

Técnicas básicas de procesamiento del lenguaje natural en el análisis de menciones a mujeres en lecturas graduadas de ELE

Basic natural language processing techniques for the analysis of mentions of women in Spanish graded readers

Iban Mañas Navarrete

Universidad de las Islas Baleares,
Filología Española, Moderna y Clásica
iban.manas@uib.es

María-Valle Sell

Universidad Autónoma de Barcelona,
Departamento de Didáctica de
la Lengua y la Literatura, y de las
Ciencias Sociales
valle.sell@gmx.de

Recepción: 29 de junio del 2023

Aceptación: 19 de septiembre del 2023

doi: 10.22201/enallt.01852647p.2024.78.1084

Resumen

Los análisis de la representación de los géneros en materiales de ELE se han visto limitados hasta ahora por los métodos de recuento manual. El presente estudio abandona estos métodos y analiza la representación de los géneros en un corpus compuesto por 47 lecturas graduadas para estudiantes de ELE aplicando técnicas básicas de procesamiento del lenguaje natural basadas en el trabajo de Lucy, Demszky, Bromley y Jurafsky (2020). Se examinan las diferencias en el espacio dedicado a cada uno de los géneros y su evolución en el tiempo. Asimismo, se analiza el valor explicativo del sexo de los protagonistas sobre la distribución de menciones de hombres y mujeres. Los resultados muestran que, incluso cuando el protagonismo es compartido, las mujeres aparecen representadas en menor medida. Esta investigación aporta además un set de herramientas de análisis automático de textos que queda a disposición de la comunidad para el análisis de otros textos o para su adaptación al análisis de diferentes aspectos de la lengua con el fin de profundizar en la caracterización de los materiales de ELE.

Palabras clave: estudios de género; análisis de materiales; lecturas por niveles; procesamiento del lenguaje natural; Español como Lengua Extranjera; el texto como práctica social

Abstract

Analyses of gender representation in textbooks of Spanish as a Foreign Language (SFL) have been limited by manual-count methods. The present study adopts a different approach and investigates gender representation in a corpus of 47 graded readings for students of L2 Spanish by applying basic Natural Language Processing techniques based on the work by Lucy, Demszky, Bromley, and Jurafsky (2020). The differences in the space devoted to each gender and its evolution over time are examined. Likewise, the explanatory value of the gender of the main characters on the distribution of mentions of men and women is explored. The results show that women are under-represented, even in books where the main characters are both women and men. This work provides the results of the analysis and a set of tools for text analysis that is available for the community to apply it to the analysis of other texts or to adapt it to analyze different social aspects of language use and to deepen the characterization of the teaching materials used in SFL.

Keywords: gender studies; textbooks; graded readers; NLP; Spanish as Foreign Language; text as social practice

1. Introducción

El número de investigaciones centradas en el análisis de materiales de español como lengua extranjera (ELE) desde una perspectiva de género ha experimentado un crecimiento notable en las últimas décadas (Angulo Blanco, 2010; Filipović & Kuzmanović, 2020; Kuzmanović, 2022; Kuzmanović & Đuričić, 2016; Queiro Pérez, 2019; Robles Fernández, 2005; Taboada, Cejas, Sánchez, Spada, Taboada, Caire, Olivera, Pais, Ábrigo & Martín, 2021); hecho que refleja un importante paso en el proceso de toma de conciencia del contenido ideológico de cualquier material didáctico (Kramsch & Vinall, 2015). El presente estudio entronca con esta línea de trabajo y recurre a las ventajas de las técnicas de procesamiento del lenguaje natural y el tratamiento cuantitativo de datos para el análisis de materiales didácticos.

El procesamiento del lenguaje natural (NLP por sus siglas en inglés) puede resultar de gran ayuda en la obtención de datos relevantes para la investigación en segundas lenguas (Meurers & Dickinson, 2017) y, concretamente, en el análisis de materiales, puesto que ofrece la posibilidad de manejar corpus extensos de textos de manera automática y sistemática. En la investigación educativa, no obstante, la adopción de técnicas de NLP es todavía infrecuente (Fesler, Dee, Baker & Evans, 2019); especialmente escasos resultan los trabajos sobre su aplicación en la búsqueda de respuestas de corte sociológico en la educación (Lucy *et al.*, 2020: 2).

En este trabajo se analizan las menciones de mujeres en un corpus de 47 lecturas graduadas de ELE con ayuda de técnicas básicas de NLP para la identificación automática de menciones de hombres y mujeres. Las lecturas graduadas son materiales didácticos, concretamente, textos ficcionales diseñados para ser usados en el proceso de aprendizaje de una lengua (Acquaroni, 1996). Las herramientas de NLP desarrolladas durante el estudio están disponibles tanto para su libre uso, a fin de replicar el estudio con otros corpus, como para su libre modificación, ampliación o mejora. Además de esta contribución, los resultados obtenidos suponen

un aporte cuantitativo al cuerpo de trabajos sobre la cuestión del género y su representación en los manuales de ELE, que generalmente se desarrollan dentro de un marco cualitativo–descriptivo. Se profundiza, así, sobre un mismo fenómeno desde una perspectiva que desvela aspectos poco explorados.

2. El pensamiento crítico y la enseñanza de ELE

Existe un amplio consenso entre especialistas de la educación acerca de la importancia del pensamiento crítico en cualquier proceso educativo (Castillo Cuadra, 2020). La práctica del pensamiento crítico, entendido este como “un pensamiento que facilita el juicio porque se basa en criterios de razón, es autocorrectivo y sensible al contexto” (Lipman, 2014: 37), contribuye a la formación de ciudadanos más preparados para vivir en una democracia plena (Nussbaum, 2010). Esto supone, entre otros aspectos, aprender a relacionarse con los otros desde el respeto mutuo y ser menos vulnerable a la manipulación mediática y a la presión del pensamiento único.

El aula de lenguas extranjeras, como espacio de enseñanza/aprendizaje, resulta un contexto propicio para desarrollar estas habilidades (Ascaso & Atienza, 2018); es, de hecho, un marco idóneo para la reflexión sobre cómo el uso de la lengua influye en la manera de ver el mundo. Desde la tradición lingüística sociocultural, el discurso no solo modela la cultura en la que se inscribe, sino que es modelado por esta y, en tal sentido, es representante de los rasgos de dichas culturas y sus sociedades (Cameron, 1992; Fernández Martínez, 2012). En el Marco Común Europeo de Referencia para las Lenguas (MCER), que comparte en muchos aspectos la visión de la lengua como transmisor de cultura, se subraya la importancia de “propiciar métodos de enseñanza de lenguas modernas que fortalezcan la independencia de pensamiento, de juicio y de acción, combinados con las destrezas sociales y la responsabilidad social” (Consejo de Europa [CE], 2002: 4). En consonancia con esto, el MCER describe a los aprendientes de lenguas principalmente como agentes sociales (CE, 2002: 9) y el Plan Curricular del

Instituto Cervantes (PCIC), a su vez, define los objetivos generales de los niveles de referencia para el español partiendo de una concepción del aprendiente como “agente social, hablante intercultural y aprendiente autónomo” (Instituto Cervantes, 2006: 9). Es desde esta dimensión sociocultural del hablante que se puede fomentar especialmente el espíritu crítico con los estudiantes (Fuertes Gutiérrez, Soler Montes & Klee, 2021; Leeman, 2014). Aprovechando las facilidades que ofrece el contexto de aprendizaje de una lengua extranjera (LE) para reflexionar sobre las creencias y los prejuicios (Guerrero, 2003), Lacorte y Atienza (2019) recomiendan sobre el uso de propuestas metodológicas orientadas a la acción en el aula de ELE, por ejemplo, el aprendizaje por proyectos o el aprendizaje basado en problemas, como propuestas que facilitan la integración del pensamiento crítico en los contenidos curriculares.

