

Aspectos estadísticos del español oral

Análisis del corpus sociolingüístico de Santander (España)

Inmaculada Martínez (Centro Internacional de Estudios Superiores del Español. Fundación Comillas)

Hiroto Ueda (Universidad de Tokio)

1. INTRODUCCIÓN

El propósito de este estudio es presentar el corpus sociolingüístico PRESEEA-Santander y analizar sus aspectos cuantitativos. La razón por la que utilizamos este corpus está en la representatividad que ofrece a la hora de estudiar los fenómenos lingüísticos, garantizada por la estratificación equitativa de los factores sociales. Abordaremos la frecuencia que refleja, por un lado, las características gramaticales generales y, por otro, las cuestiones temáticas específicas de cada texto. La hipótesis fundamental con la que trabajamos es la siguiente: los rasgos gramaticales obedecen a la tendencia caracterizada por el alto grado de regularidad, mientras que los tópicos específicamente tratados en distintos textos presentan frecuencias peculiares cuya regularidad es limitada. Al mismo tiempo, reflexionaremos sobre dos aspectos básicos de la gramática española: modo-tiempo y persona-número en la morfología verbal, y género y número en la morfología del sustantivo y del adjetivo. Los trataremos desde el punto de vista de la lingüística cuantitativa que proponemos al lado de la cualitativa.

2. EL CORPUS SOCIOLINGÜÍSTICO PRESEEA-SANTANDER

2.1. *El proyecto PRESEEA*

El Proyecto para el Estudio Sociolingüístico del Español de España y América (PRESEEA) nace en abril de 1993, en el *X Congreso Internacional de la Asociación de Lingüística y Filología de la América Latina* (ALFAL). En la Comisión de Sociolingüística de esta Asociación se decide iniciar un proyecto que aborde el estudio sociolingüístico de las ciudades iberoamericanas y españolas. Las intenciones generales del proyecto incluyen “llegar a ser algo tan *valioso* para el futuro conocimiento de la lengua española, como *útil* para las personas que se ocupan de ella” (Moreno Fernández 2021:5).

El proyecto PRESEEA se concreta en la creación de corpus sociolingüísticos sincrónicos de lengua española hablada, representativos del mundo hispánico en su variedad geográfica y social. En la actualidad agrupa a cerca de 50 equipos de investigación sociolingüística coordinados en torno a una misma metodología –que se detallará más adelante– y a idénticos principios teóricos de carácter sociolingüístico: la concepción del dialecto como propiedad de una comunidad de habla, la variabilidad como rasgo caracterizador de la lengua, la cuantificación como método analítico y la representatividad de las muestras de habla (Moreno *et al.* 2000: 1).

La adhesión al proyecto PRESEEA por parte del equipo de Santander tuvo lugar en noviembre de 2013 con la firma del acuerdo de colaboración que incluye los siguientes compromisos (Moreno Fernández 2021: 28): seguir las pautas metodológicas generales propuestas para el proyecto; poner los materiales recogidos dentro del proyecto a disposición del resto de los equipos de la red; ofrecer

información permanentemente actualizada sobre el desarrollo de la investigación del equipo; mencionar en todas las publicaciones derivadas de la investigación la vinculación del equipo a PRESEEA.

2.2. Caracterización social y lingüística

Santander es una ciudad situada en el norte de España, capital de la comunidad autónoma uniprovincial de Cantabria. Con 173 375 habitantes (2020), es la urbe más poblada de la comunidad autónoma (56,1 % del total de habitantes de Cantabria). Caben destacarse como rasgos identitarios sus límites con la meseta y con el mar, pues ambos determinan en el siglo XIX la verdadera expansión urbana de Santander: es notable el auge del comercio de harinas procedentes de Castilla, así como el aumento del tráfico en el puerto de Santander hasta el punto de llamársele el “Liverpool de España”.

Santander muestra el rostro de una ciudad dedicada al sector terciario, convertida en centro administrativo, comercial, financiero, turístico y cultural; es, por tanto, escasa la implantación industrial dentro de los límites municipales, ya que la mayoría de las empresas han ubicado sus factorías en los municipios vecinos. Es básicamente una ciudad del siglo XX, pues es en este siglo cuando se construye la mayor parte del tejido urbano actual (Ayuntamiento de Santander 2020: 14). Es, por tanto, una ciudad relativamente joven, en comparación con otras capitales españolas. El proceso de terciarización se ha realizado a base de cambios morfológicos y demográficos.

