Estudio del género gramatical en un corpus de aprendices de español como segunda lengua y lengua de herencia: aproximación al lenguaje inclusivo

Miguel Hidalgo García
miguel.hidalgo@usal.es
Universidad de Salamanca

Resumen: Este trabajo analiza el uso del género gramatical en las producciones escritas de estudiantes matriculados en cursos de español como segunda lengua y lengua de herencia en la Universidad de California Davis. Estas producciones han sido extraídas del *Corpus of Written Spanish of L2 and Heritage Speakers* (COWS-L2H), aunque solo se han seleccionado aquellas en las que los estudiantes se describen a sí mismos. Esto permite analizar, en primer lugar, los errores más comunes en cuanto a la asignación de género gramatical en este tipo de producciones. En segundo lugar, se plantea si existen diferencias en cuanto a este asunto entre las producciones de estudiantes de español como segunda lengua y estudiantes de español como lengua de herencia. Finalmente, se analiza si las distintas identidades de género que manifiestan estos estudiantes afectan también a este asunto, centrándonos en el estudio del lenguaje inclusivo que estos utilizan en sus textos.

Palabras clave: género gramatical; COWS-L2H; lenguaje inclusivo; identidad de género; español como segunda lengua/lengua de herencia.

1. Introducción

"El género es uno de los rasgos gramaticales de análisis más controvertido y polémico en las lenguas naturales", aduce Fábregas (2022, p. 40) en uno de sus últimos trabajos sobre la caracterización sintáctica del género del sustantivo en español. Y es cierto lo que dice. En español, concretamente, el género gramatical no solo posee rasgos

morfológicos, sintácticos e incluso léxicos que lo alejan de otras lenguas como el inglés, que carece de este marcador (Fábregas, 2022, p. 40), sino que en la sociedad actual se está poniendo en duda el férreo binarismo de este fenómeno gramatical, pues entra en conflicto con las nuevas identidades de género que están surgiendo (Giammatteo, 2020, p. 179). Todas estas cuestiones serán tratadas en el presente estudio que, centrado en el análisis del género gramatical en las producciones escritas de estudiantes de español como segunda lengua y lengua de herencia, no dejará de lado su visión aplicada e investigadora.

Con este cometido, se tratará de buscar el origen de los errores más comunes de género gramatical que comenten estos estudiantes de español, como los que registran Lozano y Fernández-Mira sobre la concordancia de sustantivos masculinos terminados en *-a* con el artículo *la*: **la clima*, **la día* o **la idioma* (2022, p. 201). Estos son solo algunos de los múltiples ejemplos que cualquier profesor de español puede encontrarse en el aula. Y es que esta problemática viene documentándose desde hace ya un tiempo, y no solo en niveles iniciales de español, sino en todos los niveles (Pastor Cesteros, 2001, p. 23; Ruiz Martín, 2005, p. 777; Orejudo, 2020, p. 158; entre otros). De ahí nace la principal motivación de este trabajo, además de la colaboración que se ha llevado a cabo entre la Universidad de Salamanca y la Universidad de California Davis en el recién acabado proyecto de corrección y anotación del corpus de aprendices de español COWS-L2H, en el que se han corregido y anotado numerosos errores de concordancia y atribución indebida de género gramatical sobre distintas clases de palabras como determinantes, sustantivos o adjetivos.

2. Estado de la cuestión

2.1. El género gramatical en español

El género gramatical en español hunde sus raíces en el indoeuropeo como recurso morfológico para expresar la distinción entre sexos; por tanto, hacía referencia, en un primer momento, al género natural de los seres animados (González Calvo, 1979, p. 51). Al mismo tiempo, el género gramatical empezó a usarse también para designar a objetos inanimados, de una forma arbitraria e incoherente (González Calvo, 1979, p. 54). La tensión entre género natural y género gramatical sigue, en verdad, patente en la sociedad actual, como se manifestó en la introducción del presente trabajo. Por eso, la Academia[1] ha intentado, desde hace un tiempo, desligar el criterio semántico, es decir, el de género natural, del propio significado de género gramatical (González Calvo, 1979, pp. 53-54).

[1] Entiéndase *Academia* como Real Academia Española (RAE), que también suele trabajar conjuntamente con la Asociación de Academias de la Lengua Española (ASALE).

Consideremos, entonces, la definición que se propone desde el *Diccionario de la lengua española* de la palabra *género* referida a la gramática. La octava acepción dice así: "Categoría gramatical inherente en sustantivos y pronombres, codificada a través de la concordancia en otras clases de palabras y que en pronombres y sustantivos animados puede expresar sexo" (RAE, 2023, definición 8). De esta definición se deducen varias ideas importantes. En primer lugar, que el género gramatical concierne únicamente a categorías nominales. Esta es una idea que no solo señalan las Academias de la lengua española (RAE y ASALE, 2009, p. 81), sino también otros autores como González Calvo (1979, p. 73), para quien "Solo los sustantivos implican género", o Picallo (2015, p. 633), quien añade que el género gramatical es una "clase formal que se atribuye a un sustantivo y es independiente de la entidad que este denota". Esta definición vuelve a poner de relieve la distinción que se hacía con anterioridad entre género natural y género gramatical: es necesario desligar a este último de cualquier denotación semántica, pues puede llevar a error, sobre todo, a los aprendices de español como lengua adicional.[2]

Si tenemos presente la mencionada definición del *Diccionario* académico, en ella se constata otro aspecto fundamental del género gramatical en español: la concordancia. Se trata de un fenómeno sintáctico obligatorio en español que sirve para correlacionar el sustantivo o pronombre con los demás elementos que lo acompañan; en este caso, los determinantes y adjetivos (RAE y ASALE, 2009, p. 81; González Calvo, 1979, p. 72; Picallo, 2015, p. 634). En español existen únicamente dos marcas de género gramatical: el masculino y el femenino, que, aunque en los seres animados pueden expresar sexo, como también recoge la definición académica, en los inanimados es una simple marca formal arbitraria cuyas consecuencias sintácticas se ven reflejadas en la concordancia (RAE y ASALE, 2009, p. 82; González Calvo, 1979, p. 72; Picallo, 2015, p. 634). Y estas consecuencias no son siempre positivas para su empleo por parte de los aprendices de español como lengua adicional, entre los que se registran numerosos errores de concordancia.