Ahora bien, dado que un texto es un elemento clave de la práctica social, este es siempre susceptible de ser analizado críticamente en un entorno social como es el aula. Lo fundamental no es tanto “lo que el texto dice, sino cómo este se deconstruye en clase” (Lacorte & Atienza, 2019: 146). Para deconstruir un texto es necesario conocerlo a fondo y un análisis cuantitativo como el presente puede revelar propiedades inapreciables desde una inspección cualitativa de los contenidos. En este sentido, las lecturas graduadas permiten introducir dicha práctica desde los niveles iniciales, contribuyendo así a formar el espíritu crítico del aprendiente desde el comienzo. Además, para muchos estudiantes en contexto de no inmersión representan un producto cultural más que les permite acercarse a la sociedad de la lengua meta. Las herramientas que ofrece el análisis crítico del discurso (Van Dijk, 2016) capacitan a los estudiantes para posicionarse ante el material didáctico como miembros de la sociedad (Lacorte & Atienza, 2019), con el objetivo de analizar el lenguaje como práctica social para descifrar las ideologías y relaciones de poder subyacentes (Amari, 2015; Fairclough, 1989; Trujillo, 2009). Una de las relaciones de poder que pueden estudiarse en los materiales didácticos de ELE son las diferencias de género.

3. La cuestión de género y la enseñanza de ELE

Este análisis se encuadra en la educación con perspectiva de género (EPG), cuyo objetivo es contribuir a erradicar desigualdades, así como “al empoderamiento de las mujeres y las niñas y a fomentar el análisis crítico de los roles que se han concedido a los distintos sexos a lo largo de la historia” (Martínez Lirola, 2022: 85).

Desde los inicios de los estudios en lenguaje y género (Lakoff, 1973) se han realizado innumerables trabajos sobre la representación de los géneros en los libros de texto (Cameron, 1992; Chisholm, 2018), entendiéndose *género* como el constructo social que define cómo ha de ser un hombre y cómo una mujer.¹

En la perpetuación o impugnación de la desigualdad que genera la distinción de ciudadanos por género, el sistema educativo, en cualquiera de sus modalidades y etapas, tiene un papel central. De ahí que desde las instituciones públicas se impulsen iniciativas dirigidas a promover la educación en igualdad, un enfoque pedagógico basado en la “reivindicación de la igualdad [...] que propone una reformulación del modelo de transmisión del conocimiento desde una perspectiva de género en los espacios de socialización destinados a la formación y el aprendizaje” (Instituto de la Mujer, 2008: 17), para una educación libre de estereotipos sexistas. En la misma línea, la IV Conferencia Mundial sobre las Mujeres recoge la importancia de “elaborar planes de estudio, libros de texto y material didáctico libres de estereotipos sexistas para todos los niveles de enseñanza” (UN Women, 1996: 56).

En los primeros niveles educativos esto es especialmente importante ya que los niños empiezan desde muy temprano a construir las nociones de feminidad y masculinidad en su interacción con el entorno. Pero los sesgos de género en los materiales didác-

¹ Esta interpretación de género responde al concepto elaborado desde el pensamiento filosófico del feminismo de la igualdad, que denuncia el desequilibrio estructural que el binomio sexo/género produce en sociedades patriarcales (Lerner 1986).

ticos se encuentran de forma transversal en todas las etapas y todos los contextos (Blumberg, 2007), incluida la literatura que se trabaja en aula de ELE (Chover Lafarga, 2007).

El informe de Blumberg (2007) destaca la infrarrepresentación sistemática de las mujeres en los materiales didácticos, así como el grado altamente estereotipado de las actividades que realizan. Asimismo, se afirma que los avances que han tenido lugar para las mujeres en el mundo laboral en las últimas décadas no se ven reflejados en los materiales didácticos y que únicamente se detecta una cierta tendencia de reducción del sesgo de género. En el ámbito de la educación en España, según Vaíllo Rodríguez, los estudios comparativos “han alertado sobre la insuficiencia de los cambios acaecidos” (2016: 107), cuando no de una invisibilización sistemática y constante en los libros de texto de la educación obligatoria (López Navajas, 2014; Peñalver Pérez, 2003).

En el área de ELE específicamente, el análisis de seis manuales por parte de Kuzmanović y Đuričić indica tan solo “una intención clara de evitar (o, al menos, minimizar) la presencia de contenidos o representaciones estereotipados de géneros” (2016: 117). A conclusiones similares llegaron otros análisis de materiales elaborados desde la perspectiva de género (Robles Fernández, 2005; Queiro Pérez, 2019, entre otros). Molina Gil (2022), en su estudio sobre la representación de mujeres escritoras en los cursos de literatura de ELE, confirma diferencias de género tradicionalmente asentadas en la historiografía literaria. Por tanto, es necesario continuar reivindicando la necesidad de introducir la perspectiva de género también en el contexto ELE.

En cuanto a la metodología aplicada, para el análisis del contenido se suele emplear el recuento manual de ocurrencias. Galiano Sierra (1993) realiza un recuento de las profesiones de las mujeres en las imágenes de nueve manuales de ELE. En otros trabajos, como el de Morales Vidal y Cassany (2020), los autores examinan hasta siete parámetros de desigualdad social en cinco manuales de ELE, entre ellas la categoría de género; si bien los autores se refieren a estereotipos de género y lo incluyen en la categoría de

diversidad sexual e identidades de género. En todos estos estudios, el tiempo necesario para realizar un recuento manual limita las dimensiones de los corpus.

Esta observación coincide con Lucy *et al.* (2020: 2), cuando afirman que “large textbook coding efforts may reduce their tasks to counting or identifying simple indicators”.² También Vaíllo desarrolla este tema y apunta que “las investigaciones producidas en la última década no abandonan la cuantificación, aunque evidencian un claro esfuerzo por perfeccionar los instrumentos de análisis en relación con los aspectos de corte cualitativo” (2016: 105).

Aunque se ha estudiado la cuestión de género en manuales de ELE, las lecturas graduadas están mucho menos exploradas. No se ha encontrado ningún análisis centrado en un corpus de este tipo. En el ámbito del inglés como LE, Muto–Humphrey (2005) analiza el lenguaje utilizado en dos lecturas graduadas y confirma un desequilibrio en las representaciones de género. Hunter y Chick, luego de analizar 47 lecturas simplificadas para niños nativos de inglés encontraron que los personajes masculinos adultos, además de aparecer con una frecuencia mayor estadísticamente significativa, ejercen una profesión con mucha más frecuencia que los femeninos (2005: 1). La principal dificultad en estos estudios es la cantidad de recursos humanos necesaria: Hunter y Chick (2005) contaron con la ayuda de 42 asistentes para la recolección de los datos. Resulta evidente que las investigaciones se ven limitadas por el sistema de recuento manual.

Algunos estudios, tomando conciencia de esta limitación, han aplicado técnicas de la lingüística de corpus, automatizando el análisis de datos en un tiempo razonable (McEnery & Hardie, 2011). Entre el conjunto de estos trabajos, cabe destacar el estudio de Hunt (2015) sobre las representaciones de género en los libros de *Harry Potter*, en el que la autora pondera el grado de agencia

² “[L]os grandes esfuerzos de codificación de libros de texto pueden reducir sus tareas al recuento o la identificación de indicadores simples” (la traducción de las citas es nuestra).

de los diferentes personajes en las situaciones en las que aparecen representados. Hunt expone que su estudio analiza aspectos más sutiles del texto y el discurso que estudios anteriores similares en el ámbito de la literatura infantil y llega a la conclusión de que en las novelas de *Harry Potter* “females [...] are systematically excluded from acting on the world in powerful or meaningful ways and the message is clear: agency is not for girls”³ (Hunt, 2015: 266).

