

# ¿Qué se dice de la mujer y el hombre en el español de Chile?: estudio exploratorio de las combinaciones frecuentes de los vocablos *mujer* y *hombre* en un corpus de referencia estratificado

*María Natalia Castillo Fadić\**  
*Pontificia Universidad Católica de Chile, Chile*

## Resumen

El vocablo *mujer* presenta alta frecuencia ( $F_i$ ), dispersión (D) y uso (U) en el *Corpus Básico del Español de Chile*<sup>1</sup> ( $F_i = 340$ ,  $D=0,73$  y  $U=248,2$ ). Si se lo compara con el vocablo *hombre*, éste aparece con índices mayores ( $F_i = 502$ ,  $D=0,8$  y  $U=401,6$ ), lo que podría atribuirse a su posibilidad de representar tanto a los individuos humanos masculinos como a los humanos en general; esta hipótesis parece confirmarse por el hecho de que la distribución de ambos vocablos por mundo es similar en cuatro de cinco mundos (Drama, Narrativa, Técnico-Científico y Prensa), pero disímil en Ensayo; más aún, el 9% de las ocurrencias de *hombre* y de *mujer* tiene lugar en Drama y el 14 % de las ocurrencias de cada vocablo se presenta en Técnico-Científico, con una distribución idéntica; en Ensayo, en cambio, tiene

\* Para correspondencia, dirigirse a: María Natalia Castillo Fadić (mcastilf@uc.cl).  
Facultad de Letras, Pontificia Universidad Católica de Chile, Campus San Joaquín. Av. Vicuña Mackenna 4860, Macul. Santiago, Chile. Teléfono: (+56) 22 3547859.

<sup>1</sup> María Natalia Castillo Fadić, 2012 ©.

lugar un 26% de las ocurrencias de *hombre*, el doble de las de *mujer* en el mismo mundo, justamente por tratarse del género donde con mayor prominencia *hombre* se emplea con un significado que abarca tanto a hombres como a mujeres de la especie humana.

La diferencia en los índices estadísticos de uno y otro, por tanto, no puede interpretarse directamente como una manifestación de la disímil representación de la mujer y el hombre en el español de Chile. No obstante, las unidades léxicas con las que cada uno de estos vocablos se combina, especialmente cuando se trata de unidades predicativas, sí pueden dar indicios respecto de cómo concibe la comunidad a hombres y a mujeres.

Palabras clave: estadística léxica, léxico-sintaxis, lingüística de corpus, español de Chile, estudios de género.

WHAT DO GENRES SAY ABOUT WOMEN AND MEN IN CHILEAN SPANISH? AN EXPLORATORY STUDY OF THE FREQUENT COMBINATIONS OF THE WORDS *WOMAN* AND *MAN* IN A STRATIFIED REFERENCE CORPUS

Abstract

The term *mujer* occurs with a high frequency (if), dispersion (D) and use (U) in the *Basic Corpus of Chilean Spanish* ( $F_i = 340$ ,  $D = 0,73$  and  $U = 248,2$ ). The word *man* on the other hand, presents a higher rate of occurrence ( $F_i = 502$ ,  $D = 0,8$  and  $U = 401,6$ ), which is likely to be since both males and females are referred to as men when denoted as humans beings. This hypothesis seems to be confirmed by a similar distribution in four out of five textual genres (Drama, Narrative, Technical-Scientific and Press), but quite a disparate one in the Essay. 9% of the *man* and *woman* occurrences are recorded in Drama and 14% of the occurrences of each term appear in Technical-Scientific texts, with an identical distribution. Unlike these, the occurrences of *man* reaches 26% in the Essay, twice the rate of *woman*, precisely because in this genre the term *man* is more frequently used to encompass both as members of the human species. These statistic differences cannot be directly interpreted as an uneven representation of men and women in Chilean Spanish. However, the combinations these lexical units establish, especially when it comes to predictable units, may give some hints to the conception of men and women in society.

Keywords: Lexical statistics, lexico-syntaxis, corpus linguistics, Chilean Spanish, gender studies.

Recibido: 10/11/18

Aceptado: 11/01/19

## 1. INTRODUCCIÓN

En el marco del actual interés público por la inclusión social de la mujer a través del lenguaje (*cf.*, por ejemplo, Marín 2018 y Manrique Sabogal 2012) y en medio de la discusión acerca de si la lengua española la invisibiliza o no al usar el masculino como género no marcado, nos hemos propuesto hacer una contribución al tema desde la lingüística de corpus.

Siendo la lengua espejo de la cosmovisión de los pueblos, un cambio lingüístico requiere de un cambio cultural que se produzca y se asiente en el tiempo; y como nos muestran también los datos empíricos históricos y la observación de procesos homólogos contemporáneos, hay niveles más permeables al cambio y otros menos permeables: el morfológico, que se busca modificar, negando la posibilidad de usar el masculino como género no marcado, ha sido históricamente y en diferentes lenguas el menos permeable, junto con el fonológico; el léxico, en cambio, no solo es el más permeable, sino también el más cercano a la cosmovisión de los hablantes; los usos de ese léxico, sus combinatorias, las predicaciones, no solo entregan luces sobre la mirada de un pueblo acerca de una realidad, sino que también pueden ser el punto de más fácil intervención cuando se desea erradicar una asociación conceptual indeseada o políticamente incorrecta.

Ahora bien, una intervención en este nivel supone conocer previamente el estado de las cosas. Es por ello que nos preguntamos en este artículo qué se dice del hombre y la mujer o, en términos lexicosintácticos, qué se predica de ambos cuando funcionan como sustantivos. Este estudio lexicostatístico y lexicosintáctico, por tanto, se centra en las combinaciones frecuentes de los vocablos *mujer* y *hombre* en un corpus de referencia estratificado: el *Corpus Básico del Español de Chile* (Castillo Fadić 2012). El foco está en la combinación de cada uno con los modificadores nominales ubicados a la derecha de cada uno de estos sustantivos: adjetivos, complementos y locuciones.

Se trata de un trabajo descriptivo y predominantemente cuantitativo, exploratorio, centrado en los niveles léxico y sintáctico, sincrónico y sintópico, de naturaleza empírica, finalidad pura e implicancias aplicadas, que utiliza herramientas de lingüística de corpus y lingüística computacional y que considera el factor social. Su objetivo general es precisar qué modificadores nominales determinan a los vocablos *mujer* y *hombre* en el español de Chile. Los específicos son: a) describir estadísticamente la distribución de *mujer* y *hombre* por mundo; b) describir estadísticamente las combinaciones léxicas frecuentes de cada vocablo; c) abstraer las construcciones sintácticas más frecuentes que modifican a los vocablos *mujer* y *hombre*, a partir del

análisis por *clusters*; d) precisar qué modificadores nominales determinan al vocablo *mujer* con  $F_i \geq 2$ ; e) precisar qué modificadores nominales determinan al vocablo *hombre* con  $F_i \geq 2$ ; f) comparar y contrastar los modificadores nominales frecuentes de ambos vocablos; g) analizar las combinaciones lexicosintácticas frecuentes para verificar la existencia de unidades fraseológicas, colocaciones o restricciones léxicas que incluyan o modifiquen a uno de estos vocablos o a ambos; h) contribuir al estudio del español de Chile; i) sentar bases para contribuir a los estudios de género desde la lingüística.

