achilles Morozzi. [Se encuentra en *Reglas a cerca de la lengua toscana o italiana*; seudónimo de Esteban de Terreros].
- Tomasi, Pietro (1779), *Catálogo de nombres más necesarios para hablar el italiano*, Madrid, Manuel Martín. [Contenida en *Nueva y completa gramática italiana explicada en español*].
- Vergani, Angelo (1826), *Colección de los nombres más usuales*, Madrid, Miguel de Burgos. [Se encuentra en *Gramática italiana simplificada y reducida a 20 lecciones con sus respectivas prácticas, diálogos, y una colección de trozos históricos en italiano*].
- Vocabolario (s.f.), Milano, Oreste Ferrario. [Se encuentra en *Il perfetto dialoghista italiano e spagnuolo: manuale della conversazione dello stile epistolare ad uso del viaggiatore e degli studiosi della lingua spagnuola*. en Google boks en mi biblioteca
- Vocabulaire Europeen, Six Langues par Abc [...] français, anglais, allemand, espagnol, italien, portugais* (1878), Paris, Catrony.
- Zumarán, Juan Ángel de (1621), *Nomenclatura en cuatro lenguas de las cosas más necesarias y comunes* [Se encuentra en *Libro muy provechoso para aprender las lenguas*], München, Witwe von Berg.

## Referencias citadas

- Alvar Ezquerro, Manuel (1993), “Diccionarios ideológicos”, *Lexicografía descriptiva*, Barcelona, Biblograf: 289-301.
- Alvar Ezquerro, Manuel (2008), “Presentación del «Nuevo tesoro lexicográfico del español» (S. XIV-1726)”, *Real Academia Española*, [11/12/2024] <<https://www.rae.es/noticia/presentacion-del-nuevo-tesoro-lexicografico-del-espanol-s-xiv-1726>>.
- Alvar Ezquerro, Manuel (2013), *Las nomenclaturas del español. Siglos XV-XIX*, Liceus.
- Ayala Castro, Marta Concepción (1992), “Nomenclatures de l’espagnol (1526-1800). Considérations générales sur la nature et la fonction des nomenclatures”, *Cahiers de lexicologie*, 61/2: 127-160.
- Carranza Torrejón, Ana M.<sup>a</sup> (2017), “Nomenclatura”, *Diccionario de historia de la enseñanza del francés en España (siglos XVI-XX) (DHEFE)*, ed. Juan García Bascuñana. Tarragona, Universitat Rovira i Virgili [3/12/2024] <<http://www.grelinap.recerca.urv.cat/ca/projectes/diccionario-historia-ensenanza-frances-espana/entradas/174/nomenclatura>>.
- Castillo Peña, Carmen (2004), “La lengua española en la *Nomenclatura italiana, francesa y española* de Guillaume Alexandre de Noviliers Clavel”, *Lexis*, 28: 105-138.
- Castillo Peña, Carmen (2005), “La *Nomenclatura italiana, francesa y española* de Guillaume Alexandre de Noviliers Clavel: La cuestión de las fuentes”, *Quaderni del CIRSIL*, 4:

1-14.

- Gallina, Annamaria (1959), *Contributi alla storia della lessicografia italo-spagnola dei secoli XVI e XVII*, Firenze, Leo S. Olschki.
- García Aranda, María Ángeles (2003), *Un capítulo de la lexicografía didáctica del español: nomenclaturas hispanolatinas (1493-1745)*, Tesis Doctoral, Universidad Complutense de Madrid.
- García Aranda, María Ángeles (2004), “La Ianua linguarum de Jan Amos Comenio, un capítulo de la lexicografía onomasiológica”, *Nuevas aportaciones a la historiografía lingüística: actas del IV Congreso Internacional de la SEHL*, eds. Cristóbal Corrales Zumbado et al. Madrid, Arco/Libros: 551-560.
- García Aranda, María Ángeles (2007), “La clasificación conceptual del léxico en repertorios de los siglos XVI y XVII. Origen y evolución”, *Historia de la lexicografía española*, eds. M. Campos Souto; R. Cotelo García, J. I. Pérez Pascual. A Coruña, Universidade da Coruña: 69-76.
- García Aranda, María Ángeles (2011), “La lexicografía menor ítalo-española”, *Ideolex. Estudios de lexicografía e ideología*, eds. Félix San Vicente; Cecilio Garriga; Hugo E. Lombardini. Monza, Polimetrica: 293-303.
- Garriga, Cecilio (2003), “La microestructura del diccionario: las informaciones lexicográficas”, *Lexicografía española*, ed. Antonia María Medina Guerra. Barcelona, Ariel: 103-126.
- Gili Gaya, Samuel (1960), *Tesoro lexicográfico, (1492-1726)*, Madrid, Consejo Superior de Investigaciones Científicas.
- Lombardini, Hugo E. (2023), “La entrada lexicográfica en el diccionario bilingüe”, *Lexicografía hispánica / The Routledge Handbook of Spanish Lexicography*, eds. Sergi Torner; Paz Battaner; Irene Renau. London, Routledge: 251-273.
- Marello, Carla (1980), *Lessico ed educazione popolare. Dizionari metodici italiani dell'800*, Roma, Armando Armando.
- Nieto Jiménez, Lidio; Manuel Alvar Ezquerro (2007), *Nuevo Tesoro Lexicográfico del español*, Madrid, Arco Libros.
- Pablo Núñez, Luis (2008), *Lexicografía hispano-francesa de los siglos XVI y XVII: catálogo y estudio de los repertorios*, Tesis Doctoral, Universidad Complutense de Madrid.
- Peñín Fernández, Natalia (2023), *Catálogo analítico crítico de la lexicografía italoespañola*, Bologna, CLUEB.
- Quemada, Bernard (1967), *Les Dictionnaires du français moderne: 1539-1863*, Paris, Didier.
- Redondo Rodríguez, María Jesús (2007), “El *Quinque / Lingvarum, Latinae, / Teuthonicae, Gallicae, Hispanicae, Italicae, / Dilucidissimus Dictionarius* (...) de J. Steelsius, heredero del *Introito e Porta* de A. Rottweil y base de los anónimos plurilingües del siglo XVI”, *Cuatrocientos años de la lengua del Quijote: estudios de historiografía e historia de la lengua española. Actas del V Congreso Nacional de la Asociación de Jóvenes Investigadores de Historiografía e Historia de la Lengua Española* (Sevilla; 31 de marzo, 1 y 2 de abril



- de 2005), eds. Marta Fernández Alcaide; Araceli López Serena. Sevilla, Universidad de Sevilla: 533-546.
- Rossebastiano Bart, Alda (1984), *Antichi vocabolari plurilingui d'uso popolare: la tradizione del «Solenissimo vocabuolista»*, Alessandria, Edizioni dell'Orso.
- Salgado, Ana *et al.* (2024), "The Morais Dictionary: Following Best Practices in a Retro-digitized Dictionary Project", *International Journal of Humanities and Arts Computing*, 18: 125-147.
- Tasovac, Toma *et al.* (2001), "TEI Lex-0 — A baseline encoding for lexicographic data", *DARIAH Working Group on Lexical Resources*, [24/11/2024] <<https://dariah-eric.github.io/lexicalresources/pages/TEILex0/spec.html#TEI.seg>>.
- Vázquez, Gloria; Ana Fernández-Montraveta (2023), "Los diccionarios ideológicos", *Lexicografía hispánica / The Routledge Handbook of Spanish Lexicography*, eds. Sergi Torner; Paz Battaner; Irene Renau. London, Routledge: 551-564.

**Carmen Castillo Peña** es catedrática de Lengua española y traducción en el Departamento de Studi Linguistici e Letterari de la Universidad de Padua. Se dedica a la historiografía lingüística del español en Italia con particular atención a la gramaticografía y a la historia de la lexicografía. Sus últimos proyectos se centran en la aplicación de las humanidades digitales a la edición de gramáticas antiguas y al patrimonio lexicográfico.  
**carmen.castillo@unipd.it**



---

## SECCIÓN GENERAL

**Maria Vittoria Ambrosini**, *¿Debo hacerlo o Tengo que hacerlo?*

Reflexionar sobre la polisemia modal  
desde una perspectiva interlingüística

**Sonia Bailini**, *I diari della motocicletta*,

un viaggio nella traduzione intersemiotica e interlinguistica

**Elena Liverani**, Lanzarote e Saramago nelle parole di una traduttrice

d'eccezione: *La intuición de la isla* (2022) di Pilar del Río

e i problemi di resa in lingua italiana

**Marisa Martínez Pérsico**, *Embodiment* y metáforas somáticas en la autotraducción

bilingüe. El caso de la poeta vasca Miren Agur Meabe

desde el enfoque de los *Cognitive Translation Studies*



# MARIA VITTORIA AMBROSINI

## ¿DEBO HACERLO O TENGO QUE HACERLO? REFLEXIONAR SOBRE LA POLISEMIA MODAL DESDE UNA PERSPECTIVA INTERLINGÜÍSTICA

Università del Salento

### Resumen

El estudio parte de un análisis contrastivo donde se observa que el verbo italiano *dovere* puede expresar modalidad deóntica o dinámica, modalidades que — en español — se reparten entre dos verbos, *deber* y *tener que*, según el posicionamiento del hablante. Se analizan, en clave adquisicional, las interferencias que estas premisas pueden ocasionar en la interlengua de estudiantes de español itálofonos y se propone un "mapa modal" como herramienta para el desarrollo de su *competencia modal* en L2.

palabras clave: modalidad, español L2, *deber*, *tener que*, interlengua

### Abstract

¿Debo hacerlo or Tengo que hacerlo? *Reflecting on modal polysemy from an interlinguistic perspective*

The study adopts a contrastive approach, showing that the Italian verb *dovere* may express either deontic or dynamic modality, whereas in Spanish these values are distributed between *deber* and *tener que*, depending on the speaker's stance. From an acquisitional perspective, it examines potential interlanguage interferences in Italian-speaking learners of L2 Spanish and proposes a "modal map" as a tool to support the development of their L2 modal competence.

keywords: modality, L2 Spanish, *deber*, *tener que*, interlanguage

## 1. Introducción

Este estudio nace a partir de una pregunta que queda significativamente reflejada en su título: ¿en qué consiste la diferencia entre dos frases como *Debo hacerlo y Tengo que hacerlo*? Está claro que se trata de dos perífrasis obligativas y no cabe duda de que cualquier hablante nativo sabría cuál de las dos escoger, según un contexto dado. Sin embargo, podría resultarle menos fácil explicar el criterio que le habría llevado a quedarse con una y a descartar la otra. Al formar parte de su repertorio de lengua materna (de ahora en adelante, L1), es probable que nunca le haya surgido esa pregunta y que aplique esa regla de manera automática; pero muy distinto es el caso de quien se enfrente a estas fórmulas como parte de un repertorio de lengua segunda (de ahora en adelante, L2), más todavía si solo una de las dos cuenta con una correspondencia formal en su L1. Consideremos los ejemplos *debo hacer lo correcto* y *tienes que salir ya o llegarás tarde* y pongámonos en la perspectiva de un estudiante de español italoéfono que necesite entender por qué la L2 requiere el uso de *deber* en la primera frase y de *tener que* en la segunda. Es muy posible que la presencia de la forma *debo*, en la primera frase, le recuerde automáticamente la forma *devo*, en su L1, llevándole a una traducción literal del mensaje de partida; pero *dovere* también se prestaría a ser empleado en una traducción de la segunda frase. Ahora bien, el resultado serían dos frases en L1, *devo fare la cosa giusta* y *devi partire subito, o arriverai in ritardo*, en las que quedarían “neutralizadas”, a través del uso unificado del verbo *dovere*, las distintas funciones modales que en la L2 quedan reflejadas, a nivel formal, en la alternancia entre *deber* y *tener que*. Así pues, en este caso, una traducción hacia la L1 no le aportaría al estudiante ningún elemento de reflexión útil para entender la L2. Es más, una estrategia de traducción podría llegar a ser contraproducente si se invirtiera la dirección, es decir a la hora de actuar en L2: por influencia de la L1, el estudiante podría convencerse de que *deber* funciona como *dovere* y puede ser empleado en los mismos contextos. Por lo tanto, podría traducir el ejemplo anterior –*devo fare la cosa giusta*– como *debo hacer lo correcto*, pero también una frase como *mi spiace doverti disturbare* como *siento deber molestarte\**, una solución evidentemente problemática –a oídos de un nativo– desde un punto de vista pragmático. Eso se debe a que *debo hacer lo correcto* es una expresión de obligación deóntica, vehiculada por el mismo verbo modal (*deber* y *dovere*) tanto en español como en italiano; en cambio, *tienes que salir ya* y *devi partire subito* expresan más bien necesidad dinámica, vehiculada –en las dos lenguas– por verbos modales distintos (*tener que* y *dovere*). Estos pocos ejemplos ya señalan unas cuestiones in-

terés desde un punto de vista contrastivo, que previsiblemente pueden tener un impacto relevante en los procesos de interlengua. El objetivo de este estudio, por lo tanto, es ofrecer unas claves de respuesta a la pregunta que lo intitula y que podrían (y tal vez deberían) ser objeto de una reflexión específica en el aprendizaje de español L2 por parte de italófonos.

El presente estudio está estructurado de la manera siguiente: en §2, se ofrecen algunas aclaraciones teórico-terminológicas preliminares sobre modalidad deóntica y modalidad dinámica (Palmer 1990; 2001; Nuyts 2006; Lyons 1981; Van der Auwera, Plungian 1998; Conte 1995); en §3, nuestra atención se centra en los verbos modales *deber* y *tener que* y en un análisis de sus empleos deónticos y dinámicos (Fernández de Castro 1999; Yllera 1980; Real Academia Española [RAE] 2010; Garachana Camarero 2017; Flores Salvo, Jaque Hidalgo 2024); sigue, en §4, una reflexión sobre el verbo italiano *dovere* y sobre sus usos deónticos y dinámicos (Bertinetto 1979; Skytte *et al.* 1991; Samu 2005). Las premisas teóricas presentadas en §2, §3 y §4 constituyen el eje en torno al cual se desarrollan las reflexiones de carácter contrastivo y adquisicional (Myers 2004; Amenós-Pons *et al.* 2017; Ambrosini, Della Putta 2021; Ambrosini, 2018; Carrera Díaz 1979; Schmid 1994), que se presentan en §5: dicho análisis se propone arrojar luz sobre las interferencias que podrían producirse, en la interlengua de un estudiante de español con L1 italiana, a la hora de enfrentarse a una L2 en la que los valores deónticos y dinámicos del verbo *dovere* se “recodifican” –dependiendo del contexto dado– mediante dos exponentes distintos: o *deber* o *tener que*. Este análisis contrastivo se concluye con un pequeño *mapa modal*, que traduce las observaciones que se han desarrollado a lo largo de este estudio en una propuesta de reflexión práctica pensada para un contexto didáctico de español L2 para italófonos. Cierran el artículo, en §6, algunas consideraciones finales sobre la importancia de proponer, en tal contexto, un enfoque que combine plano formal y plano semántico (Venier 1991; Skytte 1995), ya que solo un enfoque de este tipo puede permitir el desarrollo, en el estudiante, de una *competencia modal* que le permita “manejar” las interferencias interlingüísticas y acercarse de manera más eficaz a un uso nativo de los modales *deber* y *tener que*.

## 2. Modalidad deóntica y modalidad dinámica: premisas teóricas

Los seres humanos evalúan continuamente el estado de la realidad que les rodea o empujan a sus interlocutores a modificarlo (o les dan permiso para hacerlo),

traduciendo constantemente estas actividades en palabras: en este sentido, según Kiefer (1987: 73), la modalidad podría verse como la actitud cognitiva, emotiva o volitiva del hablante hacia una situación. Cabe señalar que la literatura científica todavía no cuenta con una definición unívoca del concepto de modalidad, ni con un consenso sobre la composición u organización interna del dominio modal. Dicho esto, abrimos este estudio con unas aclaraciones terminológicas básicas<sup>1</sup>.

## 2.1 *Modalidad deóntica*

La modalidad deóntica indica el grado de deseabilidad moral de la situación expresada en el enunciado (Nuyts 2006: 4). Como observa Palmer (2001), la modalidad deóntica se expresa a través de actos comunicativos directivos (Searle 1975) que pueden adquirir forma de permiso (1) u obligación (2):

- (1) You may/can go now. (Palmer 2001: 71)
- (2) You must go now. (Palmer 2001: 71)

Lyons (1981: 235-42) propone una distinción entre modalidad deóntica subjetiva y objetiva: en el primer caso, la fuente de autoridad coincide con el hablante (3), mientras que, en el segundo, la fuente de la obligación o permiso es una autoridad externa (4)<sup>2</sup>:

- (3) You may go now.
- (4) You may not smoke in a public space.

## 2.2 *Modalidad dinámica*

El término *dinámica* es adoptado por Palmer (1990) para abordar todos los contextos que expresan capacidades del sujeto (5) o su expresión de una voluntad personal (6). Todos ellos forman parte de la que Palmer define modalidad dinámica *subject-oriented*:

1 Por razones de espacio, en esta sección se omite el análisis de las modalidades epistémica y aléctica, remitiendo al lector a Nuyts (2006), Von Wright (1951), Lyons (1981) y Pietrandrea, Cervoni (2015) para una discusión más detallada al respecto.

2 Entre otros autores que sugieren una distinción según la fuente del deóntico, cfr. Bybee *et al.* (1994) y Verstraete (2007).



(5) I can make or break my life and myself. (1990: 85)

(6) I'll help you. (1990: 111)

Según Palmer (1990: 911), también confluye en la modalidad orientada al sujeto “the indication of a need or necessity for the first-argument participant”, es decir una necesidad inherente del sujeto, la que podríamos apreciar en un ejemplo como (7):

(7) I have to take less sleeping medication.

Junto a la modalidad orientada al sujeto, Palmer incluye en el dominio dinámico otro tipo de modalidad, que caracteriza una posibilidad o necesidad inevitables inherentes al contexto proposicional: se trata de la modalidad *circunstancial*, ejemplificada en (8):

(8) I've got to be at London airport at fourish. (1990: 114)

La modalidad circunstancial está por lo tanto conectada tanto con la modalidad dinámica – en cuanto expresa una característica inherente no del participante, sino de la situación – como con la modalidad deóntica, ya que en muchos contextos circunstanciales también es posible inferir alguna forma de obligación. Van der Auwera y Plungian (1998), de hecho, consideran la modalidad deóntica como un caso especial de modalidad circunstancial (que definen como *externa al participante*), con la diferencia de que –en la primera– la posibilidad o necesidad son impuestas por una persona o una institución, mientras que –en la segunda– son inherentes a las circunstancias físicas del contexto. Conte (1995: 3-9), por su parte, sugiere distinguir entre modalidad deóntica y modalidad *anankástica*. La segunda caracterizaría aquellos contextos –típicos del ámbito normativo– en los que realmente no se expresa ninguna obligación o prohibición, sino que simplemente se establece una condición requerida por las circunstancias (coincidiendo, por ende, con la modalidad circunstancial de Palmer). Obsérvese el ejemplo siguiente (1995:7):

(9) Il registro di laurea *deve essere* firmato dagli 11 membri della commissione. **Questo requisito** per la validità del verbale è ineliminabile.

Nótese que, como subraya Conte, la primera frase del ejemplo puede dar lugar a una lectura de tipo deóntico si aparece encapsulada en un contexto diferente, como en (10):

(10) Il registro di laurea *deve essere* firmato dagli 11 membri della commissione. Nessun commissario può sottrarsi a **quest'obbligo**.

Estas observaciones sugieren que hay contextos en los que la frontera entre modalidad deóntica y dinámica puede hacerse tal vez un tanto borrosa, como veremos más en detalle en §3.

### 3. *Deber y tener que*: en la frontera entre deonticidad y dinamicidad<sup>3</sup>

Según observa Fernández de Castro (1999: 185), existe cierta tendencia a considerar las distintas construcciones obligativas disponibles en español como variantes prácticamente sinonímicas y, por lo tanto, intercambiables entre ellas:

En el caso de *tener que* + infinitivo, sin embargo, es unánime la observación de que expresa una necesidad u obligación especialmente intensa o ineludible, lo cual la diferenciaría de sus cercanas *deber* o *haber de* + infinitivo. Ahora bien, ¿es oportuno hablar de una diferencia de grado, también en el eje semántico de la necesidad dinámica o deóntica? ¿Se puede decir que los hablantes señalen unos hechos como *más* necesarios que otros?

Estas dudas introducen el asunto que se intentará aclarar en el presente apartado: ni estos dos verbos modales son intercambiables, en la mayoría de los contextos obligativos, ni es cierto que uno de los dos sea más “fuerte” que el otro. En cambio, como veremos, cada uno presenta su propia gama de usos modales, su propio *mapa modal* –por así decirlo– que procede de operaciones semánticas distintas y estrechamente relacionadas con el posicionamiento del hablante frente a lo que –caso por caso– se *deba* o *tenga que* hacer.

3 Conviene precisar que cuando en este trabajo se hace referencia al uso de *deber* + infinitivo y *tener que* + infinitivo desde una perspectiva normativa se alude a las descripciones tradicionalmente ofrecidas por las gramáticas y los manuales de referencia. Dichas descripciones representan estándares de la lengua que no siempre se corresponden plenamente con las prácticas discursivas de todos los hablantes, entre los cuales cabe esperar una variación interindividual ligada a factores diatópicos, diastráticos y diafásicos. Asimismo, los ejemplos analizados corresponden al español peninsular escrito (y al italiano); no se abordan aquí las variedades del español americano, cuya consideración se reserva para futuros estudios.

### 3.1 *Deber*

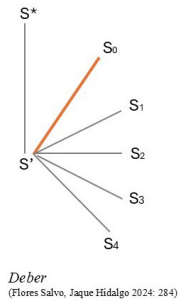
Como señala Yllera (1980: 128), *deber* ya es empleado –solo, o en combinación con la preposición *de*– en el español del siglo XIII para expresar obligación “entendiendo este concepto en sentido amplio”: por ejemplo, obligación moral (11) u obligación impuesta por una ley, un acuerdo o una conveniencia (12)<sup>4</sup>:

(11) nin servir a Dios según *deben de façer* [...]. (1980: 128)

(12) ca de otra guisa nin âmigo nin âmiga con *deuedes de decir* vuestra poridat. (1980: 128)

Yllera también observa que –a diferencia de otras– ‘*deber* + infinitivo’ non constituye una perífrasis obligativa propiamente dicha, ya que, dentro de la misma, no se produce ninguna deslexicalización del auxiliar *deber*, que conserva intacto su valor semántico original, *inherentemente* directivo. También la RAE (2010: 539) reconoce que “la obligación expresada por ‘*deber* + infinitivo’ está en consonancia con lo generalmente deseable o con normas aceptadas que así lo imponen”, lo que corresponde a la definición de modalidad deóntica ilustrada en §2.

Basándose en la propuesta de Ramchand (2018), Flores Salvo y Jaque Hidalgo (2024: 278-294) sugieren una lectura en clave lógico-semántica de *deber* como operador de *selección exclusiva*, según ilustra el Esquema 1:



<sup>4</sup> La norma tradicional recomienda el uso de *deber* + infinitivo con valor deóntico (*obligación*) y de su variante preposicional *deber de* + infinitivo con valor epistémico (*conjetura*), si bien también reconoce que usos de *deber* + infinitivo para expresar conjetura (RAE 2010: 539) y/o de *deber de* + infinitivo en sentido obligatorio (Fernández de Castro 1999: 173-176) están ampliamente extendidos entre los hablantes. Diversos autores sostienen, por ende, que tal distinción entre ambas formas constituye una separación forzada, introducida artificialmente por los gramáticos modernos (Yllera 1980: 128; Fernández de Castro 1999: 173-176; Rosemeyer 2017: 147-196). Sin embargo, por motivos de brevedad y en atención al formato del presente estudio, no analizaremos esta alternancia en la lengua actual y nos centraremos únicamente en los usos obligativos de la variante no preposicional.

El verbo *deber* selecciona como cierre del evento de referencia (S') –desde la perspectiva evaluativa del hablante (S\*)– la alternativa S0 (señalada por la línea de selección naranja), que corresponde a la más altamente valuada a partir de un conjunto de alternativas distintas y ordenadas jerárquicamente según un parámetro deóntico (representadas en el esquema por S0, S1, S2, S3, S4). Por poner un ejemplo, en la frase *Marta debe perdonar a Juan*, *deber* asigna a la alternativa *perdonar a Juan* la valoración más alta a nivel deóntico, entre las disponibles para Marta (por ejemplo, *no hacer nada*, *dejar a Juan*, *vengarse*, etc.). Distinta es la selección de alternativas operada por *tener que*, como veremos a continuación.

### 3.2 *Tener que*

Según la RAE (2010: 539), también *tener que* expresa obligación, aunque esta es “generalmente presentada como necesidad externa inevitable o impuesta por las circunstancias, de tal forma que el sujeto no puede sustraerse a ella: *Si no se comporta tendré que medicarle* (Cabal, Briones)”. En este sentido, *tener que* parece expresar más bien una necesidad dinámica (§2.2), que una obligación deóntica (§2.1). Esta es, efectivamente, la modalidad vehiculada por *tener que* no solo en contextos circunstanciales, sino también en contextos orientados al sujeto. Significativamente, mientras que, según lo explicado en §3.1, la obligatoriedad expresada por *deber* siempre está en línea con cierto criterio ético-normativo, no vale lo mismo para la necesidad expresada por *tener que*: “Podría decirse, por ejemplo, *Las autoridades a veces tienen que cometer actos ilegales* si se quiere expresar que se ven forzadas a ello, pero sería más extraña la variante ... *deben cometer actos ilegales*, porque expresaría que cierta norma superior exige una conducta ilícita” (2010: 549).

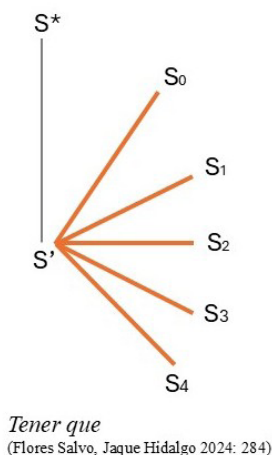
Según Garachana Camarero (2017) el valor dinámico<sup>5</sup> de la perífrasis *tener que* + infinitivo nace en construcciones que se remontan al siglo XIII, en las que *tener* se combina con verbos plenos de acción/actividad tales como *dar*, *comer*, *despender* y va típicamente antecedido por una negación o por un elemento cuantificador. En estos contextos, se activa una relectura metonímica de *tener* desde un significado original de *posesión* hacia uno de *finalidad*, como en (13):

5 Dado que Garachana Camarero considera *tener que* una perífrasis deóntica, consideramos recomendable precisar – cfr. la falta de consenso terminológico en el ámbito modal ya señalada en §1 – que la modalidad a la que la autora se refiere con *necesidad deóntica* corresponde – según la terminología adoptada en este estudio – a la modalidad *dinámica* de Palmer (§2).

(13) ca non como agora, que **nin tenemos pan que comer nin otra cosa ninguna**, nin solamiente ell agua que beber, nin vemos logar del mundo dond lo podamos aver. (Alfonso X, General Estoria I, siglo XIII) (2017: 259)

Con el paso del tiempo, en estas estructuras se produce una descategorización de *tener*, que se convierte en simple verbo de soporte, y de la palabra *que*, que deja de ser un relativo que introduce una cláusula con valor final y se convierte en un simple nexos entre las formas. Es así como se completa el proceso de gramaticalización de la nueva perífrasis verbal *tener que*.

Estas observaciones sugieren que el valor dinámico de *tener* y el valor deóntico de *deber* tienen orígenes semánticos muy distintos, que repercuten de manera evidente en sus funciones modales, como también se desprende de Flores Salvo y Jaque Hidalgo (2024), que oponen a la función de *deber* como operador de *selección exclusiva*—presentada en §3.1— la de *tener que* como operador de selección *exhaustiva*, ilustrada en el Esquema 2:



En un contexto de necesidad, no existe ningún tipo de jerarquía entre las alternativas abiertas o, mejor dicho, tan solo hay una alternativa (S0): eso significa que en todas las situaciones derivadas de S' (todas, indistintamente, marcadas por una línea de selección naranja) se produce el cierre existencial del evento de referencia. Retomando el ejemplo *Marta debe perdonar a Juan* y substituyendo *deber* por *tener que*, el resultado será un contexto en el que, de hecho, Marta “no tiene alternativa”: en todos los escenarios posibles, al final, ella perdona a Juan.

### 3.3 Zonas de solapamiento y posicionamiento del hablante

Completamos nuestro análisis de *deber* y *tener que* con algunas reflexiones que Fernández de Castro (1999: 176-189) propone en respuesta a sus interrogantes ya adelantados en §3. Como explica el autor, puede haber contextos en los que la presión causal ejercida por las circunstancias o cierta necesidad inherente a la naturaleza del sujeto son tales que este los percibe como obligativos, empleando el verbo *deber* en vez de *tener que*:

(14) Esto nos obliga ya, (...), a adoptar una serie de medidas de carácter especial que se corresponden con los análisis y estudios (...) de las diferentes variantes de lo que *debemos* hacer en cada momento, de acuerdo a la cantidad de combustible que dispongamos. (1999: 178)

A la inversa, si el hablante considera que determinada norma “por el prestigio de quien la impuso o por su pretendida conveniencia, forma parte de la mecánica del universo mundo, los juicios modales en que esa norma se manifieste ya no serán deónticos sino dinámicos, al menos para ese sujeto”. Como es fácil de entender, pues, habrá casos en los que el límite entre interpretación deóntica y dinámica se hace tan difuminado que casi llega a desaparecer, como en (15), donde se puede observar un uso casi sinonímico de los dos verbos modales:

(15) Y apostilló que si el ministro se va, todos los que apoyaron la ley también *tendrían* que dimitir, e igualmente *deberían* hacerlo los que la rechazaron [...]. (1999: 188)

Pero para que un mismo contexto pueda prestarse tanto a una interpretación deóntica como a una interpretación dinámica tiene que darse una de estas dos condiciones: o bien la fuente de autoridad es interna (es decir, coincide con el hablante mismo), o bien –aun siendo externa– cuenta con el respaldo del hablante<sup>6</sup>.

A la luz de las propuestas de Flores Salvo y Jaque Hidalgo ilustradas en §3.1 y §3.2, en estos casos es como si la selección exhaustiva operada por *tener que* y la selección exclusiva operada por *deber* pudieran ser ambas asociadas al mismo cierre del evento: dicho de otra forma, este cierre sería interpretable –alternativamente– o como *objetivamente* necesario (*tener que*) o como el más altamente valorado en una escala deóntica de alternativas posibles (*deber*). Efectivamente, en

6 Palmer también (1990: 106-107) propone distinguir los contextos de *obligación externa* de los contextos de *obligación interna*, asociando operativamente el verbo inglés *have (got) to* a los primeros y *must* a los segundos. Hay *obligación interna* no solamente cuando el hablante es la fuente de la misma, sino también cuando ha *interiorizado* una *obligación externa* (es decir, cuando la avala).

casos como (14) o (15), lo *que se tiene que hacer* coincide con lo que —según el hablante— también se *debe hacer*. ¿Pero qué pasa si no se da esta segunda condición?

Pues, en ese caso, solo puede seleccionarse el modal *tener que*, que entonces señalará bien una postura neutral del hablante con respecto a cuanto es requerido (por ley o por las circunstancias), como en (16), bien una postura no conforme con ello (17) (cfr. RAE, §3.2):

(16) Juanatey se volverá a sentar hoy en el banquillo [...]. Es la última cuenta que *tiene que* pagar a la justicia por los nueve años que fue drogadicto. (Fernández de Castro 1999: 187)

(17) Siento *tener que* molestarle. (1999: 187)

En (16) el periodista simplemente da a conocer lo dispuesto por una autoridad externa, manteniéndose neutral con respecto a ello. En este caso, el uso de *deber* en lugar de *tener que* habría marcado su posicionamiento deóntico con respecto a los acontecimientos, poniéndole en una postura probablemente poco acorde a su ética profesional. En (17) la selección de *tener que* exime al hablante de la responsabilidad por la molestia que este ya sabe que podría ocasionar, justificándola pragmáticamente como imputable a circunstancias externas. Según Fernández de Castro (1999: 188):

Este contenido es el realmente distinguido por *tener que* + infinitivo, y de ahí la tendencia a interpretar esta perífrasis como más intensamente obligativa que otras. Sin embargo, no se trata de una diferencia de grado, sino cualitativa: la obligación adquiere un relieve especial cuando se enfrenta a la voluntad del sujeto, relieve que desaparece cuando no hay tal discrepancia por ser la exigencia no conflictiva, aceptada a priori o asumida como propia.

Todo esto nos lleva a una conclusión muy importante: puede resultar poco útil (o hasta contraproducente) reducir la amplia gama de contextos que hemos estado observando en este apartado a una categorización dicotómica de tipo tradicional entre obligaciones deónticas y necesidades dinámicas. Nos parece más adecuado imaginar un *continuum* entre dos polos: el deóntico, cuya fuente es el hablante (2), por un lado, y el dinámico circunstancial (8) por el otro, entre los que se distribuye una serie de contextos intermedios en los que realmente es el posicionamiento del hablante, su actitud frente al contexto dado, el elemento que establece dónde colocar ese contexto entre los dos polos y, consecuentemente, si tal contexto requiere el uso de *deber* o *tener que*. Un panorama tan complejo, como ya se adelantaba en §1, merecería ser objeto de una reflexión específica en un contexto de español L2, más todavía si la interlengua se desarrolla a partir de

una L1 afín. En este sentido, presentamos, en el siguiente apartado, una reflexión sobre algunos valores modales del verbo italiano *dovere*.