El estudio de Lucy *et al.* (2020) es una de las investigaciones de referencia en la aplicación de técnicas de NLP para el análisis de cuestiones sociales en materiales didácticos. Los autores apuntan que “los recientes desarrollos metodológicos [...] son muy prometedores a la hora de arrojar luz sobre cuestiones clave de los aspectos políticos y sociales de la educación” (Lucy *et al.*, 2020: 1). Ahora bien, advierten que los enfoques computacionales no sustituyen los métodos de investigación clásicos (Lucy *et al.*, 2020: 3) y que lo recomendable es combinar ambos enfoques para aprovechar las fortalezas y compensar las debilidades de cada uno. En su estudio recogen un corpus de libros de texto de historia en Texas, Estados Unidos, y realizan un análisis de contenido para describir la imagen de grupos sociales históricamente marginados por cuestiones de sexo, raza y etnia. En sus resultados destacan la prevalencia del hombre blanco en los manuales, además de una infrarrepresentación de mujeres y personas negras.

Teniendo en cuenta los resultados de las investigaciones previas, el presente estudio se propone analizar la presencia de hombres y mujeres en las lecturas graduadas de ELE. Entre los estudios sobre esta cuestión es habitual encontrar líneas de trabajo dirigidas a modificar la presencia de la mujer en los materiales didácticos. Por ejemplo, Salido Machado (2017) destaca la importancia de tratar la imagen de la mujer en el aula de ELE desde una perspectiva intercultural. Ahora bien, para que esto se pueda plantear es

³ “[S]e excluye sistemáticamente a los personajes femeninos [...] de actuar en el mundo de manera poderosa y relevante, y el mensaje en claro: la agencia no es para chicas.”

imprescindible que, previamente, los docentes tengan a disposición materiales didácticos con suficiente presencia de mujeres en sus contenidos. Por este motivo, en este estudio se plantean las siguientes preguntas de investigación:

- ¿Cómo ha evolucionado la presencia de las mujeres, en términos cuantitativos, en las lecturas graduadas de ELE durante las dos últimas décadas?
- ¿Cómo se distribuyen las frecuencias de menciones a mujeres y hombres en las lecturas graduadas de ELE en función del sexo de los protagonistas?

4. Metodología

4.1. Descripción del corpus

El corpus se compone de 47 lecturas graduadas distribuidas en once colecciones, con un total de 472 768 tokens, 341 856 excluyendo la puntuación y 3 477 types. El corpus incluye 80% de las lecturas de la editorial Difusión. Se eligió dicha editorial por varios motivos. En primer lugar, se trata de una editorial de referencia en el área de la enseñanza de ELE (Kuzmanović & Đuričić, 2016) y ha sido objeto de estudio en diferentes análisis de materiales relacionados con el tratamiento del género y la diversidad sexual (Morales Vidal & Cassany, 2020; Ramírez, 2020; Taboada *et al.*, 2021). Por otro lado, se cuenta entre las primeras que incluyó en su catálogo lecturas graduadas de ELE, por lo que permite analizar la evolución en el tiempo de los parámetros estudiados.

En estas lecturas graduadas se encuentra un número suficiente de personajes principales femeninos para el análisis comparativo propuesto; en concreto, el corpus se compone de 18 lecturas con mujer protagonista, 12 lecturas con hombre protagonista y 17 con hombre y mujer como protagonistas. La incorporación de esta variable en el estudio y análisis de su efecto en la representación de mujeres y hombres contribuye al carácter innovador de esta inves-

tigación, en tanto que no se ha encontrado hasta la fecha trabajos similares que tengan en cuenta el sexo del protagonista. Este estudio se centra en cuantificar la presencia de hombres y mujeres, y no las diversas expresiones de género posibles. La flexión binaria de género en español permite rastrear únicamente las menciones a mujeres y hombres.

Para la creación del corpus, se digitalizaron todas las lecturas con la aplicación Scanner Pro y se convirtieron en texto mediante un programa de reconocimiento óptico de caracteres, ABBYY FineReader. No se incluyeron las ilustraciones ni el texto correspondiente a ejercicios y glosarios. A continuación, se eliminaron notas al pie, parte de la puntuación y caracteres especiales innecesarios. Cada libro fue etiquetado con su título, colección a la que pertenece, año de publicación y sexo de los protagonistas (véase el Anexo para conocer la lista completa de títulos incluidos en el análisis).

4.2. Las herramientas de NLP

El código desarrollado para el análisis se basó en el trabajo de Lucy *et al.* (2020). Los *scripts* programados están disponibles en GitHub para su libre uso con una licencia GNU General Public License v3.0. Con el fin de cuantificar el espacio destinado a hablar de hombres y mujeres fue necesario identificar términos referidos a unos y otras. Para ello se crearon dos listas, una de nombres propios masculinos y femeninos y otra de sustantivos con moción de género en la flexión para hombres (*camarero*), para mujeres (*jefa*) y sustantivos comunes en cuanto al género (*periodista*).

El listado de nombres propios se creó combinando datos extraídos de Wikipedia y del Instituto Nacional de Estadística de España. Se añadieron manualmente algunos nombres y también algunos apellidos, ya que en España se utilizan con frecuencia para referirse a hombres, y tan solo de manera anecdótica a mujeres (Guerrero, 2001). El listado final no incluyó nombres compuestos,

ya que el programa fue diseñado para ignorar la segunda parte del nombre, y contiene 2622 entradas.

Para la elaboración de la segunda lista se recurrió a WordNet (Fellbaum, 1998), una base de datos léxica en inglés. Se creó una lista de hipónimos de *woman* y *man* y se realizó una traducción automática al español. La Figura 1 ilustra la lógica simplificada del *script*. El algoritmo parte de una palabra y crea una lista que incluye dicha palabra y todos sus hipónimos. Después busca los hipónimos de cada una de las palabras contenidas en la lista y los añade a la misma. Este ciclo se repite tantas veces como se haya especificado en los parámetros de la función. La comprobación de si se trata de un término referido a personas se realiza mediante restricción a la categoría lexicográfica de sustantivos referidos a personas. El número de niveles o profundidad de la búsqueda se puede modificar cambiando el número de iteraciones ($n = 4$ en el código, *default* = 1). Asimismo, es posible introducir la categoría lexicográfica o dejar el argumento en blanco (*default* = *none*).

A continuación, se realizó una traducción automática al español del listado resultante utilizando la interfaz de programación de aplicaciones (API) de Google Translate (Han, 2020). Se revisó el resultado, primero por ordenador, para eliminar duplicados, y después manualmente, borrando sustantivos no adecuados y añadiendo también algunos términos a mano tras comprobar que faltaban muchas formas femeninas de sustantivos masculinos, como *ingeniera*. Se modificó la categoría de algunas palabras, por ejemplo, *turista*, de la categoría *hombres* a *sustantivos comunes en cuanto a género*. También se realizaron algunos ajustes manuales, como eliminar la palabra *amazonas* tras comprobar que en los libros aparece en su sentido geográfico, o *puta* y *puto*, porque aparecen en numerosas ocasiones como sustantivos adjetivados que forman parte de expresiones malsonantes. El resultado fue una lista de 459 términos, 186 de la categoría *mujeres*, 194 de la categoría *hombres* y 79 de la categoría *sustantivos comunes en cuanto al género*.

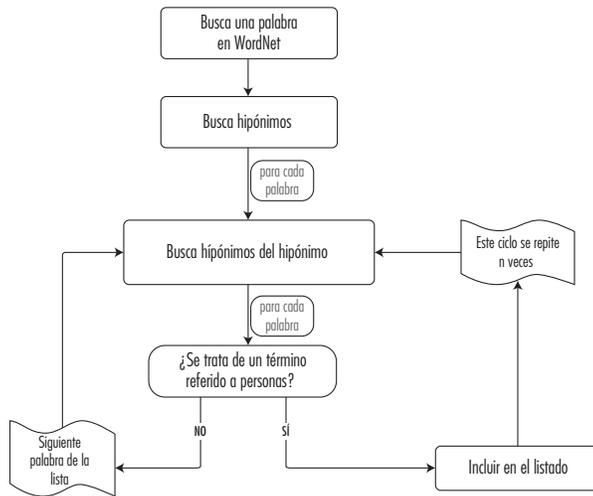


FIGURA 1. Lógica simplificada del algoritmo de búsqueda de hipónimos en WordNet

Por último, se añadieron los determinantes masculinos y femeninos y los pronombres y se combinó con el listado de nombres propios. Para comprobar que esta lista de términos era completa y, por tanto, que los resultados que se obtuvieran serían fiables, se extrajeron todos los núcleos nominales de las 47 lecturas que aparecen más de diez veces utilizando la biblioteca de Python spaCy (Honnibal & Montani, 2017), siguiendo a Lucy *et al.* (2020). Tras esta comprobación, se añadieron siete nombres propios (0.2% del total) y siete sustantivos (1.5% del total). Se creó un *script* modificando el de Lucy *et al.* (2020) para obtener el recuento de menciones de hombres y mujeres. La Figura 2 muestra la lógica simplificada del *script* modificado.