Dado el alto uso de los vocablos *hombre* y *mujer* en el español de Chile (cf. Castillo Fadić, en referato), ambos deberían presentar escasas restricciones combinatorias y mínima dependencia de variables externas. Se espera, por tanto, y a modo de hipótesis de base, un alto número de modificadores nominales con  $F_i = 1$  y un bajo número de modificadores nominales con  $F_i > 1$ . Como segunda hipótesis, se espera obtener más predicaciones con valor semántico-pragmático positivo para *hombre* que para *mujer* y, a la inversa, como tercera hipótesis, se espera obtener más predicaciones con valor semántico-pragmático negativo para *mujer* que para *hombre*.

La investigación pretende ayudar a identificar la imagen de la mujer y el hombre en la sociedad chilena contemporánea, a partir de sus representaciones lingüísticas. Esta identificación puede ser el punto de partida para abordar la posibilidad de introducir mejoras en las políticas de inclusión, equidad e igualdad de género. Desde las políticas lingüísticas, esta mejora podría considerar las estructuras adjetivas en las políticas de género y no solo las centradas en flexión de género gramatical.

## 2. FUNDAMENTO TEÓRICO

### 2.1. LEXICOESTADÍSTICA

La *lexicoestadística* o *estadística léxica* estudia cuantitativamente el nivel léxico. Puesto que opera a nivel de lengua, no se detiene en las unidades de texto o *palabras*, sino que se enfoca en las unidades de léxico o *vocablos*. Mientras las palabras –variantes– corresponden a las realizaciones a nivel textual, los vocablos –invariantes– representan un paradigma de palabras que comparten base léxica y contenido semántico. En el caso de las formas nominales, el vocablo suele corresponder a la unidad léxica en masculino,

en singular y en su forma primitiva o no derivada; las palabras, en cambio, se forman a partir de la flexión o derivación de una base léxica común. En lo concerniente a las unidades léxicas en estudio, en la medida en que *mujer* y *hombre* no presentan flexión de género –puesto que la noción de femenino o masculino está incluida en la base respectiva–, las palabras por medio de las cuales se realizan ambos vocablos solo pueden experimentar flexión de número o derivaciones. Para hacer factible el conteo estadístico, metodológicamente se emplea también el concepto de *tipo*, unidad léxica intermedia que funciona como etiqueta para todas las ocurrencias de palabras que, agrupadas bajo un mismo vocablo por sus características morfológicas y semánticas, comparten un mismo significante.

La estadística léxica supone la medición de índices estadísticos. El más simple es el de la *frecuencia*: esta puede ser *absoluta* ( $F_i$ ) si da cuenta del número de ocurrencias de una unidad léxica o *relativa* ( $f_i$ ) si se calcula sobre el tamaño de la muestra ( $F_i/n$ ). En este trabajo consideraremos ambas frecuencias.

Otro índice relevante es el de *dispersión compleja* (D), que permite conocer cuán equilibrada es la distribución de las frecuencias en distintos subcorpus o *mundos* (*Worlds*, en Juilland y Chang-Rodríguez 1964: XXVI) que, en esta investigación, representan diferentes géneros discursivos. Para calcular la dispersión, hemos usado la fórmula de Juilland, Traversa, Beltramo, y Di Blasi (1973), que considera el número de mundos ( $n=5$ ), la suma de las frecuencias por mundo ( $x_i$ ) y la suma de las frecuencias totales (T):

$$D = 1 - \frac{\sqrt{nx_i^2 - T^2}}{2T}$$

Esta fórmula arroja resultados que van desde 0 (dispersión mínima) a 1 (dispersión máxima).

A partir de los índices de frecuencia y dispersión es posible obtener el de *uso* (U), consistente en el producto de la frecuencia –absoluta o relativa– por la dispersión. Puesto que los índices de dispersión y uso que empleamos en este trabajo han sido extraídos de Castillo Fadić (en referato), ambos se calculan a partir de frecuencias absolutas por dos razones: a) la primera, metodológica, radica en que se trabaja con subcorpus de tamaño homogéneo, lo que vuelve irrelevante estadísticamente la consideración del tamaño del mundo ( $n$ ); b) la segunda, lexicográfica, estima que las frecuencias relativas, expresadas en decimales sin una parte entera, resultan menos amigables para los usuarios de un diccionario que las frecuencias absolutas, expresadas en números enteros.

Los dos vocablos en estudio –*mujer* y *hombre*–, forman parte del *Léxico Básico del Español de Chile*, constituido por los alrededor de cinco mil vocablos de mayor uso en la comunidad, que se caracterizan por su distribución estable y por su tendencia atemática. Esta última característica es directamente proporcional al *rango* de uso o posición en que se ubican las unidades léxicas al ordenarlas de mayor a menor uso; este rango, a su vez, es inversamente proporcional a la dependencia de variables externas, como el género discursivo, el tema, el estilo, etc.

Tanto el vocablo *mujer* como el vocablo *hombre* presentan alta frecuencia ( $F_c$ ), dispersión (D) y uso (U) en el *Corpus Básico del Español de Chile*, por lo que ambos se ubican en los primeros rangos entre los sustantivos de mayor uso, como se aprecia en la Figura 1.

Índices	Mujer	Hombre
Frecuencia absoluta ( $F_c$ )	340	502
Dispersión compleja (D)	0,73	0,8
Uso (U)	248,2	401,6
Rango entre los sustantivos de mayor dispersión	392	150
Rango entre los sustantivos de mayor uso	12	7

Figura 1. Índices estadísticos de *mujer* y *hombre*

## 2.2. LEXICOSINTAXIS

Si la estadística léxica permite establecer los índices estadísticos de las unidades léxicas, esta disciplina en interrelación con la *lexicosintaxis*, que observa las relaciones que se establecen entre un argumento y sus predicaciones, puede precisar también los índices combinatorios de dichas unidades. Esto es relevante, por cuanto las combinaciones frecuentes pueden revelar la presencia de locuciones u otras estructuras pluriverbales más o menos lexicalizadas. Al respecto, es importante precisar que incluso las estructuras de menor lexicalización, las llamadas *colocaciones*, que antaño se consideraron meras combinaciones frecuentes sin mayor intencionalidad y altamente transparentes, se entienden hoy mayoritariamente como “asociaciones entre un predicado y un conjunto restringido de argumentos que determinan una subclase léxica de mayor o menor extensión” (Bosque 2001: 17), puesto que la frecuencia de la combinatoria suele no ser arbitraria, sino estar sujeta a distintos factores lingüísticos y extralingüísticos. De hecho, “[f]uera de la fraseología existe un gran número de restricciones

sintácticas y semánticas, incluso en las relaciones predicado-argumento”. (Barrios Rodríguez 2016: 18)

