

LA FLEXIÓN DE GÉNERO EN ELE. APORTACIONES E IMPLICACIONES DIDÁCTICAS DE UN ESTUDIO DE CORPUS

Ignacio Miguel Palacios Martínez

Universidad de Santiago de Compostela

Ana Fernández-Dobao

University of Washington

Resumen: Tomando como referencia las principales gramáticas descriptivas y pedagógicas del español, tratamientos didácticos y estudios previos sobre el procesamiento del género, este trabajo analiza las dificultades que tienen los estudiantes de español de diferentes L1 en el aprendizaje de la flexión de género de nombres y artículos. Con este fin, se analizan más de 45 000 ejemplos extraídos del CAES (Corpus de Aprendices de Español). Los resultados obtenidos muestran que, si bien los alumnos de los niveles más bajos son los que cometen un mayor número de discordancias, estas se encuentran presentes incluso en los niveles más avanzados. Los estudiantes de habla inglesa, chino mandarín y árabe son los que manifiestan mayores dificultades mientras que una tendencia contraria se observa entre los alumnos que tienen el portugués y francés como lengua materna. Para concluir se discuten las implicaciones didácticas de estos resultados y se proponen ejemplos de aplicaciones al aula.

Palabras clave: español como lengua extranjera, género, corpus de aprendices.

INFLECTION OF GENDER IN SPANISH AS A FOREIGN LANGUAGE. CONTRIBUTIONS AND DIDACTIC IMPLICATIONS OF A CORPUS STUDY

Abstract: Taking as a reference the main descriptive and pedagogical grammars of Spanish, instructional treatments, and previous studies on gender processing, this paper examines the difficulties that Spanish students of different L1s have in the learning of the noun inflection for gender with articles. To this end, more than 45 000 examples extracted from the CAES (Corpus of Spanish Learners) are analysed. Results show that, although students at the lowest levels show the highest number of agrammatical cases, these are even found among the most advanced learners. Students who speak English, Mandarin and Arabic show the greatest difficulties, while a contrary trend is observed among students who have Portuguese and French as their mother tongue. To conclude, we discuss the pedagogical implications of our results and propose examples of classroom applications.

Key words: Spanish as a foreign language, gender, learner corpus.

1. INTRODUCCIÓN

En este trabajo se combinan tres aproximaciones diferentes pero complementarias al tratamiento de la flexión de género del sustantivo en el ámbito del español como lengua extranjera (ELE). Por un lado, se emplea una perspectiva teórica pues se parte de las descripciones de este fenómeno que se encuentran en las gramáticas de referencia del español al uso (Alarcos Llorach, 1994; Bosque y Demonte, 1999; Alcina y Blecua, 2001; RAE, 2009) y de los trabajos previos sobre el procesamiento de esta marca gramatical (Franceschina, 2001, 2005; Barber y Carreiras, 2005; Lew-Williams y Fernald, 2010; Alarcón, 2011; Martínez-Gibson, 2011; Acuña et al., 2014; Scontras et al., 2018; Fuchs, 2019; Sá-Leite et al., 2020), para a continuación investigar en qué medida la flexión de género supone una dificultad en el aprendizaje de ELE, utilizando evidencia proporcionada por el Corpus de Aprendices de Español CAES. En función de los resultados, se hace una propuesta de intervención didáctica que intenta ofrecer solución a los problemas identificados. Se entrecruzan aquí enfoques de carácter teórico con otros de naturaleza aplicada, introduciendo al mismo tiempo un componente de investigación en acción (Schön, 1991), pues el profesorado trata de dar respuesta con su propia investigación a una dificultad identificada en el proceso de enseñanza.

To cite this article: Palacios Martínez, I.M. y Fernández-Dobao, A. (2024). "La flexión de género en ELE. Aportaciones e implicaciones didácticas de un estudio de Corpus". *Revista de Lingüística y Lenguas Aplicadas*, 19, 132-147. <https://doi.org/10.4995/rlyla.2024.19337>

Correspondence author: ignacio.palacios@usc.es



El artículo comienza por presentar el marco teórico en el que se fundamenta el estudio empírico para seguidamente formular los objetivos generales y describir la metodología. En el apartado siguiente se presentan los resultados del análisis del CAES. A continuación, se aborda la flexión del género del sustantivo desde un punto de vista pedagógico, haciendo una propuesta que pueda servir como referencia para ser adaptada en función del nivel y características de los estudiantes. Para formular esta intervención didáctica utilizamos ejemplos del propio corpus, de modo que este recurso nos proporciona tanto los datos para la investigación como el material para la enseñanza, aproximación didáctica hasta el momento apenas explorada. El trabajo finaliza con una serie de consideraciones generales.

2. MARCO TEÓRICO

La flexión de género en los sustantivos en español se ha venido estudiando de forma reiterada y desde perspectivas muy diversas. Por un lado, nos encontramos con tratamientos teóricos en las gramáticas de referencia y uso indicadas anteriormente, donde se destaca su importancia como fenómeno lingüístico y se describe su naturaleza semántica, formal y pragmática. Entre otras cuestiones se aborda la semántica del género en los nombres animados y no animados, así como su expresión formal, y la cuestión de la concordancia que también afecta a otras clases de palabras como los determinantes, adjetivos, pronombres y participios. Se tratan también los procesos de asignación del género, el género como mecanismo para distinguir aspectos diferenciales del sustantivo relativos al tamaño o forma (*jarro/jarra*), el caso de los nombres heterónimos, cuyo género viene marcado por el artículo (*el artista/la artista*), los sustantivos ambiguos en cuanto al género (*el mar/la mar*), los nombres epicenos que denotan seres vivos de uno u otro sexo mediante un único género gramatical, ya sea este masculino (*el rinoceronte*) o femenino (*la lechuza*), la distinción entre el género marcado y no marcado, etc.

Frente a estas descripciones nos encontramos con tratamientos de la flexión de género desde el punto de vista de su procesamiento, en hablantes nativos y no nativos (Franceschina, 2001, 2005; Barber y Carreiras, 2005; Lew-Williams y Fernald, 2010), bilingües (Fuchs, 2019; Sá-Leite et al., 2020) y de herencia (Alarcón, 2011; Martínez-Gibson, 2011; Scontras et al., 2018). Así, por ejemplo, Lew-Williams y Fernald (2010) demuestran con experimentos basados en la técnica del rastreo ocular cómo los hablantes que no son nativos no hacen el mismo uso de la información referencial proporcionada por el género de los sustantivos que los nativos. Mientras los primeros tienden a descansar en reglas proporcionadas en el proceso de enseñanza, los segundos interiorizan todo lo relativo al género de los sustantivos de forma natural a través de la exposición al input y la interacción. Sá-Leite et al. (2020), centrándose en estudiantes bilingües, demuestran con un metaanálisis cómo la teoría del efecto de congruencia de género se manifiesta de forma leve pero significativa en el fenómeno del procesamiento, es decir, el que la L1 y L2 de los hablantes muestren similitudes en el género de algunos nombres tiene efectos positivos en su procesamiento. Los hablantes de herencia son objeto de estudio, entre otros, por parte de Scontras et al. (2018), quienes contrastan cuestiones de procesamiento de número y de género entre hablantes nativos de español y de herencia. Sus resultados indican que las dificultades de concordancia de los hablantes de herencia no son de carácter accidental, sino que son debidas a cambios estructurales. Previamente Alarcón (2011) investigó el género en hablantes de herencia y estudiantes avanzados de español L2. De los resultados obtenidos se desprende que tanto los hablantes de herencia como los estudiantes avanzados de L2 poseen el concepto de género en sus gramáticas subyacentes y que los errores que de forma ocasional comenten se deben a dificultades para procesar las manifestaciones formales de las propiedades abstractas del género. Junto a estos trabajos existen otros en los que se aborda el procesamiento conjunto del género y número en hablantes nativos de español (Vigliocco et al., 1996; Acuña et al., 2014).

Especialmente en fechas recientes, la cuestión del género en el uso del español ha cobrado mayor protagonismo a propósito de lo que se ha venido a llamar el lenguaje inclusivo y no sexista. En esta línea podemos destacar, entre otros, los trabajos de Calero Fernández (1999), Portal Nieto (1999) y, más recientemente, Grijelmo (2019) y la propia RAE (2020). Portal Nieto (1999) reclama una mayor atención a la perspectiva de género en ELE como instrumento para desarrollar la competencia sociolingüística de los estudiantes.

A las aproximaciones anteriores, debemos añadir perspectivas aplicadas al aprendizaje y su enseñanza como son, por ejemplo, los trabajos de Ruiz Martín (2004), Clegg (2011) y Wang (2016), junto a su tratamiento didáctico en libros de textos (ej., Castro Viúdez et al., 2016; Alonso Arijá et al., 2019; Corpas et al., 2020), gramáticas pedagógicas (Matte Bon, 1992; Gómez Torrego, 2007). y manuales de dudas y ejercicios (ej., Garnacho López y Martín Acosta, 2014; Alonso Raya et al., 2021; Mellado Jurado y Klein, 2022). La concordancia de género de los sustantivos aparece de forma reiterada en todos estos materiales pedagógicos, junto a toda una serie de otros problemas asociados a los diferentes niveles lingüísticos como, por ejemplo, el uso del subjuntivo, la distinción entre los verbos *ser* y *estar*, los falsos amigos, la colocación de tildes, el uso de fórmulas de cortesía, etc.