En cualquier caso, parece conveniente realizar también una clasificación de los sustantivos atendiendo a las diferentes formas en las que expresan su género gramatical. Una de las clasificaciones más útiles es la que ofrece la *Nueva gramática de la lengua española* (RAE y ASALE, 2009), y que se explicitará a continuación. Por un lado, algunos sustantivos marcan el género añadiendo un morfema a la raíz, como los pares *gato/gata* o *duque/duquesa*. En esta primera clasificación se deduce un criterio puramente morfológico de flexión nominal en la formación de estos pares. El segundo grupo es el que utiliza raíces diferentes para marcar la distinción de género, como los pares *toro/vaca* o *yerno/nuera*: son los *heterónimos*, que poseen un claro matiz léxico en su formación. Un tercer tipo de sustantivos animados son los denominados *comunes en cuanto al género*:

[2] Se usa el término *lengua adicional* para abarcar tanto a los estudiantes de español como *lengua de herencia* como a los estudiantes de español como *segunda lengua*.

no experimentan cambios en su forma, sino que manifiestan su género a partir de los determinantes o adjetivos que los acompañan. Forman parte de este tipo pares de sustantivos como *el artista/la artista* o *este testigo/esta testigo*. Una cuarta división la formarían los sustantivos *ambiguos en cuanto al género*, normalmente inanimados y que pueden aparecer en masculino o femenino designando en ambos casos la misma entidad. Ejemplos de este tipo de nombres serían los pares *el mar/la mar*, o *el tizne/la tizne*. Esta variedad atiende, generalmente, a factores dialectales. Y, finalmente, se registra otro tipo de sustantivos llamados *epicenos*, que se refieren a personas o animales mediante un único género gramatical, sea este masculino, como *el personaje*, o femenino, como *la lechuza*. Para especificar el sexo de estos sustantivos epicenos se suelen emplear los términos *macho/hembra* en el caso de los animales, como *la lechuza macho/la lechuza hembra*, y *masculino/femenino* para las personas, como *el personaje masculino/ el personaje femenino* (RAE y ASALE, 2009, pp. 83-84).

2.2. Género gramatical y lengua nativa

Como se ha podido deducir de lo expuesto anteriormente, el género gramatical "no es un universal lingüístico" (González Calvo, 1979, p. 57). Existe, en realidad, mucha variedad en torno al concepto de género gramatical en las diferentes lenguas del mundo. A este respecto, Scotto y Pérez (2020, pp. 17-18) establecen tres grandes grupos. Por un lado, se encuentra la mayor parte de las lenguas indoeuropeas, como el español, el francés o el italiano, que, como ya se ha mencionado, poseen carga de género. Por otro lado, están las lenguas que, como el alemán, el ruso o el polaco, distinguen tres géneros: el masculino, el femenino y el neutro. Y, por último, existen lenguas como el inglés, el japonés o el mandarín en las que ningún sustantivo inanimado posee género gramatical (Scotto y Pérez, 2020, pp. 17-18).

A esta gran diversidad hay que añadir otro factor que ya se comentó en párrafos anteriores: la arbitrariedad en la asignación[3] de género a los sustantivos inanimados, que también cambia de una lengua a otra. Lamas (1996, p. 2) documenta, por ejemplo, el caso de los sustantivos *sol* y *luna* en alemán, que son femenino (*die Sonne*) y masculino (*der Mond*) respectivamente, justo al contrario de los que ocurre en español. Incluso la autora se pone en la mente de un anglohablante que aprenda español, y se pregunta si no le resultará extraño tener que asignar el género femenino al sustantivo *silla*, o el género masculino al nombre *espejo* (Lamas, 1996, p. 2). En inglés, por ejemplo, solo poseen marca de género algunos pronombres de tercera persona, como los pares *he/she* o *him/her*, así como algunos sustantivos animados que marcan léxicamente estas diferencias de género, como *king/queen* o *bull/cow* (Scotto y Pérez, 2020, p. 23; Orejudo, 2020, p. 157).

[3] Se utilizan los términos *asignación* y *atribución* de género gramatical como sinónimos.

Estos aspectos no hacen más que dificultar el aprendizaje del género gramatical en español a aquellos estudiantes cuyos idiomas nativos carecen de esta marca lingüística. De hecho, se argumenta que el origen de estos problemas podría tener su explicación en la propia adquisición de este rasgo gramatical. Mientras que los hablantes nativos de español ya dominan el sistema de asignación y concordancia de género a la edad aproximada de cuatro años, muchos de los hablantes de español como lengua adicional poseen problemas de asignación y concordancia género gramatical incluso en niveles avanzados de competencia lingüística (Montrul et al., 2012, p. 88). A partir de los cinco años, los mecanismos de asignación y concordancia de género gramatical se atrofian o fosilizan, por lo que resulta muy complicado que un estudiante de español cuya lengua nativa carece de un sistema de género sea íntegramente competente en estos procesos después de la pubertad (Montrul et al., 2012, p. 89).

En el ámbito del español como segunda lengua, se han realizado estudios tanto a grupos de aprendices de diversa lengua materna como de una sola lengua nativa.[4] Dentro de los estudios de hablantes de diversa lengua materna destaca el de Fernández López (1997), en el que se señalan algunas conclusiones interesantes en cuanto a los errores de género gramatical: las discordancias ocurren, principalmente, cuando el sustantivo y el elemento con el que se concuerda no se sitúan contiguamente; y, además, el elemento discordado suele aparecer generalmente en masculino, al tratarse del género no marcado en español (Pastor Cesteros, 2001, pp. 11-12). Por otra parte, los estudios realizados con aprendices de español que comparten una misma lengua nativa se han dirigido principalmente a anglohablantes, como el de Ruiz Martín (2005, p. 777), que señala los sustantivos masculinos acabados en *-ma* y los femeninos acabados en *-o*, *-ión* y *-e* como los más problemáticos en cuanto a la correcta asignación y concordancia de género gramatical.