Las características de esta estructura demográfica son, principalmente: la escasa proporción de población correspondiente a grupos de edad jóvenes (0-14 años), unida al marcado peso de la población perteneciente al grupo de personas mayores de 65 años; ello provoca un índice de envejecimiento municipal levemente superior a la media regional y claramente superior a la media del país (Ayuntamiento de Santander 2020: 221). El relevo generacional, por tanto, corre un serio riesgo.

Todos estos rasgos parecen presentar una caracterización lingüística de la ciudad de Santander determinada, en primer lugar, por una fuerte conciencia de habla santanderina, así como por la homogeneidad y la ausencia de mezclas lingüísticas. Su situación geográfica en el punto central de la provincia y mirando al mar ha provocado el hecho de que no se produzcan trasvases lingüísticos significativos desde las comunidades limítrofes (Asturias y País Vasco).

2.3. Metodología

La muestra para el estudio sociolingüístico de la ciudad de Santander se ha conformado siguiendo los requisitos mínimos contemplados en la metodología común del proyecto internacional. En primer lugar, los sujetos seleccionados para el corpus pueden ser monolingües o bilingües. Por la idiosincrasia particular de nuestro núcleo urbano, tan solo se han contemplado los informantes monolingües. En segundo lugar, los núcleos de población en los que se recogen los datos deben corresponderse con núcleos de población asentada: Santander posee conciencia de comunidad de habla acreditada, en torno a la cual surge el hablante santanderino.

Por último, el muestreo debe ser representativo del universo que sirve de base al estudio sociolingüístico. Se trabaja con “muestras por cuotas con afijación uniforme” (López Morales 1994), a

partir de tres variables: sexo, edad y nivel de instrucción. Por “sexo” la población queda agrupada en Hombres (H) y Mujeres (M); en la estratificación por “edad” los informantes se distribuyen en E1 (de 20 a 34 años), E2 (de 35 a 54 años) y E3 (de 55 años en adelante). La estratificación por grado de instrucción se delimita en tres niveles: N1 (educación básica, hasta la edad de 10); N2 (educación secundaria hasta la edad de 16-18) y N3 (educación superior, hasta la edad de 21-22).

2.4. Materiales: la encuesta

Con respecto a los materiales que componen la entrevista, se ha seguido, en todo momento, el modelo trazado –y depurado– por el equipo central de la Universidad de Alcalá, las directrices establecidas para su realización, así como los módulos temáticos predeterminados. De esta manera, se adoptó el criterio de que entrevistador y entrevistado fuesen siempre del mismo sexo. Las grabaciones se realizaron mayoritariamente en los domicilios de los entrevistados, los cuales amablemente ofrecían sus viviendas como el lugar idóneo donde conseguir la confianza y el silencio necesarios para la correcta grabación y posterior transcripción¹. Las conversaciones tuvieron lugar con la grabadora a la vista, cerca de los interlocutores, y todas ellas sobrepasaron con creces el tiempo mínimo exigido de 45 minutos. Ello reveló que el entrevistado hablaba sintiéndose cómodo en la conversación sin alterar, por tanto, su registro habitual.

A través de los módulos temáticos que integran la entrevista, y en el marco de la sociolingüística variacionista, se obtienen todo tipo de variables para el estudio en los planos fonético-fonológico, léxico-semántico y morfosintáctico. A modo de ejemplo, destacamos a continuación la relación entre estos módulos y algunas de las variables morfosintácticas que provocan:

Módulo	Variables morfosintácticas
Saludos	Formas de tratamiento, con especial atención al uso de <i>tú / usted</i> .
El tiempo	Usos de los verbos <i>haber</i> y <i>hacer</i> .
Lugar donde vive	Usos de los verbos <i>haber</i> y <i>estar</i> .
Familia y amistad	Usos de los verbos <i>ser</i> y <i>estar</i> .
Costumbres	Impersonalidad.
Peligro de muerte	Valores y usos del sistema verbal: indicativo, subjuntivo / condicionales
Anécdotas importantes en la vida	Valores y usos del sistema verbal: presente / pretérito con valor de pasado.
Deseo de mejora económica	Valores y usos del sistema verbal: indicativo, subjuntivo / condicionales

Por último, solo cabe resaltar que el seguimiento escrupuloso de los criterios metodológicos establecidos por el equipo central se ha llevado a cabo con el mayor grado de exigencia posible. Solo de

¹ El proceso de transcripción ha podido llevarse a cabo gracias a la financiación recibida de la Fundación Comillas.

esta manera se conseguirá el rigor en los estudios de cotejo (estudios coordinados) de la variedad santanderina con el resto de variedades del español representadas en el proyecto PRESEEA².