En el caso de los sustantivos comunes en cuanto al género, el algoritmo consultaba la palabra anterior y, si se trataba de un determinante incluido en el listado, añadía el sustantivo en la categoría correspondiente. Si la palabra anterior no era un determinante, confirmaba si se trataba de un adjetivo y, en caso afirmativo, comprobaba su flexión de género. Dado que muchos nombres propios

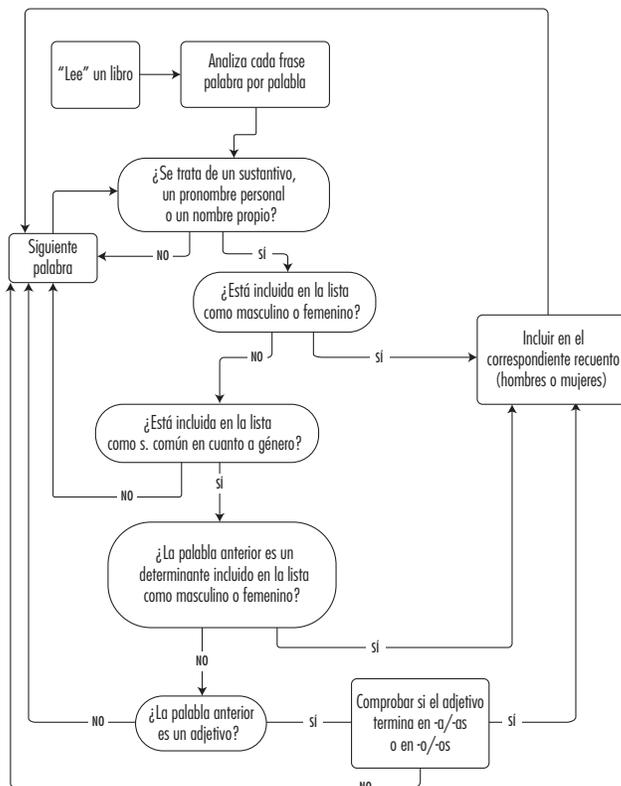


FIGURA 2. Lógica del recuento de menciones de hombres y mujeres

en español tienen carga semántica, algunas palabras, como *esperanza* entrarían en el recuento como nombre de mujer. Para evitar este resultado, se incluyó un listado de nombres propios con carga semántica, para que el algoritmo comprobara si los nombres estaban escritos con mayúscula inicial. Tras esta última modificación del código se llevó a cabo el proceso de recuento de menciones en el corpus.

4.3. Análisis

Se aplicaron dos análisis de correlaciones sobre el número de menciones a mujeres y a hombres. En el primero, se exploró la relación entre el número de menciones a mujeres y los años en que se publicaron las lecturas a través del índice de Spearman. En el segundo, se analizó la distribución de menciones a hombres y a mujeres dentro de cada subconjunto de lecturas según el tipo de protagonista (hombre, mujer o ambos sexos) a partir de la prueba de independencia χ^2 -cuadrado. Posteriormente, se exploró el alcance explicativo que tiene el sexo del protagonista sobre el número de menciones a hombres y mujeres. Tras confirmar sobredispersión de los datos en los dos casos (menciones a mujeres: $X^2 = 490,74$; $p < 0,001$; menciones a hombres: $X^2 = 589,24$; $p < 0,001$), se ajustaron sendos modelos de regresión binomial negativa con el paquete MASS en R. En los dos modelos se introdujo el número total de menciones como covariable centrada por la media y el tipo de protagonista como factor, ajustando como nivel de referencia la categoría de protagonistas de ambos sexos (MH).

5. Resultados

Una vez obtenido el recuento de menciones a mujeres y a hombres, se explora la relación entre el número de menciones a mujeres y el año de publicación de las obras. El intervalo temporal abarca aproximadamente las dos últimas décadas; desde 1999 (*El secreto de las flores* y *La chica del tren*) hasta 2020 (*Amor en línea*). Asumiendo que el número de menciones a mujeres en cada lectura está correlacionado con el número de menciones a personas, se computó un índice de correlación parcial entre las menciones a mujeres y el año de publicación, controlada por el total de menciones, calculado a partir de la suma de menciones a hombres y a mujeres en cada lectura (Figura 3).

Dado que no se cumplía la condición de distribución normal bivariada ($HZ = 0,89$; $p = 0,043$), se calculó el índice de correlación

de Spearman. El resultado apuntó a una ausencia de correlación entre el año de publicación y el número de menciones a mujeres ($r(45) = 0,19$; $p = 0,184$). Se confirma así una distribución uniforme de la presencia de mujeres en las lecturas graduadas durante los últimos 20 años.

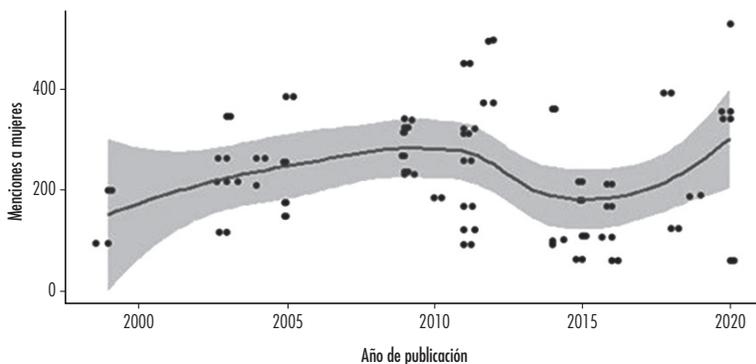


FIGURA 3. Correlación parcial del número de menciones a mujeres y año de publicación

La siguiente condición para el análisis de las menciones a mujeres fue el sexo del protagonista. Las lecturas se dividieron en tres grupos: protagonista hombre (H), mujer (M) y ambos (MH), y se observó la distribución de menciones a hombres y a mujeres en cada grupo a partir de una tabla de contingencia (véase Cuadro 1). En las lecturas con protagonista femenina las menciones a hombres representaron 47.4%, mientras que las menciones a mujeres supusieron 52.6% del total; a diferencia de los otros dos grupos, en los que se observó un desequilibrio pronunciado a favor de las menciones a hombres (61.5% en el caso de protagonistas de ambos sexos y 68.5% en las lecturas de protagonista masculino). El resultado confirmó la existencia de diferencias significativas en las distribuciones de menciones entre H y M según los protagonistas de las obras ($X^2 = 769,18$; $df = 2$; $p < 0.001$; $V = 0.17$, IC95% [0.14;0.18]).