### 2.3. IDEOLOGÍA, CAMBIO LINGÜÍSTICO Y POLÍTICA LINGÜÍSTICA

Concordamos con Ignacio Bosque (2012) –y con el pleno de la Real Academia Española que suscribe su informe– en que la lengua no es en sí sexista o no sexista, sino que el uso que se hace de ella es el que puede serlo o no. “El lenguaje puede usarse, en efecto, con múltiples propósitos. Puede emplearse para describir, ordenar, preguntar, ensalzar o insultar, entre otras muchas acciones, y, desde luego, también puede usarse para discriminar a personas o a grupos sociales” (Bosque 2012). Coincidimos con él también en discrepar de quienes estiman que “el léxico, la morfología y la sintaxis de nuestra lengua han de hacer explícita sistemáticamente la relación entre género y sexo, de forma que serán automáticamente sexistas las manifestaciones verbales que no sigan tal directriz, ya que no garantizarían “la visibilidad de la mujer””. Nuestros reparos a esta postura se basan en la necesidad de no confundir sexo natural con género gramatical, en el principio de economía de las lenguas y, por último, en que intentar imponer una construcción morfológica de laboratorio a los hablantes, sea cual sea, implica, más allá de las buenas intenciones, caer en un prescriptivismo a ultranza, donde un grupo decide entre cuatro paredes cuál es la forma correcta de hablar, desestimando las otras y pretendiendo exiliarlas del ámbito público; algo similar, pero esta vez no en lo relativo al exilio de una variedad, sino al de una lengua, fue lo que intentó Franco con el catalán, y a la vista tenemos los resultados; no en vano ya Cervantes (2004: 872) ponía en boca de Don Quijote que “cuando algunos no entienden estos términos, importa poco, que el uso los irá introduciendo con el tiempo, que con facilidad se entiendan; y esto es enriquecer la lengua, sobre quien tiene poder el vulgo y el uso”. En efecto: el vulgo y el uso; no individuos ni grupos particulares. De hecho, toda intervención lingüística que pretenda impactar en y desde las políticas públicas debe partir del conocimiento acerca de cómo evolucionan las lenguas y cómo se generan los procesos de cambio lingüístico en la historia (Cooper 1997). Como bien señala Alvar (2006),

[...] la lengua sigue siendo independiente por más que la sociedad pueda condicionarla; no se olvide que las motivaciones sociales actúan lentamente, sin llegar a romper lo que la lengua tiene de inmanencia. Cuando Stalin se enfrentaba con la realidad lingüística del ruso señalaba los cambios sociales que se habían producido en su país en algo más de

cien años: desaparición del sistema feudal, liquidación del capitalismo, nacimiento del régimen socialista; se habían superado dos bases con sus supraestructuras y había surgido otra nueva con la suya propia, pero el ruso actual apenas si difiere en su estructura del ruso de Pushkin y poco después interroga: “¿Cómo se puede destruir la lengua existente y crear en su lugar, en unos pocos años, una lengua nueva sin llevar la anarquía a la vida social y sin provocar a la propia sociedad el riesgo de su disgregación?”

El cambio lingüístico, en efecto, no solo es resistente a las imposiciones, sino que además requiere de un tiempo considerable para asentarse. En esta línea, es altamente ilustrativo el caso del francés, mencionado por Moreno Fernández (2009: 118), en el que la pérdida de la /-n/ postnuclear y la nasalización de la vocal precedente tardó nada menos que cinco siglos.

Resumiendo, manifestamos tanto la ineficacia de intentar generar un cambio lingüístico rápido (cf. Alvar 2006 y Moreno Fernández 2009), como la esterilidad de estos esfuerzos. Concordamos con Lemus (2001), en que el sexismo es un problema sociocultural, no lingüístico, por lo que su solución no pasa por intentar modificar la estructura de un idioma.

Proponemos, por último, la conveniencia de trasladar la atención del nivel morfológico al léxico; este nivel, el más cercano a la cosmovisión de los hablantes, es también el menos resistente al cambio lingüístico, por lo que puede ser el mejor comienzo no solo para estudiar la brecha de género, sino también para evaluar su potencial reducción. Es en el léxico, su uso y combinatoria, donde más claramente podemos observar tanto el pensamiento manifiesto como el que se esconde en el inconsciente colectivo.

#### 2.4. EL FACTOR SOCIAL

En el caso de esta investigación, las combinaciones frecuentes de los sustantivos *mujer* y *hombre* podrían estar motivadas por factores extralingüísticos de corte social: específicamente, por la disímil valoración social de mujeres y hombres en Chile, que redundaría en una mayor presencia masculina en la vida pública, no solo por sus mayores posibilidades de obtener y mantener un empleo bien valorado y remunerado, sino también por sus mayores ingresos y su mayor tiempo libre:

The gender gap in terms of jobs and earnings is large in Chile. Women earn on average around 20% less than men, are twice as likely to be in low-paid jobs, have a lower level of employment, and are more likely to be unemployed. However, Chilean girls fare slightly better than boys



on tests of cognitive skills at 15 years of age. And while women feel less safe at night than men do, the gap between Chilean men and women is smaller than the OECD average gap. (OECD 2017: 4)

Adicionalmente, mientras los hombres en Chile dedican 34,6 horas semanales al trabajo remunerado y solo 19,0 al no remunerado, las mujeres solo cuentan con 19,8 horas de trabajo remunerado y destinan 42,1 a labores no remuneradas, especialmente en el hogar. En total, los hombres trabajan 53,6 horas a la semana, mientras las mujeres trabajan 61,9 horas, lo que supone no solo menos ingresos para ellas, sino también menos tiempo libre para el descanso, la vida social y la realización de actividades recreativas o de perfeccionamiento laboral, lo que disminuye su visibilidad en la vida pública. (CEPAL - Naciones Unidas 2016)

### 3. METODOLOGÍA

#### 3.1. OBTENCIÓN DEL CORPUS

El *Corpus Básico del Español de Chile* es un corpus de referencia etiquetado de 501.211 palabras en contexto. Este corpus, que recibe este nombre porque fue la base a partir de la cual se obtuvo el *Léxico Básico del Español de Chile* (Castillo Fadić en referato), se distribuye en cinco mundos de alrededor de cien mil palabras cada uno, como se aprecia en la Figura 2. Ha sido obtenido de alrededor de cuatro mil textos de autores chilenos, publicados por primera vez en español como lengua original entre 1981 y 2006, distribuidos en casi mil libros y más de tres mil ejemplares de diarios de circulación nacional, que corresponden al menos al 15% del total de publicaciones por mundo o género discursivo en ese período. Para la obtención del corpus se realizaron diversos muestreos estratificados por mundo: a) primero, para precisar cuántas oraciones se requerían para obtener cien mil palabras por mundo; b) segundo, para determinar aleatoriamente de qué fuentes se extraerían las oraciones; c) tercero, para precisar las páginas; d) y cuarto, para seleccionar las líneas. En el caso de Prensa, los muestreos se estratificaron también por sección, para lo cual se consideraron las cuatro secciones comunes a todos los medios de circulación nacional del período, a saber, Opinión, Nacional, Deportes y Espectáculos. Para mayores informaciones sobre la obtención y constitución del corpus, cf. Castillo Fadić (2015: 3-5).

Mundos	Número de palabras
Drama	100220
Narrativa	100323
Técnico-Científico	100578
Ensayo	100033
Prensa	100057
<b>Total</b>	<b>501211</b>

Figura 2. N.º de palabras por mundo

### 3.2. PROCESAMIENTO DEL CORPUS

El corpus fue convertido a \*.TXT, codificación ANSI, y se procesó con AntConc 3.2.4w.

Se realizó una búsqueda de las concordancias de *mujer* y *hombre* en los cinco mundos. Dado que las realizaciones de ambos vocablos contienen la base léxica *mujer* y *hombre*, respectivamente, para la búsqueda se seleccionó la opción *Regex*. Esto permitió obtener las concordancias de ambos vocablos (*cf.* Figuras 3 y 4).

### 3.3. DESCRIPCIÓN Y ANÁLISIS DE LOS DATOS

Los índices de frecuencia, dispersión y uso de ambos vocablos se obtuvieron de Castillo Fadić (en referato).

En AntConc 3.2.4w, se buscaron las combinaciones más frecuentes mediante un análisis por *clusters* en opción *Regex*, centrado en combinaciones de  $F_i \geq 2$  de mínimo dos y máximo cinco unidades, situando a los dos vocablos en estudio a la izquierda del *cluster*, como se aprecia en las Figuras 5 y 6. Esto permitió generar un archivo *output* para cada vocablo, con las combinaciones léxicas más frecuentes de cada uno. A partir del análisis de las concordancias y de las combinaciones léxicas frecuentes obtenidas de la búsqueda por *clusters*, se abstraieron las estructuras sintácticas en las que participa preferentemente cada vocablo en estudio, se evaluaron las posibles unidades fraseológicas o restricciones que les atañen y se determinó qué adjetivos, complementos o locuciones son los que más los determinan.