Algo similar ocurre con documentos oficiales de carácter curricular como es el caso del *Marco Común de Referencia para la Enseñanza de Lenguas* (Consejo de Europa, 2001, 2020) y el *Plan Curricular* del Instituto Cervantes (2006), donde la concordancia del género tiene una amplia presencia. En este documento, al nivel A1 se asocia la distinción entre masculino y femenino definida por las terminaciones canónicas en *-o/-a*. También se incluyen aquí casos de sustantivos que terminan en vocal distinta de *-o/-a* o en consonante (*muerte, árbol, catedral*) y los procedentes de abreviación (*foto/fotografía*), así como los heterónimos (*hombre/mujer*). Frente a esto, en el nivel C1 se incluyen sustantivos donde el cambio de género implica cambio de significado, ya sea desde una denotación individual a un colectivo, (*el fruto/la fruta*), diferencias referidas al tamaño (*el cuchillo/la cuchilla*), o implique un significado totalmente diferente (*el frente/la frente*). Se asocian también a este nivel el grupo de nombres ambiguos masculinos y femeninos como *el mar/la mar*. Se dejan para el nivel C2 las tendencias que se observan en determinadas variedades del español que distinguen el género de ciertos sustantivos que no tienen marca de género en español estándar creando nuevas formas femeninas en *-a* (*estudiante, tigre, sapa*) o masculinas en *-o* (*pianista, maquinista*).

3. EL ESTUDIO

3.1. Objetivos

Es nuestro propósito examinar con datos extraídos del CAES en qué medida la concordancia de género supone un problema para los estudiantes de ELE. Es importante dilucidar hasta qué punto nuestra percepción subjetiva, derivada de la observación en nuestra actividad docente y de su tratamiento en los materiales de enseñanza, se confirma con el análisis del corpus. De este modo pretendemos también poner en valor la utilidad de los corpus de aprendices para la enseñanza y aprendizaje de lenguas. A este respecto es preciso señalar que en la mayoría de estudios anteriores se han empleado métodos de investigación basados en el rastreo ocular (Lew-Williams y Fernald, 2010; Acuña et al., 2014; Fuchs, 2019), listados de nombres extraídos de diccionarios (Bull, 1965; Techsner y Russell, 1984), diccionarios de frecuencias (Clegg, 2011) o tareas experimentales (Alarcón, 2011; Martínez-Gibson, 2011; Scontras et al., 2018; Sá-Leite et al., 2020), mientras que en este estudio utilizamos muestras escritas espontáneas que reflejan un uso auténtico de la lengua y que han sido convenientemente organizadas en un corpus. Por otra parte, es nuestra intención investigar hasta qué punto las dificultades de aprendizaje varían en función del contexto lingüístico y tipo de sustantivo, y del nivel y L1 del estudiante. Es esperable que el alumnado de los niveles más bajos tenga mayores problemas que el de los niveles más avanzados, pero es necesario investigar hasta qué punto estas dificultades persisten en estudiantes con un dominio elevado del idioma, tal como han puesto de manifiesto estudios previos (cf. apartado 2). Se espera asimismo que la L1 juegue un papel importante y que los casos de discordancia sean más frecuentes cuando la flexión de género no exista o sea muy diferente en la L1 del estudiante. A tenor de los resultados obtenidos, esbozaremos algunas implicaciones didácticas.

3.2. Metodología

El CAES es un corpus de aprendices de español como lengua extranjera compilado por un equipo de investigación de la Universidad de Santiago de Compostela y financiado por el Instituto Cervantes. En su versión 2.1, publicada en marzo de 2022, contiene más de un millón de palabras y muestras escritas de aprendices de 5 niveles de competencia lingüística (A1, A2, B1, B2 y C1) y con 11 lenguas iniciales o L1 diferentes (alemán, árabe, chino mandarín, francés, griego, inglés, italiano, japonés, polaco, portugués y ruso). Están representados un total de 2.544 participantes, con edades entre 15 y 83 años, de más de 60 países distintos.

Los participantes dieron su consentimiento para el uso de sus datos con fines de investigación y completaron una serie de tareas escritas de acuerdo con su nivel de español. Utilizaron para este fin una aplicación informática que automáticamente registraba, filtraba y organizaba las muestras recogidas. Estas fueron anotadas morfosintácticamente de forma automática con la herramienta Freeling y a continuación sometidas a un proceso de desambiguación manual destinado a subsanar las limitaciones de la anotación automática.

De forma paralela se diseñó una aplicación de consulta que permite hacer búsquedas simples o combinadas de elementos gramaticales, palabras ortográficas y lemas de acuerdo con variables como la L1 del estudiante, su edad, género, país de residencia o nivel. También es posible realizar búsquedas por tipo de texto y tema. La aplicación devuelve el número de casos del elemento buscado, los ejemplos que contienen ese elemento y una estadística completa de su distribución de acuerdo con las variables estudiadas. La frecuencia total viene acompañada de la frecuencia normalizada por millón de palabras, con lo que es posible establecer comparaciones entre grupos, ya que tanto el número de participantes como el de palabras varía considerablemente de un grupo a otro. Para más información se puede consultar la página web del corpus: <https://galvan.usc.es/caes>.

En este trabajo analizamos los datos recogidos de los hablantes de 5 de las 11 lenguas iniciales representadas en el corpus: árabe, chino mandarín, francés, inglés y portugués. Se incluyen así dos lenguas romances, francés y portugués, de la misma familia lingüística que el español y con un sistema de género muy similar; una lengua

indoeuropea, el inglés, que no asigna género gramatical al sustantivo; y dos lenguas de dos familias no europeas, el chino mandarín que tampoco marca el género de los sustantivos, y el árabe que tiene un sistema binario de género pero muy diferente al del español.

Para este estudio, centrado en el género de sustantivos femeninos y masculinos, se ejecutaron búsquedas de frecuencias completas y concordancias a través de la aplicación de consulta. Desde la casilla de etiqueta se seleccionó la búsqueda correspondiente a Determinante + Artículo + Masculino/Femenino para luego añadir, también desde la casilla de etiqueta, la búsqueda Sustantivo + Común + Masculino/Femenino. Las Figuras 1 y 2 ilustran los resultados de esta búsqueda para los estudiantes de L1 árabe. La búsqueda se repitió para todas las L1 incluidas en el estudio.

The screenshot shows the CAES (Corpus de aprendices de español) application interface. At the top, there is a navigation bar with 'CAES', 'Datos', 'Búsqueda', 'Listados', 'Ayuda', and 'Contacto'. Below this, there are sections for 'Búsqueda' (search type: 'Elem. gramaticales'), 'Resultado' (result type: 'Frecuencias completa', page size: '50'), and 'Sensibilidad' (checkboxes for 'Acentos' and 'Mayúsculas'). A 'Filtros' section includes dropdowns for 'Nivel de español' (Cualquiera), 'L1' (Árabe), 'País' (Cualquiera), 'Sexo' (Cualquiera), 'Edad desde' (Cualquiera), 'Edad hasta' (Cualquiera), 'Tipología textual' (Cualquiera), 'Tema' (Cualquiera), and 'Estudiante'. Below the filters, there are two 'Etiqueta' (tag) input fields, each with a search icon and a 'Lema' field. The first field shows '4 etiquetas seleccionadas' and the second shows '9 etiquetas seleccionadas'. At the bottom right of the filter area are buttons for 'Volver', 'Limpiar', and 'Buscar'. The main results area contains two tables: one for 'Nivel de español' and one for 'L1'. The 'Nivel de español' table has columns for 'Resultados', 'Estudiantes', and 'Frec. norm.' with rows for A1, A2, B1, B2, and C1. The 'L1' table has columns for 'Resultados', 'Estudiantes', and 'Frec. norm.' with a row for 'Árabe'. At the very bottom, there is a footer with copyright information: '© 2013-2023 CAES' and links for 'Política de protección de datos' and 'Política de privacidad'.

Nivel de español			L1				
	Resultados	Estudiantes	Frec. norm.		Resultados	Estudiantes	Frec. norm.
A1	152 / 55 230	88 / 203	2752 / millón	Árabe	464 / 170 475	233 / 502	2722 / millón
A2	197 / 48 144	75 / 127	4092 / millón				
B1	66 / 32 068	36 / 90	2058 / millón				
B2	32 / 19 322	22 / 52	1656 / millón				
C1	17 / 15 711	12 / 30	1082 / millón				

Sexo			
	Resultados	Estudiantes	Frec. norm.
Hombre	182 / 66 427	86 / 188	2740 / millón
Mujer	282 / 104 048	147 / 314	2710 / millón

Figura 1. Búsqueda de frecuencias completas en la aplicación de consulta del CAES.