#### 4. La polisemia modal del verbo *dovere*

De los dos verbos modales que hemos analizado en §3, solo *deber* cuenta con un exponente formalmente correspondiente en lengua italiana, el verbo modal *dovere*, que puede ser utilizado para expresar tanto modalidad dinámica (orientada al sujeto o circunstancial) como modalidad deóntica. Es lo que observa Bertinetto (1979: 121), inspirándose —entre otros— en Palmer, explicando que *dovere* presenta dos acepciones no-epistémicas:

1. necesidad lógica (“è (logicamente) necessario che”, “esiste la necessità che”)
2. obligación (“essere obbligati (o tenuti) a”, “aver l’obbligo di”)

En la necesidad *lógica* a la que se refiere Bertinetto, ejemplificada en (18) y (19) se ve efectivamente reflejada la modalidad circunstancial de Palmer (§2.2):

(18) Un buon cestista *deve* essere alto due metri. (1979: 89)

(19) Per appartenere a questo club *si deve* assolutamente discendere da una famiglia aristocratica. (1979: 90)

Según Bertinetto, el verbo *dovere* se presta a activar una lectura de este tipo en combinación con verbos estativos y en contextos “di tipo finale, in cui vengono indicate le condizioni necessarie, o sufficienti, per conseguire un dato effetto o risultato”. Esa acepción de *finalidad* inherente ejemplificada en (18) y (19) —en la que resuena la modalidad anankástica de Conte (1995) (§2.2)— explica que, según observan Skytte *et al.* (1991: 524) y Serianni (1988: 335), el verbo modal *dovere* pueda ser substituido en estos contextos por otra perífrasis de significado afín (aunque menos utilizada): *avere da* + infinitivo:

(20) Luigi aveva ancora da leggere la risposta di Giorgio. (Skytte *et al.* 1991: 524)

Esto recuerda lo observado por Garachana Camarero (§3.2) sobre el origen del valor dinámico de *tener que* a partir de una posible relectura metonímica de sus valores posesivos iniciales en clave final: la misma lectura que puede asociarse al uso de *dovere* en (18) y (19).

Sin embargo, si *dovere* se asocia a un verbo principal no estativo, no se activa este tipo de lectura: en tal caso, *dovere* expresa modalidad deóntica:



(21) Giovanni *deve* uscire. (1979: 90)

También Samu (2005: 63) observa que el verbo italiano *dovere* puede ser empleado para vehicular modalidad deóntica (22), orientada al sujeto (23) o circunstancial (24):

(22) *Devi* partire subito!

(23) *Devo* sedermi, mi gira la testa.

(24) *Devi* andare ora, se vuoi prendere l'autobus.

La amplia polisemia modal del verbo italiano *dovere* y la falta de una estructura que pueda considerarse formalmente correspondiente a *tener que* en lengua italiana confirman la presencia de unas divergencias importantes entre las dos lenguas. Divergencias que se analizarán, desde una perspectiva interlingüístico-adquisicional, en §5.

## 5. Del italiano al español: construyendo un mapa modal en lengua meta

Las observaciones de Conte (1995) y Van der Auwera y Plungian (1998) presentadas en §2 ya sugerían lo difícil que puede resultar establecer, en distintos contextos, si en estos se expresa modalidad deóntica o más bien dinámica. En §3, además, hemos estado analizando cómo explicar el funcionamiento de *deber* y *tener que* no puede reducirse a considerar el primer modal como operador exclusivamente deóntico y el segundo como operador exclusivamente dinámico, puesto que existen muchos contextos que no pueden considerarse ni puramente deónticos ni puramente dinámicos, sino más bien cercanos a una u otra polaridad modal, dependiendo del posicionamiento del hablante. En este sentido, en español, la alternancia entre *deber* y *tener que* puede considerarse una marca que indica “cómo de deóntico” y “cómo de dinámico” cierto contexto es percibido por el hablante, lo que no emerge –como se ha ilustrado en §4– en el uso del verbo italiano *dovere*. Efectivamente, *dovere* puede desempeñar tanto funciones deónticas como dinámicas, sin marcar superficialmente ninguna diferencia entre las dos. También Sánchez Montero (1993: 135-138; 145-149) observa que *dovere* puede ser empleado tanto para traducir el valor de “obligación o necesidad” de *tener que* (25) como el “valor obligatorio” de *deber* (26):

(25) (ESP) Como mi padre *tiene que estar* muchas noches con el operador y con otra gente del campo [...].

(ITA) Poiché mio padre *deve rimanere* molte sere con il fattore e altra gente di campagna [...].

(26) (ESP) No ignoro que los varones religiosos y los santos, que *deben servirnos* de ejemplo [...].

(ITA) So bene che gli uomini religiosi e i santi, che *debbono servirci* d'esempio [...].

Dadas estas premisas, está claro que traducir del español al italiano puede resultar bastante sencillo, pero el proceso inverso podría presentar bastantes más elementos de criticidad. Consideremos ahora, por lo tanto, la perspectiva de un estudiante de español con L1 italiana que tenga que enfrentarse al uso de los verbos modales *deber* y *tener que* como operadores de modalidad deóntica y dinámica y veamos qué dificultades podrían plantearse.

### 5.1 Modalidad en la interlengua

La afinidad formal entre el sistema de la L1 y el de la L2, típicamente, acelera la transición a través de las primeras fases de la interlengua, produciendo –al mismo tiempo– cierta sensación de confianza en el estudiante (Schmid, 1994); en el caso del aprendizaje del italiano por parte de hispanohablantes o del español por parte de itálofonos, tal confianza puede resultar, como observa Carrera Díaz (1979) un tanto “engañoso”, ya que si por un lado parece favorecer una asimilación bastante rápida de las propiedades morfosintácticas de la L2, por el otro puede llevar al estudiante a preconcebir un grado excesivo de congruencia entre los dos sistemas y –por lo tanto– a transferir potencialmente y de manera automática a la L2 cualquier elemento o regla de funcionamiento de su L1 que presente una correspondencia formal en lengua meta (Corder 1984). Es así como pueden producirse fenómenos de transferencia negativa, potencialmente sujetos a una posterior fosilización (Ambrosini 2018; Ambrosini, Della Putta 2021; Myers 2004). Lo que principalmente genera problemas en el aprendizaje de una L2 muy similar a la L1, como en el caso del aprendizaje del español por parte de itálofonos, no es tanto la construcción de nuevas categorías gramaticales, sino más bien la correcta expresión gramatical de categorías afines a las dos lenguas (Amenós-Pons *et al.* 2017). En nuestro caso, la modalidad deóntica y dinámica existen tanto en italiano como en español, pero en la L1 del estudiante hay un verbo modal –*dovere*– que puede

expresar indiferentemente ambas modalidades, las cuales –en cambio– se expresan en L2 mediante dos verbos modales diferentes, *deber* (correspondiente formal de *dovere*) y *tener que* (sin correspondiente formal en L1).

¿Qué interferencias podrían producirse, dadas estas premisas, en el proceso de interlengua? A continuación, se presentan algunas hipótesis.

### 5.1.1 Transferencias positivas

Desde el análisis llevado a cabo en §3 y §4 se desprende que tanto *deber* como *dovere* pueden vehicular modalidad deóntica en los casos en los que la fuente del deóntico es el hablante mismo o una fuente de autoridad externa interiorizada por él. En estos casos, por lo tanto, la *estrategia de la congruencia* (Schmid, 1994) podría guiar hacia una transferencia positiva en lengua meta, como muestran los ejemplos siguientes:

ITALIANO	ESPAÑOL	
(27) <i>Devi comportarti bene</i>	<i>Debes</i> portarte bien	<i>Tienes que</i> portarte bien
(28) <i>Noi cittadini dobbiamo pagare le tasse</i>	Los ciudadanos <i>debemos</i> pagar impuestos	Los ciudadanos <i>tenemos que</i> pagar impuestos <sup>2</sup>

Tanto *deber* como *dovere* actúan, en estos casos, como operadores de selección exclusiva: es decir, definen la acción representada por el verbo principal como la más altamente valorada en el plano deóntico; en estos casos, de todos modos, no sería incorrecto traducir *dovere* con *tener que*; sin embargo, el significado modal cambiaría, ya que –como explica Fernández de Castro (1999: 186-187)– el uso de *tener que* transmite más bien la idea de una imposición no establecida por el hablante, sino surgida de circunstancias externas u orientada hacia finalidades concretas (ej: *tienes que* portarte bien... o te quedas sin postre / si quieres que Papa Noel te traiga regalos). Efectivamente –como ya se ha explicado en §3.3– en casos como (27) *lo correcto* (*Debes*) también puede interpretarse como *necesario* (*Tienes que*) en presencia de circunstancias externas especialmente concretas o relevantes. Obsérvese, sin embargo, que en (28) la solución *tenemos que pagar impuestos* deja abiertas dos posibles interpretaciones: bien las circunstancias externas adquieren –desde el punto de vista del hablante– una connotación fina-

lístico-positiva o neutra –como en (27)– en cuyo caso la selección de *tener que* en lugar de *deber* sería aceptable, bien adquieren una connotación negativa de constricción (cfr. §3.2), que es totalmente ajena a la solución *debemos pagar*. Por lo tanto, seleccionar el operador modal *tener que* en lugar de *deber* en este tipo de contextos puede tener implicaciones importantes en el plano pragmático, de las que el estudiante necesitaría estar al tanto.

5.1.2 Transferencias negativas

Como se ha observado en (27), la selección del modal *tener que* en lugar de *deber* redefine la modalidad del contexto, desplazándolo del plano deóntico al plano dinámico. Esto es posible, según lo ilustrado en §3.3, en los casos en los que no haya un conflicto evidente entre lo requerido por una obligación externa y la voluntad del hablante, con lo cual la primera es interpretable tanto como una necesidad circunstancial, como como un deber interiorizado.

Muy distinto es el caso de una obligación externa en conflicto con la postura del hablante, que –como hemos visto en (28)– puede activarse a través del modal *tener que*: si el hablante necesita manifestar declaradamente tal divergencia, seleccionar *tener que* es la única solución posible, como se desprende del ejemplo (29):

ITALIANO	ESPAÑOL	
(29) Troppe tasse <i>dobbiamo</i> pagare	Demasiados impuestos <i>debemos</i> pagar*	Demasiados impuestos <i>tenemos que</i> pagar

En este contexto, la falta de aval por parte del hablante convierte la obligación en constricción externa, expulsando automáticamente el contexto de la dimensión deóntica y recolocándolo en el plano dinámico, en el que una eventual selección del modal *deber* resulta paradójica desde un punto de vista pragmático e imposible de aceptar en L2.

Así como resultaría inaceptable en contextos de modalidad circunstancial (30) u orientada al objeto (31), ya que implicaría la formulación implícita de un juicio deóntico por parte del hablante en un contexto de pura necesidad:

ITALIANO	ESPAÑOL	
(30) <i>Devi</i> prendere l'autobus 7	<i>Debes</i> coger el autobús 7*	<i>Tienes que</i> coger el autobús 7
(31) <i>Devo</i> mangiare qualcosa	<i>Debo</i> comer algo*	<i>Tengo que</i> comer algo

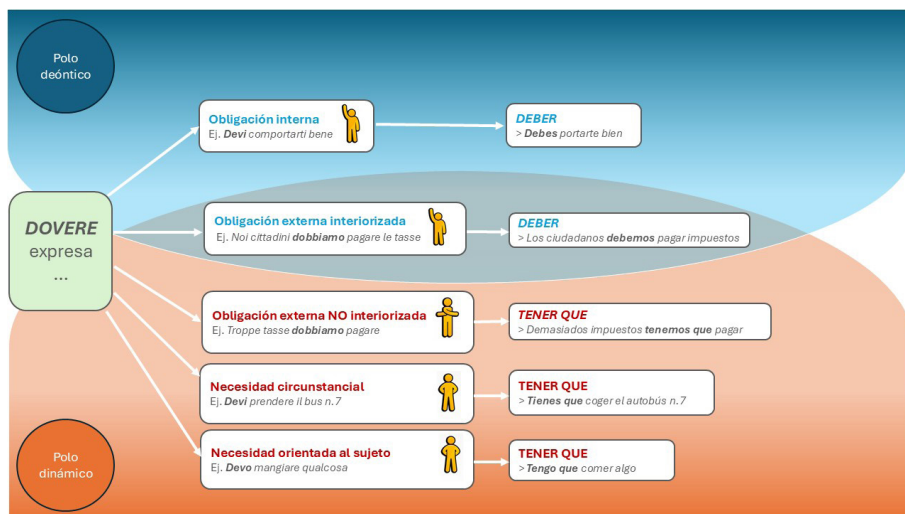
En (29), (30) y (31), por lo tanto, la transposición de los valores dinámicos de *dovere* a *deber* llevaría a un calco formal y *tener que* es la única solución aceptable en L2.

## 5.2 Un mapa modal para orientar la interlengua

En la actualidad, la diferencia entre las funciones obligativas de *deber* y *tener que* es tan difícil de identificar que, en muchas ocasiones, simplemente se resuelve en considerarlos como dos soluciones sinonímicas (Matte Bon 2002: 136), o matizando –como ya se ha adelantado en §3– que *tener que* representa una versión más intensamente obligativa de *deber*; otra creencia tan ampliamente respaldada como engañosa (Fernández de Castro 1999: 185). En una perspectiva adquisicional, además, los ejemplos que se han ilustrado en §5.1 muestran claramente hasta qué punto las divergencias entre italiano y español analizadas en §3 y §4 pueden hacer todavía más complicada una correcta asimilación –por parte del estudiante italófono– de las funciones modales de *deber* y *tener que* en lengua meta.




Todo ello (supuestamente) conlleva cierta responsabilidad por parte del docente de español. Por ejemplo, un estudiante italófono podría sentirse legitimado –por la influencia del sistema de su L1 y a falta de reglas más dirimentes en L2– a transferir integralmente al verbo *deber* toda la gama de valores modales de *dovere*, tal vez contemplando la posibilidad de remplazar *deber* por *tener que* para expresar, cuando lo considere oportuno, una mayor fuerza ilocutiva. Y nada de todo esto –como se desprende de los ejemplos analizados en §5.1– parece representar una estrategia exitosa en L2.

Por esta razón, y como fruto de las reflexiones llevadas a cabo en este estudio, nos parece importante que el estudiante desarrolle una estrategia que le permita –ante todo– descodificar los usos modales deónticos y dinámicos que puede vehicular el verbo *dovere* en su L1, según el contexto dado, y –posteriormente– seleccionar correctamente entre *deber* y *tener que* para recodificar estos usos modales en L2. Con ese propósito, se ha diseñado el mapa modal que figura en el Esquema 3, cuyo uso en clave didáctica comentaremos brevemente a continuación:



Teniendo en cuenta el hecho de que trabajar sobre la modalidad en perspectiva interlingüística representa un desafío muy importante y que los aspectos contrastivos en que se ha profundizado en este estudio podrían resultar demasiado difíciles de entender para cualquier estudiante de español que no esté, aunque sea mínimamente, familiarizado con unos fundamentos teóricos que abarquen dominio modal, adquisición lingüística y lingüística contrastiva, el primer propósito es intentar ponérselo lo más fácil posible. En este sentido, el punto de partida podría ser una pequeña introducción teórica sobre el significado de modalidad deóntica y modalidad dinámica, que no tiene por qué ser muy extensa, ni adentrarse demasiado en la literatura científica sobre modalidad, ya que lo que nos interesa es –ante todo– que el estudiante se entrene en reconocer, en los distintos contextos de su L1 en que puede emplearse el verbo *dovere*, si este expresa una u otra modalidad. Para esta primera reflexión, se puede utilizar como *input* la mitad izquierda del Esquema 3, presentándole al estudiante varios ejemplos (también añadiendo unos cuantos más, a partir de los que figuran en el esquema) de uso modal de *dovere* y pidiéndole a continuación que él mismo proponga algunos más. Es especialmente importante llevar su atención hacia el hecho de que *dovere* puede ser empleado en contextos de obligación tanto externa-interiorizada, como externa-no interiorizada, ya que es aquí donde, en lengua meta, se produce un desdoblamiento de exponentes modales. En el Esquema 3, cada contexto es marcado por un icono que señala simbólicamente el posicionamiento del hablante,

como referencia de soporte para el estudiante:

	El hablante avala
	El hablante no avala (se opone)
	El hablante es neutral / su posicionamiento no influye

Ahora, pasando a la mitad derecha del esquema, es posible mostrarle al estudiante para qué ejemplos de la L1 la lengua meta prevé el uso de *deber* y para qué otros, en cambio, es necesario el uso de *tener que*. El esquema está distribuido de manera tal que los contextos deónticos estén agrupados en el área azul, arriba, y los dinámicos en el área naranja, abajo; sin embargo, hay una zona de solapamiento entre azul y naranja, correspondiente a los contextos de obligación externa interiorizada, como en el ejemplo (28); como se ha observado en §5.1, en esta zona pueden colocarse contextos deónticos potencialmente interpretables también como dinámicos y en los que una selección de *tener que* en vez de *deber* como operador modal puede tener repercusiones significativas. Por lo tanto, podría ser muy útil e interesante explorar junto al estudiante una serie de contextos de este tipo.

El Esquema 3, el punto de llegada de nuestra reflexión contrastivo-adquisicional, solo se propone como un punto de partida para la implementación de otras actividades didácticas que trabajen en esta dirección: por ejemplo, actividades de traducción guiada hacia la L2, en las que siempre será importante llevar la atención del estudiante hacia cuál es la *intención* semántica con la que el hablante utiliza el verbo *dovere* en L1, la misma que se tendrá que transmitir, escogiendo entre *deber* o *tener que* (o, eventualmente, optando por otros tipos de reformulación), en L2.

Este mapa modal *físico* que hemos elaborado se propone, pues, como un soporte a través del que el estudiante pueda ir construyendo y reforzando su propio mapa modal *interior* en lengua meta.

## 6. Conclusiones

Las reflexiones que se han llevado a cabo en este estudio sugieren que la capacidad de utilizar correctamente los verbos *deber* y *tener que* por parte de un estudiante italófono requiere una reflexión profunda, que va mucho más allá de una simple competencia morfológica en L2. Como hemos visto, un mismo verbo modal de la L1 puede desempeñar funciones modales que en la L2 pueden (y necesitan, en la gran mayoría de los casos) ser vehiculadas por dos verbos modales distintos, que ni pueden considerarse sinonímicos, ni intercambiables.

Dadas estas premisas, para que el estudiante no quede totalmente expuesto a las interferencias generadas por la afinidad entre sistema de la L1 y sistema de la L2, podría ser de provecho plantear en ámbito didáctico una propuesta de reflexión metalingüística en línea con la que se ha presentado en §5.2 y que –por lo tanto– lleve primero la atención del estudiante hacia la relación que conecta la forma lingüística *dovere* a sus significados semánticos en L1, para ver a continuación como pueden ser recodificados –en L2– por dos formas verbales distintas: *deber* y *tener que*. Resumiendo, sin que el estudiante desarrolle una *competencia modal* en lengua meta, es muy difícil que los elementos de L2 tratados en este estudio sean objeto de un aprendizaje realmente eficaz. No se trata de una competencia fácil de desarrollar; sin embargo, propuestas como la de una confrontación entre lenguas romances que parta de la semántica para llegar a la forma (Bolinger 1968) o la de una exposición, en didáctica de lenguas, a una gramática de la L2 que también tenga en cuenta el plano semántico (Venier 1991) no son nada nuevas. En este sentido, consideramos que las siguientes reflexiones de Skytte (1995: 99-100) se encuentran especialmente en línea con el propósito de nuestra propuesta:

Considero ideale la descrizione grammaticale che combini in modo operativo contenuto e forma, offrendo un'esposizione alla modalità in cui si dimostrino le diverse possibilità della sua espressione linguistica, combinata con un'esposizione dei modi esplicitante i loro contenuti modali – in entrambi i casi con riferimento a forme modali e espressioni di modalità di altre lingue, che possano servire a corroborare il processo di apprendimento. È questa una forma di esposizione che, secondo la mia esperienza didattica, serve inoltre a rendere meno misteriose forme della lingua straniera inesistenti nella lingua di partenza.

Es este un tipo de enfoque que, como docentes de lengua, *deberíamos* adoptar.

Y tal vez *tengamos que* hacerlo.



## Bibliografía citada

- Ambrosini, Maria Vittoria (2018), *La expresión de la modalidad continuativa en italiano, español y catalán: Analogías, divergencias e interferencias*, colección *Linguistica delle differenze* 3, Roma, Aracne Editrice.
- Ambrosini, Maria Vittoria; Della Putta, Paolo (2021), “Problemi di apprendimento di tre perifrasi fra spagnolo e italiano. Osservazioni acquisizionali e proposte pedagogiche”, *Cuadernos de Filología Italiana*, 28: 11-44.
- Amenós-Pons, José; Ahern, Aoife; Guijarro-Fuentes, Pedro (2017), “L1 French learning of L2 Spanish past tenses: L1 transfer versus aspect and interface issues”, *Studies in Second Language Learning and Teaching*, 7: 489-515.
- Bertinetto, Pier Marco (1979), “Alcune ipotesi sul nostro futuro (con osservazioni su potere e dovere)”, *Rivista di grammatica generativa*, 4: 77-138.
- Bolinger, Dwight (1968), “Postposed main phrases: an English rule for the Romance subjunctive”, *The Canadian Journal of Linguistics. La Revue linguistique*, 14/1: 3-30.
- Bybee, Joan L.; Perkins, Revere; Pagliuca, William (1994), *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*, Chicago, The University of Chicago Press.
- Carrera Díaz, Manuel (1979), “Italiano para hispanohablantes: la engañosa facilidad”, *Rassegna Italiana di Linguistica Applicata*, 11: 23-40.
- Conte, Maria-Elisabeth (1995), “Epistemico, deontico, anankastico”, *From Pragmatics to Syntax: Modality in Second Language Acquisition*, eds. Anna Giacalone Ramat; Grazia Crocco-Galêas. Tübingen, Narr: 3-9.
- Corder, Pit (1984), “La lingua dell'apprendente”, *Interlingua, aspetti teorici e implicazioni didattiche*, eds. Enrico Arcaini; Bernard Py. Roma, Istituto della Enciclopedia Italiana: 49-72.
- Fernández de Castro, Félix (1999), *Las perífrasis verbales en el español actual*, Madrid, Gredos.
- Flores Salvo, Camila; Jaque Hidalgo, Matías (2024), “Negación y modales deónticos de obligación en español: el caso de deber y tener que”, *LEXIS*, 48/1: 268-301.
- Garachana Camarero, Mar (2017), “Perífrasis formadas en torno a tener en español: ser tenudo/tenido ø/a/de + infinitivo, tener a/de + infinitivo, tener que + infinitivo”, *La gramática en la diacronía. La evolución de las perífrasis verbales modales en español*, ed. Mar Garachana Camarero. Madrid-Fráncofort, Iberoamericana-Vervuert: 229-286.
- Kiefer, Frene (1987), “On defining modality”, *Folia Lingüística*, 21/1: 76-94.
- Lyons, John (1981) [1977], *Semantics*, vol. 2, segunda edición, Cambridge, Cambridge University Press.
- Matte Bon, Francisco (2002), *Gramática comunicativa del español. Tomo I: De la lengua a la idea*, nueva edición revisada, Madrid, Edelsa.
- Myers, Marie (2004), *Modalités d'apprentissage d'une langue seconde*, Louvain-la-Neuve, De Boeck Supérieur.

- Nuyts, Jan (2006), “Modality: Overview and linguistic issues.”, *The Expression of Modality*, ed. William Frawley. Berlin & New York, De Gruyter Mouton: 1-26.
- Palmer, Frank (1990) [1979], *Modality and the English Modals*, segunda edición, London & New York, Longman.
- Palmer, Frank (2001) [1986], *Mood and Modality*, segunda edición, Cambridge, Cambridge University Press.
- Pietrandrea, Paola; Cervoni, Valerio (2015), “Guide d’annotation sémantique projet MODAL”, [04/07/2025] <[http://paolapietrandrea.altervista.org/guidelines\\_decision\\_trees\\_for\\_French\\_with\\_examples.pdf](http://paolapietrandrea.altervista.org/guidelines_decision_trees_for_French_with_examples.pdf)>.
- Ramchand, Gillian (2018), *Situations and Syntactic Structures. Rethinking Auxiliaries and Order in English*, Cambridge/London, MIT Press.
- Real Academia Española [RAE] (2010), *Nueva gramática de la lengua española*, Madrid, Espasa Calpe.
- Rosemeyer, Malte (2017). “La historia de las perífrasis *deber / deber de + infinitivo*: variación, norma y géneros textuales”, *La gramática en la diacronía. La evolución de las perífrasis verbales modales en español*, ed. Mar Garachana Camarero. Madrid-Fráncfort, Iberoamericana-Vervuert: 147-196.
- Samu, Borbala (2005), *L’acquisizione della modalità deontica in italiano L2* [Tesi di dottorato, Università per Stranieri di Perugia].
- Sánchez Montero, María del Carmen (1993), *Perífrasis verbales en español e italiano. Estudio contrastivo*, Trieste, Edizioni Lint.
- Schmid, Stefan (1994), *L’italiano degli spagnoli. Interlingue di immigrati nella svizzera tedesca*, Milano, Franco Angeli.
- Searle, John (1975), “Indirect speech acts”, *Syntax and semantics*, New York, Academic Press, 3-*Speech acts*: 59-82.
- Serianni, Luca (1988), *Grammatica Italiana. Italiano comune e lingua letteraria. Suoni, forme, costrutti*, Torino, UTET.
- Skytte, Gunver; Salvi, Giampaolo; Manzini, Rita (1991), “Frase subordinate all’infinito”, *Grande grammatica italiana di consultazione*, vol. 2, *I sintagmi verbale, aggettivale, avverbiale. La subordinazione*, eds. Lorenzo Renzi; Giampaolo Salvi. Bologna, Il Mulino: 483-570.
- Skytte, Gunver (1995), “Modalità e modi in una grammatica italiana europea. Principi di esposizione”, *From Pragmatics to Syntax: Modality in Second Language Acquisition*, eds. Anna Giacalone Ramat; Grazia Crocco-Galès. Tübingen, Narr: 97-101.
- Van der Auwera, Johan; Plungian, Vladimir A. (1998), “Modality’s semantic map”, *Linguistic Typology*, 2/1: 79-124.
- Venier, Federica (1991), *La modalizzazione assertiva. Avverbi modali e verbi parentetici*, Milano, FrancoAngeli.
- Verstraete, Jean-Cristophe (2007), *Rethinking the Coordinate-Subordinate Dichotomy. Interpersonal Grammar and the Analysis of Adverbial Clauses in English*, Berlin, Mouton De Gruyter.

Von Wright, Georg H. (1951), *An Essay in Modal Logic*, Amsterdam, North Holland.  
Yllera, Alicia (1980), *Sintaxis histórica del verbo español: Las perífrasis verbales*, Universidad de Zaragoza.

**Maria Vittoria Ambrosini** es doctora en Lenguas y Culturas Románicas por la Universidad Autónoma de Barcelona (UAB) en 2018. Docente de Lengua Italiana y Traducción en la UAB y en la Universidad de Barcelona (UB) de 2019 a 2023 y becaria de investigación en la Universidad de Bolonia de 2023 a 2025 (proyecto: *Towards linguistic and intercultural complexity: a corpus-based study*). Desde 2025 es docente de Didáctica de lenguas modernas y Tecnologías para la didáctica de lenguas en la Universidad del Salento.

**[mariavittoria.ambrosini@unisalento.it](mailto:mariavittoria.ambrosini@unisalento.it)**



**SONIA BAILINI**

# ***I DIARI DELLA MOTOCICLETTA, UN VIAGGIO NELLA TRADUZIONE INTERSEMIOTICA E INTERLINGUISTICA***

**Università Cattolica del Sacro Cuore di Milano****Abstract**

Questo contributo analizza, da un lato, le modalità di adattamento delle fonti scritte su cui si basa il film *Diarios de motocicleta* (2004) di Walter Salles e, dall'altro, la resa in italiano di alcuni elementi culturospecifici. La selezione dei contenuti nella traduzione intersemiotica e la resa nel doppiaggio italiano nella traduzione interlinguistica restituiscono un'immagine del protagonista che ne riflette solo in parte la biografia reale perché rispondono più a criteri di fruibilità che di fedeltà storica.

palabras clave: traduzione intersemiotica, traduzione interlinguistica, adattamento testuale, doppiaggio, Ernesto Che Guevara

**Abstract****The motorcycle diaries, a journey into intersemiotic and interlinguistic translation**

*This paper analyzes, on the one hand, the adaptation of the written sources on which Walter Salles' film Diarios de motocicleta (2004) is based and, on the other, the rendering in Italian of some cultural-specific elements. The selection of contents in the intersemiotic translation and the rendering in the Italian dubbing in the interlingual translation paints a picture of the protagonist that only partially reflects his real biography because they respond more to criteria of usability than of historical fidelity.*

*keywords: intersemiotic translation, interlinguistic translation, textual adaptation, dubbing, Ernesto Che Guevara*

## 1. Dai diari al film: in viaggio tra memoria e finzione

Questo contributo si prefigge, da un lato, di analizzare la traduzione intersemiotica risultante dalla trasposizione cinematografica dei diari di viaggio di Ernesto Guevara e Alberto Granado e, dall'altro, di osservare la resa di alcuni elementi culturospecifici nel doppiaggio dallo spagnolo all'italiano. *I diari della motocicletta* (Salles 2004b) ripercorre le tappe del viaggio di sette mesi che Ernesto Guevara de la Serna e Alberto Granado realizzarono, dal 29 dicembre 1951 al 26 luglio 1952, a bordo di una Norton 500 del 1939, la *Poderosa II*. Il testo audiovisivo è una combinazione tra road-movie, documentario e racconto di formazione (Lie 2023: 26) ed è il risultato di un'articolata traduzione intersemiotica che utilizza come fonti le *Notas de viaje por América Latina* di Ernesto Guevara de la Serna e *Con el Che por Sudamérica* di Alberto Granado. In entrambi i testi quel viaggio viene presentato come un'esperienza che cambia la vita dei protagonisti, in particolare quella di Ernesto, quando ancora lo chiamavano Fúser e quando ancora non era diventato il Che.

Le *Notas de viaje*, scritte durante il viaggio del 1952, riordinate da lui stesso a distanza di un anno e pubblicate per la prima volta solo nel 1993, già nelle prime pagine danno conto del cambiamento interiore che quel viaggio determinò nel giovane Fúser ed è proprio il lavoro di rielaborazione a posteriori che permette di definire questa fonte un diario di secondo livello, come lo è, di fatto, qualsiasi testo autobiografico (Duno-Gottberg 2005, Peris Blanes 2011, Arreola 2016) e come dimostrano le parole con cui si apre il diario: “El personaje que escribió estas notas murió al pisar de nuevo tierra argentina, el que las ordena y pule, ‘yo’, no soy yo; por lo menos no soy el mismo yo interior. Ese vagar sin rumbo por nuestra ‘Mayúscula América’ me ha cambiado más de lo que creí” (Guevara 2004: 34).

*Con el Che por Sudamérica*, pubblicato per la prima volta nel 1978 dalla *Agencia Literaria de Letras Cubanas*, ripubblicato in Argentina con la casa editrice Marela nel 2007 e nel 2013 in una nuova edizione, è il diario di quel viaggio scritto da Alberto Granado, Mial, l'amico di una vita di Ernesto, che, otto anni dopo, fu invitato proprio dal Che a lavorare come scienziato nella Cuba della rivoluzione. Vi rimase per tutta la vita e morì nel 2011, a L'Avana. È un diario scritto con la freschezza e la spontaneità di un giovane che ha sognato quel viaggio sin da quando era adolescente, che osserva e commenta tutto ciò che vede, a volte con uno spirito molto più critico di quanto non appaia nelle *Notas* di Ernesto. Le motivazioni che hanno spinto Granado a pubblicare quel diario non lasciano

alcun dubbio sulla volontà di restituire una dimensione umana alla figura del Che per evitarne la mitizzazione e l'idealizzazione:

En realidad 'El Diario' no tenía nombre. La Agencia Literaria, la gente que trabajó conmigo en Letras Cubanas, le puso ese nombre [Ndr: Con el Che por Sudamérica], un poquito en contra de mi voluntad. Por la palabra Che, porque el viaje no fue con el Che Guevara, fue con Ernesto, pero yo estaba muy interesado en que se publicara y aspiraba a que no siguieran apareciendo 'amigos' del Che que nunca habían sido verdaderos amigos. Ni quería que se siguiera hablando del Che como de un hombre lleno nada más que de virtudes, sino como de un muchacho de 24 años, lo que era en realidad [...]. Yo solo pensaba en evitar que al Che lo siguieran endiosando, que a mi amigo Ernesto Guevara lo siguieran transformando en un hombre sin ningún defecto (Granado 2013: 9-10).