3. ANÁLISIS ESTADÍSTICO

Los 18 informantes estratificados por sexo, edad y niveles educativos (sección 2.4) presentan distintos números de palabras en la encuesta correspondiente. Los debemos tomar en cuenta a la hora de calcular la frecuencia normalizada, que presentamos en este estudio³:

Sexo:	H: hombre			H.total	M: Mujer			M.total	Total
Edad:Nivel	E1	E2	E3		E1	E2	E3		
Nivel: N1	6 380	6 395	8 160	20 935	6 347	6 971	5 141	18 459	39 394
Nivel: N2	4 250	8 319	7 633	20 202	3 337	10 278	5 459	19 074	39 276
Nivel: N3	7 699	5 437	5 888	19 024	4 343	3 710	9 352	17 405	36 429
Total	18 329	20 151	21 681	60 161	14 027	20 959	19 952	54 938	115 099

3.1. Verbo

En primer lugar, presentamos la frecuencia de las categorías gramaticales en su totalidad:⁴

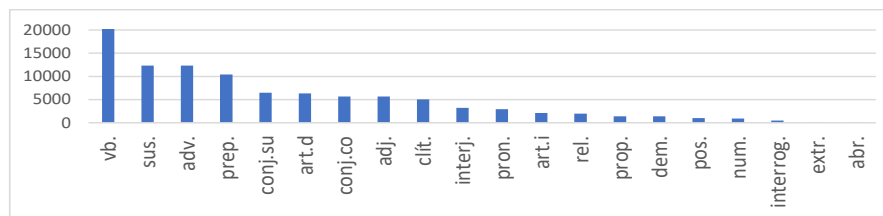


Fig. 1. Categorías gramaticales (total).

De esta manera, sabemos a ciencia cierta que la categoría gramatical más frecuente es el verbo (vb.), seguido del sustantivo (sus.), adverbio (adv.), así sucesivamente. Este gráfico muestra la tendencia gramatical, general y regular, del uso de cada categoría dentro de la totalidad de palabras.

Para confirmarlo desglosamos la frecuencia en distintos grupos de informantes:

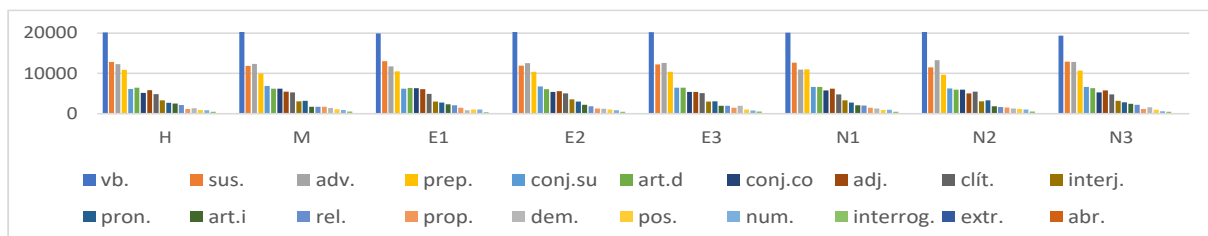


Fig. 2. Categorías gramaticales (grupos).

² Véanse los sitios siguientes: <https://preseea.linguas.net/>

<https://lecture.ecc.u-tokyo.ac.jp/~cueda/lyneal/preseea.htm>

³ Utilizamos la frecuencia normalizada (FN) por 100 000 palabras: $FN = FA \text{ (frecuencia absoluta)} * 100\ 000 / \text{Suma}$

⁴ abr.: abreviación, adj.: adjetivo, adv.: adverbio, art.d: artículo definido, art.i: artículo indefinido, clít.: clítico, conj.co: conjunción coordinada, conj.su: conjunción subordinada, dem.: demostrativo, extr.: extranjerismo, interj.: interjección, interrog.: interrogativo, num.: numeral cardinal, pos.: posesivo, prep.: prep, pron.: pronombre, prop.: nombre propio, rel.: relativo, sus.: sustantivo, vb.: verbo.

Efectivamente, se observa casi total coincidencia de las curvas descendentes en todos los grupos de sexo (H, M), edad (E1, E2, E3) y nivel educativo (N1, N2, N3). Esta coincidencia no puede ser accidental sino más bien necesaria por tratarse de rasgos gramaticales, que deben de funcionar de manera homogénea y dispersa a través de las diferencias extralingüísticas.