The screenshot shows the CAES search interface. At the top, there is a header with the CAES logo and navigation links. Below the header, there are search filters for 'Búsqueda' (search type: 'Elem. gramaticales'), 'Resultado' (type of result: 'Concordancias'), 'Ordenación' (sorting: 'Coincidencia'), and 'Tamaño de página' (page size: '50'). There are also filters for 'Sensibilidad' (accent and uppercase), 'Filtros' (level of Spanish, L1, sex, age, text type, theme, and student), and 'Elem. gramaticales' (grammatical elements). The search results section shows 'Resultados 1 a 50 de 464' and a list of concordance examples. The first example is: 'de trabajar en Dubai , y envíe mi CV a este empresa que trabaja en educación , hoy he recibido el acepción para trabajar con ellos .'. The second example is: 'otras cosas me gusta mucho trabajar y aprender el niños y una veces tabaja como director de el actividades con niños .'. The third example is: 'interna de instalación , esto mismo me ha pasado el año anterior cuando tenía una escape en las tuberías de el agua'. The fourth example is: 'Las mujeres no deben dejar se tocar hasta qu el agua viene a el pueblo.'.

Figura 2. Búsqueda de concordancias en la aplicación de consulta del CAES.

Es importante señalar que en este trabajo nos limitamos a analizar casos de concordancia y discordancia de género entre el artículo, definido e indefinido en su forma singular y plural, y el sustantivo, común singular y plural. No incluimos casos de concordancia entre sustantivos y adjetivos, o con otros determinantes. Tampoco establecemos distinción entre la forma singular y plural de un mismo lema o entre la forma definida e indefinida del artículo, y excluimos del estudio los nombres propios.

De los casos recuperados por la aplicación fue necesario filtrar manualmente aquellos que no entraban en los objetivos del estudio al no ilustrar una falta de concordancia. Eliminamos aquellos casos con sustantivos que, según la variedad de español, pueden aparecer tanto con artículo masculino como femenino, por ejemplo, *el/la radio*, así como los usos de sustantivos en una lengua distinta al español, como *un maîtrise* o *un scholar*, o creaciones léxicas que no existen en la lengua española como, por ejemplo, *un tourna* (*un tourna de la ciudad*). Se eliminaron también las secuencias del artículo definido masculino *el/un* seguido de sustantivo femenino con /a/ tónica inicial como, por ejemplo, *el/un área* o *el/un aula*. La forma *el* del artículo precede por razones morfológicas a los sustantivos que comienza con /a/ tónica y por lo tanto “no existe discordancia de género en secuencias como *el alma dormida* o *el agua embalsada*” (RAE, 2009:82). Se mantuvieron sin embargo las formas plurales que sólo aceptan la forma femenina del artículo, por ejemplo, *las/unas áreas*. Por último, se revisaron y codificaron manualmente, atendiendo al contexto lingüístico, los usos de formas léxicas que varían de significado según el género, como *el capital/la capital*, *el frente/la frente* o *el orden/la orden*.

Una vez completadas estas primeras búsquedas de frecuencias y concordancias, e identificados los sustantivos que causan dificultades de asignación de género, se realizaron búsquedas individualizadas que permitieron analizar y comparar el uso de cada uno de estos sustantivos por parte de los distintos grupos de estudiantes. La herramienta de búsqueda del corpus permite combinar las formas del artículo determinado/indeterminado, masculino/femenino, singular/plural con el lema de un sustantivo concreto y así, por ejemplo, analizar y comparar los casos agramaticales **la/las/una/unas programa/s* con usos gramaticales del mismo lema, *el/los/un/unos programa/s*. La aplicación de consulta es sensible a variaciones ortográficas, de manera que, junto a formas estándar como *programa*, *ciudad* o *vacación*, recupera variantes ortográficas como **programma*, **cuidad*, o **vacaion* / **vacacion* / **vacaciones* / **vacaciones*.

3.3. Análisis y resultados

En una primera fase de nuestro análisis nos centramos en los sustantivos de género femenino, analizando todos los casos en los que estos sustantivos aparecen precedidos de artículo. Una vez depurado el corpus, siguiendo el proceso descrito anteriormente, recuperamos un total de 22.773 casos. En 1.294 de estos casos el sustantivo femenino aparece acompañado de un artículo masculino. Hay por lo tanto discordancia de género entre el artículo y el sustantivo femenino en un 5.68% de los casos analizados.

En la Tabla 1 presentamos los resultados de esta búsqueda de frecuencias para cada grupo de estudiantes, clasificados según su L1 y nivel de dominio del español. La tabla presenta (i) el número de estudiantes por grupo, (ii) el número total de casos por grupo, es decir, de secuencias de sustantivo femenino precedido de artículo, femenino o masculino, (iii) el número de casos con discordancia de género, es decir, de usos de un sustantivo femenino precedido de artículo masculino y (iv) el porcentaje de casos con discordancia. En la tabla se puede observar que hay una gran variación entre los grupos, no solo en el número de casos analizados, sino también en el número de participantes. Es por ello que en nuestros análisis cuantitativos comparamos grupos de estudiantes definidos por nivel y L1, atendiendo a los porcentajes de discordancia.

Tabla 1. Usos de la secuencia artículo y sustantivo de género femenino, por L1 y nivel.

		A1	A2	B1	B2	C1
Árabe	Núm. de estudiantes	202	127	90	52	30
	Núm. total de casos	1388	1811	905	781	743
	Núm. de casos con discordancia de género	142	188	65	30	11
	% de casos con discordancia de género	10%	10%	7%	4%	1%
Chino mandarín	Núm. de estudiantes	67	51	80	30	2
	Núm. total de casos	462	778	1429	706	11
	Núm. de casos con discordancia de género	51	49	76	35	2
	% de casos con discordancia de género	11%	6%	5%	5%	18%
Francés	Núm. de estudiantes	59	49	72	38	22
	Núm. total de casos	401	638	654	881	539
	Núm. de casos con discordancia de género	24	32	13	21	10
	% de casos con discordancia de género	6%	5%	2%	2%	2%
Inglés	Núm. de estudiantes	41	126	52	55	26
	Núm. total de casos	312	1786	650	1448	783
	Núm. de casos con discordancia de género	58	229	61	127	29
	% de casos con discordancia de género	19%	13%	9%	9%	4%
Portugués	Núm. de estudiantes	167	95	43	56	27
	Núm. total de casos	1077	1640	753	1298	899
	Núm. de casos con discordancia de género	11	17	5	6	2
	% de casos con discordancia de género	1%	1%	1%	>1%	>1%

La Tabla 1 confirma, como predecíamos, que los estudiantes de los niveles más bajos son los que tienen mayores dificultades en la asignación de género a los sustantivos femeninos. El porcentaje de discordancia desciende o se mantiene al subir de nivel para todas las L1s, con una única excepción, el grupo de L1 chino mandarín y nivel C1 que tiene un porcentaje excepcionalmente alto, 18%. Dado que en este grupo solo hay dos participantes, sus resultados no pueden considerarse representativos y serán ignorados en nuestros análisis cuantitativos. El análisis estadístico, comparando grupos de estudiantes de distintos niveles que comparten la misma L1, confirma que estas diferencias son significativas para todas las L1 excepto el portugués: L1 árabe ($X=87.459$, $df=4$, $p < 0.0001$), L1 chino mandarín ($X=22.209$, $df=3$, $p < 0.0001$), L1 francés ($X=24.881$, $df=4$, $p < 0.0001$) y L1 inglés ($X=77.648$, $df=4$, $p < 0.0001$). En el caso de la L1 portugués el porcentaje de agramaticalidad es extremadamente bajo en todos los niveles, no superando nunca el 1%, lo que explica la falta de diferencias significativas ($X=7.981$, $df=4$, $p=0.09$).

Si pasamos a considerar la L1 de los participantes, advertimos que son los estudiantes que tienen el inglés como L1 los que producen formas no estándar con mayor frecuencia, con porcentajes de discordancia de hasta un 19%. A los estudiantes de L1 inglés le siguen los de L1 árabe y L1 chino mandarín. Como era de esperar, dada la proximidad entre el portugués y el español, y las similitudes entre los sistemas de género de las distintas lenguas romances, los estudiantes de L1 portugués, seguidos por los estudiantes de L1 francés con tan solo una excepción para el nivel C1, son los que menos dificultades tienen en todos los niveles. Los resultados del test estadístico confirman que las diferencias entre los estudiantes de las distintas L1 son significativas en todos niveles: A1 ($X=138.355$, $df=4$, $p<0.0001$), A2 ($X=194.359$, $df=4$, $p<0.0001$), B1 ($X=77.832$, $df=4$, $p<0.0001$), B2 ($X=126.281$, $df=4$, $p<0.0001$) y C1 ($X=29.857$, $df=3$, $p<0.0001$). La siguiente gráfica ilustra estas diferencias de nivel y L1.