Sin embargo, también cabe aludir a otra serie de estudios realizados con otro tipo de hablantes de español que apenas habían sido considerados hasta hace relativamente poco tiempo (Acosta Corte, 2013, p. 2). Se trata de los hablantes de español como lengua de herencia, para quienes la bibliografía sobre esta cuestión señala la definición general de Valdés (2001, p. 38) como la más aceptada, teniendo presente su contexto estadounidense: un hablante de herencia es una persona que "se ha criado en un hogar donde se habla una lengua distinta al inglés, que habla o al menos comprende el idioma y que es, hasta cierto punto, bilingüe en esa lengua y en inglés" (traducción de Alonso-Marks, 2015, p. 527). Aunque hay hablantes de herencia que residen en países donde el inglés no es la lengua predominante, interesa especialmente en este trabajo el caso de los hablantes de español como lengua de herencia en Estados Unidos, para los que Alonso-Marks (2015, p. 534) expone una descripción bastante exhaustiva y realista:

[4] Se utiliza el término *español como segunda lengua* para abarcar a aquellos estudiantes que poseen una lengua nativa distinta del español. Los términos *lengua materna* y *lengua nativa* se utilizan como sinónimos para hacer referencia a la primera lengua adquirida por los estudiantes.

> Dentro del ámbito norteamericano, los hablantes de herencia se caracterizan como individuos que se han escolarizado en la lengua mayoritaria de su comunidad, el inglés en este caso. Han hablado español especialmente durante la primera infancia con la familia y el cambio del español al inglés ha ocurrido de forma rápida, típicamente antes de la adolescencia. Es notable que se inclinen por la lengua inglesa incluso cuando el intercambio se inicia en español. (Alonso-Marks, 2015, p. 534)

Muchos son los comentarios que podrían derivar de esta caracterización de los hablantes de español como lengua de herencia en Estados Unidos. No obstante, para esta investigación nos concierne particularmente la adquisición del español por parte de estos hablantes de herencia en este contexto geográfico. Dos son los procesos que la literatura, en general, explicita a este respecto: la *adquisición incompleta* y el *desgaste* o *atrición* (Martínez-Gibson, 2011, p. 181; Acosta Corte, 2013, p. 3; Gill, 2016, pp. 2-3). Por *adquisición incompleta* se entiende lo que pasa "cuando no se han adquirido determinados rasgos de la lengua materna debido al cambio repentino hacia el idioma mayoritario en la sociedad, produciéndose esto cuando el hablante tiene unos 4 o 5 años y comienza la educación formal" (Acosta Corte, 2013, p. 3); en lo concerniente al término *desgaste* o *atrición*, este hace referencia "al hecho de que lo que se adquirió en un momento determinado del desarrollo del lenguaje se puede perder o debilitar a lo largo de un tiempo al reducirse drásticamente el input" (Acosta Corte, 2013, p. 3).

Los factores de adquisición incompleta y desgaste o atrición poseen un claro impacto en los errores de género gramatical que existen también en las producciones de los hablantes de español como lengua de herencia. Sobre estos hablantes, varios estudios documentan, por ejemplo, más errores de concordancia entre el sustantivo y el adjetivo que entre el determinante y el sustantivo (Martínez- Gibson, 2011, p. 191; Acosta Corte, 2013, p. 4; Gill, 2016, p. 3; entre otros). Se mostró, de hecho, que la adquisición de la concordancia entre el sustantivo y el adjetivo se producía después que la que ocurre entre el determinante y el adjetivo, lo que podría explicar el motivo por el que los hablantes de herencia comenten más errores en estas secuencias (Martínez-Gibson, 2011, p. 191). Al igual que los hablantes de español como segunda lengua, los hablantes de herencia también muestran más errores de género gramatical en los sustantivos no canónicos, tendiendo a usar más el masculino que el femenino, por defecto (Acosta Corte, 2013, p. 4; Montrul, 2018, pp. 147-148). Sin embargo, los estudios que comparan las producciones de estos dos tipos de hablantes avalan que los de herencia cometen muchos menos errores que los aprendices de español como segunda lengua (Montrul et al., 2012, p. 105).

2.3. Género gramatical e identidad de género

Mucho menos numerosos son los estudios dedicados a analizar la forma en la que los aprendices de español se asignan a sí mismos el género gramatical dependiendo de su propia identidad de género. Únicamente se encuentran algunas conclusiones al respecto en la investigación llevada a cabo por Acosta Corte (2013), en la que el

autor analiza el habla de una informante de lengua de herencia: constata que esta tiene "una ligera tendencia a utilizar el artículo del género masculino, el género no marcado, incluso cuando se refiere a sí misma" (p. 14). Otras líneas al respecto se encuentran en Scotto y Pérez (2020), al argumentar que "los niños que hablan una lengua con *carga de género* adquieren una identidad de género más tempranamente en el desarrollo que los niños que hablan una lengua que utiliza pocos o ningún marcador de género" (p. 23). A continuación, se tratará brevemente acerca de la *identidad de género* y del conocido como *género no marcado*.

Por *identidad de género* se entiende "la autopercepción que cada persona tiene de sí misma" (Scotto y Pérez, 2020, p. 7). Se trata, por tanto, de un concepto subjetivo que se deriva de la clásica distinción general que se hace entre los términos *sexo* y *género*. Mientras que la palabra *sexo* abarca "las diferencias biológicas entre los seres humanos", por *género* se entiende "la dimensión social que acompaña usualmente a las distinciones biológicas" (Scotto y Pérez, 2020, p. 7). Muchos autores afirman que en la actualidad no existen solo dos sexos, por lo que biológicamente no se puede afirmar que haya dos clases exhaustivas y excluyentes con las que categorizar a todas las personas; existen, por ejemplo, las personas intersex (Scotto y Pérez, 2020, p. 7) y, como bien explica Bengoechea (2015), "inscribir el YO en la lengua implica, entre otros procesos, inscribir el género y el sexo, la propia corporalidad" (p. 4). Entonces, ¿cómo se actúa ante la imperiosa necesidad que existe en la sociedad por saber si se interactúa con un hombre, con una mujer, o con alguno de esos casos "ambiguos"? (Hawkesworth, 1999, p. 36).