A continuación, veamos las formas del verbo en los distintos modos y tiempos⁵:

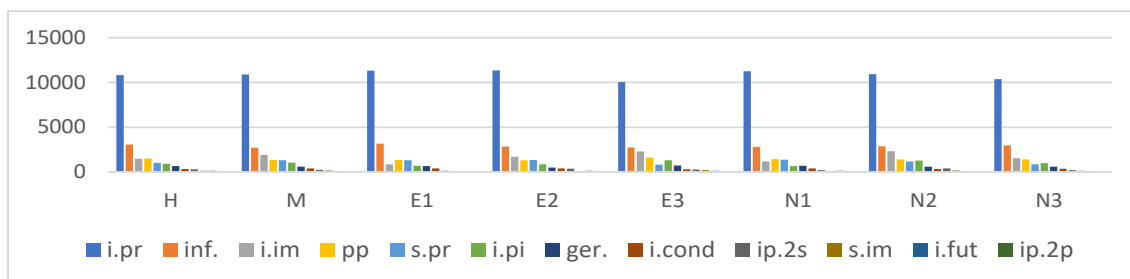


Fig. 3. Verbos. Modo y tiempo.

En todos los grupos, con escasa excepción, se mantiene la misma curva descendente, desde presente de indicativo (i.pr) hasta segunda persona del plural del imperativo (ip.2p), lo que demuestra también la tendencia gramatical.

El orden de frecuencia de [i.pr] (indicativo, presente) → [i.im] (indicativo, imperfecto) → [s.pr] (subjuntivo, presente) → [s.im] (subjuntivo, imperfecto) parece indicar el carácter bimembre de {modo} (indicativo, subjuntivo) y {tiempo} (presente, imperfecto) a expensas de la constitución unitaria {modo-tiempo} concebida por los gramáticos tradicionales y teóricos (RAE y AALE 2009: 182). El mismo orden demuestra que el indicativo es el modo no marcado frente al subjuntivo como modo marcado; y el presente es el tiempo no marcado frente al imperfecto que sería el marcado.

Se observa la misma tendencia en los datos de obras escritas ofrecidos por Bull (1947):

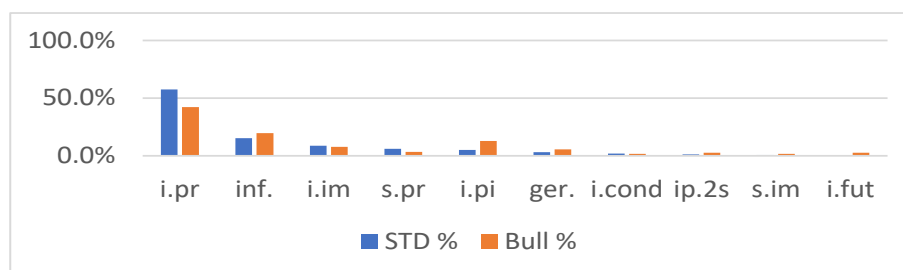


Fig. 4. Verbos. Persona y número. PRESEEA-Santander (STD) y Bull (1947).

En cuanto a la persona y número de todas las formas conjugadas, obtenemos los datos representados en el siguiente gráfico. Demuestran, también, la similitud de la curva y un alto grado de dispersión: desde la tercera persona del singular (3.s) hasta la segunda persona del plural (2.p):

⁵ cond: condicional, fut: futuro, g: gerundio, i.: indicativo, im: imperfecto, inf: infinitivo, ip: imperativo, pi: pretérito indefinido, pp: participio pasado, pr: presente, s.: subjuntivo; 2s: segunda persona singular, 2p: segunda persona plural.

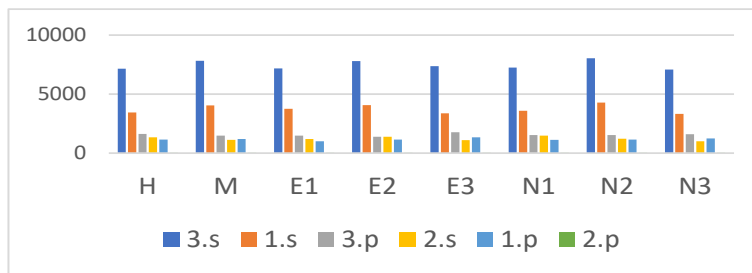


Fig. 5. Verbos. Persona y número.