L'idea di riprendere il diario di viaggio di Ernesto Guevara e costruire una storia cinematografica è nata grazie all'interesse di Walter Salles per il road-movie, interesse già manifestatosi in altri suoi film, come per esempio *Foreign Land* (1995) e *Central do Brasil* (1998). L'identificazione con questo genere cinematografico è esplicita, oltre che nel titolo, anche all'inizio del film, quando la voce in off di Ernesto recita: "Il veicolo: La Poderosa II; il metodo: l'improvvisazione" (Salles 2004a: 00:55-01:15) e in una delle scene iniziali del viaggio, quando i due giovani, a bordo della moto, intraprendono una gara con due *gauchos* a cavallo (Salles 2004a: 07:25-07:42), omaggio a *Easy Rider*.

Il regista brasiliano ha dichiarato di aver fortemente voluto questo film perché profondamente affascinato dalla lettura delle *Notas de viaje*, dall'idea del viaggio come autoscoperta e momento di crescita interiore e, infine, dalla ricchezza umana dei due protagonisti (Minà 2003). Per la realizzazione della pellicola Salles ha scelto di utilizzare tecniche proprie del cinema verità, cioè di un tipo di cinema "di autenticità totale, vero come un documentario ma col contenuto di un film romanzesco, cioè col contenuto della vita soggettiva" (Morin 1960 citato in Fofi, Morandini, Volpi 1988: 343). Non si trattava, pertanto, solo di ricostruire il viaggio di Fúser e Mial, ma di riviverlo in condizioni simili durante le riprese. Con questo proposito, gli attori venivano stimolati a improvvisare conversazioni con la gente del posto che via via incontravano e questi interventi casuali, non contemplati nel copione, sono stati poi inseriti nella struttura filmica. Come spiega lo stesso Salles:

optamos por emplear una gramática cinematográfica sencilla y directa [...] La mayor parte del tiempo, procuré evitar imponer una 'puesta en escena', y me dejé llevar por lo que íbamos encontrando en el camino, sin imponer ideas preconcebidas. También

intentamos hacer lo contrario de lo que podría llamarse un ‘documental inducido’. Básicamente, intentamos filmar la historia como si estuviese desarrollándose frente a nuestros propios ojos (<http://www.labutaca.net/films/26/diariosdemotocicleta1.htm> citato in Peris Blanes 2011).

Tra le scelte operative riconducibili a questa tecnica cinematografica, almeno due sono fondamentali:

1. la realizzazione dello stesso percorso e delle stesse tappe del viaggio originale, facendo coincidere il tragitto fisico della troupe cinematografica con quello dei due protagonisti storici, uno dei quali, tra l'altro, era presente sul set. A questo proposito è bene ricordare che il film è stato girato nel 2002, esattamente 50 anni dopo il viaggio originale (1952), le riprese sono durate 84 giorni in 33 scenari diversi, sono state riprodotte cinque copie identiche della Norton 500 del 1939 (Minà 2003) e Alberto Granado, 82 anni, ha partecipato alle riprese, contribuendo, con i suoi ricordi, alla ricostruzione di alcune scene (Moratal Ibáñez, Carli, Kennel 2005);
2. la *full immersion* degli attori nei ruoli dei personaggi che dovevano rappresentare da un punto di vista più profondo rispetto a quanto accade normalmente. Gael García Bernal e Rodrigo de la Serna hanno dovuto imparare a guidare una moto del '39, a navigare sul fiume su una zattera e hanno persino dovuto documentarsi sulle letture di Guevara all'epoca del viaggio (Peris Blanes 2011).

Durante le riprese è stato realizzato anche il documentario *In viaggio con Che Guevara* (Minà 2003), una sorta di *backstage* nel quale la presenza di Alberto Granado sul set funziona come *trait d'union* tra il passato e il presente. Questo continuum è garantito anche dall'inserimento di foto originali del viaggio del 1952, che fanno sì che finzione e memoria si sovrappongano, poiché la storia raccontata nel film coincide con i ricordi di Granado.

## 2. L'adattamento cinematografico come traduzione intersemiotica

Qualsiasi traduzione intersemiotica e, in generale, qualsiasi tipo di narrazione, implica necessariamente un processo di selezione di ciò che si intende raccontare e di ciò che si decide di escludere (Dusi 2003, Fois 2012, Hutcheon 2021). Se, da un lato, *I diari della motocicletta* ha ricevuto moltissimi premi e riconoscimenti (Goya 2005 e Bafta 2005 tra i più prestigiosi), dall'altro non bisogna dimenticare che la sua distribuzione nei circuiti nordamericani, dove la figura di Ernesto



Guevara è legata alla delicata questione delle relazioni politiche tra Cuba e gli USA, ha generato anche recensioni negative (De Ferrari 2005, Duno-Gottberg 2005, Fandiño 2005). Nonostante il regista abbia dichiarato in varie occasioni la sua intenzione di porre il focus sulla biografia del giovane Guevara prima della svolta ideologica e politica che lo trasformò nel Che e sul ruolo di quel viaggio del 1952 come elemento attivatore del cambiamento, il dibattito gira intorno a due domande: in che misura l'immagine di Ernesto Guevara proposta da Salles ne rispecchia la vera personalità? Che cosa ha guidato la selezione degli eventi da inserire nel film rispetto alle vicissitudini reali raccontate nei diari?

Come afferma Fumagalli, una biografia, per poter essere oggetto di una traduzione intersemiotica che ne garantisca la buona riuscita, deve possedere alcune caratteristiche ben precise. Innanzitutto, deve poter essere rappresentata visivamente; in secondo luogo, deve avere un protagonista affascinante ed empatico, capace di instaurare relazioni significative almeno con un altro personaggio determinante per la sua evoluzione psicologica. Infine, la traiettoria vitale o il frammento di essa che si racconta deve avere una linea drammatica ascendente, con un evento principale che aiuti a focalizzare il tema e che segni una svolta verso un climax di grande intensità (Fumagalli 2004: 181-185).

I diari di viaggio di Ernesto Guevara e Alberto Granado riuniscono queste caratteristiche e sono, di fatto, due autobiografie che si fondono per un breve tratto della vita dei due amici. Per entrambi quel viaggio ha determinato un cambiamento interiore che ha segnato il loro futuro. I paesaggi spettacolari del Sudamerica si prestano benissimo alla rappresentazione visiva di quell'esperienza fisica e spirituale, e gli spettatori attraversano insieme ai protagonisti le strade polverose di quell'America ancora tutta da scoprire, sperimentando insieme a loro quel senso di libertà e quella sete di conoscenza che permea le riflessioni impresse nei loro diari. Due personaggi così sono interessanti a prescindere, e lo sono ancor di più per lo spettatore che si appresta a guardare un frammento della loro vita con la consapevolezza di quello che poi sono diventati. Entrambi intessono relazioni significative con molte delle persone che incontrano: con i due contadini nei pressi della cava di Chuquicamata, con una povera vecchia malata a cui Ernesto dà alcune delle sue medicine contro l'asma; con Hugo Pesce, luminare della lebbra in Perù, cattolico, marxista, pacifista, che per Ernesto rappresenta un mentore sia dal punto di vista scientifico che ideologico; con un contadino a cui hanno sottratto le terre; con i malati del lebbrosario di San Pablo. Tutti questi incontri fanno crescere gradualmente l'intensità della trama fino al discorso di Ernesto nel giorno del suo compleanno e la traversata del fiume a nuoto per unirsi ai malati

in isolamento sull'altra sponda, che rappresenta il climax e il momento in cui il cambiamento interiore di Ernesto viene prima verbalizzato e poi messo in pratica con un gesto eroico.

L'evento che segna un cambio di passo è la rottura definitiva della moto: da quel momento in poi, l'attenzione si sposta dalla scoperta del paesaggio alla scoperta delle persone: la strada diventa un luogo d'incontro, dove proprio l'assenza di veicoli veloci propizia un maggior contatto con chi è diverso dai protagonisti per classe sociale, etnia e grado di istruzione. Il film assume, pertanto, un tono documentaristico, caratterizzato dall'uso della telecamera a mano e dalla volontà di dare testimonianza della triste realtà di tutti coloro che Fúser e Mial incontrano sul loro cammino: privazione delle terre, sfruttamento del lavoro, mancanza di protezione da parte della polizia, analfabetismo, povertà, soprattutto della popolazione indigena. Emerge così la stretta relazione tra road-movie e documentario (Salles 2007: 70, citato in Peris Blanes 2011). Il cambio di linguaggio cinematografico coincide con il cambio di paesaggio geografico: dopo i deserti del Cile e dell'Argentina, i due amici arrivano sulle Ande e da questo momento in poi il film si concentra più sulle persone che sui paesaggi (Díaz 2005, Lie 2023: 57).

Le scelte di traduzione intersemiotica di Salles sono riconducibili a tre procedimenti: riferimenti intertestuali alle fonti, amplificazione di alcuni eventi e omissione di altri.

Nonostante una traduzione intersemiotica da fonte scritta a trasposizione cinematografica implichi di per sé un cambio di genere testuale (Rutelli 2004), ne *I diari della motocicletta* il rimando al genere delle fonti è ripreso in varie occasioni, sia attraverso il codice visivo sia attraverso la voce in off di Ernesto. Infatti, in alcune scene i due protagonisti sono intenti nella scrittura dei rispettivi diari (i.e. quando sono ospiti da Hugo Pesce e a Machu Picchu), mentre in altre vengono riprodotti frammenti delle riflessioni che Ernesto appunta nel suo diario sotto forma di lettere che scrive alla madre. Anche se è vero che Ernesto durante quel viaggio mantenne una fitta corrispondenza con lei, la fonte della traduzione intersemiotica è il diario, non le lettere, che quindi servono a Salles per evidenziare i cambiamenti interiori che il protagonista sta vivendo. Questo procedimento mostra la volontà del regista di andare oltre la ricostruzione storica al fine di coinvolgere lo spettatore facendogli sperimentare le stesse emozioni che sta provando il protagonista. Entrambi stanno realizzando insieme lo stesso percorso: molte scene riprendono la moto sulla strada, il viaggio è reale ma anche metaforico, è un cammino verso la consapevolezza (Sánchez Zapatero 2021). Se il diario offre una relazione piatta e cronologica degli eventi, il film segue una traiettoria ascendente,

che raggiunge il suo apice con il discorso di Ernesto al lebbrosario nel giorno del suo compleanno e la traversata del fiume a nuoto. Questa parabola di intensità graduale, se da un lato rende la narrazione più dinamica e fruibile, dall'altro impone una modifica dell'ordine temporale degli eventi rispetto ai fatti reali e, di conseguenza, l'amplificazione di alcuni di essi e l'omissione di altri.

Questa breve analisi si limiterà solo a tre amplificazioni e a tre omissioni particolarmente rilevanti tra le molte presenti nella pellicola: tra le prime, il discorso al lebbrosario nel giorno del compleanno di Ernesto, la traversata a nuoto del Rio delle Amazzoni e, in generale, il ruolo di quel viaggio nella traiettoria vitale del protagonista; tra le seconde, alcune tappe del viaggio, le descrizioni dettagliate di paesaggi, aspetti culturali e procedimenti tecnici (i.e. estrazione del rame, cultura inca) e le riflessioni di carattere politico.

Nel diario, il paragrafo in cui è inserito il discorso è introdotto in modo ironico con il titolo 'El día de San Guevara' ed Ernesto minimizza l'evento, descrivendolo con queste parole:

Ya picaditos todos los ánimos [por el pisco] el director de la colonia brindó por nosotros en una manera muy simpática y yo, 'pisqueado', elaboré más o menos lo que sigue: [...] 'Aunque lo exiguo de nuestras personalidades nos impidan ser voceros de su causa, creemos, y después de este viaje más firmemente que antes, que la división de América en nacionalidades inciertas e ilusorias es completamente ficticia. Constituimos una sola raza mestiza que, desde México hasta el estrecho de Magallanes, presenta notables similitudes etnográficas. Por eso, tratando de quitarme toda carga de provincialismo, brindo por Perú y por América Unida' (Guevara 2004: 177).

Ernesto ne parla anche alla madre, in una lettera che le scrive da Bogotá il 6 luglio 1952:

el 14 me organizaron una fiesta con mucho pisco, una especie de ginebra que se trepa de lo lindo. El médico director brindó por nosotros y yo, que me había inspirado por el trago, contesté con un discurso muy panamericano que me mereció grandes aplausos del calificado y un poco piscado público asistente (Guevara 2004: 202).

Nel film, invece, il discorso al lebbrosario rappresenta, insieme alla nuotata che il protagonista intraprende poco dopo per raggiungere i malati sull'altra sponda del fiume, il climax, e offre l'immagine di un Ernesto serio e per nulla alticcio di *pisco*, quasi avesse timore di mostrare che anche lui sa divertirsi, ballare, ubriacarsi, lasciarsi andare, come farebbe un normale ragazzo di 24 anni nel giorno della sua festa. Subito dopo il discorso, di notte, si getta nel fiume, in preda a un impeto umanitario, deciso a raggiungere i lebbrosi sull'altra sponda, ma questa scena

non corrisponde alla realtà dei fatti. Granado racconta che attraversare il Rio delle Amazzoni a nuoto era un sogno di Fúser e che lo realizzò su un braccio secondario del fiume, di giorno, mentre erano a pesca. Alberto ricorda che, sapendo quanto Ernesto fosse testardo, non aveva nemmeno voluto provare a dissuaderlo; si era limitato a farsi promettere che, se si fosse ferito con uno dei molti rami che c'erano nel fiume, sarebbe salito subito a bordo della barca che lo seguiva. La precauzione era dovuta al fatto che Alberto, pochi giorni prima, aveva sperimentato un morso di piranha dopo essersi tuffato nel fiume con una piccola lesione a un polpaccio, che aveva attirato immediatamente i feroci pesci (Granado 2013, Minà 2003). Nonostante anche nelle *Notas* di Ernesto questo evento non venga riportato con particolare enfasi, il film unisce temporalmente l'esperienza della nuotata di notte con il discorso del compleanno, rendendo questo evento il punto più intenso e drammatico del film. Inoltre, in una lettera alla madre, nuotatrice esperta, Ernesto scrive che non aveva voluto buttarsi in acqua per riprendere una gallina che portavano come scorta di cibo sulla Mambo Tango perché aveva paura di nuotare al buio (Guevara 2004: 204). Nel film, lo stesso ragazzo attraversa il fiume in uno slancio eroico di fratellanza, tra le imprecazioni del suo amico Alberto e le braccia aperte dei lebbrosi pronti ad accoglierlo sull'altra sponda. L'attraversamento del fiume a nuoto è un rito di passaggio che risulta un po' forzato: si mostra l'esperienza umana dell'ingiustizia e si lascia intuire la trasformazione che suscita nel personaggio. Come afferma Duno-Gottberg (2005), è possibile che questa scelta risponda alla necessità di rendere gradita l'immagine del Che a un pubblico, come quello nordamericano, con il quale è prudente evitare il discorso politico e ideologico.

Infine, nella realtà dei fatti il viaggio che cambiò la vita di Ernesto non fu quello del 1952 con Mial, bensì quello dell'anno successivo, tra il 1953 e 1954, intrapreso con il proposito di seguire da vicino i cambiamenti politici in corso in alcuni paesi latinoamericani. Fu proprio durante quell'esperienza che assistette, in Guatemala, al colpo di stato che rovesciò il governo di Arbenz, partecipando alla sua difesa; fu sempre durante quel viaggio che conobbe, in Messico, Fidel e Raúl Castro, ai quali si unì per far cadere il regime di Batista a Cuba. Quel primo viaggio di Mial e Fúser in un continente in cui dominavano la povertà, la sofferenza e la disuguaglianza fece emergere il lato umano e sensibile di Ernesto, ma non ne determinò la svolta politica. Le testimonianze dei familiari e degli amici così come alcune biografie (Sinclair 2003, Castañeda 2007, Anderson 2017, Taibo II 2017) raccontano che la scelta di partecipare alla rivoluzione cubana ebbe come conseguenza la rinuncia a compiere un viaggio della durata di dieci anni, come era nelle sue intenzioni. Suo padre spiega che la decisione di rinunciare a

quell'idea maturò in seguito all'impegno legato agli sviluppi politici di alcuni paesi latinoamericani, al matrimonio con la prima moglie, Hilda Gadea, alla nascita della prima figlia e al coinvolgimento nel gruppo degli esiliati cubani in Messico (Guevara Lynch 1988). Il film, invece, facendo credere che la trasformazione di Fúser in Che sia la conseguenza diretta di quel primo viaggio, crea una sorta di continuum tra quell'esperienza e la rivoluzione cubana. Tale enfasi risponde alla necessità di rendere fruibile la pellicola nella società attuale facendo leva sullo spirito umanitario di cui è impregnata, anche a costo di sacrificare la fedeltà ai fatti storici: il viaggio del 1952 potrebbe essere rifatto oggi, ma non quello del 1953-1954 (Lie 2023: 37-38, Garzia 2004).

Tra le omissioni sono da segnalare salti geografici rispetto alle tappe originali del viaggio, soprattutto nella parte finale, poiché, dopo la permanenza presso il lebbrosario di San Pablo e la partenza in zattera verso Leticia (Colombia), le immagini ci portano subito al commiato tra i due amici all'aeroporto di Caracas. Mancano, pertanto, tutte le destinazioni finali: Leticia, Bogotá, Cúcuta in Colombia, e San Cristóbal, Barquisimeto e Caracas in Venezuela. Le dettagliate descrizioni sulla cultura inca, sulle rovine di Macchu Picchu e Ollantaytambo, sui procedimenti di estrazione del rame e del salnitro, che entrambi i protagonisti inseriscono nei loro diari, sono sostituite, nel film, da poche immagini, anche se molto suggestive. Infine, le riflessioni sulle disuguaglianze sociali, sulla povertà, sullo sfruttamento lavorativo, sulla politica dell'Argentina peronista e sulla coscienza di classe, che nei testi scritti sono frequenti e testimoniano la sensibilità dei due protagonisti verso gli emarginati e i deboli come caratteristica della personalità di entrambi, sono completamente assenti nel film. Tuttavia, l'intenzione di quest'ultimo è quella di presentare una tappa della vita di Ernesto Guevara prima della sua conversione alla causa rivoluzionaria, quindi porre il focus sulla biografia e non sull'ideologia. Questo spiega perché i riferimenti alla politica sono stati accuratamente selezionati. In questo senso, pertanto, il regista "no permitió que la historia pasada estuviera impregnada del futuro de los personajes" (Martínez Llamas 2005) e lascia solo intravedere, con rapide pennellate, quello che Ernesto diventerà informando lo spettatore, nella scena finale, con una nota testuale, che i due amici si sarebbero reincontrati otto anni dopo a Cuba, dove il Comandante Che Guevara avrebbe offerto all'amico Alberto di fondare una scuola di medicina.

Se le prime due scelte di omissione sono vincolate a questioni tecniche determinate dall'adattamento cinematografico di testi scritti, la terza, invece, è chiaramente legata alla volontà del regista di non politicizzare il protagonista e di rendere la sua figura gradita anche a un pubblico che lo conosce solo per il suo impegno

politico. La forzatura della pellicola, quindi, è quella di restituire un’immagine di Ernesto Guevara più poetica e romantica di quella reale grazie a scene in cui se ne mostra la tenerezza nell’assistere i malati, la sincerità portata alle estreme conseguenze e la vulnerabilità attraverso gli attacchi d’asma che lo affliggono (Díaz 2005, Duno-Gottberg 2005).

3. Il doppiaggio italiano, oltre l’adattamento linguistico verso la perdita dei riferimenti culturali

Benché la traduzione audiovisiva privilegi la fruibilità alla fedeltà testuale e benché sia particolarmente sensibile a modifiche per ragioni tecniche di sincronizzazione tra codici diversi (Logaldo 2021), nel caso dei *Diarios de motocicleta* il doppiaggio in italiano risente di un eccessivo addomesticamento (Venuti 2013 [1995]), poiché sono stati eliminati, senza una reale necessità, alcuni riferimenti culturali particolarmente rilevanti, restituendo una versione ancora più patinata di quella in spagnolo. Per ragioni di spazio citeremo solo tre esempi rappresentativi. Il primo riguarda l’eliminazione dell’interiezione *che*, caratteristica degli argentini, nella scena dell’incontro al bar con le ragazze cilene con cui Alberto ed Ernesto tentano un approccio chiedendo di sedersi al tavolo con loro. Nel frammento qui sotto riportiamo la trascrizione del dialogo nelle due lingue:

<i>Diarios de motocicleta</i> (Salles 2004a: 40:21- 41:20)	<i>I diari della motocicletta</i> (Salles, 2004b: 40:21- 41:20)
Alberto: Epa, epa, epa, epa... hola. Se rumorea que las chilenas son las más atrevidas de todo el continente. Ernesto: Los rumores no tienen rigor científico, doctor Granado. Hay que hacer una prueba de campo [...] Jazmín: Oye, y ¿ustedes son argentinos? Alberto: ¿Cómo te diste cuenta? Jazmín: Ay, no sé, por el acento, porque dicen ‘che’, qué sé... Alberto: No me había dado cuenta, che. Ernesto: No, para nada, che.	Alberto: Ferma, ferma, ferma ferma... buongiorno. Gira voce che le cilene siano le più sfacciate del continente. Ernesto: Le voci non hanno valore scientifico, Dottor Granado, si impone una ricerca sul campo [...] Jazmín: Voi due siete argentini, vero? Alberto: Come l’ha capito? Jazmín: Qui si dice che gli argentini sono un po’ birichini... è solo una voce... Alberto: Cosa dicevi tu prima delle voci, Fuser? Ernesto: Che vanno verificate. Jazmín: Va beh.

L’interiezione argentina *che* ha determinato il soprannome del futuro Comandan-

te Che Guevara, attribuitogli dai rivoluzionari cubani in Messico. Nel film, il riferimento esplicito da parte delle ragazze cilene viene risolto con una sostituzione totale che travisa il senso del testo originale (“Qui si dice che gli argentini sono un po’ birichini... è solo una voce”). Inoltre, dal punto di vista pragmatico emerge un errore nell’uso degli allocutivi: il testo in spagnolo dice “¿cómo te diste cuenta?” mentre la resa in italiano è “¿Come l’ha capito?” usando la terza persona singolare, come se Alberto si stesse dirigendo alla ragazza in modo formale. L’immagine indica chiaramente che Mial si sta rivolgendo a lei e non a Fúser, quindi non può essere interpretata come una domanda che uno dei due amici rivolge all’altro. La scelta di eliminare l’interiezione *che* impone l’introduzione, nella versione italiana, di due battute completamente assenti nell’originale in cui viene inserito, per dare coerenza al discorso e per questioni di sincronismo labiale, un riferimento al commento che Alberto ed Ernesto hanno espresso poco prima di entrare nel bar sulla necessità di verificare le voci con il lavoro sul campo. Infine, nella versione italiana di questa scena è presente anche una resa infelice: “atrevidas” viene tradotto con “sfacciate” anziché “audaci”, sicuramente più pertinente con il contesto.

La necessità di compensare l’eliminazione dell’interiezione *che* obbliga a una scelta discutibile poco dopo, quando il capo dei pompieri presenta i due amici ai suoi ragazzi:

<i>Diarios de motocicleta</i> (Salles 2004a: 43:17-43:30)	<i>I diari della motocicletta</i> (Salles 2004b: 43:17-43:30)
Bombero: Muchachos, Cañaco... vengan. Les voy a presentar a dos voluntarios argentinos que se van a alojar con ustedes aquí. Les presento al Doctor Che Flaco y al Doctor Che Gordo.	Pompieri: Ragazzi, venite qua... Ragazzi: eccoci... voglio presentarvi due giovani volontari argentini che stanotte dormiranno qui con voi. Vi presento il Dottor Empanada Guevara e il Dottor Empanada Granado.

Nel testo originale il riferimento ironico è chiaro, poiché l’interiezione argentina è nota in tutto il continente, mentre nella versione italiana il riferimento scherzoso alla grande quantità di *empanadas* che hanno mangiato nella scena precedente al bar con le ragazze è coerente con il contesto, ma non suscita lo stesso effetto umoristico e, anzi, è un po’ straniante. Poiché *che* è il soprannome con cui Ernesto Guevara de la Serna è noto in tutto il mondo, mantenere il riferimento culturale nel doppiaggio avrebbe garantito la fedeltà testuale e avrebbe potuto essere l’occasione per spiegare l’origine del soprannome allo spettatore che non ne fosse a conoscenza. La versione italiana avrebbe potuto essere riformulata con un’espressione del tipo: “Come l’hai capito?” “Beh, per l’accento, perché usate

‘che’ quando parlate...» e si sarebbe potuta mantenere la risposta ironica dei due, con un po’ di enfasi sulla parola nell’oralità del doppiaggio: “Ah, CHE, non me ne ero reso conto” e “Neanch’io, CHE, per niente”. Questo avrebbe permesso di mantenere lo stesso effetto scherzoso anche nei nomignoli delle presentazioni ai pompieri: «*Che Stanlio*» e «*Che Ollio*» inserendo l’opportuno adattamento italiano al riferimento intertestuale ai due personaggi comici, la cui pronuncia in italiano avrebbe creato anche un gioco linguistico che avrebbe rafforzato l’effetto umoristico.

Il terzo esempio riguarda la mitigazione del registro volgare che caratterizza alcuni dialoghi tra i due protagonisti, soprattutto in occasione dei vari incidenti con la moto o quando stanno litigando, come nella scena in cui devono portare a mano il loro mezzo a causa dell’ennesimo guasto:

<i>Diarios de motocicleta</i> (Salles 2004a: 29:07-29:40)	<i>I diari della motocicletta</i> (Salles 2004b: 29:07-29:40)
Ernesto: Che, dijiste que entraríamos a Chile como conquistadores, no así como un par de pelotudos. Alberto: Si no te gusta, andá a cagar, Ernesto. Ernesto: <i>¿Y si abandonamos esta mierda que nos da tanto trabajo?</i> Alberto: Ah sí, ¿para andar a patas por todo el continente, genio? Ernesto: Sí, veríamos más, conoceríamos más gente... bajarías de peso... Alberto: Pará, maricón, si con un par de dólares de Chichina esto se arregla enseguida. Ernesto: Y dale con esa historia, olvidate de esa plata. Alberto: Pues, si te enojás, andate vos hasta Venezuela caminando y dejame a mí aquí tranquilo... vos con la plata y la patrona.  Ernesto: ¿Sabés que llevaría la mitad de tiempo si dejaras de meterte en lo que no te importa? Alberto: ¡Concha de tu madre!	Ernesto: Avevi detto che saremmo entrati in Cile come conquistatori, non a piedi come due pezzenti. Alberto: Se non ti sta bene, vaffanculo, Ernesto. Ernesto: Perché non l’abbandoniamo questa merda? Ci dà solo guai! Alberto: Ah sì? Vuoi farti tutto il continente a piedi? Ernesto: Sì, vedremmo più cose, conosceremmo più gente... e tu perderesti qualche chilo. Alberto: Piantala, rompiscatole, con un paio di dollari della tua Chichina, la sistemiamo, la negra. Ernesto: Ancora con questa storia? Quei soldi non esistono! Alberto: Ah sì? Bene, perfetto, in tal caso puoi andartene in Venezuela a piedi e lasciarmi in santa pace, tu e i soldi della tua padroncina. Ernesto: Magari! Ci metterei la metà del tempo se potessi viaggiare come dico io! Alberto: Che figlio di...

Nella versione italiana “pelotudos” diventa “pezzenti” preferendo creare un vincolo semantico con “conquistatori”, anziché rispettare il significato, nella variante diatopica argentina, di “persona ingenua, que es o actúa de modo poco inteli-



gente”. Pertanto, la traduzione avrebbe potuto essere “scemi, cretini, deficienti, stupidi, imbecilli, coglioni”. In questo caso, sarebbe stata coerente, per contrasto, con l’uso ironico di “genio” nella battuta successiva, che, invece, nella versione italiana viene omessa. La riduzione del registro volgare è evidente anche nella traduzione di “maricón” con “rompiscatole” e l’omissione parziale dell’insulto dell’ultima battuta.

#### 4. Conclusioni

*I diari della motocicletta* sono senza dubbio un esempio ben riuscito di traduzione intersemiotica, soprattutto per la capacità di riprodurre lo spirito d’avventura, la sensibilità verso il prossimo e il potere trasformativo di quel viaggio, che permettono una identificazione completa dello spettatore con le emozioni dei protagonisti. Il film rivendica la forza di valori come la solidarietà, la lealtà, l’altruismo, l’uguaglianza. Fúser, prima di diventare il Che, era un ragazzo pieno di ideali che sono ancora condivisibili, così come le disuguaglianze sociali e lo sfruttamento economico sono ancora problemi irrisolti dell’America Latina. Tuttavia, nell’intento di mostrare il giovane Fúser e non il Che rivoluzionario, il testo audiovisivo pone troppa enfasi sui lati positivi del suo carattere e non ne mostra alcun aspetto negativo, creando un personaggio meno autentico di quanto ci si aspetterebbe. Questa impressione è rafforzata dalla contezza che lo spettatore ha di ciò che il protagonista divenne dopo. In altre parole, la buona intenzione del film di voler mostrare Ernesto come un giovane di sani principi diventa un’arma a doppio filo, poiché l’accentuazione dei tratti positivi della sua personalità può far sì che la traduzione intersemiotica venga interpretata come un’azione di recupero strategico dell’immagine del Che Guevara.

Il doppiaggio italiano, inoltre, riflette la tendenza generalizzata della traduzione audiovisiva all’addomesticamento (Venuti 2013 [1995]) e alla standardizzazione linguistica (Liverani 2021) che, in questo caso, si manifestano attraverso la neutralizzazione di alcuni tratti culturospecifici, che avrebbero potuto restituire un’immagine più autentica del personaggio, e attraverso la riduzione sistematica del registro volgare, ricorrendo a omissioni o a sostituzioni lessicali. Queste scelte traduttive contribuiscono a mostrare, nella versione italiana, un Ernesto ancora più ingessato di quello della versione filmica originale, seppur splendidamente interpretato.

## Bibliografia citada

- Anderson, John Lee (2017), *Che. Una vita rivoluzionaria*, Milano, Feltrinelli.
- Arreola, José (2016), “Ernesto Guevara y sus ‘Diarios de motocicleta’. El viaje narrativo del Fúser hacia el Che”, *De raíz diversa: Revista especializada en estudios latinoamericanos*, 3, 5:141-62.
- Castañeda, Jorge (2007), *Compañero. Vita e morte di Ernesto Che Guevara*, Milano, Mondadori.
- De Ferrari, Guillermina (2005), “Diarios de motocicleta: lo que los ojos de Ernesto Guevara le contaron a Walter Salles”, *A Contracorriente: Revista de Historia Social y Literatura en América Latina*, 3, 1: 148-61.
- Díaz, Rosana (2005), “El viaje como desintegración y fundación ideológica en ‘Y tu mamá también’ y ‘Diarios de motocicleta’”, *CiberLetras: revista de crítica literaria y de cultura. El cine hispanoamericano*, 13: 53-63.
- Duno-Gottberg, Luis (2005), “Notas sobre ‘Los Diarios de Motocicleta’ o las travesías de un Che globalizado”, *Revista Latina de Comunicación Social*, 8, 59: 687-706.
- Dusu, Nicola Maria (2003), *Il cinema come traduzione. Da un medium all'altro: letteratura, cinema, pittura*. Torino, Utet.
- Fandiño, Roberto (2005), “Diario de motocicletas”, *Revista hispano cubana*, 21: 213-16.
- Fofi, Goffredo; Morandini, Morando; Volpi, Gianni (1988), *Storia del cinema*, Milano, Garzanti.
- Fois, Eleonora (2012), “Traduzione audiovisiva: teoria e pratica dell’adattamento”, *Between*, 2, 4: 1-17.
- Fumagalli, Armando (2004), *I vestiti nuovi del narratore. L'adattamento da letteratura a cinema*, Milano, Il Castoro.
- Garzia, Aldo (2004), “Perché non muore il mito del Che”, *Limes*, 4, [10/06/2025] <<https://www.limesonline.com/rivista/perche-non-muore-il-mito-del-che-14610021>>
- Granado, Alberto (2013) [1978], *Con el Che por Sudamérica*, Buenos Aires, Marea.
- Guevara de la Serna, Ernesto (2004) [1993], *Diarios de motocicleta. Notas de viaje por América Latina*, Buenos Aires, Seven Stories Press/Ocean Sur.
- Guevara Lynch, Ernesto (1988), *Mi hijo, el Che*, La Habana Editorial Arte y literatura, La Habana.
- Hurtado Albir, Amparo (2001), *Traducción y traductología*, Madrid, Cátedra.
- Hutcheon, Linda (2021), *Teoria degli adattamenti: I percorsi della storia tra letteratura, cinema, nuovi media*, Roma, Armando Editore.
- Lie, Nada (2023), *The motorcycle diaries. Youth, travel and Politics in Latin America*, London/New York, Routledge.
- Liverani, Elena (2021), “Los lunes al sol (2002): il turpiloquio e le difficoltà di resa nella versione doppiata e sottotitolata”, *Mediazioni. Special issue: Tradurre l'oralità. Aspetti pragmatici e culturali*, 31: 191-219.
- Logaldo, Mara (2021), *Teoria e tecnica della traduzione audiovisiva. Tradurre, adattare,*

- sottotitolare per lo schermo, Roma, Audino.
- Martínez Llamas, David (2005), “Diarios de motocicleta. La forja del ‘Che’ Guevara”. *Filmhistoria*, 15 [10/06/2025] <<https://revistes.ub.edu/index.php/filmhistoria/article/view/13307>>
- Minà, Gianni (2003), *In viaggio con Che Guevara*, Documentario, Italia.
- Moratal Ibáñez, Laura María; Carli, Alberto J.; Kennel, Beatriz (2005), “Desigualdad social. Un viaje por Latinoamérica: *Diarios de motocicleta* (2004)”, *Revista de medicina y cine*, 1, 4: 103-08.
- Peris Blanes, Jaume (2010), “Viaje, experiencia, narración: de las notas viajeras de Guevara a Diarios de la Motocicleta”, *Espéculo: Revista de Estudios Literarios*, 46, [10/06/2025] <[https://www.researchgate.net/publication/277270955\\_Viaje\\_experiencia\\_y\\_narracion\\_de\\_las\\_notas\\_viajeras\\_de\\_Guevara\\_a\\_Diarios\\_de\\_la\\_Motocicleta](https://www.researchgate.net/publication/277270955_Viaje_experiencia_y_narracion_de_las_notas_viajeras_de_Guevara_a_Diarios_de_la_Motocicleta)>
- Rutelli, Romana (2004), *Dal libro allo schermo. Sulle traduzioni intersemiotiche dal testo verbale al cinema*, Pisa, ETS.
- Salles, Walter (2004a), *Diarios de motocicleta*. Film. Versione originale. Argentina, EEUU, Chile, Perú, Brasil, Reino Unido, Alemania y Francia.
- Salles, Walter (2004b), *I diari della motocicletta*. Film. Versione italiana. Argentina, USA, Cile, Perú, Regno Unito, Germania e Francia.
- Sánchez Zapatero, Javier (2021), “Los múltiples viajes de ‘Diarios de motocicleta’: de la visión de América a la forja del mito del Che”, *Así nos leen: La literatura hispánica en el cine internacional*, eds. Victoria Aranda Arribas; Tania Padilla Aguilera. Madrid, Sial Pigmalion: 153-73.
- Sinclair, Andrew (2003), *Ernesto Che Guevara*, Roma, Newton & Compton.
- Taibo II, Paco Ignacio (2017), *Ernesto Guevara: También conocido como el Che*, Barcelona, Planeta.
- Venuti, Lawrence (2013) [1995], *L'invisibilità del traduttore. Una storia della traduzione*, Roma, Armando Editore.