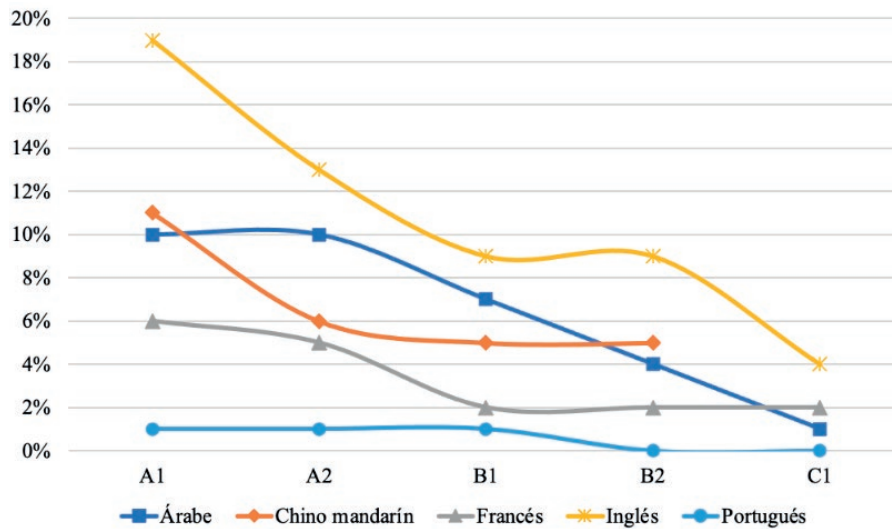


Figura 3. Porcentaje de discordancia de género entre el artículo y el sustantivo femenino por L1 y nivel.

La Tabla 2 recoge los sustantivos femeninos que plantean más dificultades a cada grupo de estudiantes, es decir, los que presentan un mayor porcentaje de discordancia. Esta tabla presenta (i) el número total de usos de cada lema, (ii) el número de casos con discordancia de género entre el artículo y el sustantivo, y (iii) el porcentaje de discordancia. Se recogen solo sustantivos femeninos utilizados con artículo masculino por más de un estudiante, ordenados de mayor a menor según el porcentaje de discordancia. No se incluyen casillas para el nivel C1 porque en este nivel ningún sustantivo femenino aparece acompañado de artículo masculino en la producción de más de un estudiante del mismo grupo.

Tabla 2. Sustantivos femeninos precedidos de artículo masculino por L1 y nivel.

	A1			A2			B1			B2		
	Lema	Núm. casos	Núm. casos discordancia % casos discordancia	Lema	Núm. casos	Núm. casos discordancia % casos discordancia	Lema	Núm. casos	Núm. casos discordancia % casos discordancia	Lema	Núm. casos	Núm. casos discordancia % casos discordancia
Árabe	parte	3	3 100%	cafetería	6	4 67%	atención	4	2 50%	costumbre	5	2 40%
	relación	3	2 67%	excursión	3	2 67%	vacación	4	2 50%	ley	16	6 38%
	madre	5	3 60%	parte	6	3 50%	llave	2	1 50%	consecuencia	6	2 33%
	salud	5	3 60%	nariz	4	2 50%	reclamación	6	2 33%	información	10	2 20%
	reunión	14	7 50%	calle	12	5 42%	historia	12	3 25%	beca	17	2 12%
	vez	4	2 50%	película	10	4 40%	indemnización	37	3 8%			
	piel	7	3 43%	condición	24	9 38%	semana	59	3 5%			
	pizza	5	2 40%	discoteca	14	5 36%						
	nariz	14	5 36%	capital	6	2 33%						
	montaña	6	2 33%	vacación	49	13 27%						
	película	53	11 21%	caja	8	2 25%						
	noche	34	7 21%	fiesta	8	2 25%						
	calle	10	2 20%	piscina	9	2 22%						
	escuela	45	6 13%	televisión	9	2 22%						
	semana	16	2 13%	reservación	17	3 18%						

Tabla 2, continúa en la página siguiente.

Tabla 2, continúa de la página anterior.

	A1			A2			B1			B2						
	Lema	Núm. casos	% casos discordancia	Lema	Núm. casos	% casos discordancia	Lema	Núm. casos	% casos discordancia	Lema	Núm. casos	% casos discordancia				
Chino mandarín	chaqueta	2	2	100%	dirección	3	2	67%	documentación	2	2	100%	clase	6	2	33%
	nariz	7	2	29%	pregunta	3	2	67%	salud	2	2	100%	condición	9	2	22%
	universidad	36	10	28%	información	5	3	60%	característica	4	2	50%	oportunidad	9	2	22%
	escuela	23	4	17%	semana	7	3	43%	imagen	4	2	50%	beca	24	4	17%
	empresa	12	2	17%	foto	6	2	33%	línea	6	2	33%				
	clase	16	2	13%	montaña	9	2	22%	culpa	7	2	29%				
	película	24	3	13%	vacación	23	2	9%	mano	7	2	29%				
					noche	24	2	8%	reunión	11	3	27%				
					habitación	140	6	4%	empresa	22	2	9%				
								maleta	47	2	4%					
Francés	vacación	5	2	40%	vacación	22	3	14%				edad	3	2	67%	
	cena	8	2	25%	habitación	65	8	12%				oportunidad	8	2	25%	
	escuela	22	2	9%								beca	19	2	11%	
Inglés	tarea	3	3	100%	bicicleta	2	2	100%	razón	6	2	33%	fuentes	2	2	100%
	afición	2	2	100%	impresión	2	2	100%	noche	16	5	31%	mente	2	2	100%
	biblioteca	2	2	100%	pared	3	2	67%	calle	10	3	30%	nube	2	2	100%
	montaña	4	3	75%	pasión	3	2	67%	situación	10	2	20%	clase	14	7	50%
	universidad	4	2	50%	región	5	3	60%	universidad	10	2	20%	área (pl.)	9	4	44%
	escuela	11	5	45%	condición	13	7	54%	maleta	21	4	19%	ley	24	9	38%
	película	13	4	31%	discoteca	4	2	50%	mujer	19	2	11%	prohibición	8	3	38%
	oficina	10	3	30%	voz	4	2	50%				calle	11	4	36%	
	hermana	15	4	27%	parte	7	3	43%				investigación	6	2	33%	
	familia	11	3	27%	esposa	5	2	40%				razón	12	3	25%	
	comida	13	2	15%	red	5	2	40%				compañía	8	2	25%	
	ciudad	15	2	13%	vacación	52	16	31%				condición	17	4	24%	
					calle	13	4	31%				salud	76	14	18%	
				madre	20	5	25%				sociedad	25	3	12%		
Portugués				mujer	16	4	25%				universidad	72	5	7%		
				postal	2	2	100%									
				habitación	182	3	2%									

Para una interpretación correcta de estos datos, es necesario tener en cuenta que el léxico empleado por los estudiantes viene condicionado por las actividades que realizaron y que, como hemos explicado anteriormente, varían según el nivel. Así, por ejemplo, los estudiantes de nivel A2 tuvieron que escribir un correo electrónico reservando una habitación, lo cual explica que sustantivos como *habitación*, *noche* o *vacación* aparezcan repetidamente en las casillas de este nivel. El hecho de que, por ejemplo, el sustantivo *habitación* no aparezca en las listas del nivel A1 no significa necesariamente que los estudiantes de este nivel no tengan dificultades con su género, simplemente no lo utilizaron porque las actividades que completaron no requerían su uso.

En español, “la mayor parte de los sustantivos que acaban en -a son femeninos” (RAE, 2009:90), mientras que los terminados en consonante o en -e pueden ser tanto masculinos como femeninos. Una primera aproximación a los datos de la Tabla 2, confirma que los sustantivos en consonante y -e, como *imagen* (ejemplo 1) o *clase* (ejemplo 2), generan dificultades a estudiantes de todos los niveles. Llama, sin embargo, la atención que palabras con la terminación canónica en -a, como *película* (ejemplo 3) o *beca* (ejemplo 2), planteen dificultades incluso a los estudiantes del nivel B2.

1. En la maleta hay un **imagen** de avión azul (B1, L1 chino mandarín)
2. Nunca me siento aburrido en ni un **clase** de universidad (B2, L1 inglés)
3. Después yo veo un **película** con mi amiga en cine. (A1, L1 árabe)
4. Quiero saber si el precio de la universidad es muy caro y si puedo tener un **beca** para ayudar me financieramente (B2, L1 francés)

Algo similar ocurre con los sustantivos terminados en *-ción*, *-sión* o *-zón*, en *-dad* y *-tad*, y en *-ez* que, salvo contadas excepciones, son femeninos, una tendencia que libros de texto y gramáticas pedagógicas enfatizan desde los niveles iniciales (ej., Matte Bon, 1992; Alonso Raya et al., 2021). Vemos, sin embargo, que estudiantes de todos los niveles utilizan sustantivos en *-ción* precedidos de artículo masculino, por ejemplo, *información* (ejemplo 5). También son frecuentes las faltas de concordancia con sustantivos femeninos terminados en *-dad*, como *universidad* (ejemplo 6).

5. También me interesa mucho los informaciones sobre la vida estudiantil (B2, L1 árabe)
6. Mi hermanita se graduó hace un año, pero no ha atendido un universidad. (B1, L1 inglés)

Si cabe todavía más llamativos son los casos de *mujer*, *hermana* y *madre*, sustantivos que se refieren a seres del género femenino y que aparecen con artículo masculino en las producciones de los estudiantes de L1 inglés y árabe. El ejemplo 7 resulta bastante curioso ya que se hace la concordancia entre el sustantivo y el adjetivo, pero no entre el sustantivo y el artículo indeterminado.