En el mismo gráfico, anotamos la preferencia de la tercera persona (término no marcado)⁶ a costa de la primera y la segunda (términos marcados), por un lado, y la forma singular por encima de la plural, por otro. Ello demuestra la composición binaria de dos morfemas {persona} y {número}, antes que la unitaria de ambas categorías {persona-número}, defendida tradicionalmente por los gramáticos teóricos (RAE y AALE 2009: 182). Si la flexión verbal de persona y número se realizara entre las seis formas unitarias, [1s, 2s, 3s, 1p, 2p, 3p], debería presentarse la frecuencia de cada forma de manera desordenada. La realidad es distinta. Tanto el orden de persona (3, 1, 2) como el de número (singular y plural) son casi constantes entre los textos. Es más, dentro de cada persona (3, 1, 2), el orden de singular a plural se mantiene, y también dentro de cada número (singular y plural) es constante el orden de persona (3, 1, 2), lo que nos conduce a pensar en el carácter bimembre de {persona} (3, 1, 2) y {número} (singular, plural), en lugar del conjunto {persona-número} (3s, 1s, ...), a pesar de la inseparabilidad formal de desinencias entre persona y número como en *-o*, *-as*, ..., *-an*; *-é*, *-aste*, ..., *-aron*, etc.

3.2. Sustantivo

En cuanto al sustantivo, la situación aquí es bastante diferente de la del verbo⁷:

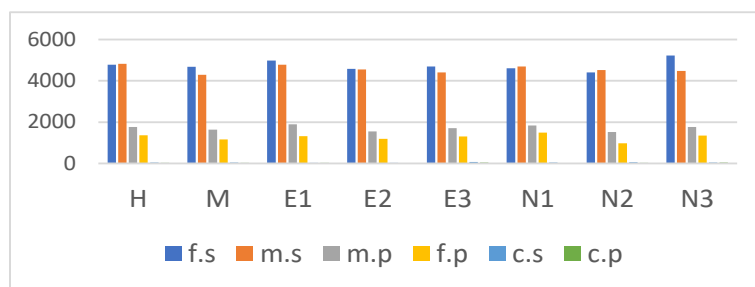


Fig. 6. Sustantivo. Género y número.

Los dos géneros, masculino y femenino, compiten en sus frecuencias, casi igualadas. Incluso, a veces, las formas femeninas superan numéricamente a las masculinas, aunque con escasa diferencia. En nuestra opinión, esta igualación numérica de uso entre los sustantivos femeninos y los masculinos se debe preferentemente a la característica léxica de género, más que a su génesis gramatical. Es decir, el

⁶ Según Alarcos Llorach (1994: 139), “se considera que hay tercera persona cuando la referencia real del sujeto gramatical no coincide ni con el hablante ni con el oyente”, lo que se contrasta con primera y segunda persona, que son términos marcados.

⁷ f: femenino, m: masculino, c: común; s: singular, p: plural.

género del sustantivo no es una cuestión de flexión nominal, sino una clase léxica propia. Si se trata de la diferencia léxica, es natural que se utilice cada palabra, tanto femenina como masculina, según la necesidad en cada acto verbal. El gráfico demuestra la vicisitud numérica de altibajos entre los textos, que tratan distintos temas, que exigen distintos léxicos y que están dotados de sus propios géneros.

El cambio de género, por ejemplo, *hijo* → *hija*, no se realiza en la flexión gramatical, sino en la derivación léxica, lo mismo que *rey* → *reina*, *actor* → *actriz*, *libro* → *librería*. Las bases de derivación, *hijo*, *rey*, *actor*, *libro*, que son términos no marcados, suelen ser más frecuentes que las derivadas, *hija*, *reina*, *actriz*, *librería*, términos marcados por la sufijación, lo que no ocurre en *pared* ni en *casa*, desprovistos del morfema de género.

En cambio, el número del sustantivo, singular y plural, sí que es una cuestión de carácter gramatical y siempre la forma singular, el término no marcado, es más frecuente que la plural, el término marcado, lo que observamos a través de los diferentes textos en el gráfico. A nuestro modo de ver, los sustantivos no flexionan en género sino solo en número⁸.

3.3. Adjetivo

La situación del adjetivo es diferente. Aquí observamos una diferencia grande de frecuencia entre las formas masculinas (no marcadas) y las femeninas (marcadas):

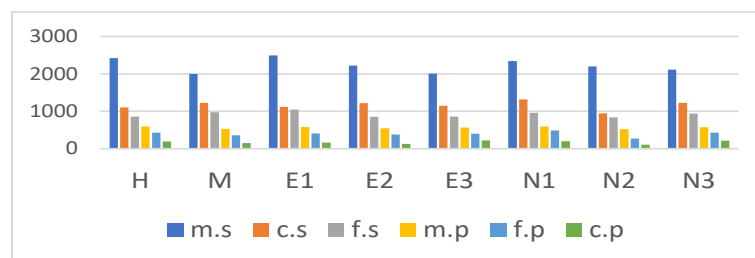


Fig. 7. Adjetivo. Género y número.