**Sonia Bailini** es catedrática de Lengua española, Lingüística y Traducción en la Facultad de Ciencias Lingüísticas y Literaturas Extranjeras de la Università Cattolica del Sacro Cuore de Milán. Sus líneas de investigación se centran en los procesos de adquisición de lenguas afines, en la didáctica del español como lengua extranjera y en la traducción de la variación lingüística y del humor en textos literarios.

**sonia.bailini@unicatt.it**



# ELENA LIVERANI

## LANZAROTE E SARAMAGO NELLE PAROLE DI UNA TRADUTTRICE D'ECCEZIONE: *LA INTUICIÓN DE LA ISLA* (2022) DI PILAR DEL RÍO E I PROBLEMI DI RESA IN LINGUA ITALIANA

Università IULM Milano

### Resumen

El artículo se propone investigar qué problemas de reescritura puede conllevar la traducción al italiano de *La intuición de la isla* (2022) de Pilar del Río —esposa de José Saramago—, texto dedicado a los años vividos en Lanzarote. La autora, efectivamente, además de ser esposa ha sido también la traductora al español del premio Nobel desde 1997, y en el *memoir* no faltan páginas dedicadas al relato de la redacción de algunas novelas por parte del autor portugués, textos cuya versión española ella inicia a menudo casi de manera simultánea. Inevitablemente, en *La intuición de la isla* se encuentran reminiscencias y referencias intertextuales, así como continuas alusiones y remisiones a la lengua y a la cultura portuguesas. Este singular cortocircuito ofrece la oportunidad de reflexionar sobre cómo la traducción italiana, con la inserción en el diálogo de una tercera lengua y de un nuevo público no necesariamente especializado en el universo Saramago, pueda acoger y albergar esa materia. En realidad, como se verá, la traducción italiana publicada por Feltrinelli en 2024 y firmada por la histórica traductora Rita Desti es la versión de la traducción al portugués, y esta es la razón por la cual en algunos fragmentos se ha llevado a cabo una neutralización que ya había sido prevista de antemano.

palabras clave: traducción, José Saramago, Pilar del Río, intertextualidad

### Abstract

**Lanzarote and Saramago in the words of an exceptional translator: Pilar del Río's *La intuición de la isla* (2022) and the problems of translating it into Italian**

*The article aims to explore the challenges posed by the Italian translation of *La intuición de la isla* (2022) by Pilar del Río—wife of José Saramago—, a work devoted to the years spent in Lanzarote. The author, in fact, was not only*

Saramago's wife but also his Spanish-language translator from 1997 onwards, and the memoir includes passages recounting the writing process of some of the Portuguese author's novels, texts which she often began translating into Spanish almost simultaneously. Inevitably, the book contains reminiscences and intertextual references, as well as frequent allusions to Portuguese language and culture. This unique interplay offers an opportunity to reflect on how the Italian translation —introducing a third language into the dialogue and addressing a new readership not necessarily familiar with Saramago's universe— can accommodate and convey such material. As will be shown, the Italian version published by Feltrinelli in 2024, signed by long-standing translator Rita Desti, is based on the Portuguese translation; this explains why some passages had already undergone a process of neutralisation.

keywords: translation, José Saramago, Pilar del Río, intertextuality

*Escrever é traduzir. Sempre o será*  
(J. Saramago, 2009<sup>1</sup>)

Il 10 novembre 2013, nell'inserto *Il club della lettura* del *Corriere della sera*, Luciano Canfora intitolava un suo contributo “Chi non traduce rinuncia a pensare”. La traduzione, diceva, “è stata il motore principale del progresso civile” e “tradurre è il più autentico dialogo del genere umano”. Per poi interpolare una riflessione sulle difficoltà manifestate da alcune culture – quella anglosassone per esempio – a tradurre:

Ma dove nasceva la difficoltà? Non solo nella profondità del pensiero di cui appropriarsi, ma soprattutto nella *lontananza*. Ed è appunto tale lontananza che fece e fa tuttora di quell'esercizio, di quello sforzo di interrogazione, un cantiere sempre aperto, sempre provvisorio, sempre passibile di prospettive prima non viste.

Tradurre come dialogo dunque. Ma con la difficoltà comportata dalla ‘lontananza’, concetto chiave anche nella visione che di tale attività Antoine Berman offrì con il titolo del celebre seminario del 1984, successivamente pubblicato nel 1991 e in un'edizione postuma del 1999, *La traduction et la lettre ou l'auberge du lointain*. Come puntualizza il curatore dell'edizione italiana, Gino Giometti, “Berman stabilisce una corrispondenza fra l'albergo della lontananza e la lingua traduce che accoglie l'estraneità della lingua straniera in una vicina lontananza” (Berman 2003: 7). E dunque, con l'istituzione di tale albergo, lo studioso si distaccava dalla prospettiva estetizzante delle teorie più tradizionali ispirate alla ricerca dell'equivalenza.

1 Cfr. Baltrusch (2003).

Un nuovo statuto della traduzione, grazie alla fortuna dei *Translation Studies*, secondo cui “tradurre è saper uscire da sé, saper riconoscere l’apertura alla differenza/alterità (*otherness*)” (Bertazzoli 2016: 7), non può non informare i processi di resa di testi legati alla dimensione del viaggio che costituzionalmente si pongono come occasioni di incontro e di conoscenza del diverso e che svolgono la preziosa funzione di prevenire la stereotipizzazione di ciò che è lontano da noi. Da qui che tradurre il viaggio implichi sempre sorvegliare la tentazione di ricondurre l’estraneità all’interno della cultura ricevente<sup>2</sup>.

In questo perimetro di riflessioni, potrebbero sembrare poco pertinenti le suggestioni etnocentriche di Calvino che in *Sul tradurre*, nel 1963, sosteneva che il traduttore doveva essere dotato di “intelligenza dello stile, nel doppio aspetto del comprendere le peculiarità stilistiche dell’autore da tradurre, e del saperne proporre equivalenti italiani in una prosa che si legga *come fosse stata pensata e scritta direttamente in italiano*” (46). Ma le considerazioni del grande intellettuale sanremese sulla traduzione rimangono un’importante bussola per chi si occupa di questo ambito di studi perché, come ricordano Ewa Nicewicz-Staszowska (2017) e Francesca Rubini (2023), il rapporto con la traduzione di Calvino percorre tutta la sua biografia letteraria e le intuizioni sono il frutto del suo lavoro di traduttore, di consulente editoriale, delle conversazioni con i traduttori delle sue opere o di quelli che lavoravano presso la casa editrice Einaudi. Celebre è il titolo di un suo intervento del 1982, in un convegno di anglisti a Roma, “Tradurre è il vero modo di leggere un testo” (78-84), in cui ci parla di un continuo esercizio di lettura, incessante, quasi che il traduttore debba essere in primis un lettore ossessivo: del testo che deve tradurre, come pure di tutto quello che ha scritto l’autore e di tutto quello che è stato scritto su di lui<sup>3</sup>.

Chi ha il privilegio di tradurre letteratura sa a quale ossessione si stia alludendo e forse intuisce anche che la consuetudine all’esercizio del tradurre talora può limitare il piacere della lettura. Perché comporta il rischio di non abbandonarsi mai completamente al testo, mantenendo la mente sempre vigile. Se si legge nella lingua straniera da cui si traduce è inevitabile continuare a chiedersi come si renderebbe una certa immagine o un certo costrutto, e se si legge in traduzione involontariamente si cerca di ipotizzare quale fosse la ‘lettera’ del testo fonte che soggiace alle pagine che abbiamo davanti; inoltre, si diventa inconsapevolmente ipersensibili agli attriti linguistici, veri e presunti, in cui si ha la sensazione di imbattersi, così come si rimane rapiti davanti a soluzioni particolarmente felici

2 Cfr. Liverani (2001 e 2010).

3 Cfr. anche il capitolo “Sulla lettura” in Cavagnoli (2012).

o che semplicemente non rientrano nell'arsenale di preferenze linguistiche che come traduttori si tendono a proporre.

È con questo filtro percettivo che ho letto il testo di Pilar del Río quando, nel 2022, Feltrinelli mi ha affidato la redazione di una scheda: con una ossessione un po' maniacale rispetto a quelli che potevano essere i problemi di resa in lingua italiana. Infatti, tradurre in italiano *La intuición de la isla. Los días de José Saramago en Lanzarote* comporta una serie di scelte sia per quanto riguarda la resa dell'alterità – asse portante di ogni progetto traduttivo che investa la narrativa di viaggio in senso lato (in questo caso un viaggio nell'isola di Lanzarote) – sia per quanto riguarda la trasposizione dei rimandi intertestuali di cui è fittamente punteggiato il testo e, non ultimo, per quanto attiene al dialogo che si instaura tra tre realtà linguistiche: lo spagnolo dell'autrice, il portoghese che è la lingua con cui scrive e si esprime Saramago, e la lingua italiana dei destinatari della traduzione.

Il testo oggetto delle nostre riflessioni è un *memoir* pubblicato nel 2022 in cui l'autrice, moglie dello scrittore portoghese, ricostruisce gli anni dell'esilio, o meglio dell'autoesilio, a Lanzarote. È lo stesso José Saramago (1922-2010), insignito del Premio Nobel per la Letteratura nel 1998, a riassumere la sua vicenda umana con queste parole:

Con excepción de otra obra de teatro, titulada *A segunda vida de Francisco de Assis* y publicada en 1987, la década de los ochenta fue enteramente dedicada a la novela: *Memorial del convento* (1982), *El año de la muerte de Ricardo Reis* (1984), *La balsa de piedra* (1986), *Historia del cerco de Lisboa* (1989). En 1986 conocí a la periodista española Pilar del Río. Nos casamos en 1988. En respuesta a la censura ejercida por el Gobierno portugués sobre la novela *El Evangelio según Jesucristo* (1991), vetando su presentación al Premio Literario Europeo con el pretexto de que el libro era ofensivo para los católicos, cambiamos, mi mujer y yo, en febrero de 1993, nuestra residencia a la isla de Lanzarote, en el archipiélago de Canarias. (Saramago 2014)

Il frammento è tratto da un volume pubblicato nel 2014 da Alfaguara, la nota casa editrice madrilena, che per i suoi 50 anni di attività (1964-2014) decide di rendere consultabile una raccolta di guide di lettura gratuite in formato digitale degli autori più rappresentativi del suo catalogo, per *abrir ventanas al mundo*. Lo spirito di tale progetto è condensato nelle parole pronunciate dallo scrittore israeliano Amos Oz in occasione del conferimento del Premio Príncipe de Asturias de las Letras, nel 2007:

Si adquieres un billete y viajas a otro país, es posible que veas las montañas, los palacios y las plazas, los museos, los paisajes y los enclaves históricos. Si te sonríe la



fortuna, quizá tengas la oportunidad de conversar con algunos habitantes del lugar. Luego volverás a casa cargado con un montón de fotografías y de postales. Pero, si lees una novela, adquieres una entrada a los pasadizos más secretos de otro país y de otro pueblo. La lectura de una novela es una invitación a *visitar las casas* de otras personas y a conocer sus estancias más íntimas<sup>4</sup>.

Queste immagini, particolarmente suggestive per chi si occupa di letteratura di viaggio, devono essere state fonte di ispirazione anche per Pilar del Río, che con il suo *memoir* sembra quasi voler offrire una guida di lettura ulteriore, più intima, consentendo ai lettori del maestro portoghese di visitare *A Casa*, la vera protagonista del suo libro.

La quarta di copertina sintetizza in modo chiaro la filosofia che anima la stesura del volume:

En *La intuición de la isla. Los días de José Saramago en Lanzarote*, Pilar del Río construye un mosaico de momentos vividos, emociones compartidas y los libros escritos bajo la luz de la isla que el escritor portugués escogió para vivir. Una forma de compartir con los lectores momentos singulares vividos en *A Casa* y cómo era la vida para José Saramago mientras escribía sus obras: los paseos por Lanzarote, las ideas de las que surgieron sus novelas, la convivencia con sus perros, los encuentros en la isla con amigos como Carlos Fuentes, Ernesto Sábato, Susan Sontag o Bertolucci, las experiencias que traía de los viajes y las amistades forjadas. Un libro irremplazable que busca continuar la respiración que se siente en *A Casa* y compartirla. Un libro para amigos y amigas.

Per le riflessioni che ci riproponiamo di svolgere, è certamente utile ripercorrere, in ordine cronologico, la storia editoriale dei più noti testi di Saramago in lingua portoghese e nelle traduzioni in spagnolo e italiano.

<i>Terra do pecado</i> , Minerva, Lisboa (1947)	<i>La viuda</i> Trad. A. Sáez Delgado Alfaguara (2021)	<i>La vedova</i> . Trad. R. Desti Feltrinelli (2022)
<i>Manual de Pintura e Caligrafia</i> , Moraes Editores, Lisboa (1977)	<i>Manual de pintura y caligrafía</i> . Trad. B. Losada Alfaguara (2004)	<i>Manuale di pittura e calligrafia</i> . Trad. R. Desti Bompiani (1996) Einaudi (2003)

<sup>4</sup> <<https://www.penguinlibros.com/uy/tematicas/14098-ebook-una-guia-para-leer-a-jose-saramago-9788420418087/fragmento>>.

<i>Levantado do chão</i> , Caminho, Lisboa (1980)	<i>Levantado del suelo</i> Trad. B. Losada Alfaguara (1988)	<i>Una terra chiamata Alentejo.</i> Trad. R. Desti Bompiani (1992) Einaudi (2006) Feltrinelli (2010)
<i>Memorial do Convento</i> , Caminho, Lisboa (1982)	<i>Memorial del convento.</i> Trad. B. Losada Alfaguara (1998)	<i>Memoriale del convento.</i> Trad. R. Desti e C. Radulet Einaudi (1993) Feltrinelli (1999)
<i>O Ano da Morte de Ricardo Reis</i> , Caminho, Lisboa (1984)	<i>El año de la muerte de Ricardo Reis</i> Trad. B. Losada Seix Barral (1985) Alfaguara (2003)	<i>L'anno della morte di Ricardo Reis.</i> Trad. R. Desti Feltrinelli (2010)
<i>A jangada de pedra</i> , Caminho, Lisboa (1986)	<i>La balsa de piedra</i> Trad. B. Losada Seix Barral (1987) Alfaguara (1998)	<i>La zattera di pietra</i> (1988) Trad. R. Desti Einaudi (1997) Feltrinelli (2010)
<i>História do cerco de Lisboa</i> , Caminho, Lisboa (1989)	<i>Historia del cerco de Lisboa</i> Trad. B. Losada Seix Barral, (1989) Alfaguara (1998)	<i>Storia dell'assedio di Lisbona</i> , Bompiani (1990) Einaudi (2000)
<i>O Evangelho segundo Jesus Cristo</i> , Caminho, Lisboa, 1991	<i>El evangelio según Jesucristo</i> Trad. B. Losada Seix Barral (1992) Alfaguara (1998)	<i>Il Vangelo secondo Gesù Cristo.</i> Trad. R. Desti Bompiani (1998) Einaudi (2002) Feltrinelli (2010)
<i>Ensaio sobre a cegueira</i> , Caminho, Lisboa (1995)	<i>Ensayo sobre la ceguera</i> , Trad. B. Losada Seix Barral (1996) Alfaguara (2007)	<i>Cecità</i> Trad. R. Desti Einaudi (1996) Feltrinelli (2010)
<i>O conto da ilha desconhecida</i> , Assírio & Alvim, Lisboa (1997)	<i>El cuento de la isla desconocida</i> Trad. P. del Río Alfaguara (1996)	<i>Il racconto dell'isola sconosciuta</i> Trad. e cura di P. Collo e R. Desti Einaudi (1998)
<i>Todos os nomes</i> , Caminho, Lisboa (1997)	<i>Todos los nombres</i> (1998) Trad. P. del Río Alfaguara (1998)	<i>Tutti i nomi.</i> Trad. R. Desti Einaudi (1997) Feltrinelli (2010)

<i>Viajem a Portugal</i> , Círculo de Leitores (1981) Caminho (1990)	<i>Viaje a Portugal</i> , Alfaguara (1995) Trad. B. Losada Alfaguara (2001)	<i>Viaggio in Portogallo</i> Trad. R. Desti Bompiani (1996) Einaudi (1999) Feltrinelli (2011)
<i>O homem duplicado</i> , Caminho, Lisboa (2002)	<i>El hombre duplicado</i> (2003) Alfaguara (2003) Trad. P. del Río	<i>L'uomo duplicato</i> Trad. R. Desti Einaudi (2003) Feltrinelli (2010)
<i>Ensaio sobre a lucidez</i> , Caminho, Lisboa (2004)	<i>Ensayo sobre la lucidez</i> Trad. P. del Río Alfaguara (2004)	<i>Saggio sulla lucidità</i> Trad. R. Desti Einaudi (2004) Feltrinelli (2013)
<i>As intermitências da morte</i> , Caminho, Lisboa (2005)	<i>Las intermitencias de la muerte</i> Trad. P. del Río Alfaguara (2005)	<i>Le intermittenze della morte</i> Trad. R. Desti Einaudi (2005) Feltrinelli (2013)
<i>As Pequenas Memórias</i> , Caminho, Lisboa (2006)	<i>Las pequeñas memorias</i> Trad. P. del Río. Alfaguara (2007)	<i>Le piccole memorie</i> Trad. R. Desti Einaudi (2007) Feltrinelli (2015)
<i>A viagem do elefante</i> , Ca- minho, Lisboa (2008)	<i>El viaje del elefante</i> Trad. P. del Río. Alfaguara (2008)	<i>Il viaggio dell'elefante</i> Trad. R. Desti Einaudi (2010) Feltrinelli (2015)
<i>Caim</i> , Caminho, Lisboa (2009)	<i>Cáin</i> Trad. P. del Río. Alfaguara (2009)	<i>Caino</i> Trad. R. Desti Feltrinelli (2010)
<i>Alabardas, alabardas, espín- gardas, espingardas</i> , Porto Editora (2014)	<i>Alabardas</i> . Trad. P. del Río Alfaguara (2014)	<i>Alabarde alabarde</i> Trad. R. Desti Feltrinelli (2014)
<i>O último caderno de Lan- zarote. O diário do ano no Nobel</i> , Porto Editora (2018)	<i>El cuaderno del año del Nobel</i> Trad. A. Sáez Delgado Alfaguara (2018)	<i>Diario dell'anno del Nobel, L'ul- timo quaderno di Lanzarote</i> Trad. R. Desti Feltrinelli (2021)

I dati appena offerti sono desunti dai cataloghi disponibili in rete e dalla bibliografia dell'autore riportata sul sito della Fundação José Saramago (<https://www.josesaramago.org/es/>). Pur nella loro provvisorietà, l'impressione è che i riscontri di questa ricognizione offrano un quadro abbastanza chiaro. Quando Saramago si afferma come autore, inizia a essere pubblicato con regolarità dalla casa editrice

di Lisboa Caminho. In Spagna, se si fa eccezione per qualche libro uscito per i tipi di Seix Barral, è Alfaguara a pubblicare tutti i suoi testi. Tuttavia, a un certo punto si interrompe la collaborazione con il traduttore storico, Basilio Losada – noto studioso, insignito anche con il premio nazionale proprio per la sua versione di *História do cerco de Lisboa*<sup>5</sup> – che viene sostituito da Pilar del Río.<sup>6</sup>

Per quanto riguarda la situazione italiana, un dato evidente balza all'occhio: l'autore viene dapprima pubblicato da Bompiani; successivamente acquisisce i diritti delle sue opere Einaudi – che ripubblica anche i testi precedenti – e infine il catalogo viene integralmente assorbito da Feltrinelli, che intorno al 2010 inizia a ripubblicare tutti i testi precedenti e a far uscire quelli nuovi. Ma la traduttrice rimane sempre la stessa, Rita Desti, che a buon diritto si guadagna il titolo di voce italiana di Saramago<sup>7</sup>. Non stupisce, dunque, che la casa editrice Feltrinelli abbia deciso di affidare alla nota lusitanista nonché prolifica traduttrice dal portoghese la traduzione del testo di Pilar del Río, redatto in lingua spagnola – ma di cui quasi in contemporanea sono uscite due versioni in portoghese, una per Porta Editoria e una per il mercato brasiliano<sup>8</sup> –, giacché la conoscenza integrale e approfondita, quella che solo una grande traduttrice come lei può avere dell'opera completa dell'autore che *La intuición de la isla* omaggia, rappresenta certamente una garanzia per la resa di un testo che, a dire il vero, per l'incedere narrativo a prima vista non sembra porre grandi problemi di resa, soprattutto per una penna tanto addestrata come la sua a ben altre andature. A riprova di ciò, paradigmatiche sono le parole che le dedica Fabio Stassi (2021):

5 Si veda il sito del Ministerio de Cultura Spagnolo: <[https://www.cultura.gob.es/premiado/busquedaPremioParticularAction.do?cache=init&layout=PremioNacMejorTraduccionLibro&params.id\\_tipo\\_premio=98&language=es&TOTAL=49&POS=45&MAX=15&action=goToPage&PAGE=2](https://www.cultura.gob.es/premiado/busquedaPremioParticularAction.do?cache=init&layout=PremioNacMejorTraduccionLibro&params.id_tipo_premio=98&language=es&TOTAL=49&POS=45&MAX=15&action=goToPage&PAGE=2)>.

6 Fino a poco tempo fa, sul sito della Fondazione si indicava un unico caso in cui si attribuiva a entrambi la traduzione di un testo, *Viajem a Portugal*, che per la verità in Spagna esce in numerose edizioni. Peraltro, anche in Portogallo la casa editrice Caminho tarda a editarlo ed esce dapprima per i tipi di Círculo de Leitores nel 1981 e nove anni più tardi in una nuova edizione con la prefazione di Claudio Magris, circostanza cui si accenna anche ne *La intuición de la isla*. Attualmente (ottobre 2025) la pagina dedicata a questo testo e alle sue traduzioni è stata rimossa.

7 Solo in due casi, tra quelli da me analizzati, si riscontra la presenza di un'altra figura, o nella veste di curatore o in quella di cotraduttrice.

8 *A intuição da ilha*, Porto, Porto Editora 2022; questo testo non è indicato come traduzione, mentre la versione brasiliana dallo stesso titolo, edita dalla Companhia das Letras nello stesso anno, specifica che si tratta di una traduzione a firma di Sérgio Machado Letria.

Ho sempre pensato al traduttore come a un musicista. L'interprete che si accosta a uno spartito già scritto ma a cui manca lingua, timbro, intonazione e tocco. Il suono e il fiato. In una parola, la voce. Il suo compito è quello di restituirne la musica, e restituire è un verbo difficile, in letteratura, perché ha a che fare con la natura più intima del raccontare. Scrivere è anche, e forse soprattutto, un tentativo di restituire qualcosa a qualcuno (un'altra versione della verità? un segreto mai detto?), nell'incoscienza di sanare delle ferite, in una necessità quasi infantile di riparazione. [...]

Ho letto, riletto e amato José Saramago dal primo libro che fu tradotto in Italia, e tutta la mia gratitudine va a Rita Desti capace di riconsegnare con la stessa potenza dell'originale la fluvialità della sua prosa e il primato della virgola nella nostra lingua. Ne ho smontato e studiato a lungo le costruzioni sintattiche: ogni nuovo libro era per me un appuntamento e un esercizio entusiasmante. Volevo indovinare il tempo interno di ogni testo, se suonava in levare o in battere, se era dispari o pari, il ritmo, insomma, che la traduttrice era riuscita a mettere in salvo e a riprodurre. Come se si potesse tradurre il respiro di un autore, il suo giro di fiato.

Come si diceva, Pilar del Río subentra come traduttrice di Saramago tardivamente, in occasione di un racconto, *O conto da ilha desconhecida*, che esce nel 1997 per i tipi della lisboneta Assírio & Alvim. Forse fu la misura del testo a spingere Caminho a cederne i diritti a un altro editore, mentre in Spagna e in Italia sono quelli tradizionali a pubblicarlo nell'anno successivo, e ciò poté determinare il cambio di traduttore; oppure furono le circostanze di vita che a breve ricorderei.

Pilar del Río e José Saramago – protagonisti del bel documentario *José y Pilar* di Miguel Gonçalves Mendes del 2010 rinvenibile in rete<sup>9</sup> – si conoscono nel 1986, quando lui aveva sessantatré anni e lei trentasei, e dopo poco, nel 1988, si sposano. Cinque anni più tardi, nel 1993, si trasferiscono a vivere a Lanzarote.

È comprensibile dunque l'avvicendamento tra traduttori voluto da Alfaguara che fino al 1996, per *Ensayo sobre la ceguera*, si era rivolta a Basilio Losada: dal 1997, invece, quasi tutti i libri in spagnolo di Saramago sono a firma della moglie<sup>10</sup>. Senza nulla togliere alla sua professionalità, viene da pensare che la fortuna editoriale di Pilar del Río sia strettamente legata alla dimensione umana della vicinanza con il noto scrittore. Che il traduttore precedente fosse molto titolato è un dato incontrovertibile, come lo è che l'attuale attività dell'autrice del nostro *memoir* presso la Fundação José Saramago di Lisbona, di cui è presidentessa, e più

9 <<https://www.youtube.com/watch?v=7gtRxfFi0>>.

10 Fanno eccezione *La viuda* del 2021, traduzione di *Terra do pecado* (1947) a firma di A. Sáez Delgado, che traduce anche *El Cuaderno del año del Nobel* (2019).

in generale tutto il lavoro che svolge per mantenere viva la memoria, i progetti, l'eredità culturale morale del premio Nobel siano assolutamente meritevoli.

Le difficoltà di resa in italiano di *La intuición de la isla* che avevo ipotizzato si sarebbero presentate mi pareva che si concentrassero soprattutto in due ambiti: la fitta rete di rimandi intertestuali da restituire nella lingua della comunità d'arrivo e la frequente presenza di termini in lingua portoghese che punteggiano il *memoir*.

Per quanto riguarda la traduzione dei rinvii intertestuali<sup>11</sup>, ciò che per un traduttore meno avvezzo a dialogare con l'opera di Saramago avrebbe potuto richiedere tempo e pazienza, con molte ore dedicate alla consultazione delle opere nella traduzione in lingua italiana per il reperimento delle citazioni o delle allusioni<sup>12</sup> – giacché non avrebbe nessun senso tradurre i riferimenti dalla versione in spagnolo, vieppiù in presenza di versioni accreditate –, avevo supposto che per Rita Desti non avrebbe rappresentato una sfida impegnativa. Persino quando il richiamo non prende le forme della citazione diretta, come avviene in molti punti del testo, ma si configura come mera allusione, per la storica traduttrice di Saramago riconoscere che dietro la sostanza linguistica spagnola si cela una forma diversa in portoghese – che in taluni casi va ripristinata – non poteva certamente costituire un inciampo. *La intuición de la isla*, che in fondo è un omaggio postumo all'autore e alla sua opera, si snoda alternando affondi biografici a richiami alla vasta produzione dello scrittore esule dai cui testi sono spesso tratti gli incipit dei numerosi capitoli. Semmai, il vero punto nodale sarà stato decidere quale fosse il progetto editoriale, cioè a quale pubblico si rivolgesse la traduzione italiana e quanto si presupponeva che il lettore a cui è destinato dovesse conoscere dell'universo saramaghiano<sup>13</sup>. A titolo di esempio, si osservi il seguente frammento:

[...] dejando embelesados a los comensales con la fuerza de la historia, la audacia de sus protagonistas, el hallazgo de las puertas de las peticiones y de los agradecimientos,

11 Cfr. il capitolo “Far sentire il rinvio intertestuale” in Eco (2003: 48), e Osimo (2010: 50).

12 Cfr. Genette e Bernardelli. Si ricorda che l'allusione, priva di dispositivi di isolamento tipografico, può essere una ripresa volontaria (come nel caso che ci occupa) o inconsapevole.

13 Una ricognizione di alcuni dei frammenti riconoscibili come citazioni o interpretabili come allusioni volontarie si trova nella bella tesi di laurea magistrale di Eleonora Gardella, che ha offerto una traduzione di alcune parti di questo testo corredata da un commento, svolta sotto la mia direzione nell'a.a. 2022/2023. Da qui estrapolo i due esempi che cito. Alla revisione del presente articolo hanno contribuito le collaboratrici alla ricerca dell'International Center for Research on Collaborative Translation (<https://www.iulm.it/speciali/International-Center-for-Research-on-Collaborative-Translation>) diretto da Francesco Laurenti e da me presieduto.

la obstinación *de uno*, la seguridad *de otra*... (del Río 2022: 110. Il corsivo è mio)

L'autrice qui sta riassumendo la trama di *O conto da ilha desconhecida*, il primo testo da lei tradotto. Mi sono domandata se il lettore italiano sarebbe stato messo nelle stesse condizioni di quello spagnolo, privato di elementi contestuali, oppure se avrebbe trovato qualche informazione in più rispetto ai due personaggi cui si fa cenno in modo generico – *uno* e *otra* –, il re e la donna delle pulizie che lo accompagna. Come era prevedibile, la versione data alle stampe in italiano si è attenuta alla lettera del testo fonte.

Parimenti, nel seguente passaggio, si pone un problema di coerenza interna. Nella carrellata di protagonisti dei testi di Saramago, Pilar del Río ricorda:

Luego serían los personajes de las distintas obras de Saramago, seres sin brillo social ni asientos en consejos de administración, que buscan a la mujer desconocida, o *se levantan del suelo*, o se empeñan en organizar una vida humana en plena época de pandemia de ceguera, o navegan hacia otros siendo balsas que transportan tierra y sueños. (2022: 161. Il corsivo è mio)

Per il lettore spagnolo le allusioni ai titoli delle opere risultano chiare, mentre quello italiano potrebbe avere difficoltà a cogliere il rimando a *Levantado do chão*, reso in spagnolo con *Levantado del suelo*, ma che in italiano ha preso il titolo *Una terra chiamata Alentejo*. In questo caso, al pubblico italiano, che leggerà la traduzione “si rialzano da terra” (del Río 2024: 110), non viene offerto nessun indizio che agevoli la comprensione del rimando.

Come si diceva, il dialogo tra la lingua spagnola e quella portoghese nel testo è fitto e continuo, quasi programmatico, come ci racconta Pilar del Río, che naturalmente riporta in corsivo tutti i lessemi propri della lingua portoghese, a cominciare dalla dimora, sempre chiamata *A Casa*. Ci si sarebbe aspettati che nella traduzione italiana queste volute incursioni di una lingua altra venissero mantenute nella forma assunta nel testo fonte, perché tali indizi di alterità, a partire dal primo Natale sull'isola quando “se escucharon dos idiomas en la mesa” (83), hanno chiaramente lo scopo di restituire un'immagine di *A Casa* – “lugar donde trabajar, juntar idiomas, recibir amigos” (34) – come microcosmo che è altresì crogiolo e intersezione di due culture che dialogano tra loro. È sorprendente, a mio avviso, che nella versione italiana si traduca ‘La Casa’ e vengano neutralizzati molti dei lessemi portoghesi o spagnoli, tendenza evidente nella resa dei riferimenti gastronomici<sup>14</sup>.