7. Físicamente, mi madre es un mujer bonita. (A2, L1 inglés)

Tanto en el caso de *mujer* como en el de *hermana* y *madre* la asociación semántica está ligada a personas de género femenino. Precisamente el hecho de que sean nombres animados pudiera tener algún efecto en el aprendizaje ya que en ocasiones las reglas de uso se aplican a este tipo de sustantivos exclusivamente, en los que existe una correlación directa entre el género gramatical y el género natural o forma del referente. Este fenómeno ha sido tratado extensamente en la bibliografía sobre el tema (Murillo, 1999; Corbett y Fraser, 2000; Ruiz Martín, 2004; Thornton, 2009; Clegg, 2011). Todo ello indica que posiblemente los estudiantes de español no se rigen únicamente por criterios formales ni tampoco de carácter semántico o discursivo, sino que puedan prevalecer en ellos otros factores internos o psicológicos relativos al propio aprendizaje.

Pasamos ahora a considerar los usos de sustantivos masculinos precedidos de artículo. Identificamos un total de 22.942 casos, un número ligeramente superior al de los sustantivos femeninos, 22.773. Sin embargo, solo 635 de estos casos, un 2.77% del total, muestran discordancia de género entre el artículo y el sustantivo. Este número es claramente inferior al observado para los sustantivos femeninos, 1.294 casos, un 5.68% del total. Esta diferencia entre los sustantivos masculinos y femeninos es estadísticamente significativa ($X = 220.741$, $df = 1$, $p < .0001$).

La Tabla 3 presenta los resultados de la búsqueda de frecuencias de uso de sustantivos masculinos precedidos de artículo. Tal y como hicimos con los sustantivos femeninos, calculamos el porcentaje de casos con discordancia de género entre el artículo y el sustantivo para cada grupo y utilizamos estos porcentajes para nuestros análisis comparativos.

Tabla 3. Usos de la secuencia artículo y sustantivo de género masculino, por L1 y nivel.

		A1	A2	B1	B2	C1
Árabe	Núm. de estudiantes	202	127	90	52	30
	Núm. total de casos	1274	1538	1115	956	701
	Núm. de casos con discordancia de género	59	75	17	10	13
	% de casos con discordancia de género	5%	5%	2%	1%	2%
Chino mandarín	Núm. de estudiantes	67	51	80	30	2
	Núm. total de casos	458	699	1712	755	16
	Núm. de casos con discordancia de género	21	33	59	20	2
	% de casos con discordancia de género	5%	5%	3%	3%	13%
Francés	Núm. de estudiantes	59	49	72	38	22
	Núm. total de casos	471	563	811	977	619
	Núm. de casos con discordancia de género	12	14	15	6	2
	% de casos con discordancia de género	3%	2%	2%	1%	>1%
Inglés	Núm. de estudiantes	41	126	52	55	26
	Núm. total de casos	307	1549	791	1487	787
	Núm. de casos con discordancia de género	27	95	32	61	10
	% de casos con discordancia de género	9%	6%	4%	4%	1%
Portugués	Núm. de estudiantes	167	95	43	56	27
	Núm. total de casos	1038	1195	828	1473	822
	Núm. de casos con discordancia de género	12	18	20	2	0
	% de casos con discordancia de género	1%	2%	2%	>1%	0%

A fin de comparar los resultados de los sustantivos masculinos (Tabla 3) y femeninos (Tabla 1), agrupamos a los estudiantes por nivel y por L1. El análisis estadístico confirma que las diferencias observadas entre estos sustantivos son significativas en todos los niveles: A1 ($X=57.032$, $df=1$, $p<0.0001$), A2 ($X=58.107$, $df=1$, $p<0.0001$), B1 ($X=34.653$, $df=1$, $p<0.0001$), B2 ($X=59.887$, $df=1$, $p<0.0001$) y C1 ($X=9.271$, $df=1$, $p=0.002$); y para todas las L1, con la única excepción del portugués: L1 árabe ($X=116.843$, $df=1$, $p<0.0001$), L1 chino mandarín ($X=24.927$, $df=1$, $p<0.0001$), L1 francés ($X=23.527$, $df=1$, $p<0.0001$) y L1 inglés ($X=111.767$, $df=1$, $p<0.0001$). De nuevo, la falta de diferencias significativas de los estudiantes de L1 portugués ($X=2.014$, $df=1$, $p=0.016$) puede atribuirse al escaso número de discordancias producidas por estos estudiantes.

Pasamos ahora a analizar el impacto de la variable nivel de competencia lingüística. Al igual que con los sustantivos femeninos, cuando consideramos los porcentajes de discordancia entre el artículo y los sustantivos masculinos, nos encontramos con que los estudiantes de los niveles más bajos son los que tienen mayores dificultades. En la Tabla 3 vemos que el porcentaje de discordancia desciende o se mantiene al subir de nivel con solo una excepción, los estudiantes de L1 portugués de nivel A1 y A2. De nuevo, dejamos de lado al grupo de L1 chino mandarín y nivel C1 por el número excesivamente limitado de participantes. Los análisis estadísticos confirman que las diferencias son significativas para los estudiantes de L1 árabe ($X=52.093$, $df=4$, $p<0.0001$), L1 francés ($X=19.722$, $df=4$, $p=0.0006$), L1 inglés ($X=42.111$, $df=4$, $p<0.0001$) y L1 portugués ($X=40.648$, $df=4$, $p<0.0001$), pero no para los de L1 chino mandarín ($X=5.738$, $df=3$, $p=0.13$).

Si analizamos el efecto de la L1 de los participantes, comprobamos de nuevo que son los estudiantes que tienen el inglés como L1 los que producen un mayor número de formas no estándar. Aunque en el nivel C1 el porcentaje de discordancia de estos estudiantes, 1%, parece superado por el de los estudiantes de L1 árabe, 2%, el test estadístico revela que esta diferencia no es significativa ($X=0.805$, $df=1$, $p=0.37$). De nuevo, a los estudiantes de L1 inglés le siguen los que tienen chino mandarín y árabe como lenguas iniciales. Los estudiantes de L1 portugués, seguidos por los estudiantes de L1 francés, son los que menos dificultades tienen. Igual que en el caso de los sustantivos femeninos, las diferencias entre los estudiantes de las distintas L1 son significativas en todos niveles: A1 ($X=47.174$, $df=4$, $p<0.0001$), A2 ($X=41.873$, $df=4$, $p<0.0001$), B1 ($X=17.297$, $df=4$, $p=0.002$), B2 ($X=83.677$, $df=4$, $p<0.0001$) y C1 ($X=19.052$, $df=3$, $p=0.0003$). La Figura 4 ilustra estas diferencias.

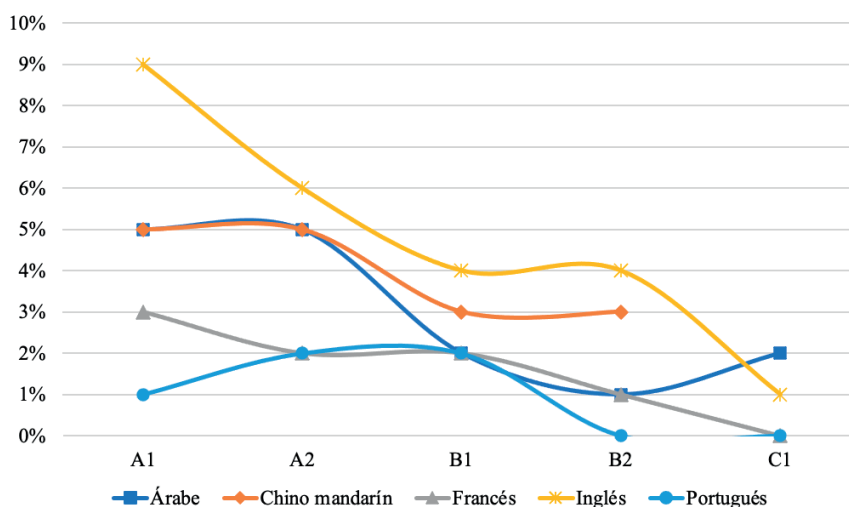


Figura 4. Porcentaje de discordancia de género entre el artículo y el sustantivo masculino, por L1 y nivel.

Se puede concluir que, en ambos casos, es decir, en las concordancias del artículo tanto con nombres masculinos como femeninos, son los alumnos que tienen como lenguas iniciales el inglés, árabe y chino mandarín los que parecen tener mayores dificultades, mientras que aquellos que tienen como L1 portugués o francés producen porcentajes sensiblemente inferiores de formas no estándar. Este resultado puede explicarse por el grado de afinidad o congruencia en relación con la marca de género de estas dos últimas lenguas con el español, al pertenecer todas ellas a la familia lingüística de las lenguas romances. Esto mismo se evidencia en estudios de procesamiento, cuando se habla de la teoría de congruencia, tal como se explicaba más arriba (Acuña et al. 2014; Sá Leite et al. 2020). Por el contrario, en chino mandarín no existen artículos y tampoco es necesario indicar si una palabra tiene género masculino o femenino. Algo similar ocurre en inglés, donde los artículos no tienen marca de género ni número. Estos estudiantes se ven por lo tanto obligados a aprender un concepto totalmente nuevo. El árabe tiene un sistema de género binario, masculino y femenino, pero las normas de asignación de género difieren considerablemente de las del español.