Como es bien sabido, el género del adjetivo no es inherente al léxico, sino que es concordante con el sustantivo al que modifica, de modo que es un elemento gramatical. El mayor uso del género masculino con respecto al femenino se debe al hecho de que el masculino es un término no marcado, normal de género, frente al femenino, término marcado, especial de género (RAE y AALE 2009: 85-94). El género femenino marcado aparece solo cuando modifica al sustantivo femenino, y en todos los demás casos, aparece el masculino no marcado, por ejemplo, *bueno* en el contexto de *lo ___ que*, o *es ___ que*, etc., además del concordante con el sustantivo masculino: *libro ___*. Se observa la homogeneidad perfecta de frecuencias de distintas formas de género entre los textos, lo que es natural por tratarse de variantes gramaticales.

⁸ Nuestro trabajo anterior (Moreno Fernández y Ueda 1986) explica el cambio de género de *hijo* a *hija* por medio de la derivación y no de la flexión; es decir, son dos palabras diferentes, una base y otra derivada, ambas dotadas de su propio género léxico. Véanse los amplios comentarios críticos de Ambadiang ofrecidos desde el punto de vista de la morfología teórica (1993: 95-101; 1999: 4879-4882). En los diccionarios de frecuencia, Juilland y Chang-Rodríguez (1964) y Davies (2006) tratan *hijo* e *hija* como vocablos diferentes, mientras que el de Garía Hoz (1951) los trata dentro de la misma forma léxica *hijo*.

3.4. Léxico

Por último, vamos a analizar las frecuencias de léxicos concretos pertenecientes a distintas categorías gramaticales para confirmar sus diferencias en cuanto a las características gramaticales y léxicas.

Comenzamos con los veinte verbos más frecuentes:

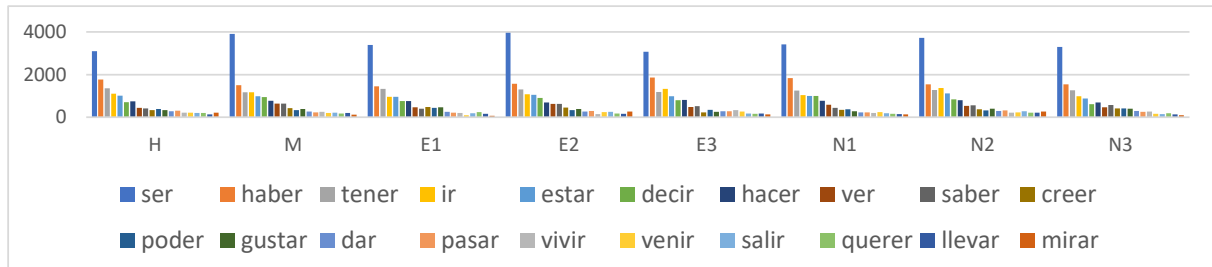


Fig. 8. Verbos frecuentes.

Es de destacar la coincidencia casi perfecta de las curvas descendentes que se observa entre los distintos tipos de textos. La razón de la coincidencia suponemos encontrarla en las propias características del léxico. Los tres verbos más frecuentes, *ser*, *haber* y *tener*, son siempre frecuentes sin distinción de tipos de textos, por ser útiles y necesarios para todo propósito de expresión. Los restantes, *ir*, *estar*, *decir*, *hacer*, ..., lo son también aunque en menor medida. Todos estos verbos frecuentes y los miles de verbos restantes pertenecen a la misma categoría gramatical de verbo. La diferencia está en su frecuencia de uso, sumamente alta, lo que demuestra el alto grado de utilidad, o instrumentalidad. Lo interesante y sorprendente es la coincidencia del orden de frecuencia y la forma de curva descendente, bastante similar entre los textos.