AntConc 3.2.4w (Windows) 2011

File Global Settings Tool Preferences About

Corpus Files

CORPUS DRAMA ansi.  
CORPUS ENSAYO ansi.  
CORPUS NARRATIVA a  
CORPUS PRENSA ansi.  
CORPUS TÉCNICO-CII

Concordance Concordance Plot File View Clusters Collocates Word List Keyword List

H#	KWIC	File
4	9 Dicen que las operarias mujeres también empezaron a hacer l	CORPUS DRAMA ansi.txt
5	sólo se aprovechaba de las mujeres pobres que le iban a pedir	CORPUS DRAMA ansi.txt
6	nculos de sangre. 28 26 La mujer que le dé el si ante el altar	CORPUS DRAMA ansi.txt
7	bre." 153 14 "Para eso soy mujer de dinero y en Chilchil, pueb	CORPUS DRAMA ansi.txt
8	Y que si no devuelven a tu mujer en el acto, viva, enferma, mu	CORPUS DRAMA ansi.txt
9	o es manera de recibir a tu mujer." 83 7 ;Rehenes? 84 8 "Vicen	CORPUS DRAMA ansi.txt
10	"castitas"." 18 14 Era una mujer inmensamente rica y quiso que	CORPUS DRAMA ansi.txt
11	que no ha visto nunca una mujer en calzones? 27 18 Más adelan	CORPUS DRAMA ansi.txt
12	mar como nunca amó ninguna mujer. 77 20 ¿Sentiste el calor de	CORPUS DRAMA ansi.txt
13	ino, lo que nos pasa a las mujeres de los pescadores." 53 24 "	CORPUS DRAMA ansi.txt
14	Tomás! 58 25 "No te apurís, mujer!" 59 26 "Hola, muchachos."	CORPUS DRAMA ansi.txt
15	xequiel." 53 14 ;Yo soy la mujer de Emilio Farias! 54 16 El pa	CORPUS DRAMA ansi.txt
16	fue ley el no desear a la mujer ajena, él sólo te pedía que l	CORPUS DRAMA ansi.txt
17	da me dio en un día/ hijo, mujer y tienda." 17 21 "¿Qué me dem	CORPUS DRAMA ansi.txt
18	uno sale el nombre de una mujer que me quiso mucho, y en el o	CORPUS DRAMA ansi.txt
19	en el corazón una risa de mujer que no puedo hacer callar con	CORPUS DRAMA ansi.txt
20	No te la puedes con una mujer de verdad. 184 20 ¿Qué te pas	CORPUS DRAMA ansi.txt
21	le el amor de verdad a una mujer 189 3 "¿Y si nos besamos, Dan	CORPUS DRAMA ansi.txt
22	o." 192 6 "Pregúntale a tu mujer, si quieres." 193 7 ;Te acuer	CORPUS DRAMA ansi.txt
23	e tocaba el piano como las mujeres." 23 14 "Se les derramó par	CORPUS DRAMA ansi.txt
24	nsiste en conquistar a una mujer, compadre?" 35 31 Y Dios siem	CORPUS DRAMA ansi.txt
25	r rumiando la rabia de una mujer indecisa ella riega sus lágri	CORPUS DRAMA ansi.txt
26	17 "La siempre magnífica mujer, detrás del gran hombre!" 19	CORPUS DRAMA ansi.txt
27	nir un poco." 34 11 "Calma, mujer" 35 12 "Pero, ¿y la policía?"	CORPUS DRAMA ansi.txt
28	a mi hija, robando a mi mujer" 67 16 "Soy Lucas Meyer, ¿sentie	CORPUS DRAMA ansi.txt

Search Term  Words  Case  Regex  Advanced  Concurrence Hits 358 Search Window Size 50

Total No. 5

Files Processed

Reset

Kwic Sort

Level 1 1R   Level 2 2R   Level 3 3R

Figura 3. AntConc 3.2.4w, recorte de pantalla: concordancias *mujer*

AntConc 3.2.4w (Windows) 2011

File Global Settings Tool Preferences About

Concordance Concordance Plot File View Clusters Collocates Word List Keyword List

Corpus Files

- CORPUS DRAMA ansi.
- CORPUS ENSAYO ansi
- CORPUS NARRATIVA a
- CORPUS PRENSA ansi
- CORPUS TÉCNICO-CIE

Hit	KWIC	File
4	bla! 64 22 "Bien lo dijo el hombre santo: "aquí en esta tierra	CORPUS DRAMA ansi.txt
5	21 Necio ¿cómo reconocer al hombre por un brazo? 86 21 "¡Que v	CORPUS DRAMA ansi.txt
6	03 16 "No te hagas el leso, hombre." 104 17 ¿Qué? 105 18 Pero	CORPUS DRAMA ansi.txt
7	comienzos. 106 19 Valiente hombre es Ud. 107 20 Sin duda tal	CORPUS DRAMA ansi.txt
8	zahumerio pa que vengan los hombres y después me llaman a mi p	CORPUS DRAMA ansi.txt
9	miento a un ser inferior al hombre y entonces habré desapareci	CORPUS DRAMA ansi.txt
10	ndo lo recobré supe que ese hombre era el hombre de mi vida. 1	CORPUS DRAMA ansi.txt
11	supe que ese hombre era el hombre de mi vida. 11 9 Parece tem	CORPUS DRAMA ansi.txt
12	ara aquilatarlo. 40 5 0 ese hombre se enamoró realmente de la	CORPUS DRAMA ansi.txt
13	rque su muchacho decide ser hombre." 66 31 Me lo merezco. 67 3	CORPUS DRAMA ansi.txt
14	onómicamente fuerte, contra hombres de mucha personalidad... q	CORPUS DRAMA ansi.txt
15	do estar sino orgulloso del hombre que hoy toma el poder." 63	CORPUS DRAMA ansi.txt
16	que hemos hallado al fin al hombre indicado para nuestros prop	CORPUS DRAMA ansi.txt
17	Viene una caravana de diez hombres y varias mulas con municio	CORPUS DRAMA ansi.txt
18	legamos al río, apareció el hombre, me vendó los ojos a la fue	CORPUS DRAMA ansi.txt
19	e tú dices y con unos pocos hombres, podremos llegar al Perú."	CORPUS DRAMA ansi.txt
20	98 10 Vamos a buscar a los hombres de más confianza y luego l	CORPUS DRAMA ansi.txt
21	ueda Flores." 13 9 "Y cinco hombres: Juan Rodulfo, el primogén	CORPUS DRAMA ansi.txt
22	io. 54 21 Vos no tenís a tu hombre en la mar. 55 22 "Ningunit	CORPUS DRAMA ansi.txt
23	volver a mi casa. 47 17 "El hombre de mi izquierda me dirá: "	CORPUS DRAMA ansi.txt
24	aparecía Farias y esos dos hombres que se lo llevaban y la az	CORPUS DRAMA ansi.txt
25	s creencias no esclaviza al hombre 62 17 "Derramará sangre tuy	CORPUS DRAMA ansi.txt
26	adón y yo sabemos que fui un hombre alegre y bueno, muy bueno."	CORPUS DRAMA ansi.txt
27	35 "Porque el hogar es al hombre / lo que la lluvia a la tier	CORPUS DRAMA ansi.txt
28	Darle la buena nueva: su hombre murió la otra semana 77 18 "So	CORPUS DRAMA ansi.txt