En la Tabla 4 podemos ver la lista de sustantivos masculinos que crean mayores dificultades a cada grupo de estudiantes, ordenados por porcentaje de discordancia, de mayor a menor. De nuevo, se incluyen solo los sustantivos que causaron dificultades a más de un estudiante del mismo grupo.

Tabla 4. Sustantivos masculinos precedidos de artículo femenino por L1 y nivel.

	A1			A2			B1			B2		
	Lema	Núm. casos	Núm. casos discordancia % casos discordancia	Lema	Núm. casos	Núm. casos discordancia % casos discordancia	Lema	Núm. casos	Núm. casos discordancia % casos discordancia	Lema	Núm. casos	Núm. casos discordancia % casos discordancia
Árabe	problema	11	7 64%	idioma	2	2 100%	avión	22	4 18%	idioma	10	2 20%
	coche	6	3 50%	sistema	2	2 100%				tema	17	3 18%
	diploma	4	2 50%	tema	2	2 100%						
	país	9	2 22%	clima	5	4 80%						
	deporte	27	5 19%	problema	18	7 39%						
	fin	29	5 17%	deporte	8	2 25%						
	café	12	2 17%	programa	8	2 25%						
	día	29	4 14%	museo	17	3 18%						
	restaurante	26	3 12%	restaurante	26	4 15%						
	cine	38	4 11%	país	28	3 11%						
				día	81	8 10%						
				amigo	31	2 6%						
				viaje	34	2 6%						
Chino mandarín	examen	3	2 67%	parque	13	3 23%	corazón	3	2 67%	tema	7	2 29%
	parque	9	2 22%	viaje	13	2 15%	mapa	3	2 67%	programa	58	3 5%
	fin	18	3 17%	mes	17	2 12%	sofá	5	2 40%			
				hotel	19	2 11%	lugar	16	3 19%			
							billete	13	2 15%			
							problema	14	2 14%			
Francés	idioma	2	2 100%	viaje	2	8 25%	valor	6	5 83%			
	día	7	3 43%									
Inglés	fin	8	6 75%	orfanato	3	3 100%	detalle	3	2 67%	tema	12	4 33%
	bar	7	3 43%	paisaje	6	3 50%	problema	14	3 21%	problema	23	7 30%
	día	7	3 43%	programa	6	3 50%	día	32	3 9%	interés	10	2 20%
				gimnasio	5	2 40%				programa	138	15 11%
				problema	13	5 38%				lugar	50	5 10%
				jardín	8	2 25%				país	40	3 8%
				restaurante	26	6 23%						
				mes	24	5 21%						
				pueblo	20	4 20%						
				sur	13	2 15%						
				lugar	31	4 13%						
				norte	18	2 11%						
				hotel	48	4 8%						
Portugués	viaje	11	5 45%	paisaje	7	3 43%	color	4	2 50%			
				viaje	30	8 27%	viaje	25	9 36%			
							equipaje	22	5 23%			

Al igual que con los sustantivos femeninos encontramos que muchos de los nombres en este listado terminan en consonante o en -e, por ejemplo, *país* (ejemplo 8) o *viaje* (ejemplo 9). Tal y como explicamos anteriormente, estas terminaciones se asocian tanto con el género masculino como el femenino. Se trata de sustantivos cuyo género no puede ser deducido ni por su morfología ni por su significado, lo que hace que planteen dificultades a todos los estudiantes, independientemente de su L1.

8. Turkia es **una país** bonita con una cultura diferente (A2, L1 árabe)
9. Este año, estamos planajando **una viaje** para Bordeaux (A2, L1 francés)

Son también frecuentes en esta tabla, como era de esperar, los sustantivos masculinos terminados en -a. Por sobregeneralización de la regla, estudiantes de todos los niveles asignan género femenino a palabras terminadas en -a, como *problema* (ejemplo 10) o *día* (ejemplo 11). En el nivel B2, los estudiantes de L1 árabe y chino mandarín, parecen tener problemas solo con este tipo de sustantivos, y únicamente los estudiantes de L1 inglés muestran todavía en este nivel dificultades con los sustantivos terminados en consonante.

10. Además tenemos **la problema** que unos de los fumadores no ponen sus cigarros en la basura cuando terminen. (B2, L1 inglés)
11. Tu tienes que venir **una día** para descansar conmigo (A1, L1 francés)

La terminación en -o es la terminación canónica de los sustantivos masculinos. Con alguna excepción, como la de *la mano*, o la de sustantivos que resultan del acortamiento de palabras femeninas, como *la foto*, estos sustantivos son siempre masculinos (RAE, 2009:90). A diferencia de lo visto para los sustantivos femeninos canónicos, que a pesar de su terminación en -a causan dificultades, no encontramos apenas casos de sustantivos terminados en -o precedidos de artículo femenino. Los únicos identificados fueron producidos por estudiantes de nivel A2 y L1 árabe e inglés (ver ejemplos 12 y 13).

12. El tercer día fui a **la museo** nacional (A2, L1 árabe)
13. Necesito **una cuarto** triple que tiene un baño muy grande (A2, L1 inglés)

Por último, para los aprendices de L1 portugués los únicos problemas de concordancia registrados son con los sustantivos *viaje*, *paisaje*, *equipaje* y *color*. En español, las palabras en -aje y -or suelen ser masculinas (Matte Bon, 1992, 178). En portugués, sin embargo, las palabras terminadas en -agem son siempre femeninas: *a viagem*, *a paisagem* y *a bagagem*. El sustantivo *color* se corresponde con *cor*, también femenino.

14. **La viaje** fue muy buena. (A1, L1 portugués)
15. En Viena me deslumbré con **las paisajes** y con sus tradiciones. (A2, L1 portugués)

Estos resultados sugieren que muchos de los errores de asignación de género cometidos por los participantes pueden ser interlingüísticos. Es decir, la lengua de partida del estudiante condiciona su comportamiento en el uso de la lengua meta. Si bien no podemos olvidar que en otros casos los errores no se pueden explicar en relación con factores interlingüísticos sino intralingüísticos. Las dificultades de asignación de género a sustantivos masculinos terminados en -a como *problema* o *día*, que se observan en todos los niveles y todas las L1, parecen estar condicionadas por la propia naturaleza del proceso de aprendizaje lingüístico. Trazar una línea divisoria entre unos y otros no es siempre fácil.

4. DISCUSIÓN E IMPLICACIONES DIDÁCTICAS

En esta sección presentamos la discusión de los resultados y sus implicaciones didácticas. Ofrecemos tres ejemplos de actividades diseñadas con los datos del CAES, que pretenden ilustrar el uso práctico que los profesores de español pueden hacer de este corpus.

Los resultados de nuestros análisis confirman que la asignación de género, es decir, la concordancia de género entre el artículo y el sustantivo, plantea dificultades a los estudiantes de todos los niveles y todas las L1, aun cuando la frecuencia y naturaleza de estos problemas sea diferente en función de estas variables. En las clases de ELE, el género de los sustantivos se trabaja de manera explícita en los niveles iniciales, sin embargo, raramente aparece incluido como tema de estudio en los manuales de nivel intermedio o avanzado. Cuando aparece, la atención suele centrarse en los sustantivos no canónicos como, por ejemplo, los masculinos de origen griego terminados en -ma. Las muestras del corpus revelan sin embargo que estos no son los únicos usos problemáticos a nivel intermedio e incluso avanzado, y que no deben por lo tanto

ignorarse otros usos, como el de los sustantivos femeninos terminados en *-ción* o *-dad* o el de los masculinos en *-aje*, más transparentes y aparentemente fáciles de adquirir pero que claramente siguen planteando dificultades.

El análisis del CAES revela asimismo que los aprendices de ELE, independientemente de su L1 y nivel, tienen más dificultades a la hora de asignar el género a los sustantivos femeninos que a los masculinos. Estos resultados coinciden con los de estudios previos de carácter experimental y apoyan la idea de que los aprendices asignan el género masculino por defecto (ej., Franceschina, 2005; Montrul et al., 2008; Klassen, 2016; Alarcón, 2011). Esta tendencia explicaría el uso del artículo masculino con sustantivos fácilmente identificables como femeninos, bien por su valor semántico, como es el caso de *madre* y *mujer*, o por su morfología, como ocurre con los nombres terminados en *-a*. Esto implica que, en el aula, los profesores necesitan prestar más atención al género de los sustantivos femeninos, para así compensar la tendencia al abuso del artículo masculino.