14 Si veda il ricorrente *bacalhau com todos* (83 e 232) reso con ‘baccalà per tutti’ (17 e 57) o il “baccalao asado con *batatas a murro*” (79) tradotto con ‘baccalà al forno con patate’ (54). Tralasciamo

Da un punto di vista metodologico è stato interessante riflettere, anche in astratto, sulle scelte traduttive che si possono o debbono operare qualora si tratti di restituire una materia culturale e linguistica (quella portoghese) che ci arriva tramite la mediazione di quella spagnola e che va proposta in una terza dimensione, quella italiana. Come già ci era capitato di osservare (Liverani 2001, 2010), al traduttore nostrano spetta il compito di sorvegliare attentamente la triangolazione linguistica e culturale e di ripristinare elementi che risultano assenti nel testo fonte laddove si siano operate neutralizzazioni a beneficio del lettore spagnolo. È il caso, in genere, della toponomastica, dei riferimenti geografici<sup>15</sup> o di menzioni puntuali come possono essere nel nostro testo il ricordo del “cine El Piojo” (del Río 2022: 271) opportunamente ripristinato in ‘O Píolho’ (186). Può sorprendere che il rimando al “Movimiento de los Sin Tierra en Brasil” (174), per cui si sarebbe potuta agevolmente recuperare la forma originaria, peraltro ben nota a tutti, ‘Movimento Sem Terra’ sia stato neutralizzato con la traduzione ‘movimiento dei senza terra in Brasile’ (119).

La dimensione plurilingue di *A Casa* e la relazione di Saramago con la sua lingua materna è ben chiarita nel seguente frammento. Come ci racconta Pilar del Río, quando nel 2007 viene istituita la Fundação José Saramago, lo scrittore è felice di poter tornare spesso a Lisbona a sentire

el dulce sonido del portugués, que tanto echaba en falta en Lanzarote, ya que en su casa se hablaba español, porque no podía soportar los esfuerzos de su entorno por expresarse en su lengua. “Dejadlo, no lo intentéis, sois incapaces” se lamentaba cuando con buena voluntad intentaban construir en portugués y la frase salía, una y otra

volutamente di occuparci dei molteplici problemi che pone la traduzione in italiano di culturemi spagnoli – quali, per esempio, i tipici piatti canari come le “papas *arrugás* [...] queso con *mojo picón*” (39) (resi con ‘patate *arrugás* [...] formaggi dell’isola con salsa *picón*’ (28) – perché esulano dal nostro campo d’analisi. Due, tuttavia, ci sono parsi particolarmente spinosi. Il primo riguarda la resa in italiano della variazione sociolinguistica argentina. Recita il *memoir*: “María Kodama reproduce el usted acariciador con el que Borges y ella se trataban” (257). La traduzione offerta è la seguente: ‘Kodama riproduce il tono carezzevole con cui lei e Borges si trattavano’ (176) che lascia all’oscuro del fenomeno il lettore italiano. Il secondo passaggio di difficile soluzione è quello in cui l’autrice, ricordando l’ateismo che caratterizzava José Saramago, puntualizza: “El autor no tiene convicciones religiosas, pero vive en una cultura cristiana, se despierta con campanadas de iglesias, oye música llamada sagrada, se comunica con personas que tienen fe en otra vida, dice adiós y ojalá.” (55). La scelta di restituire in lingua italiana le interiezioni *adiós* e *ojalá* con ‘addio e inshallah’ (38) mi pare felice e appropriata, anche se non rende la grande frequenza d’uso dell’interiezione che esprime desiderio in spagnolo.

15 È banale, ma per il lettore italiano, il fiume che per gli spagnoli è il Duero, nel tratto portoghese va riportato alla forma portoghese Douro.



vez, como si fuera una burla cada vez peor. “Cada uno hablará su idioma y todos nos entenderemos” fue la decisión. [...] Lo que sí está claro es que en los últimos años de su vida, José Saramago pasó a necesitar más su idioma, a valorarlo, a sentirse cómodo en él, a echarlo de menos cuando no lo tenía. La saudade del portugués llegó hasta el punto de... (303-304)

Era lecito domandarsi se in questo preciso contesto la traduttrice avrebbe deciso di preservare il lessema *saudade*, attestato anche in spagnolo (tanto che nel testo fonte non appare in corsivo), come pure in italiano quale derivato lusitano, che nell'immaginario comune spontaneamente si associa alla cultura portoghese, ma la scelta della casa editrice è stata di naturalizzare con un più consueto 'nostalgia' (208).

In *A Casa* si parlano più lingue e si traduce di continuo dal portoghese allo spagnolo, come avviene nel caso della prima lettura in presenza di amici e vicini di casa de *El cuento de la isla desconocida* di cui si decise che “se haría una traducción improvisada del portugués en que estaba escrito al español en que sería leído por primera vez” (del Río 2022: 109)<sup>16</sup>.

L'eccezionalità dell'esperienza di Pilar del Río risiede proprio nel fatto che da un certo momento in poi, stando alla sua testimonianza, la stesura di un romanzo veniva simultaneamente accompagnata dal processo di traduzione dello stesso in spagnolo:

José Saramago escribía en portugués, era su idioma y su bandera. Algunas veces escribía artículos directamente en francés, otras, pocas, en español, pero en su casa el idioma que se sentía era el portugués, aunque su entorno hablara en sonoro castellano. O sea, José Saramago escribía en portugués y ese era el idioma oficial de las obras literarias que se producían, aunque en simultáneo se estuvieran produciendo la ceremonia de la traducción al español. Carlos Fuentes definió el trabajo que vio hacer en *A Casa* como un interesante juego de espejos, un diálogo sin palabras, las páginas mirándose distintas, aunque iguales, ante el silencio de quienes las habían mecanografiado. (2022: 201)

Il frammento fa preciso riferimento alla stesura e alla traduzione di *O homem duplicado* ed è interessante rilevare che in questo gioco di specchi a guardarsi non sono i due autori – José Saramago e Pilar del Río, autrice della versione spagnola – e nemmeno le due lingue, ma proprio i due testi che si riflettono l'uno nell'altro. Che tale immagine venga associata alla stesura del romanzo *O homem duplicado*

16 Come ci viene raccontato, il testo era stato commissionato da Simonetta Luz Alfano per il padiglione del Portogallo dell'Expo 1998 a Lisbona.

e alla sua resa in spagnolo innesca un cortocircuito per cui la traduzione sembra assumere la forma della duplicazione.

Le parole con cui in un'intervista con Claudia Mancini (2012) l'autrice del *memoir* chiosa la peculiarità del contesto traduttivo nel quale era chiamata a operare sono certamente illuminanti:

Non mi costò molto lavoro tradurre, perché sentivo la voce di Saramago nei suoi libri e nella vita: ed è questo che manca agli altri lettori; sentirlo nella vita quotidiana è un mio privilegio. E a volte anche un mio problema, perché cercavo di avvicinarmi tanto a quella musica che rischiavo di perdere la libertà nella mia lingua.

Convinti dell'importanza di tale esperienza dovevano essere “también los traductores empeñados en su trabajo que querían sentir el portugués del autor en su cotidianidad” (del Río 2022: 279), e nella mia pratica professionale non posso che confermare quanto tale opportunità sia di estremo aiuto. In questo dialogare a distanza, la voce – fittizia, mediata, affidata alle proprie creature letterarie o reale che sia – svolge un ruolo fondamentale.

In un contributo di grande interesse per lo studio della traduzione collaborativa, Jean-Renée Ladmiral (2018) analizza i casi di traduzione al plurale distinguendo le traduzioni svolte in due o in gruppo. Nelle traduzioni di coppia individua due diverse possibilità: il binomio, cioè la coppia in cui sono coinvolte due persone con competenze diverse, ma complementari, e la traduzione in tandem in cui uno dei traduttori conosce meglio la lingua di partenza e l'altro quella di arrivo. Evidentemente la collaborazione tra José Saramago e Pilar del Río non può essere ricondotta a nessuna delle due categorie, in primis perché nulla induce a credere che lo scrittore portoghese intervenisse direttamente nella resa in spagnolo dei suoi testi. Ma, come si ricordava, tradurre comporta sempre un dialogo con la lettera da tradurre e quando è possibile anche con l'autore del testo di partenza. Forzando i termini della tassonomia, si può dunque azzardare che il cambio di traduttore abbia beneficiato di questa felice circostanza: l'opportunità di cui Pilar del Río si è avvalsa di mettere in moto quasi in tempo reale un processo virtuoso di traduzione collaborativa.

La traduzione, dunque, come esercizio di inclusione a tutto campo. Ed è proprio José Saramago, che a tale attività rendeva grande merito, a ricordarcelo con le poetiche ed efficaci parole pronunciate in una conferenza a Buenos Aires, nel 2003, intitolata “Todo son traducciones, todos somos traductores”:

Durante varios años he sido traductor, he vivido del pan que ganaba traduciendo, luego sé de qué va este oficio al que siempre rindo homenaje con una aseveración que

me parece justa y que ando repitiendo a lo largo y ancho del mundo: ‘Los escritores hacen las literaturas nacionales y los traductores hacen la literatura universal’. [...] Los traductores convierten el archipiélago incommunicable de los idiomas en un lugar de acogida. Ellos son los ingenieros que alzan los puentes necesarios para que transitemos quienes no dominamos los idiomas del mundo. Ellos nos aplanan el terreno del saber y del gozo.

## Bibliografia citata

- Berardelli, Andrea (2013), *Cos'è l'intertestualità*, Roma, Carocci.
- Bertazzoli, Raffaella (2016), *La traduzione: teorie e metodi*, Roma, Carocci.
- Berman, Antoine (2003), *La traduzione e la lettera o l'albergo della lontananza*, Roma, Quodlibet.
- Baltrusch, Burghard (2003), “Escrever è traduzir. José Saramago e a tradução”, *Convergência Lusitana*, 34: 100-35.
- Calvino, Italo (2002), *Mondo scritto e mondo non scritto*, ed. Mario Barenghi, Milano, Mondadori.
- Canfora, Luciano (2013), “Chi non traduce rinuncia a pensare”, *Corriere della sera. Il club de La lettura* (10/11/2013).
- Cavagnoli, Franca (2012), *La voce del testo*, Milano, Feltrinelli.
- Eco, Umberto (2003), *Dire quasi la stessa cosa*, Milano, Bompiani.
- Genette, Gérard (1997), *Palinsesti*, Torino, Einaudi.
- Ladmiral, Jean-René (2018), “La traduction au pluriel”, *Traduire à plusieurs – Collaborative Translation*, eds. Enrico Monti; Peter Schnyder. Paris, Orizons: 19-35.
- Liverani, Elena (2001), “Tradurre un traduttore: *Tras-os-Montes* di Julio Llamazares e la sua versione italiana”, *Orillas. Sudi in onore di Giovanni Battista De Cesare*, ed. Elena Liverani. Salerno, Paguro edizioni: 277-90.
- Liverani, Elena (2010), “Appunti per la traduzione degli elementi linguistico-culturali nei testi di viaggio: *Cuaderno del Duero* di Julio Llamazares”, *Viaggiare con la parola*, ed. Elena Liverani; Jordi Canals. Milano, Franco Angeli: 147-64.
- Mancini, Claudia (2012), “Un ricordo di Saramago: intervista a Pilar del Río”, *Le parole e le cose* [1/10/2025] <<https://www.leparoleelecose.it/?p=7443>>.
- Nicewicz-Staszowska, Ewa (2017), “Una metodica follia. Italo Calvino e sull'arte del tradurre”, *Il traduttore errante. Figure, strumenti, orizzonti*, eds. Dario Prola; Elzbieta Jamrozik. Varsavia, Università di Varsavia: 173-80.
- del Río, Pilar (2022), *La intuición de la isla. Los días de José Saramago en Lanzarote*, Ciudad Real, Itineraria Editorial.
- del Río, Pilar (2022), *A intuição da ilha: os dias de José Saramago em Lanzarote*, Porto,

- Porto Editora.
- del Río, Pilar (2022), *A intuição da ilha: os dias de José Saramago em Lanzarote*, Companhia das Letras, traducción de S. Machado Letria.
- del Río, Pilar (2024), *L'intuizione dell'isola. I giorni di José Saramago a Lanzarote*, Milano, Feltrinelli. Traducción dallo spagnolo di R. Desti.
- Osimo, Bruno (2010), *Propedeutica della traduzione*, Milano, Hoepli.
- Rubini, Francesca (2023), *Italo Calvino nel mondo*, Roma, Carocci.
- Saramago, José (2003), “Todo son traducciones, todos somos traductores”, *IV Congreso Latinoamericano de Traducción e Interpretación*, Buenos Aires, CTPCBA (CD-ROM).
- Saramago, José (2014), *Una guía para leer a Saramago. De la estatua a la piedra. El autor se explica*, Madrid, Alfaguara.
- Stassi, Fabio (2021), “Come Rita Desti suona Saramago”, *Tradurre*, 21 [01/10/2025] <<https://rivistatradurre.it/come-rita-desti-suona-saramago/>>

**Elena Liverani** es catedrática de Lengua y Traducción Española en la Universidad IULM de Milán, donde dirige el Departamento de Estudios Humanísticos. Es miembro del programa de doctorado en Lingüística, Literatura y Traducción de la Universidad de Málaga, del programa doctoral en *Visual and Media Studies* de IULM, y representante de su universidad en la Red de Universidades Lectoras. Se ha ocupado de lexicografía bilingüe, del lenguaje del turismo, de fraseología contrastiva, de traducción editorial y audiovisual; junto a la docencia y la investigación desarrolla también la actividad de traductora, por la cual en 2019 recibió el Premio Especial del Ministerio de Bienes y Actividades Culturales y en 2025 el premio a la trayectoria del *Pisa Book Translation Award*. Ha traducido una treintena de novelas de narradores españoles e hispanoamericanos (entre ellos Antonio Muñoz Molina, Julio Llamazares y Alberto Manguel) y actualmente es la voz italiana de Isabel Allende, Enrique Vila-Matas y Juan Gabriel Vásquez.

**elena.liverani@iulm.it**

# MARISA MARTÍNEZ PÉRSICO

## ***EMBODIMENT Y METÁFORAS SOMÁTICAS EN LA AUTOTRADUCCIÓN BILINGÜE. EL CASO DE LA POETA VASCA MIREN AGUR MEABE DESDE EL ENFOQUE DE LOS COGNITIVE TRANSLATION STUDIES***

Università degli Studi di Udine

### Resumen

En este artículo se estudia un corpus de metáforas somáticas acuñadas por poetas autotraductores bilingües a partir del enfoque de los Estudios Cognitivos de la Traducción y su concepto de *embodiment*. Se profundiza en el estudio de caso de la poeta vasca bilingüe Miren Agur Meabe, que se añade a una muestra de autores que practican la autotraducción endógena y exógena precedentemente estudiados. Las conclusiones de este trabajo intentan sensibilizar y dirigir la atención hacia una cuestión ética, de naturaleza sociolingüística, ligada a los derechos lingüísticos de los hablantes de lenguas minoritarias y minorizadas.

palabras clave: Miren Agur Meabe, autotraducción endógena, *embodiment*, poesía contemporánea en euskera, *in-between*

### Abstract

***Embodiment and somatic metaphors in bilingual self-translation. The case of Basque poet Miren Abur Meabe from the perspective of Cognitive Translation Studies***

*This paper examines a corpus of somatic metaphors coined by bilingual self-translating poets through the approach of Cognitive Translation Studies and its concept of embodiment. It delves into the case study of the bilingual Basque poet Miren Agur Meabe, who joins a previously studied sample of authors who practice endogenous and exogenous self-translation. The conclusions of this work aim to raise awareness and draw attention to an ethical issue, of a sociolinguistic nature, linked to the linguistic rights of speakers of minority and minoritized languages.*

keywords: Miren Agur Meabe, endogenous self-translation, embodiment, contemporary poetry in Basque, in-between

## 1. Introducción

La observación y el análisis de casos de autores y autoras bilingües es un método fértil para echar luz sobre los procesos —y la autopercepción de estos procesos— que subyacen a la adquisición de lenguas concomitantes. Dependiendo de cada biografía lingüística, tales procesos se verán influenciados por fenómenos de alcance colectivo (políticas lingüísticas, eventos históricos) y/o reflejarán las marcas de un itinerario de vida individual (desplazamientos familiares, elecciones voluntarias).

El estudio que voy a presentar en este artículo se concentra en el caso de la autora vasca Miren Agur Meabe, Premio Nacional de Poesía de España en 2021 y académica de la lengua euskera, y se enmarca en el proyecto de investigación *Il bilinguismo poetico nel XXI secolo: autotraduzione e sopratraduzione in ambito panispanico* radicado en la Universidad de Udine (2021 a la actualidad). Sus objetivos son estudiar las modalidades de bilingüismo endógeno y exógeno (Grutman 2011) en poetas contemporáneos a través del método de *case study*, analizar un corpus de biografías de autores bilingües de España, Italia y América Latina, indagar sus estrategias de autotraducción horizontal y vertical (Grutman 2011), identificar particularidades en la composición creativa y analizar el eventual impacto del contacto lingüístico en la obra literaria (*code switching*, préstamos, calcos, citas, acuñación de neologismos) en la construcción del propio idiolecto estético.

La recolección de las entrevistas que sirven como insumo para esta investigación —la cual adopta una metodología inductiva, buscando arribar a generalizaciones a partir de la descripción de un corpus de casos particulares— se realiza, casi siempre, a través de un ciclo de encuentros en línea con autoras y autores organizados desde 2021 en la Universidad de Udine, llamado *Poéticas anfibias. Escribir entre dos lenguas / Poetiche anfibie. Scrivere tra due lingue*. El título apela a la analogía de los anfibios, animales cuyo ciclo de vida se desarrolla tanto en ambiente acuático como en terrestre, por lo que pueden vivir dentro del agua (respirando a través de la piel, aunque algunos, como los ajolotes, respiren también por las branquias) o sobre la tierra (respirando por los pulmones y la piel). Hemos elegido la preposición “entre” en alusión al estado de “in-between”, de interdependencia lingüística entre la L1 y la L2 de los entrevistados, para retomar un concepto proveniente del campo de la interlingüística que estudia las interacciones entre diferentes lenguas y los efectos del contacto lingüístico, analizando la posición y relación entre códigos específicos y aparentemente separados (Cummins 1979).

La entrevista con Miren Agur Meabe se realizó el 12 de mayo de 2023 y está

disponible en el canal institucional Play Uniud<sup>1</sup>. Las informaciones más significativas recabadas durante estos encuentros (siguiendo la modalidad de entrevista dirigida) se podrían resumir en cuatro ejes temáticos:

- Decisiones lingüísticas: lengua(s) de expresión literaria, prácticas de traducción y autotraducción, direccionalidad (o simultaneidad) de la escritura bilingüe;
- Decisiones editoriales (¿en qué lengua publican antes sus libros?, ¿son ediciones monolingües o bilingües?, ¿qué idioma va primero, si son bilingües?);
- Historia familiar y ambiental: itinerario de adquisición de las lenguas y rol de la educación formal en su consolidación;
- Acuñaación de metáforas del cuerpo (somáticas) para caracterizar el propio bilingüismo.

Este último punto fue un hallazgo, una constatación derivada de las declaraciones espontáneas de varios autores incluidos en la muestra a partir de las preguntas administradas durante las entrevistas<sup>2</sup>. La apelación a somatismos es una información significativa para profundizar en la autoconsciencia del sujeto bilingüe, pues se relaciona con el concepto de *embodiment* postulado por los estudios cognitivos sobre la traducción de finales del siglo pasado (*Cognitive Translation Studies*). El enfoque de los CTS fue el primero en llamar la atención sobre el papel de los estados y experiencias corporales para la comprensión y el procesamiento del lenguaje. Si bien los Estudios de Traducción y Traductología se consolidaron como disciplina durante la segunda mitad del siglo XX, la aplicación específica de los principios de la corporeidad/corporalidad a la investigación traductológica es un desarrollo reciente dentro de este campo más amplio. Hasta donde hemos podido constatar, al día de hoy no se había aplicado el concepto de *embodiment*

1 Entrevista disponible en: <https://www.youtube.com/watch?v=7vuENQsVsBc&t=309s>

2 El tamaño de la muestra, hasta el momento, incluye a: Joan Margarit (catalán-castellano peninsular), Fabio Morábito (italiano-castellano mexicano), Gigliola Zecchin (italiano-castellano rioplatense), Yolanda Castaño (gallego-castellano peninsular), Josep Maria Rodríguez (catalán-castellano peninsular), Vanna Andreini (italiano-castellano rioplatense), Miren Agur Meabe (euskera-castellano peninsular), Bibiana Collado Cabrera (catalán valenciano-castellano peninsular), Juan Rodolfo Wilcock (italiano-castellano rioplatense), Natalia Toledo (zapoteca-castellano mexicano), Susy Delgado (guaraní-castellano paraguayo), Lionel Lienlaf (mapuche-castellano chileno), Lyz Sáez (mixe zoque-castellano mexicano), Manuel Bolom (maya tsotsil-castellano mexicano), Inaciu Galán (asturiano-castellano peninsular), Berta Piñán (asturiano-castellano peninsular) y Dafne Malvasi (italiano-castellano chileno). Catorce entrevistas de esta muestra se pueden visionar en línea en el mencionado canal institucional Play Uniud (YouTube).

a la autotraducción. En este artículo vamos a concentrarnos en añadir algunas reflexiones sobre la relación *bilingüismo-cuerpo femenino* a partir del caso de Miren Agur Meabe.

## 2. Metáforas somáticas, giro corporal y *embodiment*

En un artículo de 2022, “Poéticas anfibias. Bilingüismo y autotraducción en cinco poetas contemporáneos de expresión italiana, catalana, castellana y gallega”, llegamos a la conclusión, después de analizar las declaraciones de varios poetas entrevistados, de que los autores solían acuñar metáforas somáticas espontáneas “para explicar su propio bilingüismo: músculos, corazones, ojos, cordón umbilical, estómagos, quistes y lengua (en el sentido de órgano bucal), como si las destrezas lingüísticas fueran a la vez funciones físicas y vitales, imprescindibles para la supervivencia del hablante-organismo” (Martínez Pésico 2022: 17).

En el marco del ciclo *Poéticas anfibias*, Fabio Morábito, autor italomexicano, acudió a los músculos para hablar de su relación con la lengua materna, poniendo de manifiesto la importancia de entrenar ese músculo para que se mantenga saludable si se vive en un territorio con una lengua distinta: “el italiano definitivamente había dejado de ser mi lengua materna. O lo era desde el punto de vista meramente anagráfico, de acontecimientos de vida, pero dentro de mí era un músculo en vías de atrofia” (Martínez Pésico 2021a: en línea). Este autor usa metáforas orgánicas para hablar del idioma también en su poemario *Delante de un prado una vaca*: “escribir abre un segundo estómago / en la especie” (2013: 27). También en su ensayo *El idioma materno*, en el capítulo “Venas y arterias”, se refiere a la capilaridad, un atributo fundamental para garantizar un proceso eficaz de trasvase de la lengua de partida a la lengua meta (Morábito 2014).

En la misma línea, el poeta catalán de expresión predominantemente castellana Josep Maria Rodríguez apeló a este tipo de metáforas al hablar de sus dos lenguas, eligiendo el corazón y los ojos. Al preguntársele por su poema “Història Natural” responde:

Fue escrito originalmente en catalán, habla del bilingüismo y me define. Nombra a Quinto Ennio, que es un escritor latino que hablaba tres lenguas, y de él decían que, como hablaba tres lenguas, tenía tres corazones [...] Si me dicen *tienes que renunciar a una de tus dos lenguas* es como si me dijeran que tengo que renunciar a una forma de mirar el mundo, es para mí como perder un ojo [...] me parecería una amputación”. (Martínez Pésico 2021b: en línea)



Aunque cada encuentro se realizó en fechas distintas, sin copresencia de los autores, también la poeta italoargentina Vanna Andreini respondió en sintonía con los ejemplos anteriormente citados, y aludió a la “hermandad” de sus idiomas. Uno de sus poemas de *Fatebenefratelli* habla de las lenguas siamesas, que están “pegadas / comparten flujos sanguíneos” (2020: 16) porque siguen unidas, simbólicamente, a través del cordón umbilical primigenio. A la pregunta por esta particularidad, Andreini no usa el término “amputación” como Josep M. Rodríguez, pero alude al desgarró físico:

Vi una película hermosa de Edoardo de Angelis que se llama *Indivisibili*. La cuestión de los hermanos siameses es siempre muy interesante. Y es muy hermosa la relación de las dos protagonistas, que están juntas desde siempre pero al mismo tiempo tienen el impulso de separarse. Pero separarse es también un desgarró impresionante. Mientras la miraba pensaba en los idiomas que me habitan, pensaba en mis lenguas maternas que quieren estar juntas y separadas. (Martínez Pésico 2021d: en línea)

Por su parte, la poeta gallega Yolanda Castaño (2014) se concentra en los labios y la lengua. Al preguntársele en la entrevista por su poemario *A segunda língua / La segunda lengua*, un título que remite amargamente a un bilingüismo en desigualdad de condiciones, se refiere a los rasgos articulatorios, fonéticos, semánticos o de cosmovisión que el gallego no comparte con otros idiomas. En su poema “La palabra Galicia” afirma que para poder explicar su lugar de proveniencia tendría que sacar la lengua, jugando así con dos significados: los puntos articulatorios de las consonantes en la lengua gallega (que obligan a mostrar la lengua más de lo que sucede en otros idiomas) y la idea de burla, de desenfado. En clave irónica, relaciona este atributo sonoro con la falta de educación: hay pueblos tan educados que nunca enseñan la lengua.

En el mismo año en que se publicó el artículo “Poéticas anfibas...” en la revista *Catalonia* del laboratorio CRIMIC (Universidad de la Sorbona), salió en la revista *inTRAlinea*, editada por el Departamento de Interpretación y Traducción (DIT) de la Universidad de Bologna, el monográfico *Embodied Translating – Mit dem Körper übersetzen* coordinado por Barbara Ivancic (Universidad de Bologna) y Alexandra L. Zepter (Universidad de Colonia), con aportaciones en inglés y en alemán<sup>3</sup>. Allí, las investigadoras parten de la evidencia de que, en las últimas décadas, el cuerpo humano ha estado mucho más presente en las ciencias y las humanidades, dando pie a un auténtico “giro corporal” (Alloa *et al.* 2019), que

3 Agradezco a Fabio Regattin, colega francesista de la Universidad de Údine, la valiosa y oportuna recomendación de este monográfico apenas se publicó.

ellas buscan aplicar al ámbito de la traductología. Este monográfico nos permitió profundizar en las intuiciones y ejemplos empíricos recolectados, aplicando el concepto de *embodiment*. Hemos preferido adoptar en este artículo el préstamo puro por considerar que las opciones disponibles en español (“encarnación”, “corporeidad”, “materialización”) no son funcionales ni adherentes, y dan lugar a interpretaciones y connotaciones ajenas al concepto en cuestión. Las autoras inauguran el monográfico precisando sus intenciones de investigación:

With this special issue, we aim to contribute to the discussion of the potentials and implications of Embodiment theories for translation studies and translation didactics. We focus on perspectives concerning the factual relevance of bodily dimensions for translators’ literary translation processes. When it comes to language processing and understanding language, the embodied perspective suggests that even in the mental processing of linguistic meaning and in the understanding of concepts there are multiple domains involved – domains that concern perception, senses, social-interactive actions and/or inner feelings. The claim is that the body not only constitutes a medium for language use and communication, but language itself is grounded in bodily processes of perception and action. (Ivancic, Zepter 2022: 1)

Se trata de un enfoque que busca ir más allá de una interpretación neurobiológica de la cognición, incorporando premisas filosóficas y hallazgos de la psicología cognitiva al estudio del lenguaje. Las autoras afirman que está científicamente comprobado que la comprensión de conceptos y el procesamiento del lenguaje están arraigados en los estados y experiencias corporales, y efectúan un pormenorizado recorrido de investigaciones precedentes que confirmarían esta hipótesis (Ivancic, Zepter 2022). Plantean el concepto de *embodiment* como una perspectiva necesaria para las humanidades, en particular los estudios de traducción y la didáctica de la traducción, puesto que permite considerar que, incluso en el procesamiento mental del significado lingüístico y en la comprensión de conceptos, intervienen múltiples dominios que conciernen a la percepción, los sentidos, las acciones sociointeractivas y/o los sentimientos internos. Es decir: el cuerpo no solo constituye un medio para el uso y la comunicación del lenguaje, sino que el lenguaje mismo se fundamenta en procesos corporales de percepción y acción.

Tales afirmaciones sobre la relación entre la percepción y el procesamiento lingüístico nos permiten entender mejor las declaraciones de los autores y autoras bilingües que venimos trabajando y justifican la necesidad de habilitar un espacio para dar cauce a reflexiones éticas sobre los derechos lingüísticos de los hablantes bilingües, especialmente en el caso de aquellos que se encuentran en una situación de desigualdad sociolingüística, es decir, cuando una de sus dos lenguas está

en la posición subalterna de lengua minoritaria y minorizada. Para los autores entrevistados, si una lengua no se practica, se atrofia como un músculo. Si al sujeto se le prohíbe hablarla, es como amputarle un ojo e impedirle ver con claridad. Si las lenguas se separan, cortando el cordón umbilical entre ambas, ocurre un desgarró. En las páginas siguientes vamos a analizar qué añade Miren Agur Meabe a estos testimonios y cómo se focaliza en la experiencia del cuerpo femenino, pero antes ofreceremos algunas informaciones sobre la situación reciente y la historia del euskera, así como sobre la biografía lingüística de la autora, como forma de contextualización.

### 3. Modelos escritos no literarios del euskera: pesca y religión

Datos de las últimas décadas sobre el monolingüismo y el bilingüismo en los territorios vascoparlantes de la Península señalan una tendencia a la disminución tanto de los *erdaldunes* (monolingües castellanos) como de los *euskaldunes* (monolingües vascoparlantes), mientras que el porcentaje de bilingües pasivos se duplicó: en 2012 representaban el 14,7% mientras en 1991 suponían el 7,7%. El crecimiento más alto se produjo en la CAV y en Navarra, y en menor medida, en el País Vasco Norte. Por otra parte, el aumento del número de bilingües que señalan las encuestas entre el año 1991 y 2011 se cifra en 185.600 bilingües más. Según Arantzazu Fernández Iglesias, este incremento ha sido posible gracias a los programas de recuperación lingüística aplicados en la CAV que se pusieron en marcha desde la designación del euskera como lengua cooficial en el Estatuto de Autonomía aprobado en 1979 con la llegada de la democracia en España, puesto que “junto con la subsiguiente Ley de Normalización y Uso del Euskera de 1982 posibilitaron la implantación de un modelo educativo que contemplara la enseñanza en euskera” (2012: 156).

Fernández Iglesias se detiene, además, en los modelos de lengua utilizados como base para la compleja estandarización del idioma. Afirma que una descripción somera del entorno geográfico y administrativo de la lengua vasca, así como de la historia de sus modelos de lengua literaria en relación con las variantes dialectales, permiten comprender mejor cuáles son las dificultades que entrañan tanto el proceso de estandarización de la lengua como su recepción y difusión en la comunidad lingüística a la que se ha dirigido. Sostiene que la debilidad del sistema literario vasco lleva a incluir en su canon “obras que no son estrictamente literarias y sería más propio hablar de modelos escritos” (Fernández Iglesias 2012:

151). Sin embargo, la bibliografía suele mantener la expresión de “modelos literarios” por ser la de más amplio uso.