Todos estos conceptos son fenómenos extralingüísticos que no tienen una relación directa con el género gramatical, aunque sí indirecta (Mendívil Giró, 2020, p. 43). Una de las más claras repercusiones al respecto es el actual debate que existe alrededor del lenguaje inclusivo y la supresión, por ejemplo, del masculino genérico. Como ya se ha apuntado con anterioridad, el masculino es el género no marcado en español, mientras que el femenino es el género marcado (RAE y ASALE, 2009, p. 85). Pero ¿qué significa *no marcado*? Esta expresión "alude al miembro de una oposición binaria que puede abarcarla en su conjunto, lo que hace innecesario mencionar el término marcado" (RAE y ASALE, 2009, p. 85). Así, en la designación de seres animados, la Academias exponen que los sustantivos de género masculino aluden no solo a los individuos de ese sexo, sino también a todos los individuos de esa especie, sin distinción de sexos (RAE y ASALE, 2009, p. 85). Por ejemplo, en *Todos los empleados deben cumplir el horario*, la Academia interpretaría el sustantivo masculino plural de persona *empleados* de forma genérica para aludir a todos los individuos de la clase o el grupo que se menciona, sean varones o mujeres (RAE y ASALE, 2009, p. 85; Mendívil Giró, 2020, p. 37).

Sin embargo, como ya se ha mencionado, el sistema de género biológico no es binario ni, por tanto, reducible a varones y mujeres; y también se ha elucubrado sobre el carácter sexista del masculino genérico, pues deja sin representación a las mujeres en el habla (Scotto y Pérez, 2020, p. 27). Por eso, desde las perspectivas

feministas se ha propuesto una posible alternativa a esta problemática: el femenino genérico. Este fenómeno consiste "en el empleo de las formas de femenino en los contextos en los que se suspende la distinción de sexos basada en el género" (San Julián Solana, 2017, p. 117). Así, es habitual que en las aulas formadas por un número mayor de alumnas que de alumnos el profesor se dirija a la clase en general haciendo uso del femenino genérico, enunciando secuencias como *Que cada una de vosotras escoja un tema*, sin excluir a ningún varón de este enunciado (San Julián Solana, 2017, p. 123). Así, muchas mujeres hacen uso del femenino genérico como forma de enunciar su YO en el habla de una manera reivindicativa y con el objetivo de proyectarse así ante el mundo (Bengoechea, 2015, p. 10). En este sentido, Bengoechea (2015) recoge enunciados reales proferidos por mujeres como el que se ilustra a continuación: "como a cualquier escritora, no me gusta escribir para solamente dos o tres chicas" (p. 11).

Otro asunto interesante a este respecto es el de las profesiones femeninas. Las Academias recogen lo siguiente: "la presencia de marcas de género en los nombres que designan profesiones o actividades desempeñadas por mujeres está sujeta a cierta variación, a veces solo desde tiempos relativamente recientes" (RAE y ASALE, 2009, p. 105). Así, se destaca, por ejemplo, el caso de *jueza*, cuyo sentido de 'mujer que desempeña el cargo de juez' se ha extendido en países como Chile, el área rioplatense o el Caribe continental, mientras que en otros como México, España o Perú su uso es minoritario (RAE y ASALE, 2009, p. 105). Otros sustantivos, en cambio, que también eran comunes en cuanto al género antiguamente, hacen actualmente el femenino en *-a*, como los pares *abogado/abogada*, *arquitecto/arquitecta* o *catedrático/catedrática* (RAE y ASALE, 2009, p. 106). Es sonado, por ejemplo, el caso de Emilia Pardo Bazán, que se negaba a ser denominada *catedrática*, autoidentificándose como *catedrático*; o el de Clara Campoamor, que firmaba como *abogado* (Bengoechea, 2015, p. 19). Esto era lo habitual durante la época de finales del siglo XXI y principios del XX: las mujeres españolas que se podrían considerar "más progresistas y feministas, optaban por auto-presentarse mediante títulos profesionales en masculino" (Bengoechea, 2015, pp. 18-19).

En cambio, desde los años setenta del siglo pasado, "parece existir una tendencia por la que las mujeres de conciencia más feminista y liberadas se autodenominan en femenino, mientras otras más conservadoras y pertenecientes a profesiones muy masculinizadas y con poder social, se autonombran en masculino" (Bengoechea, 2015, p. 19). Así, por ejemplo, términos como *médico*, *abogado*, *notario* o *arquitecto* parecen haber solidificado el masculino para referirse a la profesión que ejerce tanto hombres como mujeres (Bengoechea, 2015, p. 20). Esto no es más que el reflejo de la siguiente reflexión: "la masculinidad y lo masculino evocan nociones de poder, legitimidad y privilegio, más aún en el campo de ciertas profesiones" (Bengoechea, 2015, p. 19). Por eso, la variación del género gramatical en las ocupaciones ejercidas por mujeres no depende solo de factores geográficos o puramente morfológicos, sino también sociales (RAE y ASALE, 2009, p. 105).

En relación a lo que se apuntaba en párrafos precedentes, el masculino genérico parece disgustar a ciertos sectores de la sociedad, pues invisibiliza o no abarca a todos los géneros o identidades de géneros que existen (Giammatteo, 2020, p. 179). Además del femenino genérico, ya descrito, se han propuesto otras alternativas que se comentarán someramente a continuación. En primer lugar, se documenta el desdoblamiento, que supone sustituir el masculino genérico que podemos encontrar en la secuencia *los amigos* por las secuencias *los amigos y las amigas*, *los/las amigos/amigas* o *los/as amigos/as* (Scotto y Pérez, 2020, p. 32). Otras formas más actuales son el uso de la @, la *x*, el asterisco * o la *e* como morfemas de género neutro, pues abarcan a un mayor número de colectivos como las personas intersex (Scotto y Pérez, 2020, p. 32).

3. Análisis del corpus

3.1. Aspectos metodológicos

A continuación, nos detendremos en el análisis del corpus, que se centrará en los errores de concordancia y atribución de género gramatical en los determinantes, sustantivos y adjetivos producidos por los estudiantes de español que conforman el conjunto de textos seleccionados para este estudio. Para la clasificación de estos errores se ha considerado el criterio establecido por el propio proyecto de corrección y anotación del COWS-L2H que, además, aparece también respaldado por otra bibliografía como la de Montrul et al. (2012, pp. 106-107). En este sentido, se anotará como error de *atribución* incorrecta de género gramatical si hay evidencia de, al menos, dos determinantes o adjetivos que concuerdan entre sí, pero no con el sustantivo, pues se infiere que el aprendiz piensa que el sustantivo es del género incorrecto. Los demás casos se anotarán como errores de *concordancia* de género. Así, secuencias como **la coche vieja* se analizarán como poseedoras de errores de atribución de género gramatical, mientras que **el coche vieja* o **la coche viejo* serían ejemplos de concordancia indebida de género (Montrul et al., 2012, pp. 106-107).