Tal y como mencionamos en la sección 3.3, los resultados de nuestro análisis sugieren que muchos de los errores de concordancia cometidos por los aprendices están condicionados por su L1 y que por lo tanto estudiantes con distintas lenguas iniciales necesitan distinta atención pedagógica. El corpus confirma que los estudiantes que tienen una lengua inicial romance, francés o portugués, tienen escasas dificultades con el género de los sustantivos. En el caso del portugués, estas se centran en sustantivos, como *color* o *viaje*, que tienen distinto género en las dos lenguas. El profesorado necesita, en consecuencia, incidir en estos sustantivos en el aula, valiéndose para ello de los análisis contrastivos existentes entre la L1 del estudiante y la lengua meta. En el caso de los estudiantes de L1 inglés, árabe o chino mandarín, los problemas son no solo más frecuentes sino también menos predecibles. Los hablantes de estas lenguas iniciales necesitan que se siga haciendo hincapié a lo largo de todos los niveles en ciertas reglas y tendencias introducidas en el nivel inicial como, por ejemplo, que los sustantivos terminados en *-a* suelen ser femeninos.

En el caso de las clases mixtas, que incorporan estudiantes de distintas lenguas iniciales, estas diferencias plantean un reto al profesorado. Nosotros proponemos, para este tipo de clase, el uso de actividades colaborativas en las que los estudiantes trabajan con compañeros que no comparten su L1. Este tipo de emparejamiento favorece el aprendizaje recíproco y cooperativo (Cassany, 2004; López Rúa, 2015; Ramos y Pavón, 2015). Así, por ejemplo, los estudiantes de L1 portugués pueden ayudar a los de L1 inglés con las palabras en *-dad* o *-ción*, que tienden a corresponderse con palabras en *-dade* o *-ção*, terminaciones también asociadas con el género femenino en la lengua portuguesa. A su vez, los estudiantes de L1 inglés pueden contribuir con el género de sustantivos masculinos en *-aje*, que por tener distinto género en español y portugués son problemáticos para los estudiantes de L1 portugués, pero no necesariamente para los de inglés.

Siguiendo esta línea, presentamos tres ejemplos de uso de los recursos y datos proporcionados por el CAES para crear actividades que pueden implementarse tanto de modo individual como colaborativo. Sin entrar, por razones de espacio, en consideraciones generales en torno al tratamiento de la gramática en la clase de ELE (ver Herrera y Sans, 2018), es necesario destacar que su planteamiento descansa en aproximaciones de enseñanza centradas en el aprendiz (Nunan, 1988), combinando perspectivas implícitas y explícitas (Ellis, 2005), y buscando promover la conciencia gramatical del alumnado (Rutherford, 1987).

El primer ejemplo presenta una actividad de huecos, centrada en sustantivos que, por los datos del corpus, sabemos plantean dificultades de concordancia a los estudiantes de nivel A2. El texto ha sido tomado de una muestra de este nivel extraída del CAES, adaptando la ortografía y gramática para eliminar cualquier error o forma agramatical. Como el texto ha sido parcialmente producido por un aprendiz de su mismo nivel, sabemos que el estudiante está trabajando dentro de su Zona de Desarrollo Próximo (Vygotsky, 1978), y que no plantea dificultades morfosintácticas o léxicas añadidas que puedan desviar su atención del aspecto gramatical que estamos trabajando. La actividad puede adaptarse a estudiantes de otros niveles utilizando, como punto de partida, textos del CAES de su nivel.

Actividad 1

En este texto faltan los artículos. Prestando atención al género y número de los sustantivos, añade los artículos que faltan.

Hola,

Me llamo EC y le escribo para preguntar sobre ____ habitación y cómo puedo hacer ____ reserva

Busco ____ habitación sencilla con ____ cuarto de baño independiente, aire acondicionado y, si es posible, vistas al mar.

Mi intención es tener ____ habitación desde el 3 de diciembre hasta ____ día 8, ya que estaré en ____ ciudad para ____ reunión de trabajo.

La actividad 2 es una actividad inductiva de concienciación gramatical. De nuevo se utiliza un texto extraído del CAES y, por lo tanto, producido por un aprendiz, en este caso también de nivel A2. Los pasos de la actividad sirven para guiar la atención del alumno y centrarla en el género de sustantivos como *habitación* o *reserva*,

problemáticos para los estudiantes de este nivel (ver Tabla 2). El objetivo es que el estudiante, trabajando de manera individual o colaborativa, haga un análisis lingüístico que le permita, de manera guiada, inducir ciertas reglas básicas de la asignación de género en español.

Actividad 2

Paso 1. Lee atentamente el siguiente correo electrónico. Subraya todos los nombres del texto y, atendiendo al artículo o determinante que los acompaña, clasifícalos como masculinos o femeninos.

Buenas tardes,

Soy M R. y me gustaría hacer una reserva. Por este motivo le escribo un correo electrónico. Me gustaría reservar una habitación triple con baño y aire acondicionado. No necesito un minibar tampoco una caja fuerte, pero estaría muy bien disponer de vistas al mar. Solicito, si es posible, un cuarto en el primer piso ya que me alojaré con una persona de edad. ¿Puede informarme si tiene disponibilidad para el 20 de julio de este año? ¿Cuáles son los precios y las condiciones de la reserva?

Si el 20 de julio no hay disponibilidad, podría ser el 29 de este mismo mes. ¿Podría hacer la reserva por correo electrónico después de que me confirmen o tengo que llamarles por teléfono? Indíqueme, por favor, el teléfono del hotel.

Muchas gracias por su comprensión y atención.

Nombres masculinos
(un) correo

Nombres femeninos
(una) reserva

Paso 2. Utilizando la tabla que has creado en el paso 1, completa las siguientes afirmaciones:

1. Los nombres que acaban en -o son generalmente.....
2. Los nombres que acaban en -a son generalmente.....
3. Los nombres que acaban en vocal distinta de -o/-a pueden ser o.....
4. Los nombres que acaban en -dad suelen ser.....
5. Los nombres de los meses en español son.....

Por último, presentamos una actividad de producción libre que puede completarse individualmente o en parejas, para fomentar el aprendizaje colaborativo. Los estudiantes tienen que escribir una solicitud de admisión para estudiar en el extranjero, incluyendo una serie de sustantivos relacionados con el tema. Para crear esta lista buscamos en el CAES todos los casos de discordancia de género producidos por estudiantes de nivel B2 al completar esta misma actividad. Seleccionamos los que aparecen con mayor frecuencia y porcentaje de discordancia. La actividad está por lo tanto concebida para una clase mixta de nivel B2, pero puede ajustarse a una clase en la que todos los estudiantes comparten L1 ajustando la búsqueda del CAES para filtrar los resultados por lengua.

Actividad 3

Escribe una solicitud formal de admisión para estudiar en el extranjero este verano. Puedes elegir destino, universidad, área de estudios y duración de tu estancia, pero en tu solicitud asegúrate de incluir los sustantivos siguientes:

admisión	beca	clase	condición
información	oportunidad	programa	universidad

5. REFLEXIONES FINALES

Los corpus de aprendices pueden ser un instrumento de investigación de gran valor. Su explotación no requiere una gran especialización previa y nos puede proporcionar información de gran interés sobre nuestros estudiantes, qué problemas tienen, en qué contextos concretos y qué factores pueden incidir en ellos. Esto nos permitirá planificar nuestra enseñanza en función de las deficiencias identificadas. Además, los corpus de aprendices nos pueden ayudar a la hora de diseñar actividades, ofreciéndonos ejemplos auténticos de uso de la lengua adaptados al nivel de nuestros estudiantes. Las propuestas pedagógicas aquí presentadas van precisamente en esta dirección.

Este trabajo debe considerarse preliminar, restan por estudiar casos de discordancia entre el nombre y el adjetivo, o entre el nombre y los determinantes posesivos y demostrativos. Las gramáticas pedagógicas, que inciden de manera especial en aquellas áreas que revisten mayor dificultad para nuestros alumnos, pueden ser un buen punto de partida para investigar en qué medida estas tienen un reflejo en el uso real del idioma recogido en corpus de aprendices como el CAES.