Es necesario tener en cuenta que el euskera es una lengua pequeña, que cuenta con alrededor de 700.000 hablantes. Por lo tanto, la literatura vasca en sí es también reducida. A pesar de tener casi cinco siglos de historia (el primer libro vasco se publicó en 1545) la producción en euskera ha sido bastante escueta hasta hace poco. Desde 1545 hasta 1879 se publicaron solo 101 libros (Olaziregi Alustiza 2000: 385). En la entrevista realizada en 2023, Miren Agur Meabe hace suya esta afirmación sobre la validez de sus “modelos de lengua escrita no literaria” como insumo alfabetizador y normalizador en el aprendizaje y uso público del vasco:

Muchos de los libros escritos durante tres siglos son libros de religión, oraciones para los balleneros que se iban a pasar unos meses en la pesca o para educar a los niños. [...] En euskera muchas de las palabras son palabras compuestas que se basan en elementos que tienen que ver con el mundo rural o con el mundo de la pesca. Por ejemplo, para decir hoy *reflexionar* utilizamos palabras que estrictamente pueden significar *cavar*. Hemos ido aportando palabras del mundo rural al mundo urbano para expresar conceptos que no podíamos expresar, que no sabíamos cómo expresar en el mundo moderno. (Martínez Pérsico 2023a: en línea)

Para la autora, ese mundo pesquero fue fundamental, nutrió su imaginaria poética y sus recursos léxicos porque, por parte paterna, su familia era pescadora, proveniente de la zona de Lekeitio, un pueblo de la costa de Vizcaya cuya subsistencia tradicionalmente se ha debido a la pesca y a actividades derivadas como el turismo. En su caso también los textos religiosos funcionaron como estímulo y acervo indispensable para transmitirle una tradición escrita en la lengua minoritaria local:

Mi familia es un núcleo completamente euskaldún, pero hay que tener en cuenta el contexto histórico. Todavía perduraba la dictadura de Francisco Franco, con la consiguiente represión de la lengua, y en los que a pesar de que nuestra lengua familiar, de juego y de relación en la calle era el euskera, la lengua que estaba realmente prestigiada para la cultura era el español. Con lo cual, había un pequeño polo de lengua culta o de lengua más enriquecida que nos provenía de las actividades de la parroquia de la iglesia. Es decir, en casa, euskera coloquial: canciones, nanas, juegos; en la calle también, pero en la escuela todo completamente en castellano y además una educación muy franquista, muy tradicional, muy españolista que nos hacía eclipsar el valor de lo que nosotros poseíamos de raíz. Y luego, por otra parte, la labor de la parroquia que organizaba actividades tales como teatro o excursiones en las que tomábamos contacto por primera vez con un euskera escrito. Recuerdo el día de mi primera comunión,

con seis años y medio, en la que leí por primera vez en el altar de la iglesia junto con un montón de niños y niñas de mi pueblo, fue mi primera lectura en vasco. Se me presentaba el idioma euskera por escrito por primera vez y continué en este tipo de actividades hasta que me di cuenta de que este euskera escrito más culto que yo estaba asimilando como niña en estas lecturas de la iglesia me estaba dejando un poso léxico, un poso morfosintáctico y un poso de musicalidad relacionada. El lenguaje simbólico con las actividades de la liturgia me estaba dando pie a fabular o a empezar a expresarme en la lengua vasca. A los dieciséis años aproximadamente me apunté en un curso de alfabetización en euskera que hace despertar todo esto que yo traía desde la infancia pero que estaba como aplastado o enterrado bajo el peso de la educación formal. (Martínez Pésico 2023a: en línea)

En el entorno vasco, el ámbito litúrgico se revela como espacio de resistencia y pervivencia de la lengua minoritaria durante la censura política, pues a finales del siglo XIX y principios del XX la literatura vasca había vivido una época fecunda, pero más tarde tuvo que hacer frente a la Guerra Civil y a la posterior dictadura de Franco, y “no ha sido hasta hace unos treinta años cuando ha vuelto a tomar fuerza. Por lo tanto, se puede decir que la literatura vasca es una literatura relativamente joven” (Fernández Iglesias 2012: 159). A este respecto, la autora añade que:

tenemos que ir hasta antes de la guerra porque la recuperación de la lengua no empieza en los años 60; venía desde principios del siglo XX, pero la Guerra Civil amputó todas esas ideas de culturización y de renovación del idioma y democratización [...] En el año 1979 se dicta el Estatuto de Autonomía, ya muerto Franco en el año 75 y da cierta capacidad de gestión a las comunidades históricas y el año 1982 se promulga también la ley del euskera que significa que va a convertirse en una lengua cooficial para lo cual se desarrolla un programa de acceso al idioma en la Administración tanto en la educación como en la sanidad como en otros ámbitos. (Martínez Pésico 2023a: en línea)

Como vemos, las experiencias tempranas, escolares y diglósicas, revelan una escisión espacial y temporal entre la lengua de alfabetización y la lengua materna de la autora que nos ocupa.

#### 4. *Embodiment* y cuerpo femenino en la poética de Miren Agur Meabe

Ahora vamos a concentrarnos en una serie de poemas del libro *Cómo guardar ceniza en el pecho* (Editorial Bartleby, Madrid, 2021), publicado originalmente

en vasco en 2020 con el título *Nola gorde errautsa kolkoan* por la editorial Susa Literatura de Zarautz (Gipuzkoa) y posteriormente autotraducido. Con este libro obtuvo el Premio Nacional de Poesía en 2021.

El poema “Prefijación” está incluido en una parte del libro que se llama “Viaje de invierno” y alude a un viaje de soledad, de reconstrucción personal. En la entrevista, la autora explica que “muchas veces para poder reconstruirnos tenemos que colocarnos de una manera nueva en el lenguaje y por eso sirven los prefijos; son una especie de envases gramaticales, léxicos, que nos ayudan a colocar mejor el pensamiento” (Martínez Pésico 2023a: en línea). En euskera y en castellano los sufijos de negación son idénticos y funcionan igual, *des* y *ez*. Este poema surgió tras haber escuchado una charla de una antropóloga feminista vasca que consistía en la explicación de la desocupación de los esquemas patriarcales en la mentalidad occidental que, según afirma la poeta, le arrojó luz sobre su manera de pensar y de vivir en aquel momento. Prefijación en euskera es *aurrizkien bidea*, que significa *camino de los prefijos*. El poema establece un paralelismo entre la creación verbal y el acto de dar a luz:

Nos encontramos en la edad de los prefijos.  
Los lexemas son mi hechizo en la fogata.  
Al conformar las palabras letra a letra,  
me parece que estoy a punto de parir un meteorito:  
deconstruir, desaprender, desamar, desadueñar.

Cada vocablo es un clavo y tiene su música.

Es preciso soplar sin perder el ritmo  
para traer al mundo a estas criaturas:  
recomponer, renombrar, reflorecer, reencontrar.

El código de la piel ha cumplido su periodo de prescripción.  
La metamorfosis llena su cántaro en la lengua. (Meabe 2021: 108)

Escribir palabras es traer criaturas al mundo. La piel es código. El yo lírico concibe la escritura desde la corporalidad femenina, como un trabajo de parto. Este parto de alguna forma se vincula con la regeneración ideológica, en un sistema de equidad de género, en consonancia con la “desocupación de esquemas patriarcales” mencionados por Meabe en la entrevista. Así se comprenden los verbos “deconstruir”, “desaprender” y el neologismo “desadueñar”. Como si el mismo cuerpo de la mujer fuera capaz, con su funcionamiento natural, de modificar el

orden de las cosas. Este didactismo relacionado con el parto se presenta también en la serie lírica “Cinco cartas sobre los dolores de parto”, incluidos en el mismo libro, en el que el yo se dirige a una interlocutora “amiga” más joven e inexperta para darle consejos sobre el proceso de dar a luz. Encontramos en estos poemas nuevos somatismos para nombrar la experiencia de la creación literaria: la piel y los órganos reproductivos, estos últimos no nominalizados como en el segundo poema, “Ruego a las palabras”, pero sí aludidos mediante la acción de parir. Aquí, la palabra entra en relación directa con el acto de ovular, la ovulación, y la menopausia:

No permitáis, palabras, que me aleje de la tierra,  
del aliento de las vacas, de la sangre de la sepia.

Si me sedujisteis con vuestro lunar pintado,  
inocentes como la florecilla de mi primer sostén.

No consintáis, palabras, que me olvide de la historia,  
del insomnio de la idea, del llavín de la fe.

Si me dejé manosear por vuestros múltiples dedos,  
si hervisteis mis vendas en vuestra olla *express*.

No aceptéis, palabras, que eluda mencionar mi cuerpo  
o sus reglas variables o su sabio declinar.

Si me acarreasteis a pesar de las encrucijadas,  
escarabajos bajo su tierno pastelito oficial.

No admitáis, palabras, que me aparte de este oficio  
aunque la crudeza me golpee, aunque me devore la bonanza.

Desde que ovulé por última vez me he convertido en otra.  
No me retengáis, palabras, en ninguna escena de ningún pintor. (Meabe 2021: 200)

Aquí vemos una invocación a las palabras, un ruego profano que puede leerse como arte poética: no dejar de nombrar el cuerpo, a pesar de su sabio declinar. Se apela a una palabra bisémica, la “regla”, que puede entenderse como norma o como ciclo menstrual, en este caso, “variable”, es decir, irregular, connotando la debilitación de la función reproductiva femenina. Al preguntársele por el formato de la plegaria que adopta en este poema, la autora replica que se debe

a la influencia de las tempranas lecturas de iglesia, un aspecto ya mencionado anteriormente, por eso en sus obras suelen aparecer salmos, oraciones y cadencias de tipo versicular: “es como una especie de cuestión ancestral que siento yo en relación con el idioma” (Martínez Pésico 2023a: en línea). También se refiere al cuerpo que privilegia describir en ese poema, el de una mujer que ha superado la cincuentena:

Yo concibo como una unidad mi vida y mi obra. [...] parece que cuando hablamos de cuerpo hablamos hasta los cincuenta años; parece que después de la menopausia nuestro cuerpo desaparece y no existe. Ese es un tema que se observa en mi obra de poesía, pero sobre todo en mi obra en prosa donde se va detectando cada vez más. Somos cuerpo y seguimos siendo cuerpo con *su sabio declinar* como digo aquí. Es un poema que es una especie de resumen de mi poética. (Martínez Pésico 2023a: en línea)

Si bien el libro se publica en 2020 (euskera) y 2021 (castellano), en la entrevista precisa que ese poema, aunque se encuentre al final del libro, cronológicamente es el más antiguo, pues lo escribió al acabar el anterior, hacia 2010 o 2011. Esto significa: hacia sus cincuenta años. La reivindicación de cantar a la feminidad después de la menopausia y el derecho a la toma de palabra se repiten como un *ritornello*, con su deliberada mención a los óvulos y a las señas distintivas del cuerpo femenino, contados desde la posición activa de sujeto de enunciación y no de objeto de observación, rechazando el papel de “modelo de atelier”:

la mención al óvulo que llama tanto la atención es una especie de gesto de delimitar el territorio. Para mí, escribir desde el cuerpo femenino ha sido una manera de luchar por obtener más mayores cotas de libertad porque el cuerpo femenino ha sido considerado un tabú. Hablar desde nuestra propia realidad y por eso justifica también la última frase *No me retengáis, palabras, en ninguna escena de ningún pintor* porque han sido los pintores tradicionalmente los que han dado reflejo de nuestro cuerpo, pero casi siempre en multitud de veces desde una manera esteticista o desde una manera frivolisante o desde una manera incluso humillante. Dejadme a mí que construya yo mi propia historia, con mis propias palabras, con palabras de mujer que está evolucionando porque desde que ovulé por última vez no es que haya dejado de existir, sigo teniendo voz. Entonces aquí por ejemplo es un ejemplo de los de “traiciones al original”. En la versión euskera por ejemplo la última frase no es tan generalista, no dice *no me retengáis palabras en ninguna escena de ningún pintor*, sino que dice *no me retengáis palabras en ninguna escena del viejo Renoir*. Recuerdo cuándo puse Renoir y no puse por ejemplo Leonora Carrington o a Frida Kahlo. [...] al hacer la traducción me di cuenta de que Renoir era poco, tenía que anular a todos los demás para poder dejar sitio a Clara Peeters y a las demás que aparecen. (Martínez Pésico 2023a: en línea)



La autora subraya que el castellano, autoversionado en una fase sucesiva, le permite “corregir” el sesgo de género que en su opinión tenía el original vasco, para convertir al poema en una composición más inclusiva. Lo dicho hasta el momento confirma lo que anunciamos al principio: el cuerpo no solo constituye un medio para el uso y la comunicación del lenguaje, sino que el lenguaje mismo se fundamenta en procesos corporales de percepción y acción (*embodiment*), en este caso, influyendo incluso en el proceso de autotraducción. El parto, la ovulación, la menopausia, la regla, los procesos biológicos y hormonales están intrínsecamente asociados a la escritura.

## 5. La autotraducción ecológica: una lengua que “aromatiza” a la otra

El testimonio de Miren Agur Meabe sobre el proceso de autotraducirse nos permite reflexionar sobre la complejidad de la escritura creativa de los bilingües que han “encarnado” dos idiomas para expresarse desde la niñez, es decir, que gozan de un bilingüismo precoz, como en este caso. Por lo general, tal proceso sufre modificaciones a lo largo del tiempo, no es homogéneo ni unidireccional. La autora vasca no es una excepción, y en sus declaraciones se describen dos métodos:

1. autotraducción consecutiva del euskera al castellano;
2. autotraducción en paralelo, con esporádicas ocasiones en que la escritura creativa comienza en castellano.

Respecto de la predilección por la autonomía o la fidelidad entre autotraducciones, Meabe explicita su voluntad de libertad creativa, de recreación y reescritura. Un ejemplo ya analizado es el de “Ruego a las palabras” en el que se menciona a Renoir solamente en el original vasco. Algunas variantes ponen de manifiesto un distinto posicionamiento ideológico respecto de un tema sensible para la autora y, de este modo, podríamos arriesgar la hipótesis de que la autotraducción contribuye a la construcción de distintas identidades autorales. Cuando se le pregunta si sus poemas nacen en euskera, en castellano o si se trata de un proceso híbrido y bidireccional, responde con unas sorprendentes analogías ecológicas:

Generalmente escribo en euskera, pero es verdad que depende de qué libro esté leyendo. Si la última frase o el último texto es en castellano, por ejemplo, me da el acento para ponerme y arranco tal vez en castellano. Pocas veces, pero sí me ocurre que, subyugada o fascinada por un texto en castellano, que quizás es una traducción de un poema de inglés, siento que tiene tal poder de imantación que me está arrastrando al

castellano y entonces arranco el poema en esa lengua, pero enseguida ya me cambio otra vez al euskera. Y después, como he explicado antes, voy haciendo una traducción simultánea porque veo que un texto va aromatizando el otro y va enriqueciéndolo. Pero esto me ocurre desde hace muy poco porque cuando empecé a escribir no pensaba en absoluto que a mí me iba a leer nadie en castellano. Decidí escribir en euskera porque me parecía un acto de compromiso con mis raíces, con mi familia, me parecía un acto ecológico de poder ayudar a que un idioma pequeño perviviera en la amalgama de la globalidad. Además, porque está rodeado de lenguas muy potentes como el español o el francés, o ahora con el gran velo del inglés. También he ido evolucionando como poeta y como escritora, y la necesidad de ir pasando al castellano ha sido cada vez mayor precisamente porque me he vuelto más profesional y me han empezado a conocer fuera de mi ambiente o de mi entorno habitual y entonces intento ejercitarme formalmente de las dos maneras. Pero eso es algo que ha venido mucho más tarde. (Martínez Pésico 2023a: en línea)

En una fase reciente, el proceso de escritura y de autotraducción ha ido mutando hasta convertirse en híbrido y bidireccional. En su testimonio reconocemos la autopercepción de subalternidad característica de toda autotraducción vertical, lo que confirma, con Grutman, que las autotraducciones realizadas en la España post-franquista son predominantemente asimétricas y que no puede ser azar que la mayoría de los autotraductores provenga de comunidades lingüísticas no simplemente bilingües sino diglósicas (Grutman 2011). El compromiso ecológico, de salvaguarda de una lengua pequeña rodeada de idiomas que representan una perenne amenaza de fagocitación, se evidencia en elecciones editoriales como publicar primero en la lengua minoritaria periférica y más tarde en la mayoritaria central. La autora utiliza la palabra “pervivencia”. A este respecto, resulta útil retomar la distinción tipológica acuñada por Xosé Manuel Dasilva entre *autotraducción opaca* y *autotraducción transparente* (2011, 2015), una clasificación especialmente significativa en el caso de lenguas asimétricas: la primera sería aquella traducción en la que no se indica que hay un texto escrito en otra lengua, por lo que el lector puede creer que se encuentra ante un original. Lo contrario es la autotraducción transparente, en la que no se oculta el texto de partida. En el caso de Meabe, siempre se precisa en los créditos la versión original, preexistente, en euskera.

En lo que concierne al segundo método, la autotraducción en paralelo, es interesante el verbo que Meabe utiliza para justificar su elección, “aromatizar”:

Últimamente hago la prueba de autotraducirme casi en paralelo porque noto que un vocablo en el idioma original que me plantea tal vez una dificultad de ser traducido

en la lengua de llegada me hace buscar alternativas. La traducción tiene mucho de buscar alternativas. Aunque siempre se ha dicho eso de *traduttore, traditore*, el traductor igual hace perder algunas cosas del idioma original pero también hace ganar otras en el idioma de llegada. Para mí, mi bilingüismo es ventajoso completamente porque es como si una lengua aromatizara la otra, y al buscar esa alternativa, encontrara una palabra que me hace volver al texto original porque me ha despertado alguna pulsión del tipo que sea y me hace retocar el texto original para enriquecerlo también. Yo me desenvuelvo por igual en ambas lenguas y eso de cara a mi obra es completamente positivo. (Martínez Pérsico 2023a: en línea)

A los beneficios del bilingüismo como forma de incrementar el alcance léxico se añade la libertad creativa de la autotraducción, que incluye la posibilidad de corregir(se) con posterioridad:

yo suelo actuar de una forma muy libre con la autotraducción. Cuando se autotraduce se está haciendo una reescritura [...] Cuando estamos trabajando con un original es como si tuviéramos puestas las lentes de cerca. Estás intentando construir un libro que tenga una coherencia interna [...] Pero cuando llega la hora de traducir al castellano es como si te pusieras otras lentes; a mí lo que me pasa es que en cuanto me pongo las lentes para empezar a traducirme lo primero que veo son “errores”. No son propiamente errores o erratas sino, por ejemplo, repeticiones. Al escribir este poemario en euskera *Nola gorde errautsa kolkoan*, un libro bastante extenso y con influencia de muchas lecturas que he ido realizando, no me di cuenta de que había utilizado como unas siete veces la palabra *árbol*. [...] No son erratas ni equivocaciones, son a veces repeticiones infructuosas. Al traducirme en castellano pienso en el hecho de que en teoría ahora va a leerme un público que no conoce la obra en euskera y me planteo por qué no me voy a dar la libertad de poner en un caso *sauce* en otro caso *espino* en otro caso *roble* y en otro caso *haya* si lo puedo enriquecer a través de los campos semánticos. Ese es un tipo de libertades que sí me suelo conceder porque soy la dueña del original. Si viniera un investigador ortodoxo se enfadaría muchísimo conmigo porque diría en este mismo poema de *Prefijación* “la metamorfosis llena su cántaro en la lengua”. En la versión original en euskera no está [...] hay momentos en los que tienes que traicionar el original, pero es una traducción que haces para seducir, una traducción que haces por amor al texto y también pensando en complacer más a estos lectores que te leen por primera vez en ese idioma. (Martínez Pérsico 2023a: en línea)

Dolors Poch y Jordi Julià, en *Escribir con dos voces. Bilingüismo, contacto idiomático y autotraducción en literaturas ibéricas* (2020), afirman que el escritor bilingüe escribe con dos voces porque tiene, como mínimo, dos vocablos que conoce bien para expresar aquello que percibe, que siente, y por eso mismo incluso la conceptualización de la realidad puede verse a menudo condicionada o modificada por la

existencia de otro término. Concluyen que quizá la particularidad de la escritura creativa bilingüe se deba, precisamente, a esta riqueza idiomática. Además del aspecto léxico, proponemos añadir el impacto que el bilingüismo puede tener en el plano fónico a la hora de autotraducirse. En su testimonio, Miren Agur Meabe apela al proceso de *compensación*, y que se aplicaría, en su caso, a los aspectos sonoros y rítmicos de reescritura autotraductora:

Otra cosa distinta es lo que ocurre con la métrica ya que no es un asunto de significado sino de musicalidad. A veces en el poema en euskera encuentro la manera de crear unas rimas internas jocosas. Yo sí que juego a la ironía. Por ejemplo, se nota en mis textos cuando incluyo rima. Cuando incluyo rima le estoy queriendo quitar solemnidad a algo. Son rimas internas que por nada puedo conseguir en el castellano en la misma estrofa, entonces qué hago: si es la tercera estrofa la que me posibilita crear unas rimas internas de tipo lúdico, pues lo hago en la tercera, pero a veces es imposible establecer correspondencias exactas. (Martínez Pérsico 2023a: en línea)

La compensación se refiere a la técnica de recuperación de la pérdida de características formales relevantes del prototexto, intentando aproximar sus efectos en el metatexto por medios distintos a los empleados en el prototexto (Hervey, Higgins 1992). De este modo, una hipótesis que podemos arriesgar es que el bilingüismo creativo, especialmente cuando se da entre lenguas lejanas como el castellano y el euskera, empujan al autotraductor o autotraductora a frecuentes compensaciones métricas y rítmicas para intentar reproducir los efectos fónicos del original.

## 6. Conclusiones

El enfoque de los *Cognitive Translation Studies* fue el primero en llamar la atención sobre el papel de los estados y experiencias corporales para la comprensión y el procesamiento del lenguaje: el cuerpo no solo constituye un medio para el uso y la comunicación del lenguaje, sino que el lenguaje mismo se fundamenta en procesos corporales de percepción y acción.

Esta concepción de fuerte interdependencia entre cuerpo y lenguaje (*embodiment*), aplicada al terreno de los estudios de la traducción, resulta una categoría provechosa para comprender por qué los autores y autoras autotraductores acuñan metáforas somáticas —en sus testimonios y en sus libros— para definir su bilingüismo o su relación con el lenguaje. Este artículo, que completa y complementa investigaciones precedentes y en curso, revela que una muestra significativa de

estos informantes apela a partes del cuerpo como óvulos, músculos, corazones, ojos, cordón umbilical, estómagos, quistes, piel, lengua, y a un léxico asociado al funcionamiento de órganos y aparatos corporales (atrofiarse, parir, ovular, parto, regla, menopausia, ovulación).

Especialmente en los casos de autores bilingües cuya lengua materna ocupa una posición periférica respecto de la de alfabetización, tales metáforas y apelaciones al cuerpo suelen poner en evidencia una amenaza de aniquilación: amputación, desgarrar, desaparición, declinación física, atrofiamiento, pervivencia. Las conclusiones de este trabajo intentan sensibilizar y dirigir la atención hacia una cuestión ética, de naturaleza sociolingüística, ligada a los derechos lingüísticos de los hablantes de lenguas minoritarias y minorizadas, poniendo de manifiesto que la lengua no es tan solo un medio de comunicación sino un elemento consustancial a la identidad personal y existencial, homologable a una parte del cuerpo, y que su amenaza puede ser experimentada como un atentado contra la integridad subjetiva de sus hablantes.

## Bibliografía citada

- Alloa, Emmanuel; Bedorf, Thomas; Grüny, Christian; Klass, Tobias N. (2019), *Leiblichkeit. 2. verbesserte und erweiterte Auflage*, Tübingen, Mohr Siebeck.
- Andreini, Vanna (2020), *Fatebenefratelli*, Buenos Aires, Barnacle.
- Castañó, Yolanda (2014), *La segunda lengua*, Madrid, Visor.
- Cummins, James (1979), "Linguistic Interdependence and the Educational Development of Bilingual Children", *Review of Educational Research* 49: 222-51.
- Dasilva, Xosé Manuel (2011), "La autotraducción transparente y la autotraducción opaca", *Aproximaciones a la autotraducción*, eds. Xosé Manuel Dasilva; Helena Tanqueiro. Vigo, Editorial Academia del Hispanismo: 45- 67.
- Dasilva, Xosé Manuel (2015), "La opacidad de la autotraducción entre lenguas asimétricas", *TRANS: revista de traductología* 19/2: 171-82.
- Farina, Mirko (2021), "Embodied cognition: dimensions, domains and applications", *Adaptive Behavior* 29/1: 73-88.
- Fernández Iglesias, Arantzazu (2012), "La estandarización del euskera", *Revista de lenguas y literaturas catalana, gallega y vasca de la UNED*, XVII: 151-72.
- Gallén, Enric; Lafarga, Francisco; Pegenaute, Luis (2010), *Traducción y autotraducción en las literaturas ibéricas*, Bern, Peter Lang.
- Grutman, Rainier (2009), "Self-Translation", *Routledge Encyclopedia of Translation Studies*, eds. Mona Baker; Gabriela Saldanha. London, Routledge: 257-60.

- Grutman, Rainier (2011), “Diglosia y autotraducción vertical (en y fuera de España)”, *Aproximaciones a la autotraducción*, eds. Xosé Manuel Dasilva; Helena Tanqueiro. Vigo, Editorial Academia del Hispanismo: 69-91.
- Hervey, Sándor; Higgind, Ian (1992), *Thinking translation: A course of translated studies*, London, Routledge.
- Ivancic, Barbara; Zepter, Alexandra L. (2022), “Embodiment in Translation Studies: Different Perspectives”, *inTRAlinea. Online translation journal of the Department of Interpreting and Translation (DIT) of the University of Bologna*, special issue.
- Martínez Pérsico, Marisa (2021a), “Entrevista al poeta italomexicano Fabio Morábito”, *Poetiche anfibie. Scrivere tra due lingue / Poéticas anfibias. Escribir entre dos lenguas. Incontri dedicati alla poesia, la traduzione e il bilinguismo*, Università degli Studi di Udine, [23/04/2025] <<https://www.youtube.com/watch?v=GncJRATfUo>>
- Martínez Pérsico, Marisa (2021b), “Entrevista al poeta catalán Josep Maria Rodríguez”, *Poetiche anfibie. Scrivere tra due lingue / Poéticas anfibias. Escribir entre dos lenguas. Incontri dedicati alla poesia, la traduzione e il bilinguismo*, Università degli Studi di Udine, [16/04/2025] <<https://www.youtube.com/watch?v=apcytD1ttTM>>
- Martínez Pérsico, Marisa (2021c), “Entrevista a la poeta gallega Yolanda Castaño”, *Poetiche anfibie. Scrivere tra due lingue / Poéticas anfibias. Escribir entre dos lenguas. Incontri dedicati alla poesia, la traduzione e il bilinguismo*, Università degli Studi di Udine, [18/04/2025] <<https://www.youtube.com/watch?v=jq2BUajhDck>>
- Martínez Pérsico, Marisa (2021d), “Entrevista a la poeta italoargentina Vanna Andreini”, *Poetiche anfibie. Scrivere tra due lingue / Poéticas anfibias. Escribir entre dos lenguas. Incontri dedicati alla poesia, la traduzione e il bilinguismo*, Università degli Studi di Udine, [20/04/2025] <<https://www.youtube.com/watch?v=ZW1OYsXfkq8>>
- Martínez Pérsico, Marisa (2022), “Poéticas anfibias. Bilingüismo y autotraducción en cinco poetas contemporáneos de expresión italiana, catalana, castellana y gallega”, *Catalonia. (Auto)biographie langagière, conscience linguistique plurilingue, intercompréhension en contexte de romanité*, Sorbonne Université - Laboratoire CRIMIC, 30: 1-22.
- Martínez Pérsico, Marisa (2023a), “Miren Agur Meabe: A veces hay que traicionar el original por amor al texto”, *Suplemento cultural Los Diablos Azules de infoLibre*, 22 de noviembre, [17/04/2025] <[https://www.infolibre.es/cultura/los-diablos-azules/miren-agur-meabe-veces-hay-traicionar-original-amor-texto\\_1\\_1648246.html](https://www.infolibre.es/cultura/los-diablos-azules/miren-agur-meabe-veces-hay-traicionar-original-amor-texto_1_1648246.html)>
- Martínez Pérsico, Marisa (2023b), “Entrevista a la poeta vasca Miren Agur Meabe”, *Poetiche anfibie. Scrivere tra due lingue / Poéticas anfibias. Escribir entre dos lenguas. Incontri dedicati alla poesia, la traduzione e il bilinguismo*, Università degli Studi di Udine, [11/04/2025] <<https://www.youtube.com/watch?v=7vuENQsVsBc>>
- Meabe, Miren Agur (2021), *Cómo guardar ceniza en el pecho*, Madrid, Bartleby.
- Morábito, Fabio (2013), *Delante de un prado una vaca*, Madrid, Visor.
- Morábito, Fabio (2014), *El idioma materno*, México, Sexto piso.
- Olaziregi Alustiza, María José (2000), “Encuentros y desencuentros con la literatura vasca”, *Revista de lenguas y literaturas catalana, gallega y vasca*, 7: 379-92.

Poch, Dolors; Julià, Jordi (2020), *Escribir con dos voces. Bilingüismo, contacto idiomático y autotraducción en literaturas ibéricas*, Valencia, Universitat de València.

Regattin, Fabio (2021), *Autotraduzione. Pratiche, teorie, storie-Autotraduction. Pratiques, théories, histoires*, Città di Castello, I libri di Emil.

Rodríguez, Josep. M. (2017), *Sangre seca*, Madrid, Hiperión.

**Marisa Martínez Pésico** es investigadora (RTD-B) de Lengua y Traducción Española en la Università degli Studi di Udine, acreditada a titular. Sus ámbitos de interés e investigación son la traducción poética y la autotraducción literaria, los fenómenos de contacto lingüístico, la variación diatópica del español y la didáctica del español en contexto italo parlante. En ámbito literario, su interés se focaliza en los estudios transatlánticos, la edición de poesía española e hispanoamericana ultracontemporánea y las vanguardias en España e Hispanoamérica.

**marisa.martinezpersico@uniud.it**





---

## RESEÑAS CRÍTICAS

- Sonia Bailini, Maria Vittoria Calvi, Elena Liverani (eds.);  
Roberto Ortí Teruel (coord.), *El español como lengua de mediación  
en contextos educativos y profesionales* (Alba Segura Lores)  
Ana I. Codesido García, Carlos Hernández-Sacristán,  
Victoria Marrero-Aguilar (eds.), *Lingüística clínica en el ámbito hispánico:  
un panorama de estudios* (Alessandra Donati)  
Veljka Ruzicka Kenfel, Raffaella Tonin, *Aproximación poliédrica a la cuestión del género  
en la literatura infantil y juvenil* (Chiara Albertazzi)  
Natasha Leal Rivas, *La argumentación en la escritura académica  
de investigación* (Francisco José Sánchez García)  
Nuria Pérez Vicente, *Los culturemas.  
Aprender a traducir con la literatura infantil y juvenil (LLJ)* (Rocío García Jiménez)  
Isabel Santos Gargallo; Susana Pastor Cesteros (dirs.), *Metodología de la investigación  
en la enseñanza-aprendizaje del español  
como segunda lengua (L2)/ lengua extranjera (LE)* (Rosa María García Jiménez)  
Milin Bonomi, Maria Vittoria Calvi, Marcella Uberti-Bona (eds.), *Demolingüística  
del español en Italia, con un anexo sobre Malta, San Marino  
y Ciudad del Vaticano* (Sonia Bailini)



**Sonia Bailini, Maria Vittoria Calvi, Elena Liverani (eds.); Roberto Ortí Teruel (coord.), *El español como lengua de mediación en contextos educativos y profesionales*, Milano, Ledizioni, in collaborazione con Instituto Cervantes, 2024, 208 pp. ISBN 9791256002221**

**Alba Segura Lores**  
**Universität Heidelberg**

El volumen *El español como lengua de mediación en contextos educativos y profesionales* recopila las contribuciones presentadas en el II Congreso de Español como Lengua Extranjera en Italia (CELEI) celebrado en Milán en octubre de 2023. La obra, publicada en acceso abierto por el Centro Virtual Cervantes, se erige como una referencia fundamental para comprender los desafíos y oportunidades actuales en la enseñanza del español. Su eje central es el concepto de mediación como una actividad comunicativa clave en un mundo cada vez más plurilingüe y multicultural.

La introducción, a cargo de Calvi, sitúa el tema de la mediación en el centro del debate actual por el protagonismo adquirido en el plano didáctico, en la enseñanza y en la realidad plurilingüe de la sociedad, sobre todo en las aulas y en los contextos profesionales. A partir de ahí, la obra se estructura en torno a dos capítulos introductorios y tres secciones temáticas que abordan la mediación desde distintas perspectivas.

En el primer capítulo, Moreno Fernán-

dez examina en profundidad el fenómeno de las lenguas de herencia, cada vez más presente en el ecosistema de las aulas de español y que, sin embargo, sigue constituyendo una laguna en la formación del profesorado. El autor desmitifica y sistematiza el concepto de lengua de herencia desde una perspectiva internacional aplicable al ámbito de la enseñanza, y propone un modelo pentadimensional que contempla las dimensiones educativa, étnica, comunitaria, lingüística y psicocognitiva. Además, el enfoque holístico propuesto por Moreno Fernández integra conceptos cruciales como son la identidad, la generación y la integración. Por último, destaca la influencia de las familias y, especialmente, de los docentes a la hora de asociar a las lenguas de herencia un valor social que incidirá directamente en su mantenimiento o pérdida. Este capítulo se convierte en un vademécum indispensable para el profesorado de español, llamado a ejercer su labor desde la mediación y la responsabilidad.

En el segundo capítulo introductorio, Muñoz-Basols considera el multilingüismo desde el nivel del perfil específico del hablante hasta el del contexto sociolingüístico europeo. En este marco, el autor reconoce que en muchas ocasiones los docentes, a pesar de reconocer el valor del multilingüismo como un recurso social, personal y académico, no saben cómo trasladarlo a las aulas. Con este propósito, el trabajo les atribuye un rol proactivo como agentes de cambio y propone fases concretas para la acción en el aula —desde apreciar y reconocer la diversidad lingüística hasta integrar y fomentar el multilingüismo—. Además, resulta muy novedosa la visión que se propone sobre la integración

de la IA generativa como una herramienta para el futuro desarrollo de la mediación entre lenguas.

La primera sección titulada “El español en el aula multilingüe” representa una inmersión profunda en la praxis educativa al trasladar los constructos teóricos de la mediación y el plurilingüismo, delineados en los capítulos introductorios, al ecosistema del aula contemporánea. Este bloque consta de cinco capítulos que tratan desde diversas ópticas desafíos inherentes a la gestión de la diversidad lingüística y cultural.

El capítulo de Martín Leralta destaca los valores de identidad e integración asociados a la lengua para aquellas personas que se encuentran en un contexto migratorio. La autora argumenta de manera convincente que el docente trasciende su rol de mero instructor para convertirse en un mediador intercultural fundamental, un agente de transformación social y un gestor de un espacio afectivo seguro. Por último, detalla las siete características esenciales de los profesores de español como lengua de migración y refugio.

Por su parte, la contribución de Calvi ofrece un análisis sociolingüístico y una propuesta didáctica de gran calado desde el contexto específico italiano. El capítulo diagnostica con agudeza la dualidad del español en Italia: un auge del español como lengua extranjera dotado de prestigio y un español pobre que constituye la lengua de herencia (ELH) de las comunidades hispanoamericanas. La aportación más significativa del texto es su carácter eminentemente práctico al presentar estrategias y herramientas concretas aplicables a la mediación en el aula de ELE/ELH.