Con todo, se analiza un total de 100 textos en los que los estudiantes de español escriben sobre ellos mismos. Desde los cursos de español que estos estudiantes reciben en la Universidad de California Davis, se les sugiere que escriban textos de una extensión aproximada de entre 250 y 500 palabras sobre distintos temas, y el elegido para esta investigación es el titulado *Yourself* (Yamada et al., 2020, p. 25). Además, el corpus ofrece datos lingüísticos y extralingüísticos sobre los estudiantes que redactan estos textos, como su edad, su identidad de género, el curso en el que están inscritos o su lengua nativa y la lengua que utilizan en casa, entre otros, como parte de un cuestionario inicial que todos rellenan (Yamada et al., 2020, p. 27; Lozano y Fernández-Mira, 2022, pp. 190-191).

Los estudiantes que forman parte del corpus seleccionado para este trabajo rondan los 18 y 22 años, por lo que la edad tampoco se plantea como un dato relevante a la hora de determinar diferencias en cuanto a cómo utilizan el género gramatical los aprendices: todos son jóvenes adolescentes que asisten a clases de español en la Universidad de California Davis. Sin embargo, la identidad de género sí que se considera un dato interesante para la presente investigación: en el corpus encontramos tanto a hombres como a mujeres y una minoría de estudiantes que prefieren no explicitar su identidad de género. Estos datos se explicarán con más detalle en el apartado correspondiente.

El curso en el que están inscritos los aprendices también es relevante: se han elegido a estudiantes matriculados en Español 24 (SPA 24) como representantes de los aprendices de español como segunda lengua de nivel avanzado, y a los de Español 32 y 33 (SPA 32 y SPA 33) como los representantes de los hablantes de español como lengua de herencia. La comparación contrastiva de estos tres cursos que representan a dos tipos de hablantes distintos resulta pertinente, puesto que todos ellos se dedican a la mejora de la expresión escrita (Yamada et al., 2020, pp. 23-24; Lozano y Fernández-Mira, 2022, p. 204). Además, conviene añadir que todos los textos fueron escritos en el año 2020, y recogidos durante tres trimestres distintos: el de invierno (W20), el de primavera (S20) y el de otoño (F20).

3.2. Datos lingüísticos

Ahora sí, pasamos a analizar los datos lingüísticos, es decir, nos centraremos en destacar los hallazgos realizados teniendo en cuenta únicamente la cantidad de errores de concordancia y atribución de género gramatical que los estudiantes del corpus seleccionado cometen sobre las clases de palabras elegidas en este estudio: determinantes, sustantivos y adjetivos. Considerando que se han contabilizado un total de 245 errores en todas las producciones, comenzamos por comentar los más numerosos: los 68 errores de *concordancia* de género gramatical que afectan al determinante. Entre ellos, los más numerosos son aquellos que discuerdan con sustantivos inanimados de final no canónico, como los acabados en *-ión*, (1):

1. Luego tomé Literatura española AP y siguió ***ese*** *inspiración*[5] de seguir el camino de leer libros en español y ahora me encanta aprender sobre la literatura española y sobre el lenguaje.

Después, les siguen aquellos errores de concordancia de género gramatical que afectan a los adjetivos: se contabilizan 56. En este caso, existe una igualdad de errores de adjetivos que discuerdan con sustantivos animados e inanimados: 28 casos cad1a uno. Pero, sin duda, los más numerosos son aquellos que desarrollan

[5] En todos los ejemplos se destaca en *cursiva* el sintagma nominal que presenta un error de género gramatical y, dentro de este sintagma, se marca en **negrita** la categoría gramatical que presenta este error, o sobre la que se quiere comentar algo.

discordancia entre la identidad de género señalada por el estudiante que ha escrito el texto y el género gramatical que utiliza para describirse. Curiosamente, son todos estudiantes que se reconocen como mujeres y escriben enunciados como (2):

2. Es muy importante para me quede ***activo*** porque me pongo muy ansiosa si no.

Es muy común encontrar este tipo de casos en los que chicas alternan el masculino y el femenino en sus producciones, al usar adjetivos con los que se describen a ellas mismas. De hecho, los únicos 4 casos de sustantivos que presentan discordancia de género gramatical atienden también a la misma problemática: chicas que utilizan el masculino, en todos los casos, con sustantivos que designan cargos o actividades, como (3):

3. Recientemente, me enamoré con salud y nutrición y pienso que no puedo ser una abogada, me intereses con ser *una* ***entrenador*** de salud.

En cuanto a los errores de *atribución* indebida de género gramatical, los más numerosos se vuelven a registrar con adjetivos que aparecen dos o más veces en un mismo texto sin concordar con el sujeto que lo ha escrito. Nuevamente, son sujetos que se identifican como mujeres y que utilizan reiteradas veces adjetivos en masculino para referirse a ellas mismas, como en (4):

4. Pero una día, no se cuando, pero me desperté y no me gustaba yo ***mismo***. Digo que si, a mi me gusta yo ***mismo*** pero no estoy en ***enamorado***.

Y también son mayoritarios aquellos casos de atribución incorrecta de género gramatical sobre sustantivos tanto animados, (5), como inanimados, (6), que presentan determinantes y adjetivos que concuerdan entre sí, pero no con sus respectivos sustantivos:

5. Soy *un* ***estudiante*** *universitario* de primera generación.

6. *Todas las* ***programas*** son muy *buenas*, interesantes, y *cómicas*.