En esta ocasión, hemos realizado el mismo análisis con los datos de distintas obras escritas pertenecientes al teatro, novela, ensayo, periódico y documentos técnicos, ofrecidos por Juilland y Chang-Rodríguez (1964), junto con los de nuestro corpus de Santander (STD). Aquí también se observa una tendencia muy similar, especialmente entre STD y Teatro, probablemente por el carácter coloquial que comparten ambos:

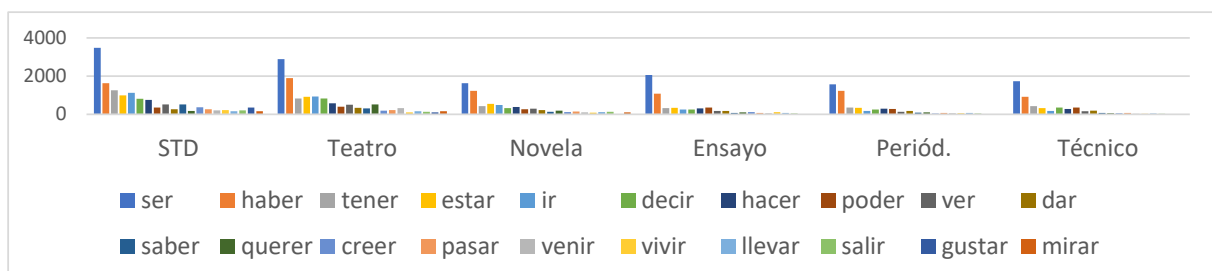


Fig. 9. Verbos frecuentes. PRESEEA-Santander (STD) / Juilland y Chang-Rodríguez (1964).

La situación en el léxico sustantivo es totalmente diferente. Aquí las frecuencias del léxico más reiterado varían considerablemente según el texto:

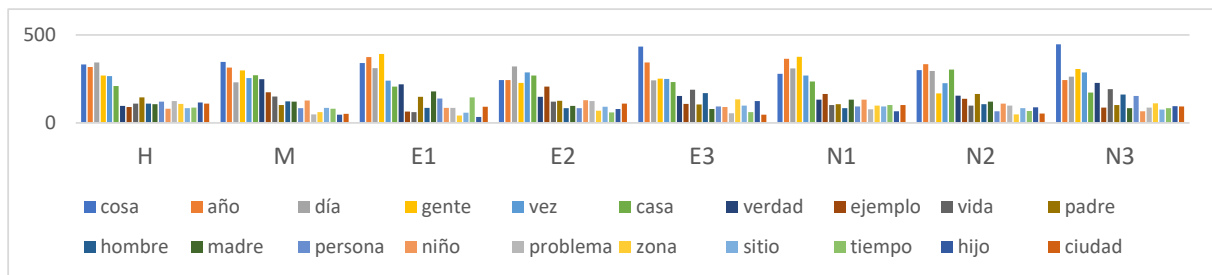


Fig. 10. Sustantivos frecuentes.

En cuanto a los adjetivos, la tendencia de frecuencia se aproxima a la de los verbos:

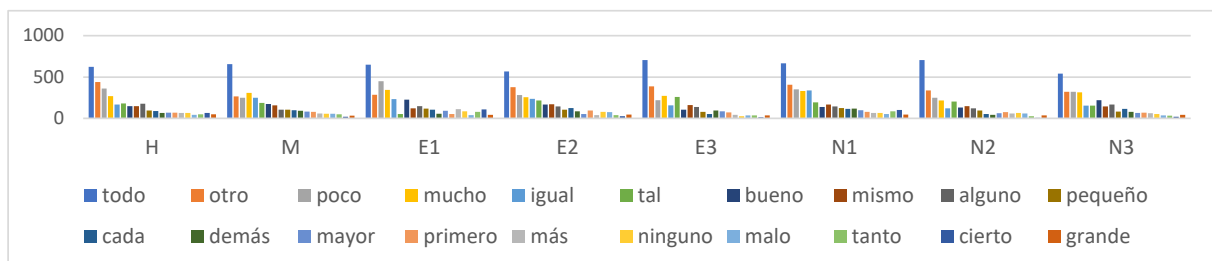


Fig. 11. Adjetivos frecuentes.

En nuestra opinión, los verbos y adjetivos más frecuentes son los vocablos instrumentales, necesarios, en general, para todo tipo de expresión; en contraposición, los sustantivos, frecuentes o poco frecuentes, son, en su mayoría, vocablos temáticos específicos cuyo uso depende del tema tratado en los diálogos o textos de distintos tipos⁹. Así, en el listado de todas las palabras, encontramos los verbos temáticos como relativamente poco frecuentes. Creemos que la distinción entre el léxico instrumental y el léxico temático es importante en el análisis práctico de la lengua.