Search Term  Words  Case  Regex  Advanced

Concordance Hits 521 Search Window Size 50

hombre

Start Stop Sort

Kwic Sort

Level 1 1R  Level 2 2R  Level 3 3R

Total No. 5

Files Processed

Reset

Figura 4. AntConc 3.2.4w, recorte de pantalla: concordancias *hombre*

AntConc 3.2.4w (Windows) 2011

File Global Settings Tool Preferences About

Corpus Files

- CORPUS DRAMA ansi.
- CORPUS ENSAYO ansi
- CORPUS NARRATIVA a
- CORPUS PRENSA ansi
- CORPUS TÉCNICO-CIE

Concordance Concordance Plot File View Clusters Collocates Word List Keyword List

Total No. of Cluster Types: 1277 Total No. of Cluster Tokens: 1410

Rank	Freq	Cluster
1	20	mujer de
2	18	mujer que
3	15	mujeres que
4	9	mujeres de
5	9	mujeres y
6	8	mujer y
7	6	mujer en
8	6	mujer no
9	5	mujeres en
10	4	mujer se
11	4	mujer, no
12	4	mujeres se
13	3	mujer de mi
14	3	mujer más
15	2	mujer a
16	2	mujer al
17	2	mujer chilena
18	2	mujer como
19	2	mujer de edad
20	2	mujer del
21	2	mujer en el
22	2	mujer en la
23	2	mujer es
24	2	mujer ha

Search Term  Words  Case  Regex  N-Grams

Cluster Size Min. Size  Max. Size

Min. Cluster Frequency

Search Term

Sort by

Search Term Position  On Left  On Right  Invert Order

Total No. 5

Files Processed

Figura 5. AntConc 3.2.4w, recorte de pantalla: análisis por *clusters* vocablo *mujer*

The screenshot shows the AntConc 3.2.4w (Windows) 2011 interface. The 'Clusters' tab is active, displaying a table of cluster results for the search term 'hombre'. The table has three columns: Rank, Freq, and Cluster. The search term is 'hombre', and the cluster size is set to 5. The search term position is set to 'On Left'.

Rank	Freq	Cluster
1	32	hombre de
2	23	hombre que
3	21	hombres de
4	20	hombre y
5	18	hombres y
6	15	hombre en
7	11	hombre es
8	9	hombre a
9	8	hombres que
10	6	hombre como
11	6	hombre con
12	6	hombre más
13	6	hombre se
14	6	hombres y mujeres
15	5	hombre de la
16	5	hombre en la
17	5	hombre, que
18	5	hombres con
19	5	hombres se
20	4	hombre común
21	4	hombre del
22	4	hombre fuerte
23	4	hombre muy
24	4	hombre por

Search Term:  Words  Case  Regex  N-Grams  
 Cluster Size: Min. Size  Max. Size   
 Search Term:  Advanced  
 Start Stop Sort by  
 Search Term Position:    
 On Left  On Right  Invert Order

Figura 6. AntConc 3.2.4w, recorte de pantalla:  
análisis por *clusters* vocablo *hombre*

## 4. RESULTADOS

### 4.1. ÍNDICES ESTADÍSTICOS

Respecto de los índices de frecuencia, como se aprecia en las Figuras 7 y 8, la  $F_i$  de *hombre* es mayor a la de *mujer* en todos los mundos, lo que podría atribuirse, por un lado, a la mayor representación del ser humano masculino en los textos y, por otro, a la posibilidad del vocablo *hombre* de referirse al ser humano en general y no únicamente al de sexo masculino. Esta posibilidad podría explicar su mayor dispersión  $-D=0,8$  sobre  $D=0,73-$ , ya que, como sabemos, la dispersión tiende a ser inversamente proporcional a la especificidad de un término.

En la Figura 7 puede observarse que tanto la  $F_i$  de *hombre* como la de *mujer* alcanzan sus mayores índices en Drama y los menores en Prensa. Sin embargo, mientras la proporción de ambos vocablos es similar en Drama, Narrativa, Técnico-Científico y Prensa, en Ensayo la  $F_i$  de *hombre* prácticamente triplica a la de *mujer*; lo que queda claramente ilustrado en la Figura 8.

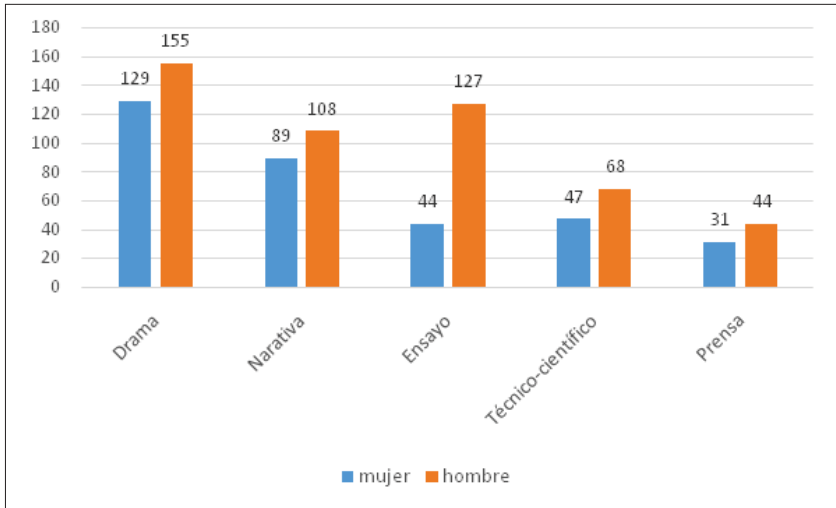


Figura 7. Distribución de vocablos *mujer* y *hombre* por mundo

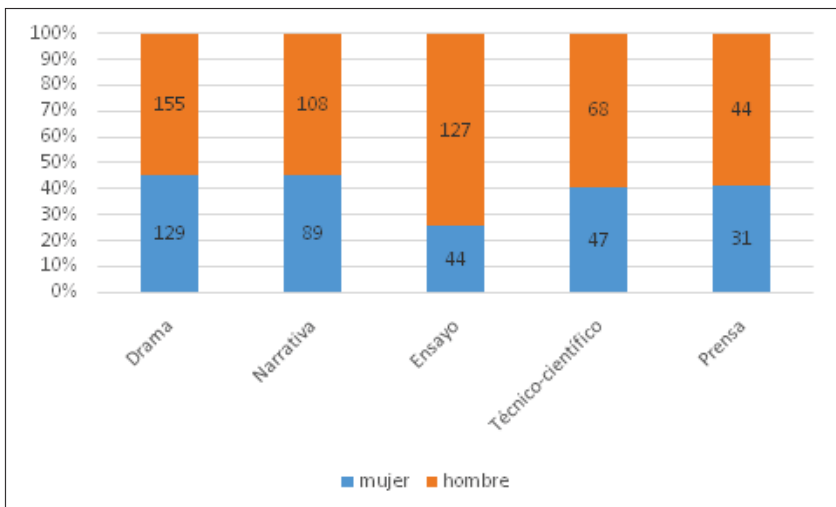


Figura 8. Proporción de vocablos *mujer* y *hombre* por mundo

En efecto, si revisamos la distribución del vocablo *mujer*, un 38% de las realizaciones aparece en Drama y un 26% en Narrativa (cf. Figura 9), lo que arroja un 64% de ocurrencias en textos de corte preferentemente literario. El porcentaje de realizaciones en Ensayo y en Técnico-Científico presenta diferencias estadísticamente marginales; ambos mundos sumados alcanzan un 27%, es decir, menos de la mitad que Drama y Narrativa juntos. Prensa, por último, contiene solo el 9% de las realizaciones del vocablo.