Tengamos en cuenta que en (5) es una chica la que se está describiendo como *un estudiante universitario*: es habitual que aparezca el sustantivo *estudiante*, que es común en cuanto al género y de final no canónico, con determinantes y adjetivos en masculino dentro de las producciones escritas por mujeres, pues estas tienden a usar el género no marcado en español: el masculino. Por su parte, el sustantivo *programa* de (6) también suele aparecer con determinantes y adjetivos femeninos, por eso se ha clasificado como un error de atribución de género sobre un sustantivo no canónico: la *-a* final de *programa* confunde a los estudiantes de español, que interpretan este sustantivo como femenino.

A continuación, se detalla el Gráfico 1 en el que se visualizan mejor los resultados obtenidos en base a las variables lingüísticas consideradas en esta investigación:

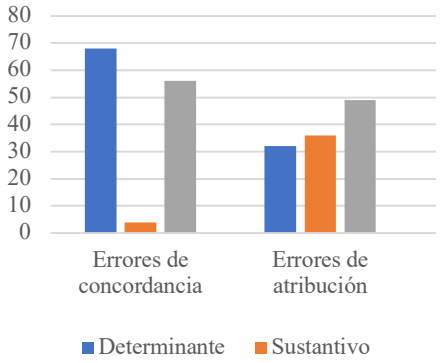

Gráfico 1. Tipología de errores según variables lingüísticas.

3.3. Datos de lengua nativa

Dejando a un lado los datos puramente lingüísticos, pasamos a comentar aquellos que proceden de la lengua nativa de los aprendientes que conforman el corpus y que, por consiguiente, podrían tener repercusiones en la interpretación de los datos obtenidos tras el recuento absoluto de los errores de género gramatical. Como ya se anunció, se han contabilizado un total de 245 errores de este tipo producidos por 100 estudiantes. De estos estudiantes, 51 pertenecen a cursos de español como segunda lengua (SPA 24), y 49 a cursos de español como lengua de herencia (SPA 32 y SPA 33). Los primeros han cometido un total de 182 errores, y los segundos un total de 63. Al hacer la media de estos errores, se comprueba que los aprendices de español como segunda lengua cometen, también de media, un número de errores superior a la media de errores totales, e incluso superior a la media de errores cometidos por los aprendices de español como lengua de herencia:

Tabla 1. Diferencias de errores de los aprendices de español como segunda lengua (L2) y los hablantes de español como lengua de herencia (LH).

	Errores	Estudiantes	Medias[6]
L2	182	51	3,57
LH	63	49	1,29
Total	245	100	2,45

Otro dato interesante es el que surge cuando se analizan detalladamente el número de errores que se contabilizan en las producciones de los hablantes de español como lengua de herencia que conforman el corpus. Durante el análisis se podían

[6] Las medias se obtienen dividiendo la cifra de la primera columna (en este caso, la de *errores*) entre la cifra de la segunda columna (en este caso, la de *estudiantes*).

observar muchas diferencias en el número de errores que cometían estos hablantes y, finalmente, se dio con el quid de la cuestión: estas discrepancias dependían, en gran medida, de la lengua nativa que estos hablantes manifestaban en sus datos sociolingüísticos. Así, se pudo comprobar que, de media, los hablantes de herencia que señalaban el español como su primera lengua cometían muchos menos errores en comparación con la media de errores totales de este grupo; mientras que los hablantes que señalaban el inglés o el inglés y el español como sus lenguas nativas cometían, también de media, un número de errores muy superior a la media total:

Tabla 2. Diferencias de errores entre los hablantes de herencia que señalan el español, el inglés o el inglés y el español como su primera lengua.

	Errores	Estudiantes	Medias
Español	18	32	0,56
Inglés	25	9	2,78
Inglés y español	19	7	2,71
Total	62	48	1,29

3.4. Datos de identidad de género

Todavía resultan más interesantes los datos que emergen del análisis detallado del número de errores que los aprendices del corpus cometen dependiendo de la identidad de género con la que se identifiquen en el cuestionario que rellenan con su información personal. Del total de 100 estudiantes que se analizan en el corpus, 80 se identifican como mujeres, 17 como hombres, y 3 no señalan ningún dato al respecto. De momento, estos 3 estudiantes de género desconocido se descartan, pues solo cometen 4 errores en total, apenas relevantes para este primer análisis. Tomamos, entonces, solo las producciones de hombres y mujeres, cuyas diferencias en el número de estudiantes se subsanan con la realización de medias estadísticas. Tras hacer la media de los errores totales y compararlos con los errores de estos dos grupos, se comprueba que las mujeres cometen un número de errores algo superior que la media de errores totales, mientras que la media de errores de los hombres es muy inferior a los dos datos anteriores:

Tabla 3. Diferencias de errores cometidos por hombres y mujeres.

	Errores	Estudiantes	Medias
Hombres	24	17	1,41
Mujeres	217	80	2,71
Total	241	97	2,49

Sin embargo, tras reflexionar acerca de los resultados generales que ofrecieron los datos de tipo lingüístico, se ha considerado rehacer las medias, eliminando aquellos errores de género gramatical cometidos a partir de sustantivos o sujetos animados

de autorrepresentación femenina, pues solo son cometidos por mujeres. Al eliminar estos casos, queda patente un dato muy interesante si se compara el número de errores que hombres y mujeres cometen sobre sustantivos inanimados. La media de errores de las mujeres ahora es algo inferior que la media total de errores cometidos sobre sustantivos inanimados, mientras que el de los hombres supera estas dos cifras:

Tabla 4. Diferencias de errores cometidos sobre sustantivos inanimados entre hombres y mujeres.

	Errores inanimados	Errores totales	Medias
Hombres	20	24	0,83
Mujeres	114	217	0,53
Errores	134	241[7]	0,56

Retomamos ahora las producciones de datos de los 3 estudiantes cuya identidad de género desconocíamos porque no apuntaban nada al respecto en sus datos personales. Estos casos suponen, a veces, un reto para correctores y anotadores de errores, pues solo nos podríamos fiar del género gramatical mayoritario que esos estudiantes utilizan en sus producciones para referirse a ellos mismos. De hecho, ese es el criterio por el que se ha optado para clasificar esos estudiantes. Todas parecen ser mujeres, pues en sus primeras oraciones ya utilizan sustantivos como *mujer*, (7), o *estudiante*, (8), concordados con determinantes femeninos; algo que, como se viene comprobando, es inusual incluso en los textos producidos por aquellas estudiantes que sí se identifican como mujeres. Además, una de ellas visualiza su futuro profesional como *abogada*, (9), algo que nos invita a reflexionar, en el siguiente párrafo, acerca del género gramatical que las mujeres utilizan al describir sus ocupaciones laborales:

7. Yo me puedo describir como ***una mujer*** alta, con ojos cafes, pelo brunet largo con un cuerpo corpulento.
8. Yo soy ***una estudiante*** en *UNIVERSITY* *UNIVERSITY* y estudio diseño.
9. Hola, tengo *AGE* años soy una estudiante universitaria y estoy estudiando para convertirme en ***abogada*** pero específicamente una fiscal penal.