CUADRO 1. Distribución de menciones a cada sexo en las lecturas graduadas según su protagonista

| | Protagonista | | |
|---------------------|----------------|--------------|--------------|
| | Mujer y hombre | Hombre | Mujer |
| Menciones a hombres | 6382 (61.5%) | 3888 (68.6%) | 4750 (47.4%) |
| Menciones a mujeres | 3988 (38.5%) | 1787 (31.4%) | 5270 (52.6%) |
| Total menciones | 10370 (100%) | 5675 (100%) | 10020 (100%) |

La exploración de los residuos estandarizados permitió determinar dónde se encontraban las diferencias entre proporciones, después de ajustarse el nivel de significación al número de comparaciones múltiples ($p = 0,008$). Tal como se esperaba, las lecturas con protagonista H presentaron una proporción mayor de menciones a hombres ($Z_{res} = 18,763$), mientras que la proporción de menciones a mujeres fue superior en las lecturas con protagonista M ($Z_{res} = 26,386$). Más destacable resultó la proporción significativamente mayor de menciones a hombres en las lecturas con protagonistas MH ($Z_{res} = 10,404$). De estos resultados se desprende que, en las lecturas donde hombre y mujer comparten protagonismo, la presencia de mujeres fue significativamente inferior a la de los hombres.

Más allá de la relación entre el tipo de protagonista (H, M o MH) y las menciones a hombres y mujeres, en este estudio se examinó el alcance explicativo que tiene el sexo del protagonista de las lecturas sobre el número de menciones a hombres y mujeres. Los resultados que arrojaron los modelos de regresión creados para cada uno de los recuentos permitieron explorar esta cuestión. En el Cuadro 2 se reportan los resúmenes de los modelos finales (Modelo 2) comparados con sus respectivos modelos basales (Modelo 1), teniendo el número total de menciones como único parámetro.

Según los resultados del test de razón de verosimilitud, la inclusión del tipo de protagonista como parámetro resultó una contribución significativa a la capacidad de ajuste del modelo sobre las menciones a mujeres ($X^2 = 3,485$; $df = 2$; $p < 0,001$) y a hombres ($X^2 = 11,616$; $df = 2$; $p = 0,003$). Además, para ambos conjuntos

de datos los índices globales pseudoR² mejoraron del Modelo 1 al Modelo 2. En el caso de las menciones a mujeres, el modelo final explicó el 76% de la varianza de los datos; para las menciones a hombres, el 74%. La exploración de los residuos de ambos modelos (menciones a mujeres: $W = 0,956$; $p = 0,078$; menciones a hombres: $W = 0,980$; $p = 0,608$) confirmó el comportamiento óptimo del Modelo 2 en ambos casos.

CUADRO 2. Resultados de los modelos de regresión para las menciones a mujeres y a hombres

| | Menciones a mujeres | | | | | | Menciones a hombres | | | | | |
|----------------------|---------------------|-------|------|----------|-------|------|---------------------|-------|------|----------|-------|------|
| | Modelo 2 | | | Modelo 1 | | | Modelo 2 | | | Modelo 1 | | |
| | Est. | SE | sig. | b | SE | sig. | Est. | ES | sig. | b | ES | sig. |
| Constante | | | | | | | | | | | | |
| Menciones total | 5,327 | ,0630 | ** | 5,389 | ,0534 | ** | 5,718 | ,0578 | ** | 5,695 | ,0388 | ** |
| Protagonista H | 0,001 | ,0001 | ** | 0,001 | ,0002 | ** | 0,001 | ,0001 | ** | 0,001 | ,0001 | ** |
| Protagonista M | -0,374 | ,1011 | ** | | | | 0,141 | ,0917 | ns | | | |
| | 0,318 | ,0867 | ** | | | | -0,175 | ,0802 | * | | | |
| pseudoR ² | | | | | | | | | | | | |
| McFadden | 0,119 | | | 0,059 | | | 0,107 | | | 0,088 | | |
| Nagelkerke | 0,768 | | | 0,517 | | | 0,744 | | | 0,673 | | |
| AIC | 516,94 | | | 547,43 | | | 540,74 | | | 548,36 | | |

Nota: ** significativo a un nivel $p < 0,01$; * significativo a un nivel $p < 0,05$; ns = no significativo.

Los estimados del Modelo 2 para las menciones a mujeres indicaron que tener una protagonista M incrementaba significativamente la tasa de incidencia de mención a mujeres (IRR = 1,374; IC95% = 1,160; 1,628), lo cual supone un 37% más comparado con las menciones en lecturas con protagonista MH. Por su parte, las lecturas con protagonista H mostraron una reducción significativa de 31% en la tasa de menciones a mujeres (IRR = 0,687; IC95% = 0,566; 0,836) en comparación con las lecturas con protagonista MH. El cambio de protagonista MH a M conllevó una tasa superior

de menciones a mujeres, a la vez que el cambio de protagonista MH a H implicó una tasa inferior de menciones a mujeres. En la Figura 4 se recoge la media de menciones a mujeres esperada en función del número de menciones total para los diferentes tipos de protagonista.

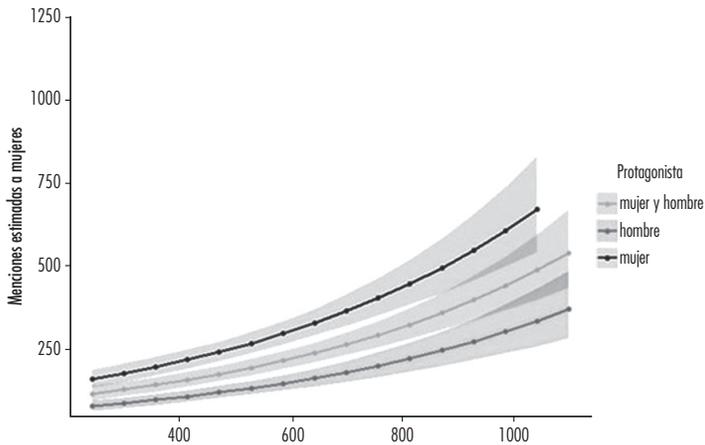


FIGURA 4. Media estimada de menciones a mujeres según el tipo de protagonista

En el Modelo 2 sobre las menciones a hombres, el estimado para la condición de protagonista M presentó el efecto opuesto al del Modelo 1 sobre menciones a mujeres. El protagonismo femenino conllevó una reducción significativa en la tasa de incidencia de mención a hombres ($IRR = 0,839$; $IC95\% = 0,716; 0,983$) en comparación con las lecturas con protagonista MH (Figura 5). En la tasa de menciones estimada para las lecturas con protagonista H, en cambio, no se presentó un aumento significativo comparado con la tasa de menciones de lecturas con protagonista MH ($IRR = 1,152$; $IC95\% = 0,958; 1,388$). Estos resultados indican que la tasa de menciones a hombres en los grupos de lecturas que tienen como protagonista MH y H son equivalentes. La ausencia de un cambio significativo entre estos dos grupos se puede explicar por la mayor presencia de hombres en las lecturas con protagonistas

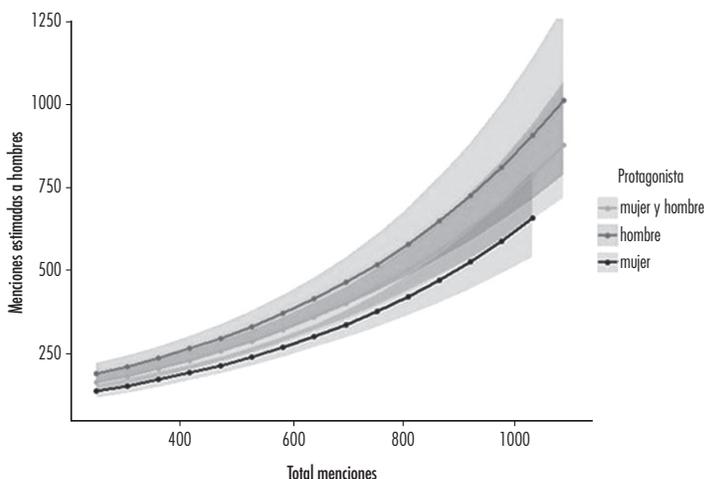


FIGURA 5. Media estimada de menciones a hombres según el tipo de protagonista

de ambos sexos, tal y como se mostró de manera descriptiva en el Cuadro 2.