En el caso del vocablo *hombre*, en cambio (cf. Figura 10), un 31% de las realizaciones tiene lugar en Drama, un 25% en Ensayo y un 21% en Narrativa. Si consideramos que en Técnico-Científico se da un 14% de las realizaciones de *hombre* –mismo porcentaje que en ese mundo alcanza el vocablo *mujer*–, podemos observar que mientras el 52% de las ocurrencias de *hombre* se presenta en textos preferentemente literarios –Drama y Narrativa–, un elevado 39% tiene lugar en textos que tienden a ser especializados –Ensayo y Técnico-Científico–, lo que contrasta con el 27% que en este tipo de textos alcanza el vocablo *mujer*.

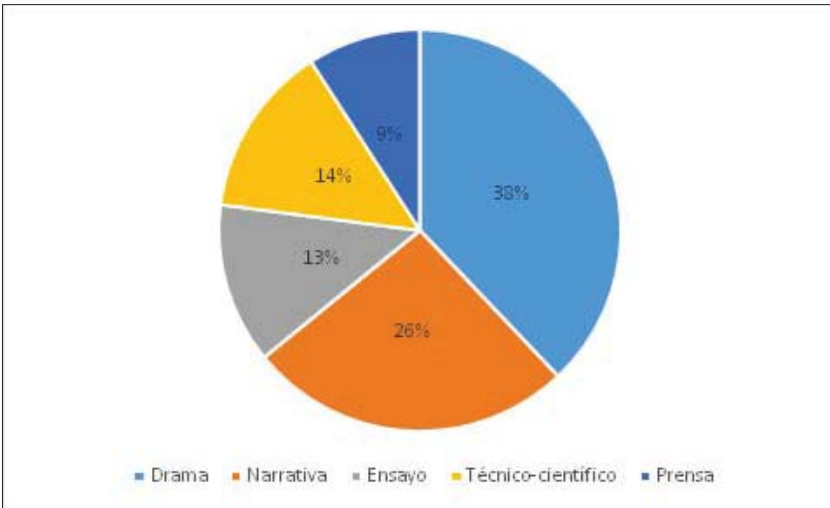


Figura 9. Distribución de *mujer* por mundo



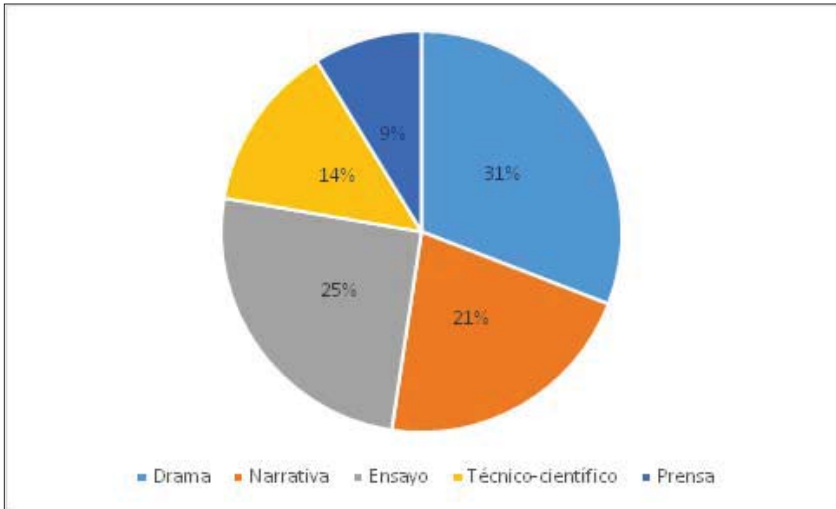


Figura 10. Distribución de *hombre* por mundo

#### 4.2. COMBINACIONES FRECUENTES

Ambos vocablos se combinan preferentemente con elementos subordinantes que introducen estructuras con función adjetiva (cf. Figura 11).

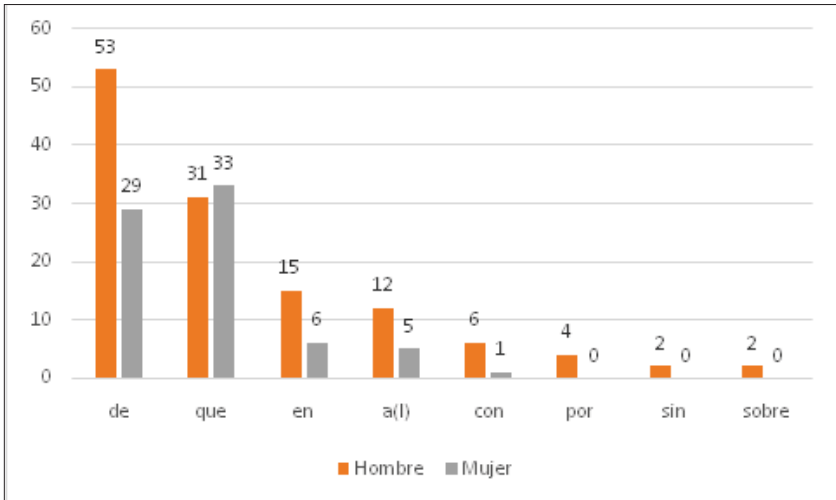


Figura 11. Elementos subordinantes que introducen estructuras con función adjetiva

Al sumar las frecuencias por tipo de elemento subordinante (cf. Figura 12), podría pensarse que mientras la combinación con prorrelativos –que introducen suboraciones adjetivas– es similar en ambos vocablos – $F_i=33$  en *mujer* y  $F_i=31$ – en *hombre*, la combinación con preposiciones –que introducen complementos del nombre– es significativamente superior para *hombre*, con una  $F_i=94$ , más del doble de la  $F_i=41$  de *mujer*.

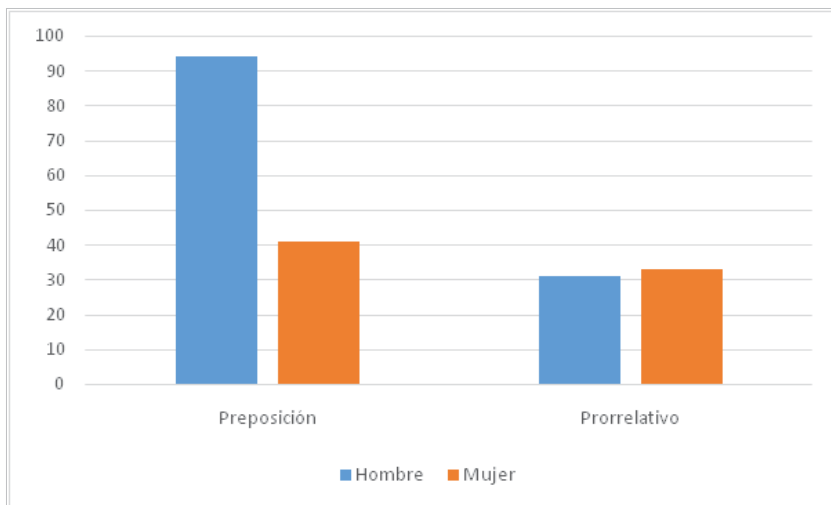


Figura 12. Combinación con prorrelativos y preposiciones ( $F_i$ )

Sin embargo, debemos recordar que la  $F_i$  de ambos vocablos es disímil, por lo que para precisar las diferencias combinatorias es necesario trabajar con frecuencias relativas ( $f_i$ ), es decir, con aquellas que se calculan sobre el total de la muestra ( $f_i = F_i/n$ ). Al dividir la  $F_i$  de la combinación de ambos vocablos con preposiciones y prorrelativos por el número de ocurrencias de *mujer* y *hombre*, podemos observar un cambio en los gráficos (cf. Figura 13).