Si tomamos el sustantivo *abogada*, encontramos que en los textos escritos por mujeres aparece hasta 13 veces en femenino cuando hablan de su futuro profesional, frente a 1 solo caso en masculino, (10). Algo parecido ocurre con el sustantivo *doctor*: 6 casos en femenino, frente a 1 solo caso en masculino, (11). De hecho, en este último ejemplo se detecta un claro caso que comentaba Bengoechea (2015, pp. 19-20): las mujeres siguen escribiendo en masculino las profesiones más masculinizadas (*doctor*, *panadero*) y en femenino las más feminizadas (*maestra*):

[7] Del total de 245 errores se han descartado los 4 de aquellos estudiantes que no señalan su identidad de género en el cuestionario de datos personales.

10. Quiero ir a la facultad de derecho y luego convertirme en ***abogado*** penal.
11. Cuando era niña dice que quiero a hacer **un doctor, maestra, panadero**, pero nunca puedo eligo una cosa que quiero a hacer.

Por último, es preciso también reflexionar acerca de si aparece algún rastro de lenguaje inclusivo en el corpus analizado. Y bien, la verdad es que el uso del masculino genérico es, como su propio nombre indica, generalizado, sobre todo, cuando los estudiantes se refieren al público en general como se puede apreciar en la fórmula de saludo de (12), muy utilizada por los estudiantes para empezar sus escritos. Sin embargo, también se pueden identificar otras variantes, como 1 posible uso del femenino genérico en (13), o 2 casos de desdoblamiento en (14-15). Aunque no se registran casos de otras fórmulas más actuales de lenguaje inclusivo como la @, el asterisco * o la *e*, sí que se registran 2 textos de 2 estudiantes que utilizan la *x* como morfema de género inclusivo, (16-17). De hecho, la voz *latinx* de (17) parece que ya está instaurada en el habla de la comunidad latina de Estados Unidos como la palabra más adecuada e inclusiva para referirse a este grupo social (Cashman y Trujillo, 2018, p. 138):

12. Hola a ***todos***.
13. Personalmente me encantan las *personas*, cada quien es única en su distinta manera y eso me llama la atención.
14. Sin embargo, personas le gusta cuando yo dicho que una persona no está vistiendo bien, o no es ***guapo/a***, pero a veces yo pienso que como yo digo es un poco cruel.
15. En mis clases, cuando ***el profesor o la profesora*** dice que todas las personas van a describirellos mismos, yo lea las personas y piensa si ellos reirían o se comportarían muy serios.
16. Cuando paso tiempo con mis mejores ***amigxs***, dejo todo del estrés atrás y río sobre cosas pequeñas como las caras extrañas de mis ***amigxs***, cómo la gente camina, los detalles en el camino,etc.
17. Espero en mi futuro no solo ser alguien con quien mis alumnos o pacientes confían, pero también se inspiran, especialmente a jóvenes en el ***latinx*** comunidad que son primera generaciónestudiantes y quieren seguir su pasión en la universidad.

4. Conclusiones

El primer objetivo de este trabajo era conocer qué estructuras lingüísticas presentan un mayor número de errores con respecto a la atribución y concordancia indebida de género gramatical. Tras el análisis de los datos lingüísticos se ha comprobado que estos errores proceden, principalmente, de sustantivos inanimados de final no canónico, como bien indicaba la bibliografía al respecto.

En cuanto al segundo objetivo de este trabajo, el análisis de corpus ha corroborado lo que ya se había planteado en el estado de la cuestión: los hablantes de herencia cometen muchos menos errores de género gramatical que los estudiantes de español como segunda lengua, quizá debido a la adquisición temprana de este rasgo por parte de los primeros. Pero los datos del COWS-L2H nos han proporcionado otro dato interesante: existen también diferentes niveles dentro de los hablantes de lengua de herencia, que se ven reflejados en el número de errores que estos exhiben dependiendo de la lengua nativa que señalan. En este sentido, los datos sociolingüísticos que recoge el corpus reflejan que aquellos hablantes que marcan el español como su primera lengua cometen, de media, menos errores que aquellos que marcan el inglés. Así, los primeros suelen añadir, en estos datos, información como "I mostly speak Spanish at home", frente a los segundos, que señalan que hablan inglés y que, además, "I did not write other".

Finalmente, los datos de identidad de género han permitido formular otra hipótesis interesante de investigación futura que apenas recoge bibliografía: las mujeres parecen cometer, de media, numerosos errores de concordancia y atribución de género gramatical cuando se describen a sí mismas, frente a los hombres, que en el corpus analizado no cometen ningún error de este tipo. Este hecho puede deberse al carácter del masculino como género no marcado en español, con los consiguientes problemas que puede ocasionar a las mujeres cuando lo utilizan junto a determinantes, sustantivos y adjetivos para describirse a ellas mismas. Sin embargo, se ha comprobado también que, de media, los hombres cometen más errores de género gramatical que las mujeres con los sustantivos inanimados, lo cual no deja de ser relevante.

Además, se ha cotejado que las nuevas fórmulas de lenguaje inclusivo son más utilizadas por mujeres que por hombres. Por ejemplo, los únicos dos estudiantes que utilizaron el morfema neutro *x* y que se mencionaron en el apartado anterior eran mujeres hablantes de español como lengua de herencia. Quizá este colectivo esté más sensibilizado con las nuevas realidades lingüísticas y de identidades de género. A pesar de que la actual investigación es simplemente una pequeña aportación, se hace necesaria la ampliación de estudios que analicen los nuevos usos del lenguaje inclusivo y cómo las nuevas identidades de género están moldeando el empleo del género gramatical en español.