Cuando las mujeres comparten protagonismo con los hombres, el espacio de las menciones a hombres fue equivalente al espacio que ocupan en obras con protagonista H. En el caso contrario, no se presentó este fenómeno. El número de menciones a mujeres aumentó significativamente de las lecturas con protagonista MH a las lecturas con protagonista M. Este efecto se debe a que las menciones a mujeres recuperan un espacio que, cuando comparten protagonismo con los hombres, poseen en una proporción inferior.

6. Discusión

En relación con la primera pregunta de investigación, sobre la presencia de las mujeres en las lecturas graduadas de ELE durante las dos últimas décadas, nuestros resultados han mostrado que el número de menciones a mujeres se mantiene a lo largo del tiempo. No se percibe, por tanto, ningún cambio en el espacio que ocupan las mujeres en el corpus. Estos resultados contrastan con las con-

clusiones que arrojan los análisis cualitativos de manuales de ELE, en los que generalmente se siguen detectando sesgos sexistas (Filipović & Kuzmanović, 2020; Kuzmanović, 2016; Kuzmanović & Đuričić, 2016), que se van reduciendo cuanto más actuales son los materiales.

La conjunción de estos resultados revela que, si bien se puede afirmar que se detecta un esfuerzo en el trabajo de editores y autores por reducir los sesgos sexistas a la hora de mostrar a las mujeres, en las últimas dos décadas no parece haberse superado el sesgo más elemental, el de la menor presencia de mujeres en los materiales didácticos. En veinte años el espacio que ocupan las mujeres en las lecturas graduadas no ha aumentado. Mientras, la imagen de las mujeres sí ha evolucionado en los manuales actuales de ELE, en concreto, se ha detectado una mayor presencia de mujeres y una reducción de lenguaje que las infravalore o subordine (Kuzmanović, 2016). Este resultado, relevante por sí mismo, se vuelve todavía más importante cuando se toma en consideración que, en el corpus, el número de lecturas con protagonista femenina ($n = 18$) es superior al de lecturas con protagonista masculino ($n = 12$). Entre los títulos de principio de siglo, donde se encuentran las lecturas de *Lola Lago*, cuya protagonista es una detective, se registra un número de menciones similar al de años posteriores, cuando las mujeres protagonistas no son tan frecuentes. Sabemos que en sociedades androcéntricas la producción cultural normalmente está copada por la presencia de hombres. Esto explicaría la inercia que opera por defecto y de manera tácita en la porción de espacio que se otorga a las mujeres, que persiste y resiste al margen del papel que ocupen las mujeres, protagonistas o no, en los relatos.

Respecto a la segunda pregunta de investigación, centrada en analizar las menciones a mujeres y hombres en función del sexo de los protagonistas, los resultados sobre la distribución de los totales señalaron una preponderancia ostensible de las menciones a hombres en las lecturas con protagonistas de ambos sexos. Una vez más, la disputa por el espacio en la narración se resuelve a favor de los hombres. Además, el análisis de la tabla de contin-

gencia desveló diferencias en las proporciones de menciones a hombres y mujeres según el tipo de protagonista. En lecturas con protagonista femenina, la presencia de hombres es notablemente superior (47.4%) a la presencia de mujeres en las lecturas protagonizadas por hombres, que no llegan a alcanzar un tercio del total de menciones (31.4%). Las lecturas con protagonista masculino prescinden en mayor medida de las mujeres que viceversa. Los resultados de los dos modelos ayudan a determinar el efecto de la variable sexo de los protagonistas. En el caso de las menciones a mujeres, el protagonismo femenino implica un incremento significativo del número de menciones, si bien tan solo supone poco más de la mitad, por lo que parecería que, incluso cuando las historias están protagonizadas por mujeres, la presencia de los hombres se antoja irrenunciable. Por otro lado, en las menciones a hombres, las lecturas protagonizadas por hombres y por ambos sexos muestran un número de menciones equivalente. Dado que el punto de referencia lo constituyen las lecturas con protagonistas de los dos sexos, en las que hay un desequilibrio patente a favor de los hombres, podemos afirmar que se observa un paso de la baja representación de mujeres en las lecturas coprotagonizadas con hombres a un cuasi-equilibrio de las menciones en las lecturas con mujeres protagonistas.

En cambio, al analizar las menciones a hombres, no se detectan diferencias en el espacio que estos ocupan entre las lecturas protagonizadas por hombres y las protagonizadas por ambos sexos. Estos resultados, leídos desde su reverso, podrían entenderse como la constatación de que las mujeres están infrarrepresentadas cuando comparten protagonismo con hombres. Estos resultados entroncan con las conclusiones que han alcanzado estudios previos sobre lecturas graduadas y simplificadas en inglés (Hunter & Chick, 2005; Muto–Humphrey, 2005).

La persistencia del patrón de dominación de los hombres sobre las mujeres se observa, por tanto, entre estudios de diferentes lenguas sobre el mismo género literario, lo cual justifica la importancia de este tipo de estudios con el propósito de alcanzar una des-

cripción más detallada de los materiales didácticos que se emplean en el aula de lenguas extranjeras. Un uso crítico de los materiales pasa inevitablemente por un conocimiento profundo de los mismos que capacite, tanto a profesores como a estudiantes, para desvelar la ideología subyacente de los discursos (Atienza Cerezo & Aznar Bertolín, 2020).

Además, los análisis de materiales desde una perspectiva crítica pueden estimular la toma de conciencia por parte de los creadores y despertar, como señalan Morales-Vidal y Cassany, un mayor “compromiso político de las editoriales del ámbito del español LE/L2 para asumir responsabilidades éticas más allá de sus legítimos intereses comerciales, con el propósito de reflejar mejor la diversidad de la comunidad y avanzar hacia un futuro con mayor equidad” (2020: 16). Para alcanzar una sociedad con mayor equidad, un paso ineludible es identificar y cuantificar la desigualdad existente derivada del sexismo.

La presencia del sexismo, como hemos podido comprobar en la revisión de estudios previos, se expresa no solo cualitativamente en la manera como se representa a las mujeres, sino también, como muestran los resultados de este estudio, cuantitativamente, limitando sistemáticamente el espacio que ocupan en relación con los hombres. En este sentido, cabe destacar el trabajo, desde el feminismo radical, humanista e ilustrado, por vindicar el derecho de las mujeres a ocupar el espacio público (de Miguel, 2004), del cual el espacio cultural y educativo forman parte. La lucha por la visibilización no tiene sentido si no se acompaña de la lucha por ocupar el espacio. Sin esta, la visibilización de la mujer corre el riesgo de convertirse en un cambio puramente estético, una estrategia de escaparate (Nogué, 2013). De esta postura se desprende una idea que no por ser evidente resulta menos reveladora: si las mujeres ocuparan más espacios, no necesitarían ser visibilizadas.

7. Conclusión

Los resultados de este trabajo han constatado que la presencia de hombres y mujeres no se distribuye de manera equilibrada en las lecturas graduadas de ELE. Análisis cuantitativos de este tipo son imprescindibles para proporcionar un escenario con datos robustos sobre el cual discutir los sesgos de género en los materiales didácticos. Al mismo tiempo, el profesor de ELE puede valerse de esta información para fomentar entre sus estudiantes una reflexión crítica acerca de las representaciones de hombres y mujeres en la sociedad. Si los sesgos de género son una de las “piedras mejor camufladas en el camino hacia la igualdad” (Blumberg, 2007: 4), información como la obtenida en esta investigación contribuye a la detección de características y patrones en los textos de ELE que, de otra manera, tal vez permanecerían camuflados.

En futuras investigaciones sería deseable aplicar un módulo de resolución de la anáfora, de manera que los sujetos elididos y los pronombres átonos sin flexión de género fueran asociados al personaje correspondiente. Si bien esto ayudaría a registrar un mayor número de menciones, la producción de sujetos nulos representa un mecanismo que afecta a ambos géneros por igual. Por ello, entendemos que los resultados del estudio no se desvirtúan por la ausencia de este módulo.