Si bien ambos vocablos continúan combinándose preferentemente con preposiciones –y, por ende, con complementos del nombre–, la brecha que se apreciaba en la Figura 12 entre *hombre* + preposición y *mujer* + preposición decrece. Por otro lado, la combinación de *mujer* con prorrelativos –y, por tanto, con suboraciones adjetivas– casi duplica a la de *hombre*.

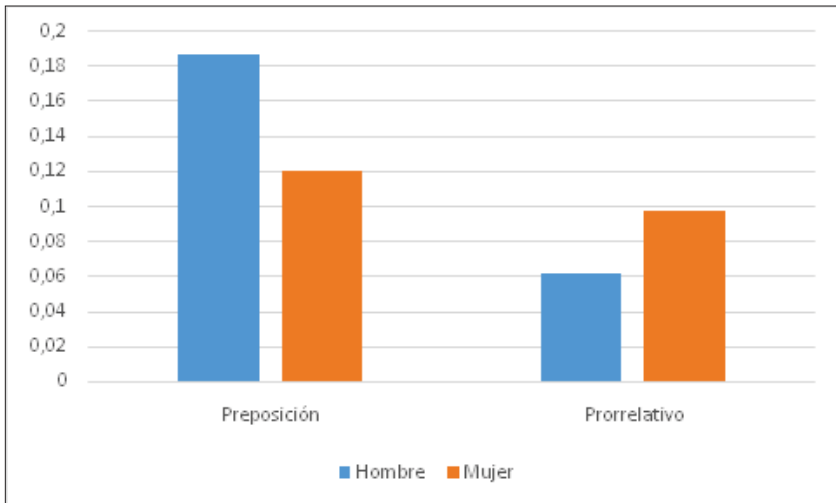


Figura 13. Combinación con prorrelativos y preposiciones (fi)

Pese a la elevada frecuencia combinatoria con elementos subordinantes que introducen complementos del nombre y suboraciones adjetivas, tanto los complementos como las suboraciones presentan alta variabilidad: mientras solo cuatro de los primeros alcanzan una  $F_i \geq 2$ , ninguna de las segundas logra una  $F_i > 1$ .

Los complementos que se reiteran predicen en todos los casos de *hombre* y solo en un caso de ambos vocablos (cf. Figura 14). A saber:

*hombre + de ciencia* ( $F_i=3$ )

*hombre + de edad* ( $F_i=2$ ) y *mujer + de edad* ( $F_i=2$ )

*hombre + con respuestas* ( $F_i=2$ )

*hombre + de buena voluntad* ( $F_i=2$ )

Los cuatro presentan cierto grado de lexicalización: *de edad* se encuentra incluso en RAE (2014) como locución adjetiva sin marca diatópica; *de buena voluntad*, otra locución adjetiva, está registrada en el mismo diccionario, pero solo como locución adverbial; *de ciencia*, no incorporada en la mencionada obra lexicográfica, es también a nuestro juicio una locución adjetiva; *con respuestas* parece tener un grado menor de fijación, por lo que podría clasificarse como colocación.

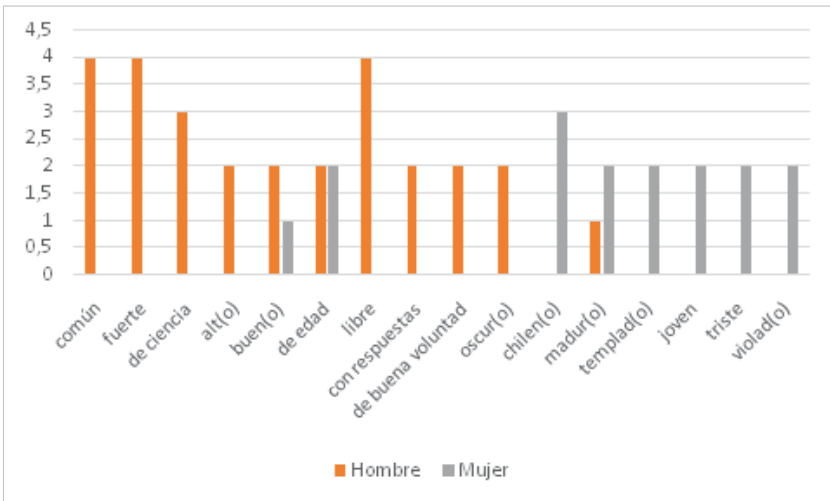


Figura 14. Modificadores del sustantivo con  $F_i \geq 2$

Al revisar las locuciones adjetivas y la colocación identificadas (cf. Figura 14), encontramos que las tres que predicen exclusivamente de *hombre* –*de ciencia*, *de buena voluntad* y *con respuestas*– son positivas, mientras la que puede predicar de *hombre* o de *mujer* –*de edad*– es neutra.

Si observamos en la misma Figura 14 los adjetivos calificativos que predicen de *hombre* y de *mujer*, podemos apreciar que solo dos de ellos se aplican a ambos vocablos, aunque con diferente  $F_i$ : *bueno* (el doble para *hombre*) y *maduro* (el doble para *mujer*). En este último caso, el contexto permite observar que *maduro* dicho de *hombre* da cuenta de una condición mental positiva, ‘Que ha alcanzado la capacidad mental propia de una persona adulta’ (RAE 2014, s. v. *maduro*, *ra*, acep. 3), mientras que *maduro* dicho de *mujer* se centra en una condición física negativa: ‘Dicho de una persona: Que ha dejado de ser joven, pero no ha llegado a la vejez’.

Los diez adjetivos restantes –la mayoría– se aplican solo a *hombre* o solo a *mujer* (cf. Figura 15). Los que predicen de *hombre* son preferentemente positivos –*fuerte*, *alto*, *libre*–. Los que predicen de *mujer* no solo son preferentemente negativos –*templada* (‘lujuriosa’), *triste*, *violada*–, sino que además incluyen dos unidades léxicas –*templada* y *violada*– que por su morfología de participio pasivo parecen dar cuenta de una mujer objeto, que recibe la acción realizada por otros.

Hombre	Mujer
común	chilena
fuerte	templada
alto	joven
libre	triste
oscuro	violada

Figura 15. Modificadores del sustantivo con  $F_i \geq 2$

## 5. CONCLUSIONES

### 5.1. ALCANCES METODOLÓGICOS

El análisis automático de nuestro corpus de referencia nos ha permitido alcanzar los objetivos planteados y comprobar nuestras hipótesis. En una futura fase, consideramos conveniente complementar este trabajo con una mayor aplicación de métodos cualitativos.

### 5.2. PRINCIPALES RESULTADOS OBTENIDOS

Se observa que los adjetivos, las locuciones y la colocación que modifican con más frecuencia a *hombre* y a *mujer* son diferentes y parecen responder a valoraciones más positivas del individuo humano masculino que del individuo humano femenino. Se confirman las tres hipótesis:

5.2.1. *Alto número de modificadores nominales con  $F_i = 1$  y bajo número con  $F_i > 1$ .*

En efecto, por su alto uso, ambos vocablos presentan mínimas restricciones combinatorias, lo que se manifiesta en el escaso número de modificadores con  $F_i \geq 2$ . No obstante, se detectan dos estructuras que modifican frecuentemente a ambos sustantivos: complementos del nombre y suboraciones adjetivas de relativo. Además, se identifican tres locuciones adjetivas, una colocación y doce adjetivos que modifican a uno o a ambos vocablos con  $F_i \geq 2$ .