El trabajo de González Luna C. y Sa-

gi-Vela González, a través de una investigación sociolingüística de corte cualitativo y cuantitativo, explora el rol de la lengua en la construcción identitaria y el proceso de integración en la sociedad italiana de los estudiantes universitarios de origen hispanoamericano. La fortaleza de este capítulo reside en su rigor metodológico y en su capacidad para dar voz a los propios sujetos de estudio. El análisis de las percepciones, actitudes y usos lingüísticos revela un cuadro de situaciones heterogéneas. En consecuencia, las autoras enfatizan la importancia de que las medidas políticas estén basadas en estudios integrales que atiendan a las características concretas de cada comunidad de inmigrantes.

Fritzler postula la necesidad de un enfoque holístico, intergeneracional e intercultural, que denomina ecología didáctica. El valor principal de su aportación es la sistematización del modelo de vitalidad lingüística COD (*Capacity, Opportunity, Desire*), como herramienta para planificar una metodología adecuada. A partir de este modelo, la autora sostiene que, en un panorama cada vez más heterogéneo desde el punto de vista lingüístico, el enfoque pedagógico adoptado debería abarcar el proceso migratorio, la enseñanza, el aprendizaje y la construcción de la identidad de los hablantes de lenguas de herencia.

Finalmente, el capítulo de Lombardini constituye una reflexión crítica sobre las bases teóricas e ideológicas que sustentan la enseñanza de las variedades del español. A través de la formulación de preguntas clave que dan nombre a cada apartado, el autor desmantela con rigor la concepción de un español estándar monolítico y centralizado y, en su lugar, se acoge a un paradigma

pluricéntrico y plurinormativo. Uno de los aportes más significativos de este trabajo es la síntesis realizada sobre la variación de distintos fenómenos lingüísticos clasificados según los niveles de la lengua.

La segunda sección del volumen, “Traducción e interculturalidad”, está formada por cuatro capítulos, los cuales se centran en la praxis de la transferencia lingüística y cultural como un potente dispositivo pedagógico para el desarrollo de competencias complejas.

En el primer capítulo, Morillas aborda de manera práctica la traducción de la oralidad. Su contribución es valiosa por su capacidad para desentrañar y sistematizar la tarea de trasladar el habla viva, un acto que a menudo se da por sentado pero que está plagado de complejidades pragmáticas. Morillas propone un modelo de traducción al cubo, que descompone el proceso en tres diálogos simultáneos: la traducción intralingüística, la interlingüística y la interactiva. Al utilizar el género detectivesco-policiaco como campo de pruebas, demuestra cómo este registro, rico en oralidad, puede servir como un ejercicio de concienciación pragmática fundamental para el desarrollo de la competencia comunicativa.

El capítulo de Bazzocchi defiende de manera convincente que la literatura infantil y juvenil es un instrumento de máxima eficacia para la formación de traductores. Su mérito principal reside en la aplicación del modelo de subcompetencias del grupo PACTE para analizar los desafíos que presentan este tipo de textos. A través del análisis de *El país de Juan* de Andruetto y las rimas de Gloria Fuertes, Bazzocchi ilustra magistralmente la necesidad de mediar con referencias extralingüísticas, de recrear

la dimensión lúdica y sonora del lenguaje y, crucialmente, de negociar con el componente visual a través de la paratraducción.

La contribución de Valero Gisbert sistematiza el uso de la Traducción Audiovisual (TAV) como recurso didáctico y explora las potencialidades del doblaje, la subtitulación y la audiodescripción. De esta manera, el visionado pasivo se transforma en una actividad de aprendizaje activa y multimodal. La propuesta de actividades concretas—desde la transcripción y análisis de registros hasta la creación de diálogos alternativos—dota al docente de un repertorio de herramientas altamente eficaces para mediar entre lenguas y culturas en el formato que define la comunicación del siglo XXI.

Finalmente, el capítulo de Landone sostiene que la verdadera mediación intercultural no puede alcanzarse a través de enfoques basados en contenidos, sino que requiere una didáctica centrada en la relación y en la concepción del estudiante como una persona en toda su complejidad. Landone ofrece un camino para cultivar la empatía, la inteligencia emocional y las actitudes que, según el propio Marco común europeo de referencia, son el fundamento de la mediación.

La tercera y última sección del volumen “La mediación cultural en la didáctica del español para fines específicos” proyecta el concepto de mediación más allá del ámbito educativo general para insertarlo en dominios profesionales y sociales de alta especialización. Los múltiples campos de estudio tratados en esta sección son un ejemplo de la extensa aplicabilidad del español como lengua de servicio.

El capítulo de Garofalo realiza un diagnóstico riguroso de la opacidad histórica

del lenguaje jurídico, cuyas raíces se hunden en una tradición discursiva que perpetúa una asimetría de poder entre la institución y el ciudadano. La principal fortaleza de su contribución es la de enmarcar el lenguaje claro como un acto de mediación fundamental para la salud democrática.

Mapelli explora la creatividad inmersa en el lenguaje periodístico del fútbol y cómo esta puede convertirse en una herramienta didáctica efectiva dentro del aula. Este capítulo pone de manifiesto que la crónica deportiva es un laboratorio perfecto para la mediación textual debido a su riqueza léxica, su plasticidad sintáctica y su densidad cultural. Estas características la convierten en un material idóneo sobre el que trabajar la comprensión, la producción creativa y, en última instancia, la capacidad de mediar entre diferentes registros y visiones del mundo.

La contribución de Morelli y Muñoz Cruz aboga por un enfoque comparativo entre Italia y comunidades indomexicanas. Su propuesta concibe la mediación entre pares como un proceso emergente y dinámico condicionado por factores sociolingüísticos, culturales y de poder. Su trabajo es crucial para comprender que formar mediadores implica, necesariamente, fomentar la reflexividad, el pensamiento crítico y la capacidad de gestionar la incertidumbre.

Para cerrar la sección, el capítulo de Ortí Teruel y Raga Gimeno ofrece un análisis a través de los Patrones de Interacción Comunicativa (PIC). De este modo, abordan un análisis transcultural de las interacciones en el ámbito sanitario y emplean como corpus algunas escenas de series de televisión. Esta metodología permite identificar diferencias sutiles en los grados de

proximidad y distancia comunicativa con implicaciones directas en la eficacia sanitaria y en la mediación intercultural.

Este volumen presenta una coherencia interna y una progresión argumental que lo convierten en un manifiesto pedagógico riguroso y profundamente anclado en la realidad contemporánea. El recorrido propuesto a partir de la polifonía de autores implicados dota al lector de una visión panorámica que pretende construir puentes: entre la teoría y la práctica, entre la investigación sociolingüística y la aplicación didáctica, y entre el aula y la sociedad. Se trata, sin duda, de una herramienta indispensable para profesores de español, formadores de docentes, investigadores o cualquier profesional interesado en la enseñanza de lenguas en el siglo XXI.

**DOI: 10.14672/2.2025.3230**

**Ana I. Codesido García, Carlos Hernández-Sacristán, Victoria Marrero-Aguilar (eds.), *Lingüística clínica en el ámbito hispánico: un panorama de estudios*, Berlin, Peter Lang, 2023, 416 pp. ISBN: 978-3-631-89395-1**

**Alessandra Donati**  
**Università degli Studi di Milano**

*Lingüística clínica en el ámbito hispánico: un panorama de estudios*, volumen editado por Ana Codesido García, Carlos Hernández-Sacristán y Victoria Marrero-Aguilar,

reúne los estudios más recientes que abordan una amplia gama de trastornos lingüísticos y comunicativos –entre ellos afasias, patologías del neurodesarrollo, degenerativas y condiciones neuropsiquiátricas– analizados desde una perspectiva lingüística, neuropsicológica y clínica. El volumen incluye tanto contribuciones teóricas como estudios empíricos, muchos de ellos basados en datos experimentales o en recopilaciones de casos clínicos, con el objetivo de fortalecer el diálogo entre la lingüística y las ciencias de la salud y de ofrecer una visión general de los avances de la lingüística clínica en el mundo hispanohablante, en un contexto en que la lingüística aplicada a la práctica clínica está desempeñando un papel cada vez más relevante en el tratamiento de los trastornos del habla.

Los objetivos principales del volumen, bien presentados en su introducción, son varios. En primer lugar, visibilizar la pluridisciplinariedad que caracteriza los estudios y aportaciones más recientes en el ámbito de la lingüística clínica en español. En segundo lugar, poner el foco en estudios actuales que, al mismo tiempo, cuenten con una base empírica sólida, alejándose así de la tendencia de muchos otros textos. En tercer lugar, destacar el estado del arte de la disciplina, dejando de lado los enfoques generales para centrarse en los temas de interés actuales en la evaluación y rehabilitación de los trastornos del lenguaje y ofrecer una orientación para investigaciones futuras. Finalmente, el volumen busca incluir un amplio abanico de estudios, epistemológicos o metodológicos, todos provenientes de congresos académicos españoles y de investigaciones hispanoamericanas. En este sentido, el Congreso Internacional

de Lingüística Clínica (CILC), celebrado cada dos años desde 2006, representa el principal punto de encuentro español para estos enriquecedores intercambios. Es precisamente en este contexto internacional donde han confluído diversas perspectivas lingüísticas, culturales y geográficas que han dado lugar a los heterogéneos estudios incluidos en este volumen. Reconocer la lingüística clínica a nivel académico internacional, crear itinerarios formativos de distintos niveles, valorar el aporte de la lingüística en el tratamiento de los trastornos del lenguaje y crear una Asociación Internacional de Lingüística Clínica del español son, de hecho, los propósitos principales difundidos por el Manifiesto de Santiago, formulado con ocasión del VI CILC 2022 en Santiago de Compostela y difundidos a través de esta publicación.

El *leitmotiv* del libro es, sin embargo, el sujeto con alteraciones del lenguaje, puesto en relación tanto con su imagen social como con sus peculiaridades lingüístico-culturales. A este respecto, la perspectiva adoptada por los editores es fiel, como ya se ha señalado en los objetivos del volumen, a un principio pragmático y a la pluridisciplinariedad. En contextos de trastornos lingüísticos, y en general en cualquier situación comunicativa, la interacción es la vía de encuentro entre las personas, por lo cual la orientación pragmática resulta ser indispensable. Por ello, el volumen reúne estudios en los que el marco interactivo y el entorno –elementos de influencia constante– desempeñan un papel protagónico. La perspectiva pluridisciplinar es otra característica común en todos los estudios recogidos en este libro. De hecho, en casos de trastornos lingüísticos, las bases teóricas

de neurología, ciencias cognitivas, sociología y, en ocasiones, ciencias de la educación sustentan los métodos evaluativos y de intervención.

El volumen está dividido en cuatro secciones, y la mayoría de los estudios cuentan con múltiples autores. Las primeras tres partes abordan distintas habilidades lingüísticas en contextos clínicos: la primera reúne artículos sobre adquisición y desarrollo lingüístico, enfocándose en niños con déficits sensoriales y cognitivos; la segunda incluye estudios sobre cognición y pragmática con ancianos que presentan alteraciones lingüístico-comunicativas; la tercera se centra en la habilidad de lectoescritura y comunicación en diversos entornos clínicos. El cuarto bloque se distingue por su enfoque metodológico.

La primera sección agrupa cuatro estudios sobre adquisición y desarrollo del lenguaje, abordados desde distintas condiciones clínicas y metodologías innovadoras. Núñez Fidalgo y Ramírez Ferreiro analizan los primeros indicadores comunicativos en bebés prematuros mediante un cuestionario semiestructurado vinculado al método canguro. Madrid Cánovas examina el desarrollo lingüístico en niños sordos con implante coclear, identificando una asincronía entre el desarrollo léxico y fonológico según el modelo neurolingüístico de doble vía, lo que plantea la necesidad de intervenciones terapéuticas específicas. Varo Varo y Sarrias Arrabal utilizan técnicas neurofisiológicas (EEG y ERP) para evaluar el procesamiento cerebral del lenguaje en niños con cromosopatías de baja prevalencia, revelando patrones atípicos útiles para el diagnóstico e intervención temprana. Finalmente, Ema López y Marrero-Aguiré presentan un estu-

dio de caso con musicoterapia adaptada en una niña con síndrome de Phelan-McDermid, mostrando mejoras en vocalización e interacción verbal.

La segunda sección aborda alteraciones lingüístico-comunicativas en la vejez mediante seis estudios. Machado de Souza y colaboradores identifican biomarcadores fonéticos –como F0, velocidad del habla y calidad vocal– asociados a la depresión en hablantes de español, con potencial para diagnósticos tempranos. López-Ruiz y Rodríguez-López destacan la observación e imitación de gestos como estrategia compensatoria en pacientes con afasia. Cano Villagrasa y Valles González analizan el impacto conjunto de déficits cognitivos y lingüísticos en pacientes con daño cerebral adquirido crónico, subrayando la necesidad de evaluaciones integradas para mejorar la calidad comunicativa y la participación social. Moreno Campos examina estrategias conversacionales en afasia, revelando que la aparente fluidez oculta dificultades estructurales, lo que implica enfocar la terapia también en la competencia interactiva. Hernández Sacristán ofrece una visión fenomenológica de la anomia, evidenciando su efecto en la identidad y autorrepresentación del paciente. Finalmente, Sainz-Pardo Sainz resalta los efectos protectores del bilingüismo en el deterioro lingüístico asociado al Alzheimer, mostrando que actúa como reserva cognitiva y neural que retrasa la aparición de síntomas clínicos.

La sección *Destrezas y contextos* incluye tres estudios sobre lenguaje en contextos educativos y sanitarios. López-Sández y Lorenzo Herrera analizan la producción escrita de alumnado con dislexia en las prue-



bas ABAU Galicia 2021, identificando la persistencia de errores ortográficos, especialmente en la acentuación, lo que subraya la necesidad de apoyo ortográfico continuado. Míguez-Álvarez y Cuevas-Alonso exploran la relación entre comprensión lectora y metacompreensión en primaria, hallando una correlación significativa entre ambas habilidades, lo que justifica el fomento de estrategias metacognitivas desde edades tempranas. Bañón Hernández y Sánchez Pérez estudian las interacciones comunicativas en la donación en asistolia controlada, destacando la complejidad de la comunicación con familias y la importancia de protocolos que integren aspectos éticos, técnicos y relacionales.

La cuarta sección presenta dos enfoques metodológicos complementarios. Russell-Clari revisa el concepto de medicina basada en la evidencia aplicado a la logopedia, integrando variables personales del paciente para favorecer decisiones más éticas y contextualizadas. Rose analiza estudios longitudinales sobre adquisición fonológica infantil en distintas lenguas, destacando la influencia de la variación lingüística y el bilingüismo. Propone un modelo multifactorial basado en la interacción entre niveles acústico-articulatorio, léxico y estructural.

Entre los puntos fuertes del volumen destacan la amplitud temática, la atención a la variabilidad lingüística en el mundo hispanohablante, la riqueza metodológica y el objetivo de construir una base teórico-metodológica común. La obra se caracteriza también por el carácter coral de las contribuciones, realizadas por especialistas de distintos ámbitos y países, lo que enriquece los enfoques y refleja la complejidad de la lingüística clínica.

El volumen presenta una estructura clara y coherente: la división en secciones facilita una lectura ordenada y progresiva. Se valora su enfoque empírico y aplicado, que refuerza la utilidad del libro en contextos académicos y clínicos. Algunos estudios aportan propuestas relevantes para la salud pública, como el de Sainz-Pardo Sainz sobre el valor del multilingüismo frente al deterioro lingüístico.

Destacan la originalidad de varios temas y enfoques, así como la innovación metodológica. Se abordan enfermedades poco tratadas, como las cromosomopatías de baja prevalencia (Varo Varo y Sarrias Arrabal), junto a patologías frecuentes que requieren nuevas herramientas diagnósticas, como la depresión, analizada desde una perspectiva fonética por Machado de Souza, Ivanova y García Meilán.

Se presentan metodologías propias (Núñez Fidalgo y Ramírez Ferreiro) y aplicaciones novedosas de técnicas existentes, como en el estudio de Varo Varo y Sarrias Arrabal. También se destaca el uso de la musicoterapia en el caso clínico de de Ema López y Marrero-Aguar, como complemento a intervenciones tradicionales, especialmente en cuadros de alta complejidad genética y lingüística.

Otros estudios abordan el lenguaje como recurso cognitivo y relacional, más allá de lo técnico. Es el caso del enfoque 'comprensivo' y holístico de Moreno Campos, que considera factores físicos, cognitivos, lingüísticos y psicológicos, y del análisis de Hernández Sacristán sobre el impacto del déficit lingüístico en la vida emocional del paciente.

Aunque la diversidad temática y metodológica representa una de las mayores

fortalezas del volumen, esta misma amplitud puede dar lugar, en algunos capítulos, a una cierta variabilidad en el grado de profundidad analítica. Sin embargo, este fenómeno refleja, como ya se ha dicho, la complejidad del campo. Por otra parte, aunque el enfoque se centra mayormente en el ámbito ibérico, podría ser interesante una aún mayor interacción con la tradición investigadora latinoamericana, que sin duda enriquecería el panorama de la lingüística clínica en español.

En conclusión, *Lingüística clínica en el ámbito hispanico* es una obra valiosa para investigadores y profesionales interesados en la lingüística aplicada a la patología del lenguaje, así como para estudiantes de lingüística, logopedia y ciencias cognitivas. Ofrece una visión clara del estado actual de la disciplina en el ámbito hispanohablante, llenando un vacío bibliográfico y contribuyendo a su consolidación. Su enfoque integrador y orientación pragmática la convierten en una referencia imprescindible. Por ello, se recomienda su lectura a quienes trabajen entre la lingüística y el ámbito clínico y deseen comprender los retos actuales del lenguaje en contextos de vulnerabilidad y cambio.

**DOI: 10.14672/2.2025.3231**

**Veljka Ruzicka Kenfel, Raffaella Tonin, *Aproximación poliédrica a la cuestión del género en la literatura infantil y juvenil*, Berlin, Peter Lang Verlag, 2024, 158 pp. ISBN 9783631907627**

**Chiara Albertazzi**

**Alma Mater Studiorum Università di Bologna**

Este volumen, coordinado por Veljka Ruzicka Kenfel, catedrática de Filología Alemana en la Universidade de Vigo, y Raffaella Tonin, profesora titular de lengua y traducción española en la Università di Bologna, reúne seis contribuciones que giran en torno a la cuestión de género en la literatura infantil y juvenil (LIJ). A través de un enfoque multidisciplinario y una mirada “poliédrica”, como reza el título del volumen, los estudios en él recogidos abordan temas tales como la identidad, la igualdad, la diversidad y la inclusión y pretenden cartografiar el espacio ocupado por la LIJ en una sociedad en constante evolución.

El volumen se abre con una introducción de las editoras en la que se hace hincapié en el papel crucial de la LIJ a la hora de forjar la identidad y el sistema de valores de jóvenes lectoras y lectores y en su impacto en la construcción de una sociedad más igualitaria.

El primer trabajo recogido en este volumen, firmado por Ana Fernández Mosquera, traza un recorrido de la representación de género que se ha proyectado en la LIJ en el pasado siglo XIX y principios del XX a la vez que proporciona una fotogra-

fía de la actualidad de la LIJ, en concreto del género fantástico y de sus protagonistas femeninas. Bajo la lupa de los estudios de género y desde una perspectiva de evolución histórica, la autora observa cómo la crítica literaria feminista ha impulsado la revisión de los roles y estereotipos de género en la ficción. Tal y como se destaca en el artículo, la escritora Nesbit marcó un antes y un después en la historia de la LIJ a finales del siglo XIX y principios del XX, ya que fue pionera en la creación de protagonistas femeninas en la novela de aventuras. En las últimas décadas, el espacio ocupado por los personajes femeninos en la literatura juvenil ha evolucionado de manera significativa. Gracias a la labor de editoriales especializadas, se ha asistido a una proliferación de títulos: desde el género fantástico, hasta las biografías de mujeres preeminentes, sin olvidar las distopías y las novelas juveniles con contenido LGB-TQ+. La autora se detiene especialmente en literatura juvenil fantástica escrita por mujeres, un campo donde las escritoras se vieron doblemente silenciadas, en primer lugar, por ser mujeres escritoras y, en segundo lugar, por tratarse de un campo ya de por sí ninguneado durante mucho tiempo. En definitiva, tal y como subraya Fernández, el elemento fantástico puede ser un valioso instrumento, ya que permite cuestionar la sociedad, construir nuevos mundos y ofrecer nuevas perspectivas.

La segunda contribución es de Ana Pereira Rodríguez y Lourdes Lorenzo García, que ofrecen una panorámica de la LIJ no sexista e inclusiva que se ha editado en España en las primeras dos décadas del siglo XXI. Para ello, las autoras se apoyan en una bibliografía de 72 títulos elaborada

por el equipo de la Universidad de Vigo dentro del proyecto *Gender Identity: Child Readers and Library Collections (G-BOOK)* (2016-2018). Dichos títulos, dirigidos a dos franjas etarias (3-5 y 6-10 años) se han agrupado en cinco distintas áreas temáticas (azul para niños y rosa para niñas, la diversidad, la igualdad de género, la familia y, por último, modelos inspiradores que rompen con los estereotipos) para analizar más en detalle cada una de ellas. En el estudio se observa cierta preferencia por dos líneas temáticas: la igualdad de género (y la subversión de los roles masculinos y femeninos) y la familia (donde se rompe con los modelos tradicionales). Se han detectado, además, algunas diferencias dependiendo de la franja etaria a la que el texto va dirigido. Por ejemplo, problemáticas como la violencia de género, el acoso escolar y el abuso sexual aparecen solamente en los textos dirigidos a un público lector más maduro. Tras llevar a cabo el análisis temático, las autoras proceden a exponer los datos relativos a la traducción de los títulos examinados a las otras lenguas del proyecto (esloveno, francés, inglés e italiano) y, a modo de conclusión, se destaca que hay más traducciones al español que títulos editados en España y posteriormente traducidos a las otras lenguas del proyecto, en línea con la que viene siendo la tendencia general del mercado.

En el tercer estudio, Beatriz Rodríguez Rodríguez y Jesús Meiriño Gómez pretenden detectar los elementos relacionados con la diversidad de género más recurrentes en la LIJ (tanto libros como álbumes ilustrados) en lengua inglesa y, a continuación, analizar cómo estos se han tratado en su traducción al español. El análisis se

fundamenta en un corpus de diez textos en lengua inglesa con sus respectivas versiones en español, en los que se han identificado varios elementos relacionados con el tema del género. Entre los elementos que se han analizado y comentado a partir de la lectura de las obras recogidas se señalan los siguientes: estereotipos de género; las niñas princesas; profesiones y actividades que suelen asociarse, bien a los hombres, bien las a mujeres; estereotipos relacionados con la vestimenta y la apariencia física; la representación numérica y la caracterización de los personajes femeninos; patrones de comportamiento y, a nivel morfológico, las marcas de género, que a menudo vehiculan estereotipos (ej.: *médicos* y *enfermeras*). Tras llevar a cabo un análisis de dichos elementos aportando varios ejemplos, se señala que, en general, las decisiones traductoras han permitido mantener gran parte de los estereotipos y su intencionalidad. Sin embargo, también se pone de manifiesto cierta tendencia a la omisión y a la modificación, que en algunos casos hizo que se perdiera la ironía y la crítica subyacente a los estándares de género.

En el cuarto trabajo, Lourdes López Roperio analiza la representación de la masculinidad en la literatura juvenil angloamericana a partir de una selección de novelas representativas de géneros tales como la novela de internado o relato escolar, el realismo social y la distopía. Tras una revisión teórica sobre el concepto de masculinidad dentro de los estudios de género y la evolución del mismo a lo largo de las últimas décadas, se pasa a examinar un corpus compuesto por cuatro obras: *The chocolate war* (Robert Cormier), *A separate peace* (John Knowles), *What happened to*

*Lani Garver* (Carol Plum-Ucci) y *How I live now* (Meg Rosoff). A través de dichos textos, se demuestra cómo la masculinidad no es sino un constructo social y cómo la narrativa juvenil puede ser una poderosa herramienta para dismantlar el “andamiaje” que existe en torno a los roles de género y proponer modelos alternativos capaces de trascender la dicotomía masculino-femenino.

El quinto capítulo del volumen consiste en un estudio de Rebeca López González que se centra en los cuentos de hadas para el cine animado, un formato artístico que tiene un gran atractivo entre el público infantil y juvenil. El planteamiento de este estudio surge a partir de la representación de la figura paterna en los cuentos de hadas tradicionales: desde Perrault hasta los hermanos Grimm, el padre suele ser una figura autoritaria y a la vez ausente, que desatiende las necesidades e inquietudes de sus hijas e hijos. Esta premisa, por tanto, lleva a la autora a preguntarse si la forma de representar la masculinidad y, más en concreto, la paternidad ha experimentado algún cambio en las películas de animación. A raíz del análisis de cuatro largometrajes estrenados a lo largo de la década pasada (*Cómo entrenar a tu dragón*, 2010; *Gru, mi villano favorito*, 2010; *Shrek: Felices para siempre*, 2010; *Onward*, 2020), la autora arroja luz sobre nuevos enfoques en la representación masculina. De todas maneras, como sugiere la autora misma, se requeriría un corpus más extenso para determinar la tendencia actual en futuras investigaciones.

Cierra este volumen la contribución de Helena Cortés Gabaudan, que pretende desmontar los prejuicios de género que hay

en torno a los cuentos tradicionales, en particular a la obra de los hermanos Grimm. El revisionismo surgido en la segunda mitad del siglo XX de la mano de los estudios de género hizo que se empezaran a cuestionar los cuentos de hadas por perpetuar modelos sexistas. A partir de entonces ha habido una proliferación de productos culturales que revierten los estereotipos de género y reinterpretan el cuento tradicional. Sin embargo, esta perspectiva se explota muy a menudo y de manera poco acertada, incluso en el ámbito académico, donde la autora detecta cierto sesgo a la hora de analizar cuentos. A continuación, Cortés lleva a cabo un análisis de los arquetipos femeninos en la obra de los hermanos Grimm y demuestra cómo, a pesar de lo que se podría esperar tópicamente, la gran mayoría de las mujeres están lejos de ser personajes pasivos y sumisos, revelando modelos extremadamente emancipadores. A modo de conclusión, la autora hace hincapié en que la verdadera función de los cuentos no es describir personalidades, sino reflejar sentimientos humanos a través de figuras simbólicas que se hacen eco de valores morales y que ofrecen, a través de sus vivencias, consuelo frente a la adversidad.

En definitiva, *Aproximación poliédrica a la cuestión del género en la literatura infantil y juvenil* constituye un valioso recurso para quien quiera aproximarse a la LIJ desde una perspectiva de género e interdisciplinar. En un mundo en constante evolución, investigación y compromiso social están cada vez más entrelazados, como bien se refleja en este volumen. Apoyándose tanto en análisis teóricos como en ejemplos concretos (esto es, obras literarias y cinematográficas), cada una de estas con-

tribuciones analiza, desafía y cuestiona los roles y estereotipos de género en la literatura infantil y juvenil, explorando su impacto en la construcción de una sociedad más plural, abierta e igualitaria. Todos los estudios, además, ponen de manifiesto otro aspecto que no se puede pasar por alto, y más teniendo en cuenta la enorme proyección internacional de la LIJ, a saber, el rol de quienes la traducen. Traductoras y traductores no solo actúan como puente a la hora de difundir nuevas representaciones y visiones del mundo, sino que también tienen que cargar con la responsabilidad de reflexionar sobre su propia lengua para utilizarla de manera inclusiva y libre de sexismo.

DOI: 10.14672/2.2025.3232

**Natasha Leal Rivas, *La argumentación en la escritura académica de investigación: perspectivas, procesos y estrategias*, Madrid, Visor Lingüística, 2023, 183 pp. ISBN 9788498952803**

**Francisco José Sánchez García**  
**Universidad de Granada**

La presente obra es bienvenida no solo por su publicación bajo el sello de la prestigiosa editorial Visor Libros, sino también porque, como señala su autora en la Introducción (págs. 15-31), responde a la necesidad actual de guiar los procesos argumentativos académicos hacia un compromiso dis-

cursivo adaptado a las exigencias del siglo XXI. La autora presenta una reflexión profunda y una propuesta teórico-metodológica para el estudio de la argumentación en los géneros académicos científicos y formativos, adoptando una perspectiva social y enunciativa del discurso. Este enfoque busca lograr una comprensión integral de las diversas funciones de la argumentación en dichos géneros, promoviendo una conciencia eficaz de la función epistémica de la escritura como medio para fomentar procesos de reflexión y transformación del conocimiento.

Con este objetivo, los fundamentos teóricos de la obra abogan por la interdisciplinariedad, presentando una propuesta metodológica y un enfoque teórico tridimensional. En este marco, la autora desglosa los principales postulados e instrumentos epistemológicos que facilitan la activación de la función epistémica y crítica en la escritura de textos argumentativos de investigación. Uno de los aspectos más destacables de esta fundamentación teórica es el valor que se le otorga al acto de argumentar como un macroacto de habla estratégico y dialógico, en el cual la propuesta de argumentación y la formulación del pensamiento crítico en la escritura argumentativa se desarrollan en un plano de interacción mutua. En este proceso, el rol consciente del argumentador/enunciador se vuelve fundamental (p. 21).

*La argumentación en la escritura académica de investigación: perspectivas, procesos y estrategias* complementa la bibliografía prolífica sobre la escritura académica en español, demostrando un profundo conocimiento de las referencias clave que sustentan el estudio y ofreciendo una actua-

lización bibliográfica con obras destacadas sobre géneros académicos. Hasta la fecha, no existía un texto de referencia enfocado en la construcción argumentativa de los denominados géneros híbridos, donde convergen tanto los géneros formativos, como trabajos de fin de grado y tesis doctorales, como los géneros científicos, tales como artículos de investigación y capítulos de publicaciones científicas. En la construcción argumentativa de estos géneros híbridos se observan características discursivas comunes que, pese a las diferencias situacionales, persiguen demostrar la validez de una propuesta argumentativa basada en un proceso de indagación de fuentes. La autora ilustra este enfoque mediante un modelo esquemático y procesual que describe cómo, en la lógica de la investigación, se activan procesos cognitivos que exigen la selección precisa de mecanismos lingüísticos y retóricos para organizar el conocimiento previo y adquirido, construir el conocimiento nuevo mediante la argumentación, así como revisar y transferir el conocimiento resultante.

El libro se estructura en tres secciones, precedidas por una introducción (págs. 9-24) que funciona como declaración de intenciones, ya que en las primeras páginas se presentan los objetivos, la justificación del volumen y una síntesis de cada uno de los cinco capítulos que lo componen. La autora anticipa, de forma concisa y precisa, algunos de los fundamentos teóricos e instrumentos epistemológicos empleados para enmarcar la dimensión dialógica y estratégica de la argumentación, resaltando su importancia para el desarrollo del pensamiento crítico en la escritura académica formativa y científica. Los enfoques y pos-

tulados teóricos centrales se arraigan en la *Retórica* de Aristóteles y se enriquecen con aportaciones destacadas de la nueva argumentación de Perelman y Olbrechts-Tyteca (1958), Toulmin (1958), Plantin (1998) y Anscombe y Ducrot (1994), complementadas por perspectivas de la lingüística textual (Halliday) y el análisis del discurso, que enfatizan las relaciones entre texto y contexto, la organización textual y las dinámicas enunciativas de la argumentación.

La primera sección incluye tres capítulos. En el primero (págs. 33-78), Leal Rivas ofrece consideraciones preliminares sobre el contexto académico universitario como un espacio institucional en el que “aprendemos nuevos modos de pensar y de construir el conocimiento a través de sus distintas prácticas sociales y discursivas” (p. 34). A continuación, aborda la noción de discurso académico en español y revisa estudios relevantes en escritura académica y literacidad crítica en la argumentación académica.

El modelo conceptual tridimensional que propone es una aproximación metodológica ambiciosa, que permite comprender el lenguaje como un instrumento semiótico tanto social como individual, orientado a la validación de argumentos en géneros académicos específicos. Este modelo se enfoca en tres dimensiones: la dimensión sociocognitiva, que permite asimilar, elaborar, construir y transformar el pensamiento a partir de la comprensión y producción de la escritura argumentativa; la dimensión lingüístico-textual, que evidencia la importancia de los mecanismos de argumentación para comprender la relación entre texto y contexto cultural y comunicativo; y la dimensión retórica-dis-

cursiva, que explora cómo se construyen las fuerzas dialógicas y argumentativas de un texto según su contexto de enunciación y su estructura argumentativa (p. 62).

En el segundo capítulo (págs. 79-110), la autora profundiza en los mecanismos discursivos y lingüísticos fundamentales para una argumentación académica efectiva. Inicia con una exposición sobre la relación entre las fuerzas enunciativas (emisor, interlocutor y el discurso) en la argumentación académica. A partir de esto, distingue la estructura de un ensayo argumentativo de investigación en el ámbito académico, caracterizado por su función de presentar un posicionamiento crítico basado en un proceso de investigación de fuentes (p. 83).

Destaca especialmente la organización de los epígrafes en este capítulo, siguiendo un modelo de estructuración textual por niveles (estructura global, macro y microestructura) basado en el análisis de las operaciones cognitivas clave para la organización semántica y pragmática en la escritura argumentativa. Además, la autora clasifica las estructuras retóricas y argumentativas, y selecciona las estrategias y elementos argumentativos más pertinentes para canalizar los propósitos de los textos académicos y científicos, basándose en un corpus de 88 textos de estudiantes universitarios de español como lengua extranjera, que luego analiza en la sección empírica.