Por otro lado, en un siguiente paso, sería recomendable incorporar un análisis de contenido de las menciones a hombres y mujeres para desvelar qué imágenes de los géneros se promueven. Para ello se puede utilizar el código programado durante este trabajo, que permite explorar cuáles son los sustantivos más comunes utilizados para referirse a hombres y mujeres, las profesiones ejercidas y qué verbos y adjetivos les caracterizan con una frecuencia significativamente distinta. Asimismo, es importante investigar la evolución en el tiempo de dichas variables.

8. Referencias

- Acquaroni, Rosana (1996). Lecturas graduadas: algunas consideraciones para su incorporación a las actividades de clase. *Frecuencia L*, 1, 18–20.
- Amari, Fatma Zohra (2015). The role of critical discourse analysis in EFL teaching/learning. *Frontiers of Language and Teaching*, 6, 87–93.
- Angulo Blanco, María Esther (2010). *La evolución del papel de la mujer en dos manuales de ELE a través de las imágenes*. Las Lagunillas: Universidad de Jaén, Fundación Universitaria Iberoamericana.
- Ascaso, Samanta, & Atienza, Encarna (2018). El desarrollo del pensamiento crítico en ELE y su evaluación. *e-AESLA*, 4, 122–132.
- Atienza Cerezo, Encarna, & Aznar Bertolín, Joan (2020). Criterios para el desarrollo de las competencias interculturales en español LE/L2 desde una perspectiva crítica. *Journal of Spanish Language Teaching*, 7(2), 137–149. doi: 10.1080/23247797.2020.1844471
- Blumberg, Rae (2007). *Gender bias in textbooks: A hidden obstacle on the road to gender equality in education*. París: Unesco.
- Cameron, Deborah (1992). *Feminism and linguistic theory*. Londres: Palgrave Macmillan.
- Castillo Cuadra, Rodrigo (2020). El pensamiento crítico como competencia básica. Una propuesta de nuevos estándares pedagógicos. *Revista Latinoamericana de Filosofía de la Educación*, 7(14), 127–148.
- Chisholm, Linda (2018). Representations of class, race, and gender in textbooks. En Eckhardt Fuchs & Annekatriin Bock (Eds.), *The Palgrave handbook of textbook studies* (pp. 225–237). Londres: Palgrave Macmillan.
- Chover Lafarga, Ana (2007). El canon ‘otro’ de la literatura hispánica. Sobre la importancia de la perspectiva de género en la enseñanza de E/LE. En Jorge Martí Contreras (Coord.), *Actas del I Congreso Internacional de Lengua, Literatura y Cultura Española: la didáctica de la enseñanza para extranjeros* (pp. 161–168). Onda: Jorge Martí Contreras.
- Consejo de Europa (2002). *Marco Común Europeo de Referencia para las lenguas: aprendizaje, enseñanza y evaluación*. Madrid: Anaya.
- de Miguel, Ana (2004). La situación de las mujeres en el espacio público. En Adela García (Ed.), *Género y ciudadanía: un debate* (pp. 19–40). Barcelona: Icaria.

- Fairclough, Norman (1989). *Language and power*. Londres: Longman.
- Fellbaum, Christiane (Ed.) (1998). *WordNet: An electronic lexical database*. Cambridge: MIT Press.
- Fernández Martínez, Dolores (2012). Critical learning: Critical discourse analysis in EFL teaching. *Journal of Language Teaching and Research*, 3(2), 283–288. doi: 10.4304/jltr.3.2.283–288
- Fesler, Lily; Dee, Thomas; Baker, Rachel, & Evans, Brent (2019). Text as data methods for education research. *Journal of Research on Educational Effectiveness*, 12(4), 707–727. doi: 10.1080/19345747.2019.1634168
- Filipović, Jelena, & Kuzmanović, Ana (2020). Critical approach to intercultural education and foreign language learning and teaching: Neoliberal gender ideologies in textbooks of Spanish as a foreign language. *Porta Linguarum*, 34, 55–74. doi: 10.30827/portalin.v0i34.16733
- Fuertes Gutiérrez, Mara; Soler Montes, Carlos, & Klee, Carol (2021). Sociolingüística aplicada a la enseñanza del español. *Journal of Spanish Language Teaching*, 8(2), 105–113. doi: 10.1080/23247797.2021.2019448
- Galiano Sierra, Isabel María (1993). La mujer en los manuales de español para extranjeros. En Salvador Montesa y Antonio Garrido (Eds.), *Actas del Tercer Congreso Nacional de Asele. El español como lengua extranjera: De la teoría al aula* (pp. 119–136). Málaga: Asele.
- Guerrero, Susana (2001). Medios de comunicación y estrategias lingüísticas no sexistas. *Isla de Arriarán*, 17, 406–417.
- Guerrero, Susana (2003). Propuesta no sexista para favorecer la interculturalidad en ELE. En Manuel Pérez Gutiérrez & José Coloma Maestre (Eds.), *Actas del XIII Congreso Internacional de la Asociación para la Enseñanza del Español como Lengua Extranjera. El español, lengua del mestizaje y la interculturalidad* (pp. 393–403). Murcia: Asele.
- Han, SuHun (2020). *Googletrans: Free and unlimited Google translate API for Python* (Ver. 3.0.0). <https://py-googletrans.readthedocs.io/en/latest/index.html>
- Honnibal, Mathew, & Montani, Inés (2017). spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing. *Sentometrics Research*. <https://sentometrics-research.com/publication/72>

- Hunt, Sally (2015). Representations of Gender and Agency in the *Harry Potter* Series. En Jonathan Crichton (Ed.), *Corpora and Discourse Studies* (pp. 266–284). Londres: Palgrave Macmillan.
- Hunter, Maxwell, & Chick, Kay (2005). Treatment of gender in basal readers. *Reading Research and Instruction*, 44(3), 65–76. doi: 10.1080/1938807050955843
- Instituto Cervantes (2006). *Plan Curricular del Instituto Cervantes: niveles de referencia para el español*. Madrid: Biblioteca Nueva.
- Instituto de la Mujer (2008). *Guía de coeducación: síntesis sobre la educación para la igualdad de oportunidades entre mujeres y hombres*. Madrid: Instituto de la Mujer (Ministerio de Igualdad).
- Kramsch, Claire, & Vinall, Kimberly (2015). The cultural politics of language textbooks in the era of globalization. En Xiao Lan, Curdt-Christian & Csilla Weninger (Eds.), *Language, ideology, and education: The politics of textbooks in language education* (pp. 11–28). Nueva York: Routledge.
- Kuzmanović, Ana (2016). Los roles de género y los manuales de E/LE: ¿Perpetuación de estereotipos o cambio de ideología? En Anđelka Pejović, Ana Jovanović, Ivana Vučina Simović, Mirjana Sekulić, Vladimir Karanović & Aneta Trivić (Eds.), *Estudios hispánicos en la cultura y ciencia serbias* (pp. 339–348). Kragujevac: Universidad de Kragujevac.
- Kuzmanović, Ana (2022). ‘Haz realidad tus sueños’: las nuevas ideologías de género en los manuales de ELE. *Revista Internacional de Lenguas Extranjeras*, 17, 24–35. doi: 10.17345/rile17.3385
- Kuzmanović, Ana, & Đuričić, Milica (2016). ¿Manuales sexistas, aula sexista? El papel de docentes en la promoción de la igualdad de géneros en el aula de ELE. *Colindancias: Revista de la Red de Hispanistas de Europa Central*, 7, 209–224.
- Lacorte, Manuel, & Atienza, Encarna (2019). Dimensiones críticas en la enseñanza del español. En Javier Muñoz Basols, Elisa Gironzetti & Manel Lacorte (Eds.), *The Routledge handbook of Spanish language teaching* (pp. 628–640). Nueva York: Routledge.
- Lakoff, Robin (1973). Language and woman’s place. *Language in Society*, 2(1), 45–79.