REFERENCIAS

- Acuña, C., Meseguer, E. y Carreiras, M. (2014). "Gender and number agreement comprehension in Spanish", *Lingua* 143, 108-128. <https://doi.org/10.1016/j.lingua.2014.01.013>
- Alarcón, I.V. (2011). "Spanish gender agreement under complete and incomplete acquisition: Early and late bilinguals' linguistic behavior within the noun phrase", *Bilingualism: Language and Cognition* 14/3, 332-350. <https://doi.org/10.1017/S1366728910000222>
- Alarcos Llorach, E. (1994). *Gramática de la Lengua Española*. Madrid: Espasa Calpe.
- Alcina Franch, J. y Blecaua J.M. (2001). *Gramática española*. Barcelona: Ariel.
- Alonso Arijá, E., Alonso Arijá, E. y Ortiz Pérez, S. (2019). *Experiencias. Internacional*. Madrid: Edelsa.
- Alonso Raya, R., Castañeda Castro, A., Martínez Gila, P., Miquel López, L., Ortega Olivares, J. y Ruiz Campillo, J.P. (2021). *Gramática básica del estudiante de español* (nueva edición). Barcelona: Difusión.
- Barber, H. y Carreiras, M. (2005). "Grammatical gender and number agreement in Spanish: An ERP comparison", *Journal of Cognitive Neuroscience* 17/1, 137-153. <https://doi.org/10.1162/0898929052880101>
- Bosque, I. y Demonte, V. (1999). *Gramática descriptiva de la lengua española*. Madrid: Espasa.
- Bull, W. (1965). *Spanish for Teachers: Applied Linguistics*. Nueva York: Ronald Print.
- Calero Fernández, M.A. (1999). *Sexismo lingüístico: análisis y propuestas ante la discriminación sexual en el lenguaje*. Madrid: Narcea.
- Cassany, D. (2004). "Aprendizaje cooperativo para ELE", en C. Pastor (coord.) *Actas del programa de formación para profesorado de español como lengua extranjera 2003-2004*. Munich; Instituto Cervantes de Munich, 11-30.
- Castro Viúdez, F., Rodero Díez, I. Sardinero Francos, C. y Díaz Ballesteros, P. (2016). *Nuevo Español en Marcha*. Madrid: SGEL
- Clegg, J.H. (2011). "A Frequency-based analysis of the norms for Spanish noun gender", *Hispania* 94/2: 303-319. <https://doi.org/10.1353/hpn.2011.a441843>
- Consejo de Europa. (2001). *Marco Común Europeo de Referencia para las lenguas. Aprendizaje, enseñanza, evaluación*. Estrasburgo: Servicio de Publicaciones del Consejo de Europa.
- Consejo de Europa. (2020). *Marco Común Europeo de Referencia para las lenguas. Aprendizaje, enseñanza, evaluación*. Volumen complementario. Estrasburgo: Servicio de Publicaciones del Consejo de Europa.
- Corbett, G.G. y Fraser, N.M. (2000). "Gender assignment: a typology and a model", en G. Senft (ed.) *Systems of Nominal Classification*. Cambridge: Cambridge University Press, 293-325.
- Corpas, J., García, E., Garmendia, A. y Sans, N. (2020). *Aula Plus. Curso de Español*. Barcelona: Difusión.
- Ellis, N.C. (2005). "At the interface: dynamic interactions of explicit and implicit knowledge", *Studies in Second Language Acquisition* 27/2, 305-352. <https://doi.org/10.1017/S027226310505014X>
- Franceschina, F. (2001). "Against an L2 morphological deficit as an explanation for the differences between native and non-native grammars." *EUROSLA Yearbook* 1/1, 143-158. <https://doi.org/10.1075/eurosla.1.12fra>
- Franceschina, F. (2005). *Fossilized Second language grammars: The acquisition of grammatical gender*. Ámsterdam: John Benjamins. <https://doi.org/10.1075/lald.38>
- Fuchs, Z. (2019). *Gender in the nominal domain: Evidence from bilingualism and eye-tracking*. [Tesis doctoral, Harvard University, Graduate School of Arts & Sciences]. Digital Access to Scholarship at Harvard. <http://nrs.harvard.edu/urn-3:HUL.InstRepos:42029638>
- Garnacho López, P. y Martín Acosta, L. (2014). *Diccionario de dudas del estudiante de español como lengua extranjera*. Madrid: SGEL.
- Gómez Torrego, L. (2007). *Gramática didáctica del español*. Madrid: Grupo SM.
- Grijelmo, A. (2019). *Propuesta de acuerdo sobre el lenguaje inclusivo*. Barcelona: Taurus.
- Instituto Cervantes. (2006). *Plan Curricular del Instituto Cervantes: Niveles de Referencia para el español*. Madrid: Instituto Cervantes. https://cvc.cervantes.es/ensenanza/biblioteca_ele/plan_curricular/default.htm
- Herrera, F. y Sans, N. (2018). *Enseñar gramática en el aula de español. Nuevas perspectivas y propuestas*. Barcelona: Difusión.
- Klassen, R. (2016). "When masculine as default supercedes L1 transfer", en A. Alba de la Fuente, E. Valenzuela, C. Martínez Sanz (eds.) *Language acquisition beyond parameters: Studies in honour of Juana M. Liceras*. Ámsterdam: John Benjamins, 73-98. <https://doi.org/10.1075/sibil.51.04kla>
- Lew-Williams, C. y Fernald, A. (2010). "Real-time processing of gender-marked articles by native and non-native Spanish speakers", *Journal of Memory and Language* 63, 447-464. <https://doi.org/10.1016/j.jml.2010.07.003>
- López Rúa, P. (2015). "Posibilidades y limitaciones del aprendizaje cooperativo en el EEES: análisis de una experiencia en el aula de inglés", *Porta Linguarum* 24, 163-177. <https://doi.org/10.30827/Digibug.53868>

- Mellado Jurado, R. y Klein Fariz, I. (2022). *Diagramática. Curso de gramática visual*. Madrid: Edelsa.
- Martínez-Gibson, E.A. (2011). "A comparative study on gender agreement errors in the spoken Spanish of heritage speakers and second language learners", *Porta Linguarum* 15, 177-193. <https://doi.org/10.30827/Digibug.31991>
- Matte Bon, F. (1992). *Gramática comunicativa del español* (Nueva edición revisada). Madrid: Edelsa.
- Montrul, S., Foote, R. y Perpiñán, S. (2008). "Gender Agreement in Adult Second Language Learners and Spanish Heritage Speakers: The Effects of Age and Context of Acquisition", *Language Learning* 58/3, 503-553. <https://doi.org/10.1111/j.1467-9922.2008.00449.x>
- Murillo, J. (1999). "La marcación del género en los sustantivos del español: entre la flexión y la derivación", *Revista de Filología y Lingüística de la Universidad de Costa Rica* 25/1, 181-192. <https://doi.org/10.15517/rfl.v25i1.20546>
- Nunan, D. (1988). *The learner-centred curriculum*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139524506>
- Portal Nieto, A.M. (1999). "ELE: Género gramatical y sexismo lingüístico", en M. Franco Figueroa, C. Soler Cantos, J. de Cos Ruiz, M. Rivas Zancarrón, F. Ruiz Fernández (eds.) *Nuevas perspectivas en la enseñanza del español como lengua extranjera. I. Actas del X Congreso Nacional ASELE*. Cádiz: Universidad de Cádiz y Centro Virtual Cervantes, t. II, 551-557.
- RAE. (2009). *Nueva gramática de la lengua española, morfología sintaxis I*. Madrid: Espasa.
- RAE. (2020). "Informe de la RAE sobre el uso del lenguaje inclusivo en la Constitución Española, elaborado a petición de la Vicepresidenta del Gobierno", *Boletín de información lingüística de la RAE*, [S.I.], 5-207. Disponible en: <http://revistas.rae.es/bilrae/article/view/397/874>
- Ramos Ordóñez, M.C. y Pavón Vázquez, V. (2015). "Developing cooperative learning though tasks in content and language integrated learning." *Multidisciplinary Journal of Educational Research* 5/2, 133-166. <https://doi.org/10.17583/remie.2015.1429>
- Ruiz Martín, M.A. (2004). "¡Houston temenos *una problema! Problemas con la concordancia de género en la clase de E/LE", en M.A. Castillo Carballo, O. Cruz Moya, J.M. García Platero, J.P. Mora Gutiérrez (eds.) *Las gramáticas y los diccionarios del español como segunda lengua: deseo y realidad. Actas del XV Congreso Nacional ASELE*. Sevilla: Universidad de Sevilla y Centro Virtual Cervantes, 777-784.
- Rutherford, W. (1987). *Second language grammar learning and teaching*. Londres: Longman.
- Sá-Leite, A.R., Luna, K., Fraga, I. y Comesaña, M. (2020). "The gender congruency effect across languages in bilinguals: A meta-analysis", *Psychonomic Bulletin & Review* 27, 677-693. <https://doi.org/10.3758/s13423-019-01702-w>
- Schön, D.A. (1991). *The reflective practitioner: how professionals think in action*. Londres: Avebury.
- Scontras, G., Polinsky, M. y Fuchs, Z. (2018). "In support of representational economy: agreement in heritage Spanish", *Glossa: A Journal of General Linguistics* 3/1, 1-29. <https://doi.org/10.5334/gjgl.164>
- Teschner, R.V. y Russell, W.M. (1984). "The Gender Patterns of Spanish Nouns: An Inverse Dictionary-based Analysis", *Hispanic Linguistics* 1, 115-32.
- Thornton, A.M. (2009). "Constraining gender assignment rules", *Language Sciences* 31/1, 14-32. <https://doi.org/10.1016/j.langsci.2007.09.003>
- Vigliocco, G., Butterworth, B. y Garret, M.F. (1996). "Subject verb agreement in Spanish and English. Differences in the role of conceptualisations", *Cognition* 51, 261-298. [https://doi.org/10.1016/S0010-0277\(96\)00713-5](https://doi.org/10.1016/S0010-0277(96)00713-5)
- Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Wang, J. (2016). *El género gramatical en una gramática del español como segunda lengua*. [Tesis doctoral, Universidad Complutense de Madrid]. E-Prints Complutense. Repositorio Institucional de la UCM. <https://eprints.ucm.es/id/eprint/40011/1/T37963.pdf>