El tercer capítulo (págs. 111-122) cierra esta sección teórica al destacar la necesidad de un “reflejo discursivo” en todo proceso de investigación, que haga explícita la lógica de la indagación. La autora esquematiza los siguientes procesos: organizar el conocimiento previo mediante intertextualidad documental; planificar el



conocimiento adquirido; construir el nuevo conocimiento mediante una argumentación coherente; y revisar y transferir este conocimiento en textos académicos o científicos, respetando criterios de redacción formal y calidad argumentativa.

La segunda sección, correspondiente al capítulo 4 (págs. 123-148), presenta un estudio empírico sobre la escritura argumentativa de ensayos de investigación breve realizados por hablantes no nativos de español en contextos de no inmersión. Este análisis cualitativo evalúa la competencia argumentativa de un corpus de 88 ensayos, destacando la pertinencia y adecuación de sus estructuras argumentativas, así como la organización de las secciones argumentativas. Los resultados revelan mejoras en los participantes gracias a una Pauta de Actuación Guiada, aunque persiste una falta de estrategias para elaborar textos argumentativos adecuados, lo cual dificulta el desarrollo del pensamiento crítico y analítico.

La última sección, el capítulo 5 (págs. 149-159), concluye con orientaciones prácticas y pautas para la escritura argumentativa de investigación. La autora subraya la importancia de la coherencia argumentativa en los procesos de comprensión y producción crítica de un texto y resalta el papel de los marcadores discursivos y de evidencialidad, que contribuyen a cohesionar los argumentos y a reforzar las voces de autoridad bibliográficas.

En conclusión, este libro de Natasha Leal Rivas representa una valiosa contribución para quienes desean comprender en profundidad los procesos argumentativos en la escritura académica. Su utilidad para estudiantes universitarios, investigadores

y lectores interesados lo convierte en una referencia fundamental en el ámbito de la escritura argumentativa, promoviendo la reflexión sobre los parámetros de calidad en la transferencia del conocimiento.

**DOI: 10.14672/2.2025.3233**

**Nuria Pérez Vicente, *Los culturemas. Aprender a traducir con la literatura infantil y juvenil (LIJ)*, Bologna, Clueb, 2024, 140 pp. ISBN 978-88-491-5803-8**

**Rocío García Jiménez**  
**Universidad de Málaga**

Nuria Pérez Vicente, profesora titular de Lengua y Traducción española del Departamento de Estudios Humanísticos de la Universidad de Macerata, se ocupa en el presente volumen de la traducción de la literatura infantil y juvenil (en adelante, LIJ) desde un punto de vista teórico y, sobre todo, práctico. La autora pretende, mediante la lectura y el análisis de traducciones en la combinación lingüística italiano-español, que el traductor se vea provisto de una competencia traductora y de una aptitud crítica que puedan orientarlo en el ejercicio de su profesión. Para ello, pivotará sobre esos elementos con enorme carga cultural que aparecen constantemente en la traducción de LIJ, estos son, los culturemas.

Ya en la presentación, Pérez Vicente hace referencia a que la LIJ suele calificarse como marginal y no canónica den-



tro del polisistema literario. Esto sucede, principalmente, porque se trata de un tipo de literatura estrechamente ligada a lo infantil y a lo femenino. Sin embargo, la investigadora pretende demostrar lo contrario y, además de reivindicar este tipo de literatura, también la ensalza desde una perspectiva traductológica, ya que, en sus propias palabras (p. 8), «las características intrínsecas de la LIJ la convierten en un laboratorio de excelencia para el traductor o aprendiz de traductor».

El libro está dividido en cinco capítulos: en el primero, Pérez Vicente establece las bases teóricas que definen la LIJ e incide en la posición de extrarradio que esta ocupa dentro del polisistema literario estándar. Esta posición está normalmente alejada de lo canónico —y no de lo clásico—, ya que lo que leen los niños suele cambiar con frecuencia, dependiendo del momento histórico en el que vivan, lo cual distorsiona a menudo las bases de la LIJ. Esta posición periférica suele influir de forma negativa cuando la traducción entra en juego, puesto que las literaturas marginales son las que más expuestas están a manipulaciones de todo tipo. En este sentido, el capítulo habla sobre la poca consideración que se le ha dado a la LIJ a lo largo de la historia desde un plano tanto ideológico como pedagógico, sobre la contraposición de esta con la literatura para adultos y, por consiguiente, sobre la existencia de un doble destinatario (adulto y niño) que condiciona su traducción a otras lenguas y desemboca, en muchas ocasiones, en diferentes tipos de intervencionismo que pueden colocar al traductor ante la clásica dicotomía entre adecuación o aceptabilidad.

El segundo capítulo está dedicado al

culturema, elemento de estudio clave en esta obra y término usado por Vermeer por primera vez en los años ochenta —aunque Nida ya había tratado los elementos culturales en la década de los cuarenta—, cuando, dentro de los Estudios de Traducción, se produce el conocido *cultural turn*. La autora encuadra las investigaciones traductológicas sobre los culturemas dentro de la “teoría de los polisistemas” o la “escuela de la manipulación”, entre otros. A lo largo del capítulo, Pérez Vicente examina las características que definen los culturemas y ofrece un amplio abanico de posibilidades a la hora de afrontar su traducción, dependiendo de los factores implicados. Además, la autora retoma la dualidad entre adecuación y aceptabilidad, es decir, entre extranjerizar o domesticar. En la actualidad, se opta por mostrar al niño otras realidades diferentes a la conocida y abrirlo a la multiculturalidad, aunque respetando siempre su conocimiento limitado del mundo y haciendo que el texto que recibe le resulte lo más natural posible.

En el tercer capítulo, el más “consistente” según la propia autora, se parte de la clasificación de culturemas, por áreas temáticas, que Lucía Molina realizó en 2006, para examinar el corpus propuesto por Pérez Vicente. Por tanto, para el análisis se tienen en cuenta: a) el medio natural, b) el patrimonio cultural, c) la cultura social y d) la cultura lingüística, presentes en un corpus de obras originales en español y su traducción o traducciones al italiano. Dichas obras se enmarcan en un arco temporal bastante amplio —casi ochenta años— que transcurre desde los años cuarenta del siglo pasado hasta la actualidad. Cabe destacar que los autores de las obras

estudiadas son españoles en su mayoría y escriben también para adultos, como Fernando Aramburu, Elena Fortún, Gloria Fuertes, Elvira Lindo, Juan Marsé o Carmen Martín Gaité, entre otros. Por último, para el análisis se tiene siempre en cuenta la edad de los niños a quienes va destinado el texto, pues se trata de un dato esencial que, a juicio de la autora, determina las estrategias de traducción de los distintos culturemas. De este modo, dentro del medio natural, Pérez Vicente estudia animales y plantas, simbologías animales o aquellas comparaciones que hacen referencia a animales. Para el patrimonio cultural, toma en consideración elementos relacionados con la literatura y el cine, con las canciones, los ripios y las oraciones, con las personas y personajes, con los juegos, con los objetos y con la gastronomía. En lo que respecta a la cultura social, se examinan los bailes y la música, las marcas, las diferentes costumbres y comportamientos y el tabú. Por último, en la cultura lingüística se analizan los nombres parlantes —aquellos apelativos cargados de connotaciones y simbolismo cuya traslación al TM es crucial—, los sonidos y las onomatopeyas, las interjecciones, exclamaciones o injurias, los juegos de palabras, el *nonsense* y los neologismos, la variación lingüística, la fraseología o el lenguaje coloquial.

El cuarto capítulo se aborda desde un punto de vista eminentemente práctico, ya que la intención de Pérez Vicente es que traductores y aprendices de traductores puedan ejercitar y reflexionar sobre sus dotes traductorales al observar aquellas dificultades relacionadas con la traducción de algunos de los culturemas que se acaban de exponer en el capítulo anterior. Para ello,

la autora presenta una selección de catorce textos originales, tanto narrativos como poéticos, junto con su traducción o traducciones comentadas desde un enfoque descriptivo y sin emitir ningún juicio de valor. Algunos de estos textos son *Gloria Fuertes os cuenta cuentos de animales*, de Gloria Fuertes; *Abuelas de la A a la Z*, de Raquel Díaz Reguera; *Olvidado rey Gudú*, de Ana María Matute; *Vida de un piojo llamado Matías*, de Fernando Aramburu, o varios textos de Elvira Lindo y su famoso personaje Manolito Gafotas (*Manolito Gafotas* o *Pobre Manolito*). Algunos de estos textos han sido trabajados por Pérez Vicente en el aula, concretamente en cursos de máster, y la autoría de las traducciones presentadas corresponden tanto a traductores de reconocido prestigio como, en otras ocasiones, a estudiantes que las han propuesto para su TFM bajo la dirección de la propia Pérez Vicente.

Por último, el quinto capítulo está dedicado a las conclusiones. En él, Pérez Vicente vuelve a tratar el intervencionismo traductor cuando hay involucrados elementos culturales y, como consecuencia, la tendencia hacia la adecuación o la aceptabilidad. De este modo, a pesar de que la domesticación está presente, sobre todo, en aquellos casos en los que existe equivalencia, las técnicas más empleadas suelen ser la biculturalidad o la neutralización y rara vez se domestica de manera drástica. Además, la autora repasa una vez más, aunque esta vez de forma sucinta, el lugar marginal que ocupa la LIJ dentro del polisistema literario, su función didáctica o moralizante —la cual ha ido separándose, con el paso de los años, de lo meramente literario— o las dificultades que plantea

la presencia de un doble destinatario. Sin embargo, a pesar de todos los obstáculos que rodean a este tipo de traducción, Pérez Vicente considera haber conseguido demostrar que la traducción de LIJ es un campo de estudio absolutamente “válido y ampliamente revalorizado”.

Nos encontramos, por tanto, ante una obra de consulta útil y valiosa desde una perspectiva teórica, ya que Pérez Vicente repasa, con simplicidad y eficacia, los hitos históricos que han marcado tanto la LIJ como su traducción y expone de forma clara los rasgos que la caracterizan, con especial hincapié en los elementos culturales que influyen en su traducción y que la convierten en un tipo de traducción marcada por el intervencionismo. Sin embargo, la dimensión práctica es, sin duda, la que diferencia esta investigación de otras que se hayan podido llevar a cabo dentro del ámbito de la traducción de LIJ. A través del análisis de un corpus con una combinación lingüística concreta, como es la de italiano-español, se pueden extraer conclusiones aplicables a cualquier otro par de lenguas, así como a las teorías generales sobre la traducción de LIJ. De este modo, la obra suscita el interés no solo de investigadores en el campo, sino también de profesores y estudiantes de traducción, así como de traductores profesionales (consagrados o principiantes) que quieran conocer, reflexionar o ampliar su visión sobre un ámbito que cada vez adquiere más fuerza tanto a nivel académico como profesional.

**DOI: 10.14672/2.2025.3234**

**Isabel Santos Gargallo, Susana Pastor Cesteros (dirs.), *Metodología de la investigación en la enseñanza-aprendizaje del español como segunda lengua (L2)/ lengua extranjera (LE)*, Madrid, Arco/Libros, 2022, 391 pp. ISBN 97884713538648**

**Rosa María García Jiménez**  
**Università di Pisa**

La obra dirigida por Isabel Santos Gargallo y Susana Pastor Cesteros, dos reconocidas especialistas con una dilatada trayectoria docente e investigadora, resulta especialmente necesaria en el campo de la investigación sobre la enseñanza del español como lengua extranjera (ELE). El volumen tiene como propósito reunir en un solo compendio los fundamentos metodológicos de la investigación en ELE, una información que hasta ahora se encontraba dispersa en múltiples fuentes, ofrecer orientaciones claras para llevar a cabo investigaciones en todas sus fases y servir como obra de estudio y consulta para quienes deseen analizar los procesos de adquisición, aprendizaje y enseñanza del ELE desde distintas perspectivas.

Tal como señalan las autoras en la introducción, el libro está dirigido a un público amplio y diverso: estudiantes de grado, máster y doctorado que se encuentran en proceso de elaboración de sus trabajos finales; directores encargados de guiar esas investigaciones; formadores de profesores; miembros de grupos de investigación, así

como docentes en activo que deseen profundizar en su formación investigadora y compartir sus resultados (p. 9).

La estructura del volumen responde a una lógica progresiva y funcional: se compone de dieciocho capítulos organizados en tres bloques. El primero está dedicado al diseño de la investigación (4 capítulos); el segundo, a los métodos e instrumentos (10 capítulos); y el tercero, a la presentación de resultados y a la transferencia del conocimiento (4 capítulos).

En el primer bloque, titulado “El diseño de la investigación”, se incluye el capítulo de Susana Martín Leralta y Anna Doquin de Saint Preux, “Fundamentos de metodología de la investigación en lingüística aplicada a la enseñanza-aprendizaje del español como lengua extranjera”. En él se ofrece un panorama amplio y detallado de las metodologías más actuales aplicadas a la investigación en adquisición de segundas lenguas y, específicamente, en ELE. Se trata de una contribución especialmente valiosa por su claridad expositiva y por la capacidad de sistematizar un campo muy complejo. Por su parte, Ana Blanco Canales, en su capítulo “El proceso de investigación: el placer por descubrir”, expone los distintos tipos de investigación científica, describe las fases del proceso investigativo y muestra cómo diseñar y planificar un trabajo de investigación de forma rigurosa. Su enfoque resulta especialmente útil para quienes se inician en la investigación por la forma accesible, estimulante y estructurada en que se organiza y transmite el contenido. En el capítulo “La documentación: la acotación del tema y el plagio”, Iria da Cunha Fanego proporciona recursos útiles para la búsqueda bibliográfica y aborda el

tema del plagio, ofreciendo una muy buena síntesis de las diferentes formas de citación. Esta primera sección concluye con el trabajo de Irini Mayrou titulado “Diseño de investigación y análisis de datos”, en el que se subraya la relevancia de formular adecuadamente las preguntas de investigación y, de manera muy completa, se exponen las diferencias y la complementariedad entre los métodos cualitativos y cuantitativos, así como su utilidad en función de los objetivos del estudio.

La segunda parte del libro, la más extensa, está dedicada a los métodos y los instrumentos de investigación desde diferentes perspectivas. En ella se incluyen dos capítulos centrados en los grandes paradigmas metodológicos. Por un lado, el capítulo de Mercedes Pérez Serrano y Jon Andoni Duñabeitia, titulado “Métodos experimentales desde el paradigma cuantitativo”, ofrece una presentación nítida de las diferencias entre investigaciones experimentales, cuasiexperimentales y preexperimentales y, además, reflexiona críticamente sobre el tratamiento de los datos y algunas prácticas discutibles dentro de la investigación cuantitativa. Por otro lado, Isabel García Parejo, tras destacar las ventajas de la metodología cualitativa, propone una guía sumamente útil para quienes se inician en este campo con todos los pasos necesarios para llevar a cabo una investigación cualitativa rigurosa. En el siguiente capítulo, Macarena Ortiz Jiménez presenta las características principales del estudio de caso, así como sus ventajas e inconvenientes. La perspectiva del aula se desarrolla en los aportes de Olga Esteve Ruecas, con “Investigar en el aula”, y de Alicia Hernando Velasco, con “El enfoque

de investigación en acción”. Cecilia Ainci-buru firma el capítulo titulado “Cuestionarios y entrevistas para la investigación de ELE”, en el que expone de forma detallada las características que deben reunir estos dos instrumentos fundamentales para la recopilación de datos. Ernesto Martín Peris, uno de los pioneros en la investigación sobre enseñanza de segundas lenguas, realiza un estudio pormenorizado de los trabajos realizados en España en los últimos treinta años sobre los materiales didácticos en el capítulo “El análisis de los materiales didácticos desde la investigación”. Por su parte, Mar Cruz Piñol y Kris Bruyse, especialistas en el uso de corpus lingüísticos, presentan el capítulo “Corpus lingüísticos para investigar sobre ELE”, que constituye una guía muy detallada y funcional. En él se describen los distintos tipos de corpus y se incluye una serie de anexos con tablas que sintetizan las características más relevantes de los principales corpus disponibles, lo que convierte este capítulo en una herramienta de consulta de gran utilidad. Cierran esta extensa sección central los capítulos de M.<sup>a</sup> del Rosario Martínez Arias y de Alberto Rodríguez-Lifante. El primero, centrado en la estadística para lingüistas especializados en ELE, ofrece una panorámica clara de los principales instrumentos disponibles para el análisis de los datos. Se trata de un capítulo muy útil, aunque se requieren conocimientos básicos sobre el tema. El segundo, se focaliza en los códigos y los protocolos que deben regir toda investigación científica, así como en la importancia de una práctica investigadora responsable y rigurosa.

La tercera y última sección de la obra se centra en la transferencia y la difusión

del conocimiento que, como señalan acertadamente las editoras del volumen, “son aspectos sin los cuales la actividad investigadora no tendría sentido”, ya que “investigamos para mejorar el conocimiento que tenemos del mundo” (p.21). Esta parte final incluye cuatro aportaciones. En primer lugar, M.<sup>a</sup> Mar Galindo Merino, en el capítulo “Cómo abordar el trabajo de fin de grado y el trabajo de fin de máster”, describe de forma clara las distintas etapas que conforman el proceso de elaboración de estos trabajos de investigación tutelados. A continuación, Javier de Santiago Guervós y Jesús Fernández González se centran en las particularidades del desarrollo de una tesis doctoral dedicada a ELE. En el capítulo siguiente, Antoni Nomdedeu Rull expone las principales características que debe tener un artículo científico en ELE, ofreciendo pautas útiles especialmente para los estudiosos más jóvenes. Cierra el volumen el capítulo “La transferencia de resultados de investigación”, de Paz Villar-Hernández y Santiago Mengual Andrés, en el que se sistematizan las principales vías de transferencia y difusión del conocimiento, subrayando su relevancia en el contexto de la nueva realidad universitaria.

En conclusión, el volumen dirigido por Isabel Santos Gargallo y Susana Pastor Cesteros constituye una aportación de indudable valor al campo del español como lengua extranjera. El libro reúne a un elenco de investigadores con amplia experiencia en ELE, cuyas contribuciones destacan por su solidez teórica y su rigor metodológico. Cada capítulo ofrece no solo una exposición comprensible y bien estructurada del tema tratado, sino también ejemplos significativos de investigaciones que

ilustran su aplicación práctica. Se trata de una obra que ofrece una panorámica completa y actualizada de los distintos ámbitos de la investigación en ELE y que, además, aborda un ámbito escasamente tratado hasta ahora en la bibliografía especializada. Por todo ello, se trata de una publicación altamente recomendable que, sin duda, se consolidará como lectura imprescindible en la formación de nuevos investigadores en didáctica del español.

**DOI: 10.14672/2.2025.3235**

**Milín Bonomi, Maria Vittoria Calvi, Marcella Uberti-Bona (eds.), *Demolingüística del español en Italia, con un anexo sobre Malta, San Marino y Ciudad del Vaticano*, Madrid, Instituto Cervantes, Universidad de Heidelberg, Universidad de Zurich, 2024, 246 pp. ISBN 9788418210570**

**Sonia Bailini**  
**Università Cattolica del Sacro Cuore - Milano**

*Demolingüística del español en Italia, con un anexo sobre Malta, San Marino y Ciudad del Vaticano* es el quinto volumen de la colección «El español en Europa», un proyecto de investigación que pretende ofrecer información sobre la situación y uso del español en los distintos países europeos.

Italia es un país con grandes vínculos históricos y culturales con el mundo hispanohablante, que se remontan a la presencia española en Sicilia, Nápoles, Cerdeña y Milán, se conectan con las migraciones italianas al área del Río de la Plata y se reflejan en la presencia de comunidades hispanoamericanas como resultado de los flujos migratorios de los años más recientes.

En la actualidad, en Italia hay más de trescientos mil inmigrantes de origen hispano y casi un millón de aprendientes de español, a los que se suman aproximadamente dos millones de usuarios con competencia limitada, lo que permite afirmar que el español es una lengua en expansión por su prestigio y difusión internacional, además de una lengua de uso cada vez más frecuente en la vida cotidiana y en entornos profesionales.

El volumen está estructurado en seis capítulos, en los que se ofrece una panorámica puntual y rigurosa sobre los hablantes de español en Italia adoptando criterios de análisis demolingüísticos y una metodología cuantitativa. En el primer capítulo se definen el concepto y la finalidad de la demolingüística, entendida como el estudio demográfico de una o varias lenguas en un territorio concreto de acuerdo con unas condiciones contextuales específicas. Con este objetivo, en el segundo capítulo las autoras distinguen a los hablantes de español en Italia en cuatro categorías: el Grupo de Dominio Nativo (GDN) con 793 528 individuos, el Grupo de Competencia Limitada (GCL) con 2 241 385 personas, el Grupo de Aprendices de Lengua Extranjera (GALE) con 922 722 personas, y, por último, el Grupo de Usuarios Potenciales (GUP), que reúne, de hecho, a los GDN,

GCL y GALE, llegando a 3 957 635 personas, para luego describir cómo emplean estos hablantes el idioma. Estos datos permiten identificar dos perfiles muy claros: el primero incluye a las personas cuyo dominio del español está vinculado con su trasfondo migratorio y el segundo agrupa a los aprendientes de español en los diferentes grados de la enseñanza secundaria y universitaria, donde el español, en los últimos años, ha tenido un crecimiento vertiginoso. El tercer capítulo se centra en el Grupo de Dominio Nativo (GDN), que incluye tanto a inmigrantes hispanohablantes no nacionalizados como a aprendientes italianos de español que han adquirido un dominio parecido al de un nativo, a hablantes de herencia de segunda generación y a hispanohablantes nacionalizados. En Italia, los inmigrantes hispanos procedentes de Perú y Ecuador alcanzan poco más del 50% del total, seguidos por dominicanos, españoles, cubanos, colombianos, salvadoreños, bolivianos y venezolanos. A estos se suman los argentinos, muchos de los cuales han adquirido la nacionalidad italiana por ascendencia y, por lo tanto, no están contabilizados entre la población extranjera. De estos datos se desprende que la variación diatópica del castellano está muy bien representada en el territorio italiano, ya que va desde la andina hasta la caribeña y desde la peninsular hasta la rioplatense. El texto presenta un análisis muy detallado de la distribución regional de la migración hispanohablante en Italia y de variables como el género, la edad, el tipo de empleo, los motivos de la migración y el tiempo de permanencia en Italia. La segunda parte del tercer capítulo se centra en los hablantes de herencia, que presentan una competencia

lingüística muy heterogénea, dependiendo de si han nacido en Italia o si han llegado por reagrupación familiar, de cómo se ha desarrollado el proceso de escolarización y del uso que hacen del español en el entorno familiar. Las autoras afirman que las nuevas generaciones de origen hispano no encuentran en Italia un contexto migratorio socialmente favorable a la conservación del español debido a la escasa atención de las políticas lingüísticas italianas hacia las lenguas de herencia y a un sistema fuertemente asimilacionista hacia la lengua italiana. Además, la cercanía estructural entre el español y el italiano y la percepción de escasa distancia lingüística que genera han determinado el ingreso del italiano en los núcleos familiares de origen hispano tanto en una modalidad más cercana al monolingüismo como en una modalidad mixta. Este aspecto está en relación directa con el tiempo de permanencia en Italia, con la configuración del núcleo familiar, sobre todo en el caso de matrimonios mixtos, y con la edad de los hablantes de herencia.

El capítulo cuatro presenta la situación lingüística de los hablantes de competencia limitada (GCL), es decir, todos aquellos que no reúnen las características del grupo nativo y que no están inmersos en procesos de aprendizaje en el aula. Forman parte de este grupo los hablantes de herencia de segunda o tercera generación que no se han socializado en entornos favorables a la transmisión y el mantenimiento del español (por ejemplo, los descendientes de matrimonios mixtos), los italianos con experiencia migratoria en un país hispanohablante, los inmigrantes de origen no hispanohablante con competencias limitadas de español (por ejemplo, brasileños) y



los estudiantes de español egresados de la enseñanza reglada y no reglada sin competencia nativa. Con respecto a la enseñanza reglada, en el sistema escolar italiano la opción de aprender una segunda lengua extranjera empieza en el ciclo de educación secundaria de primer grado (Reformas Moratti, 2004, y Gelmini, 2008 y 2010), donde, además del inglés, obligatorio, se puede elegir entre francés, alemán y español. En este ciclo escolar el nivel de salida es el A2 y es en esta etapa educativa donde el interés por el español ha crecido mucho en los últimos quince años. La opción de elegir una segunda lengua extranjera continúa en el segundo ciclo de la secundaria, cuyo nivel de salida suele ser entre B1 y B2 dependiendo del tipo de instituto. Otro grupo de usuarios con competencia limitada son los egresados de cursos de español en carreras universitarias no filológicas y los que han participado en programas de intercambio Erasmus. El grupo de hablantes de español con competencia limitada es, según las autoras, el más numeroso y en él predominan los egresados de la enseñanza reglada.

El quinto capítulo está reservado a los aprendientes de español en Italia, donde, en las universidades, es la segunda lengua más estudiada después del inglés. Este interés se debe, por un lado, a la afinidad lingüística, que permite alcanzar rápidamente cierta competencia comunicativa y, por otro, al aumento del valor económico atribuido al español y su prestigio. Las autoras señalan que la oferta de español en el sistema escolar y universitario italiano no presenta ninguna relación significativa con la inmigración hispanohablante, porque las regiones con mayor presencia de apren-

dientes de español en la secundaria de primer grado son Lacio (con más del 50%), Marcas (43%) y Toscana (32%), que no son las regiones donde hay más población migrante de origen hispano (Lombardía y Liguria). Los datos indican que en la secundaria el español muestra una constante tendencia al alza, tanto en el primer ciclo, donde es la segunda lengua más elegida después del francés, como en el segundo, donde en los liceos y los institutos técnicos se puede estudiar entre tres y cinco años y permite alcanzar un nivel de salida B1-B2. Una situación tan positiva tiene margen de mejora, que reside en la posibilidad de aumentar el número de lenguas estudiadas y en la adaptación de la oferta de profesores de español a la demanda creciente.

Por lo que atañe a la enseñanza universitaria, Italia es el segundo país en el mundo por número de estudiantes de español después de Estados Unidos. Sin embargo, faltan datos oficiales que puedan permitir una estimación puntual de su primacía en este ciclo de estudios, porque las estadísticas que se suelen manejar se basan en criterios diferentes (opciones elegidas, número de exámenes, etc.) y no se difunden públicamente; además, el ministerio no dispone de datos completos. Esta dificultad en la recopilación de este tipo de información ha llevado a las autoras a utilizar los datos de tan solo cinco universidades italianas (Milán Statale, L'Orientale de Nápoles, Padua, para Extranjeros de Siena y Ca' Foscari de Venecia) en los grados y posgrados de filología y mediación. En algunos casos se han basado en los datos relativos a los exámenes, en otros en los de los alumnos matriculados en los cursos, así que la información recogida, aun destacando una



presencia significativa del español, deben ser interpretados teniendo en cuenta los límites relacionados con su recogida. Con todo, el total de los alumnos universitarios de español alcanza los 46 536 individuos y es cercano al número que ya incluía Reche Cala en el informe de 2021 *El español, una lengua viva*, realizado por el Instituto Cervantes. Si calcular el número de alumnos en las universidades es muy difícil, en la enseñanza no reglada es una misión imposible, ya que la única institución que puede ofrecer datos al respecto es el Instituto Cervantes (3 409). De hecho, no existen datos acerca de los estudiantes de español en las universidades populares, de la tercera edad, en los cursos de formación para adultos y en las academias privadas de lenguas.

En el sexto y último capítulo se presenta una síntesis de todos los temas tratados en el volumen y, con respecto al futuro del español en Italia, se destaca que todavía quedan espacios de expansión, pero que dependen de la valoración que se le atribuya al multilingüismo en la política lingüística nacional, que debería prestar más atención al potencial de los hablantes de herencia, por un lado, y a la formación lingüística para fines profesionales, por otro.

El volumen se cierra con un anexo sobre la situación demolingüística del español en Malta, San Marino y la Ciudad del Vaticano. En Malta el grupo de hablantes de español más numeroso es el de competencia limitada (GCL) con 5 726 individuos, la mayoría de los cuales egresados de la educación secundaria. En este país el español es la cuarta lengua extranjera después del inglés, el italiano, el francés y el alemán en la secundaria de primer grado y la quin-

ta en la de segundo grado. En el ámbito universitario, pese a la variedad de programas de grado y posgrado que incluyen el español y las culturas hispanohablantes, el número de alumnos es muy reducido. La presencia del español en la República de San Marino es casi nula, ya que el único dato interesante son los 2 632 sanmarinenses residentes en Argentina como resultado de la emigración del siglo pasado, mientras que en la enseñanza secundaria el español está presente solo en uno de los cuatro liceos del microestado y en ningún instituto técnico y tampoco aparece en la oferta de la Università degli Studi di San Marino. Por último, en la Ciudad del Vaticano, el español se enseña como segunda lengua obligatoria en la única escuela media y en los tres tipos de liceo presentes.

Un aspecto muy apreciable de este volumen es el esfuerzo por reunir datos y estadísticas hasta ahora muy dispersos analizando con rigor la información procedente de fuentes distintas y cruzando los datos de manera coherente. Otro aspecto muy positivo es la claridad con la que se desenmarañan y explican datos estadísticos cuya interpretación no siempre es fácil. Se trata de una investigación innovadora y necesaria, que está destinada a convertirse en una obra de referencia imprescindible para el análisis de la situación del español en Italia tanto en la actualidad como en el futuro, cuando los datos que contiene se conviertan en históricos y contribuyan a alimentar el acervo de materiales acerca de la presencia del español en este país.

DOI: 10.14672/2.2025.3236



---

# **POLÍTICA EDITORIAL**

Instrucciones para los autores

Proceso de evaluación

## Instrucciones para los autores y política editorial

1. *Cuadernos AISPI* publica preferentemente en español. También se aceptarán trabajos en italiano, catalán, portugués y, excepcionalmente, en otras lenguas.
2. Los trabajos serán originales e inéditos. Según la tipología corresponderán a distintas extensiones. Los **artículos** tendrán una extensión mínima de 35000 y máxima de 50000 caracteres. Las **anotaciones** —entendiéndose con ello exposiciones de talante crítico-bibliográfico que, al reseñar una o más obras, no se limiten a resumir el argumento sino que tracen un estado de la cuestión sobre los fenómenos, temas, autores o géneros tratados— deberán contenerse entre los 25000 y los 35000 caracteres. A la misma extensión —de 25000 a 35000 caracteres— se ceñirán las **entrevistas**. Las **reseñas** se extenderán entre un mínimo de 8000 caracteres y un máximo de 12000. El número total de caracteres para todas las tipologías de texto detalladas incluye espacios, notas y bibliografía final.
3. El número de imágenes, esquemas y figuras se reducirá a lo estrictamente indispensable.
4. En la sección *Convocatorias* se indicará el plazo para el envío de las propuestas (título, resumen de unos 1500 caracteres + referencias bibliográficas + datos del autor/a). Los autores deberán remitir por correo electrónico su propuesta en forma de resumen a la secretaria de la revista (cuadernosaispi@gmail.com)
5. En el caso de que la propuesta sea aceptada por el Consejo de redacción de la revista, los autores utilizarán la plataforma OJS para entregar los originales:
  - un archivo informático con el título (en la lengua del trabajo y en inglés), nombre del autor/a, afiliación profesional (con su correspondiente dirección completa) y una breve nota biográfica en español (máximo 500 caracteres);
  - por separado, un archivo en formato Word (u otro programa de tratamiento de textos compatible con Word), en el que no debe figurar el nombre ni identificación alguna del autor/a, con:
    - a) el texto del original, con el título (máximo 100 caracteres con espacios) en la lengua del trabajo y en inglés;
    - b) un resumen (máximo 500 caracteres con espacios) en español y en inglés, en el que se indicarán claramente el asunto tratado, la metodología utilizada y la tesis que se mantiene;
    - c) una lista de hasta cinco palabras clave en las dos lenguas del resumen.
  - en archivos separados, eventuales imágenes (en formato TIFF, JPG o PDF, con calidad suficiente para ser reproducidas).
6. La revista *Cuadernos AISPI* está abierta a la colaboración de todos los hispanistas. Los autores afiliados a universidades italianas deben ser socios de AISPI y estar en regla con el pago de la cuota anual.
7. Los trabajos se someterán a un proceso de selección y evaluación por pares, según el procedimiento y los criterios hechos públicos por la revista.
8. Los trabajos se ajustarán a las normas de estilo detalladas en el portal de la revista.

## Proceso de evaluación

La primera evaluación de los trabajos es realizada por los miembros del Consejo de Redacción, o, cuando proceda, por los editores de la sección monográfica, para establecer si se adecúan a las líneas temáticas de la revista y si cumplen los requisitos mínimos de calidad.

El Consejo de Redacción envía los originales, de forma anónima, a dos evaluadores, que pueden ser miembros del Consejo Científico y Asesor de la revista o externos a él. En caso de juicios contradictorios, se solicita un tercer dictamen. A partir de estos informes, se comunican a los autores los resultados del proceso de evaluación, motivando la decisión final de aceptación o rechazo y solicitando, con indicaciones precisas, eventuales modificaciones. Una vez realizados los cambios por parte de los autores, los trabajos son evaluados nuevamente.