4. CONCLUSIÓN

En este estudio breve, no hemos entrado en las discusiones teóricas sobre los temas morfológicos y léxicos de español. Nos hemos limitado a exponer los datos recogidos en el corpus oral estratificado de PRESEEA-Santander, junto con los datos ofrecidos por algunos estudios previos sobre las obras escritas. Creemos que las frecuencias de uso de cada rasgo lingüístico no pueden construir una base teórica para dar una conclusión definitiva del tema en cuestión, puesto que la frecuencia de uso es efecto y no causa o razón del sistema. En este sentido, creemos haber presentado una aproximación desde el punto de vista de la lingüística cuantitativa práctica en lugar de la cualitativa teórica. Las dos orientaciones, práctica y teórica, no deben de conducir necesariamente a las mismas conclusiones. Basándonos en la práctica de la lengua, hemos apoyado la hipótesis sobre el género léxico o derivacional del sustantivo y el género flexivo del adjetivo y la interpretación bimembre de {modo} y {tiempo},

⁹ Por falta de espacio, no hemos podido incluir los datos de sustantivos y adjetivos tratados en Juilland y Chang-Rodríguez, donde hemos confirmado la misma tendencia numérica en distintos textos escritos. Los datos de todas las palabras funcionales, artículos, demostrativos, interrogativos, preposiciones y conjunciones también manifiestan la misma tendencia de los vocablos instrumentales, que son verbos y adjetivos frecuentes.

{persona} y {número} de las desinencias verbales, pero sin rechazar necesariamente otras posiciones propias de la lingüística teórica. Se trata de puntos de vista diferentes.

Otra cuestión práctica es la instrumentalidad léxica manifestada en los verbos y adjetivos frecuentes al lado de las palabras funcionales, mayoritariamente instrumentales por excelencia. Las características de los vocablos instrumentales residen en su alto grado de utilidad a través de diferentes tipologías de texto: en textos de hablantes dotados de distintos rasgos sociolingüísticos o en textos escritos con distintos estilos, entre otros. Son utilizados con una tendencia de frecuencia parecida en diferentes tipos lingüísticos en general. En cambio, los vocablos temáticos, sustantivos y palabras de contenido de reducida frecuencia y de limitada dispersión, son empleados en cada texto en particular.

Creemos, por tanto, necesario atender a los dos aspectos de la lengua, teórico y práctico, por la razón de que ambos son complementarios y no excluyentes. En realidad, el aspecto práctico ha sido difícil de tratar hasta el momento, por falta de datos estratificados y procesados de manera conveniente. Ahora, una vez conseguido el corpus idóneo y los datos fidedignos, estamos en condiciones de realizar una colaboración ideal entre sociolingüistas y analistas de datos lingüísticos.

BIBLIOGRAFÍA

- Alarcos Llorach, Emilio (1994): *Gramática de la lengua española*, Madrid, Espasa Calpe.
- Ambadiang, Théophile (1993): *La morfología flexiva*, Madrid, Santillana.
- _____. (1999): “La flexión nominal. Género y número”, en Ignacio Bosque y Violeta Demonte (dir.) *Gramática descriptiva de la lengua española*, vol 3. 4843-4913.
- Ayuntamiento de Santander (2020): *Plan estratégico Santander 2010-2020*, recuperado en <http://www.planestrategicosantander.com/documentos.php>
- Bull, William E. (1947): “Modern Spanish Verb-Form Frequencies”, *Hispania* (American Association of Teachers of Spanish and Portuguese), vol. 30, pp. 451-466.
- Davies, Mark (2006): *A Frequency Dictionary of Spanish*. New York, Routledge.
- García Hoz, Víctor (1951): *Vocabulario usual, vocabulario común y vocabulario fundamental*. Madrid, Consejo Superior de Investigaciones científicas.
- Juilland, Alphonse and E. Chang-Rodríguez (1964): *Frequency Dictionary of Spanish Words*, The Hague, Mouton.
- López Morales, Humberto (1993): *Sociolingüística*. 2ª ed., Madrid, Gredos.
- Moreno Fernández, Francisco (2021): *Metodología del “Proyecto para el estudio sociolingüístico del español de España y América” (PRESEEA)*. Universidad de Alcalá de Henares.
- Moreno Fernández, Francisco y Hiroto Ueda (1986): “El género en los sustantivos del español: Sobre su naturaleza gramatical”, *Boletín de la Academia Puertorriqueña de la Lengua Española*, vol. 14, 79-109.
- RAE y AALE. Real Academia Española y Asociación de Academias de la Lengua Española (2009): *Nueva gramática de la lengua española. Morfología, Sintaxis I*, Madrid, Espasa Libros.