### 5.2.2. Más predicaciones con valor positivo para hombre

Las tres unidades pluriverbales que predicen solo de *hombre* son positivas. La que predica de *hombre* y de *mujer* es neutra.

Respecto de los adjetivos, cinco de los que predicen de *hombre* son positivos: *fuerte*, *alto*, *bueno*, *libre* y *maduro* ('mentalmente adulto'). Sobre los que predicen de *mujer*, solo dos son positivos: *bueno* y *joven*.

### 5.2.3. Más predicaciones con valor negativo para mujer

Solo hay un adjetivo con valor negativo que predica de *hombre*: *oscuro* ('sombrio'). En el caso de *mujer*, hay cuatro: *templado* ('lujurioso'), *triste*, *violado* y *maduro* ('que ha dejado de ser joven').

## 5.3. TEMAS PARA INVESTIGACIONES FUTURAS

A la vista de los resultados obtenidos de este estudio exploratorio, parece relevante analizar los adjetivos con  $F_i = 1$  que determinan a cada vocablo estudiado y determinar no solo su carga valorativa, sino también si dan cuenta de cualidades físicas, mentales, espirituales, etc. Esto permitiría precisar la preponderancia de una actitud positiva o negativa frente a los vocablos *mujer* y *hombre* como representaciones lingüísticas de construcciones culturales de género.

Resulta relevante, además, analizar las oraciones subordinadas de relativo que predicen de *mujer* y de *hombre*, con énfasis en sus núcleos verbales. ¿Qué hacen los hombres y qué hacen las mujeres?

Estimamos de interés también determinar a qué se debe la mayor preponderancia relativa de *hombre* en el género Ensayo. ¿Obedece simplemente a un uso genérico (hombre= individuo) o da cuenta de la menor presencia de la mujer como tema *serio* o de reflexión?; esto solo podría determinarse a partir de un análisis no automático de las unidades léxicas en contexto.

Por último, esta investigación podría extenderse a otros sustantivos con flexión de género cuyo género próximo sea 'individuo de la especie humana', como *niño*, *niña*, para indagar si los resultados obtenidos aquí, en cuanto a la mayor valoración del individuo masculino, se repiten.

## REFERENCIAS BIBLIOGRÁFICAS

- ALVAR, MANUEL. 2006. Motivaciones sociológicas en el cambio lingüístico. *Biblioteca Virtual Miguel de Cervantes* [en línea]. Disponible en [http://www.cervantesvirtual.com/obra-visor/motivaciones-sociologicas-en-el-cambio-linguistico-0/html/00eda5c4-82b2-11df-acc7-002185ce6064\\_8.html#I\\_0\\_](http://www.cervantesvirtual.com/obra-visor/motivaciones-sociologicas-en-el-cambio-linguistico-0/html/00eda5c4-82b2-11df-acc7-002185ce6064_8.html#I_0_) [Consulta 12/09/2018].
- BARRIOS RODRÍGUEZ, MARÍA AUXILIADORA. 2016. Ignacio Bosque: gramático autor de diccionarios que nos hacen ser conscientes de todo lo que sabemos sobre las palabras. *Paremia* 25: 13-19.
- BOSQUE, IGNACIO. 2001. Bases para un diccionario de restricciones léxicas. *Moenia* 7: 11-52.
- \_\_\_\_\_. 2012. Sexismo lingüístico y visibilidad de la mujer. *Real Academia Española* [en línea]. Disponible en: [http://www.rae.es/sites/default/files/Bosque\\_sexismo\\_linguistico.pdf](http://www.rae.es/sites/default/files/Bosque_sexismo_linguistico.pdf) [Consulta 27/07/2018].
- CASTILLO FADIĆ, MARÍA NATALIA. 2012. *Corpus Básico del Español de Chile* ©.
- \_\_\_\_\_. 2015. Léxico Básico del Español de Chile: el proyecto. *E-Aesla. Revista digital* [en línea]. Disponible en: <https://cvc.cervantes.es/lengua/eaesla/pdf/01/51.pdf> [Consulta 24/04/2017].
- \_\_\_\_\_. (en referato.). *Léxico Básico del Español de Chile*.
- CEPAL - NACIONES UNIDAS. 2016. Tiempo total de trabajo. *Observatorio de Igualdad de Género de América Latina y el Caribe* [en línea]. Disponible en: <https://oig.cepal.org/es/indicadores/tiempo-total-trabajo>. [Consulta 15/08/2018].
- CERVANTES, MIGUEL DE. 2004. *Don Quijote de la Mancha*. San Pablo (Brasil): Real Academia Española y Asociación de Academias de la Lengua Española.
- COOPER, ROBERT. 1997. *La planificación lingüística y el cambio social*. Madrid: Cambridge University Press.
- JUILLAND, ALPHONSE Y EUGENIO CHANG-RODRÍGUEZ. 1964. *Frequency Dictionary of Spanish Words, The Romance Languages and their Structures, First Series SI*. La Haya: Mouton.
- JUILLAND, ALPHONSE, VICENZO TRAVERSA, ANTONIO BELTRAMO Y SEBASTIANO DI BLASI. 1973. *Frequency Dictionary of Italian Words*. The Hague- Paris: Mouton.
- LEMUS, JORGE ERNESTO. 2001. Sexismo en el lenguaje: mitos y realidades. *Memorias del Encuentro de la Red Centroamericana de Antropología, Asociación Salvadoreña de Antropología*, pp. 195-225.
- MANRIQUE SABOGAL, WINSTON. 5 de marzo de 2012. *¿La lengua tiene género? ¿Y sexo? El País* [en línea]. Disponible en: [https://elpais.com/cultura/2012/03/04/actualidad/1330896843\\_065369.html](https://elpais.com/cultura/2012/03/04/actualidad/1330896843_065369.html) [Consulta 08/12/2017].
- MARÍN, PABLO. 24 de junio de 2018. *¿Todes juntas? El lenguaje inclusivo como batalla cultural. La Tercera*. Disponible en: [https://www.latercera.com/tendencias/noticia/todes-juntas-lenguaje-inclusivo-batalla-cultural/218599/?fbclid=IwAR2GsS00Y\\_ao\\_CERAvzeSsdhTmgh3TMzTvUWM3kwu-wG48L6hJChSNOXK40](https://www.latercera.com/tendencias/noticia/todes-juntas-lenguaje-inclusivo-batalla-cultural/218599/?fbclid=IwAR2GsS00Y_ao_CERAvzeSsdhTmgh3TMzTvUWM3kwu-wG48L6hJChSNOXK40) [Consulta 27/06/2018].
- MORENO FERNÁNDEZ, FRANCISCO. 2009. *Principios de sociolingüística y sociología del lenguaje* (4ª ed). Madrid: Ariel.
- OECD. Noviembre de 2017. How's Life in Chile? *OECD.org* [en línea]. Disponible en: <https://www.oecd.org/statistics/Better-Life-Initiative-country-note-Chile.pdf> [Consulta 10/05/2018].
- REAL ACADEMIA ESPAÑOLA (RAE). 2014. *Diccionario de la Lengua Española* [en línea]. Disponible en: <http://www.rae.es/diccionario-de-la-lengua-espanola/la-23a-edicion-2014> [Consulta 01/05/2018].