Financiación

Este trabajo cuenta con el respaldo económico de la ayuda FPU23/02226 del Ministerio de Ciencia, Innovación y Universidades. Además, se inscribe dentro del proyecto MultiLingualGender (MSCA Staff Exchanges, Horizonte Europa, acuerdo de subvención N.º 101182959).

Referencias bibliográficas

Acosta Corte, Á. (2013). La adquisición y el desarrollo lingüístico de los hablantes de herencia de español. Un estudio de caso basado en la investigación-acción en el aula. *Revista Nebrija de Lingüística Aplicada a la Enseñanza de las Lenguas*, *13*, 1-16.

Alonso-Marks, E. (2015). Hablantes de herencia. *Enciclopedia de Lingüística Hispánica*, *2*, 527-535. https://doi.org/10.4324/9781315713441-121

Bengoechea, M. (2015). Cuerpos hablados, cuerpos negados y el fascinante devenir del género gramatical. *Bulletin of Hispanic Studies*, *92*(1), 1-24. https://doi.org/10.3828/bhs.2015.01

Cashman, H. y Trujillo, J.A. (2018). Queering Spanish as a heritage language. En K. Potowski (Ed.), *The Routledge Handbook of Spanish as a Heritage Language* (pp. 124-141). Routledge. https://doi.org/10.4324/9781315735139-9

Fábregas, A. (2022). Hacia una caracterización sintáctica del género del sustantivo en español. *Revista Española de Lingüística*, *52*(1), 39-96. https://doi.org/10.31810/rsel.52.1.2

Fernández López, S. (1997). *Interlengua y análisis de errores*. Edelsa.

Giammatteo, M. (2020). El género gramatical en español y la disputa por el género inclusivo. *Cuarenta Naipes*, *3*, 177-198.

Gill, A.M. (2016). El género gramatical en español entre los hablantes de herencia y los hablantes nativos. *Entrehojas: Revista de Estudios Hispánicos*, *6*(1), 1-19. https://doi.org/10.5206/entrehojas.v6i1.6169

González Calvo, J.M. (1979). El género, ¿una categoría morfológica? *Anuario de estudios filológicos*, *2*, 51-73.

Hawkesworth, M. (1999). Confundir el género. *Debate feminista*, *20*, 3-48.

Lamas, M. (1996). La perspectiva de género. *Revista de Educación y Cultura de la Sección 47 del SNTE*, *8*, 1-8.

Lozano, C. y Fernández-Mira, P. (2022). Designing, compiling and interrogating corpora in L2 Spanish acquisition research. *Journal of Spanish Language Teaching*, *9*(2), 190-206. https://doi.org/10.1080/23247797.2022.2157086

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: revista internacional de didáctica de las lenguas extranjeras*, *15*, 177-193. https://doi.org/10.30827/Digibug.31991

Mendívil Giró, J.L. (2020). El masculino inclusivo en español. *Revista Española de Lingüística*, *50*(1), 35-64. https://doi.org/10.31810/RSEL.50.1.2

Montrul, S., De la Fuente, I., Davidson, J. y Foote, R. (2012). The role of experience in the acquisition and production of diminutives and gender in Spanish: Evidence from L2 learners and heritage speakers. *Second Language Research*, *29*(1), 87-118. https://doi.org/10.1177/0267658312458268

Montrul, S. (2018). Morphology, syntax, and semantics in Spanish as a heritage language. En K. Potowski (Ed.), *The Routledge Handbook of Spanish as a Heritage Language* (pp. 145-163). Routledge. https://doi.org/10.4324/9781315735139-10

Orejudo, J.P. (2020). Concordancia variable en español de herencia: estudio del grupo nominal. *ELUA: Estudios de Lingüística. Universidad de Alicante*, *34*, 153-173. https://doi.org/10.14198/ELUA2020.34.7

Pastor Cesteros, S. (2001). La concordancia en la interlengua de los aprendices de español como lengua extranjera. *ELUA: Estudios de Lingüística. Universidad de Alicante*, *1*, 5-60. https://doi.org/10.14198/ELUA2001.Anexo1.05

Picallo, M.C. (2015). Género y número. *Enciclopedia de Lingüística Hispánica*, *1*, 630-639. https://doi.org/10.4324/9781315713441-56

Real Academia Española (RAE). (2023). *Diccionario de la lengua española* (edición del tricentenario) https://dle.rae.es/

Real Academia Española y Asociación de Academias de la Lengua Española (RAE y ASALE). (2009). *Nueva gramática de la lengua española*. Espasa Libros.

Ruiz Martín, M.A. (2005). "¡Houston, tenemos *una problema!" Problemas con la concordancia de género en la clase de E/LE.M.A. Castillo Carballo, O. Cruz Moya, J.M. García Platero y J.P. Mora Gutiérrez (Coord.), *Las gramáticas y los diccionarios en la enseñanza del español como segunda lengua. Deseo y realidad: actas del XV Congreso Internacional de ASELE* (pp. 777-784). Asociación para la Enseñanza del Español como Lengua Extranjera.

San Julián Solana, J. (2017). Consideraciones glotológicas en torno al femenino genérico. *Verba hispánica: anuario del Departamento de la Lengua y Literatura Españolas de la Facultad de Filosofía y Letras de la Universidad de Ljubljana*, *25*, 117-132. https://doi.org/10.4312/vh.25.1.117-132

Scotto, S.C. y Pérez, D.I. (2020). Relatividad lingüística, gramáticas de género y lenguaje inclusivo: algunas consideraciones. *Análisis Filosófico*, *40*(1), 5-39. https://doi.org/10.36446/af.2020.318

Valdés, G. (2001). Heritage language students: Profiles and possibilities. J. Peyton, D. Ranard y S. McGinnis (Eds.), *Heritage languages in America: Preserving a national resource* (pp. 37-80). Center for Applied Linguistics and Delta System.

Yamada, A., Davidson, S., Fernández-Mira, P., Carando, A., Sagae, K. y Sánchez-Gutiérrez, C. (2020). COWS-L2H: A corpus of Spanish learner writing. *Research in Corpus Linguistics*, *8*(1), 17-32. https://doi.org/10.32714/ricl.08.01.02