- Leeman, Jennifer (2014). Critical approaches to the teaching Spanish as a local–foreign language. En Manuel Lacorte (Ed.), *The Routledge handbook of hispanic applied linguistics* (pp. 275–292). Nueva York: Routledge.
- Lerner, Gerda (1986). *The creation of patriarchy*. Nueva York: Oxford University Press.
- Lipman, Mathew (2014). *Pensamiento complejo y educación*. Madrid: Ediciones de la Torre.
- López Navajas, Ana (2014). Análisis de la ausencia de las mujeres en los manuales de la ESO: una genealogía de conocimiento ocultada. *Revista de Educación*, 363, 282–308.
- Lucy, Li; Demszky, Dorottya; Bromley, Patricia, & Jurafsky, Dan (2020). Content analysis of textbooks via natural language processing: Findings on gender, race, and ethnicity in Texas U.S. history textbooks. *AERA Open*, 6(3), 1–27. doi: 10.1177/2332858420940312
- Martínez Lirola, María (2022). *El análisis crítico del discurso y la pedagogía crítica: explorando sus relaciones y sus aplicaciones didácticas*. Granada: Comares.
- McEney, Tony, & Hardie, Andrew (2011). *Corpus linguistics*. Cambridge: Cambridge University Press.
- Meurers, Detmar, & Dickinson, Markus (2017). Evidence and interpretation in language learning research: Opportunities for collaboration with computational linguistics. *Language Learning*, 67(1), 66–95. doi: 10.1111/lang.12233
- Molina Gil, Raúl (2022). Las mujeres escritoras en los cursos de literatura de ELE: análisis de los manuales de Edelsa y Edinumen. *Lenguaje y Textos*, 55, 67–81. doi: 10.4995/lyt.2022.16485
- Morales Vidal, Elena, & Cassany, Daniel (2020). El mundo según los libros de texto: análisis crítico del discurso aplicado a materiales de español LE/L2. *Journal of Spanish Language Teaching*, 7(1), 1–19. doi: 0.1080/23247797.2020.1790161
- Muto-Humphrey, Keiko (2005). Gender balance in EFL textbooks: Graded readers. *Journal of School of Foreign Languages*, (35), 75–92. <https://core.ac.uk/download/pdf/235011991.pdf>

- Nogué, Neus (2013). Sexe, gènere i gramàtica. En Maria del Carme Junyent Figueras (Ed.), *Visibilitzar o marcar. Repensar el gènere en la llengua catalana* (pp. 47–66). Barcelona: Empúries.
- Nussbaum, Martha (2010). *Not for profit: Why democracy needs the humanities*. Princeton: Princeton University Press.
- Peñalver Pérez, Rosa (2003). *¿Qué quieres enseñar? Un libro sexista oculta la mitad*. Murcia: Instituto de la Mujer.
- Queiro Pérez, María Jesús (2019). *Análisis de imágenes de tres manuales de español para extranjeros bajo el enfoque de género* (Tesis de maestría inédita). Universidad de Barcelona, Barcelona. <https://diposit.ub.edu/dspace/handle/2445/135517>
- Ramírez, Guadalupe (2020). ¿Hablamos queer? La representación de la diversidad sexual en los manuales de ELE más vendidos en el cantón de Berna, Suiza. En Pilar Taboada de Zúñiga Romero & María del Rocío Barros Lorenzo (Eds.), *Perfiles, factores y contextos en la enseñanza y el aprendizaje de ELE/EL2* (pp. 903–916). Santiago de Compostela: Universidad de Santiago de Compostela.
- Robles Fernández, María Goretty (2005). *La mujer en los manuales de español lengua extranjera: del estereotipo al personaje real e histórico. Mujeres insignes, mujeres comunes, mujeres sin más* (Tesis de maestría inédita). Universidad Nacional de Educación a Distancia, Ciudad de México.
- Salido Machado, Elena (2017). Género e interculturalidad en la enseñanza del español como segunda lengua. En Marian Blanco Ruiz & Clara Sainz de Baranda Andújar (Eds.), *Investigación joven con perspectiva de género II* (pp. 285–300). Madrid: Universidad Carlos III de Madrid.
- Taboada, María Beatriz; Cejas, María Agustina; Sánchez, Claudia Carolina D.; Spada, María Fernanda; Taboada, Miriam Noemí; Caire, María Lucrecia; Olivera, Mariana Inés A.; Pais, Laura Gisela; Ábrigo, Carolina Araceli, & Martín, Ana Paula (2021). Una mirada crítica a la construcción de la mujer y la familia en materiales didácticos para ELE. En Carlos Valentini (Ed.), *Actas del IV Congreso Internacional SICELE: desafíos teóricos y nuevas perspectivas aplicadas*. Madrid: Instituto Cervantes. https://cvc.cervantes.es/ensenanza/biblioteca_ele/sicele/sicele04/indice.htm

- Trujillo, Juan Antonio (2009). Con todos: Using learning communities to promote intellectual and social engagement in the Spanish curriculum. En Manel Lacorte & Jennifer Leeman (Eds.), *Español en Estados Unidos y otros contextos de contacto: sociolingüística, ideología y pedagogía* (pp. 369–395). Madrid: Vervuert/Iberoamericana.
- UN Women (1996). *Informe de la IV Conferencia Mundial sobre la Mujer*. Nueva York: Naciones Unidas.
- Vaíllo Rodríguez, María (2016). La investigación sobre libros de texto desde la perspectiva de género: ¿hacia la renovación de los materiales didácticos? *Tendencias Pedagógicas*, 27, 97–124.
- Van Dijk, Teun A. (2016). Análisis crítico del discurso. *Revista Austral de Ciencias Sociales*, 30, 203–222. doi: 10.4206/rev.austral.cienc.soc.2016.n30-10

9. Anexo

Lista de lecturas graduadas analizadas

Serie: Lola Lago

Autoras: Lourdes Miquel y Neus Sans

Vacaciones al sol

Una nota falsa

Poderoso caballero

Por amor al arte

La llamada de La Habana

Lejos de casa

¿Eres tú, María?

Amor en línea

Sin noticias

Serie 24 horas en español. Cultura, convivencia, lengua, herencia.

Autores: Ernesto Rodríguez y Diana Chery

24 horas en español. Nueva York

24 horas en español. Miami

Serie: Un día en... Un día, una ciudad, una historia

Autor: Ernesto Rodríguez

Un día en Bogotá

Un día en Valencia

Un día en Sevilla

Un día en Barcelona

Un día en Madrid

Un día en Málaga

Un día en Salamanca

Un día en Ciudad de México

Un día en Buenos Aires

Un día en La Habana

Serie: Grandes personajes

Autoras: Aroa Moreno Durán (Frida Kahlo y Lorca), Laura Corpa (Dalí y Picasso), Daniel Cabrera (Che), Cecilia Bembibre (García Márquez)

Picasso. Las mujeres de un genio

Dalí. El pintor de sueños

García Márquez. Una realidad mágica

Che. Geografías del Che

Frida Kahlo. Viva la vida

Lorca. La valiente alegría

Serie: Aventura joven

Autores: Elvira Sancho y Jordi Surís

Aventura en La Habana

¿Dónde está Emiliano Fuentes?

Misterio en las Alpujarras

Objetivo: Barcelona

Perdidos en el Camino del Inca

Persecución en Madrid

El fantasma del instituto

El monstruo del rock

La chica de Mar del Plata

Trimestre maldito

Serie: Pepa Villa, taxista en Barcelona

Autoras: Alicia Estopiñá y Neus Sans

Fantasma en la escalera

Apartamento en la Costa Brava

Cenizas calientes

Una operación arriesgada

Serie: Hacerse mayor

Autora: Lourdes Miquel

Cosas del amor

Amor por sorpresa

Serie: El Mediterráneo

Autor: Jordí Surís

La chica del tren

El secreto de las flores

El puesto de fruta y otros relatos

Autor: Leo Thoma (Traductoras: María José Durán y Sonia Calvo)

El Camino de las Estrellas. Un viaje por el Camino de Santiago.

Autoras: Clara Villanueva y Josefina Fernández