

III ANTECEDENTES BIBLIOGRÁFICOS

3.0 PRELIMINARES

En sentido estricto, la presente investigación carece de antecedentes, al menos por lo que respecta al ámbito del mundo hispánico. Según se indicó anteriormente, este es el primer estudio de tendencia, en tiempo real, de los posibles cambios del léxico disponible en una comunidad de habla en España o en Hispanoamérica. Ante la ausencia de trabajos directamente relacionados con el tema tratado aquí, a continuación se hará simplemente un breve repaso bibliográfico de los orígenes de los estudios de disponibilidad y de su trayectoria en los países de habla española.

3.1 LOS INICIOS DE LOS ESTUDIOS DE DISPONIBILIDAD

Las investigaciones cuantitativas y estadísticas sobre el léxico tienen ya una respetable tradición.⁷ El interés por conocer el vocabulario usual de la comunidad, de un grupo o un país, motivó los primeros análisis léxico-estadísticos que, en sus inicios, solo consideraban la frecuencia de aparición de las unidades. Existía la creencia de que la frecuencia era un reflejo fiel de la utilidad o de la importancia de una palabra. Sin embargo, pronto se descubrió que la medición exclusiva de este factor no resultaba satisfactoria.⁸ Una palabra puede aparecer con frecuencia en determinado tipo de texto y no aparecer nunca en otros textos. Se acudió entonces al análisis de la *dispersión* dividiendo el universo léxico en cinco mundos. La ponderación de la frecuencia con la dispersión permite descubrir el índice de uso de cada palabra. Así se determina qué términos pertenecen al *léxico básico* de la lengua.⁹

⁷Aunque no fue el primero en practicar este tipo de análisis, Pierre Guiraud es uno de los exponentes destacados del campo. Su libro de 1954, *Les caractères statistiques du vocabulaire*, constituye una obra de referencia obligada sobre la materia. Otro nombre asociado con estos temas es Charles Muller, de quien es muy conocida en el ámbito hispánico *Estadística lingüística*, además del artículo «Fréquence, dispersion et usage».

⁸Para una visión más detallada de estos asuntos, puede consultarse el importante estudio de H. López Morales «Lingüística estadística», que aparece en *Introducción a la lingüística actual*, una compilación de artículos de varios autores. López Morales, H. (ed.). 1983.

⁹En el mundo hispánico, uno de los primeros estudios sobre léxico básico es el conocido trabajo publicado en 1964 por A. Juilland y E. Chang-Rodríguez sobre el español de España *Frequency Dictionary of Spanish Words*. Dos décadas más tarde, Amparo Morales refina la metodología y publica en 1986 el *Léxico básico del español de Puerto Rico*. También en los 80 apareció el *Léxico básico de la lengua escrita en la República Dominicana*, de A. González, S. Cabanes y F. García.

Sin embargo, a pesar del rigor seguido en los estudios sobre léxico básico, con la utilización de muestras muy amplias de textos escritos y con la aplicación de la medida de la dispersión, se escapaban de los recuentos ciertas palabras comunes, conocidas y usadas por todo el mundo en determinadas circunstancias. Tales palabras presentan frecuencias muy bajas y, según este criterio, no califican para formar parte de la *lengua básica*, lo cual no parece sensato si en verdad se trata de vocablos conocidos y usados por todos o por la mayoría de los hablantes. Este hecho sirve para demostrar que, efectivamente, no hay una correlación directa entre frecuencia y utilidad de un término, de tal suerte que muchos elementos léxicos útiles pueden tener una frecuencia muy baja.

Varias investigaciones ilustran este fenómeno. En las listas de frecuencia que servían de base para la redacción del *Français Fondamental* en la década de los 50 no figuraban *autobus*, *fourchette* ni *lettre*, palabras usuales y familiares para cualquier francoparlante. De modo similar sucede en varios estudios realizados sobre el español en época más reciente. Están ausentes en el *Léxico básico del español de Puerto Rico* palabras como *bicicleta*, *correa*, *lápiz*, *mecedora*, *queso* y *uña*. En el *Diccionario básico del español de México*, no aparecen *enfermera*, *helicóptero*, *melón*, *motocicleta*, *salchicha*. Pasa lo mismo con *abrigo*, *bombillo*, *costilla*, *cuchara*, *culebra*, *lechuga*, *mochila*, *semáforo*, por ejemplo, que no figuran en el *Léxico básico de la lengua escrita en la República Dominicana*.

Ese tipo de limitación de los estudios de frecuencia llevó a varios investigadores franceses a elaborar el concepto de *léxico disponible*, que se entiende como el vocabulario utilizable en una situación comunicativa determinada. Las palabras disponibles son aquellas que surgen en torno a un tema o centro de interés cuando la conversación se refiere al mismo. Dicho de otra manera, una palabra se considera disponible si, en una situación particular, acude o se presenta fácilmente y de forma inmediata a la mente del hablante.

El iniciador de este nuevo enfoque en los estudios léxicos fue el investigador francés René Michéa (1953). De acuerdo con el autor, se considera disponible una palabra que, sin ser particularmente frecuente, siempre está lista para ser usada y se presenta inmediatamente a la memoria cuando el hablante la necesita. En este sentido, desde la perspectiva del usuario de una lengua se podría decir que es una palabra potencial que se actualiza cuando este se encuentra en una situación que requiere su utilización. Otros lingüistas que también han contribuido al desarrollo de estos estudios son Gougenheim (1967), Dimitrijevic (1969), Mackey (1971).

Según lo expuesto, se debe distinguir entre las palabras *frecuentes*, que aparecen en cualquier tipo de texto con independencia del tema tratado (son *atemáticas*) y las palabras *disponibles*, las que aparecen en situaciones concretas, específicas,

cuando el tema de la conversación las requiere (son *temáticas*). El primer grupo está constituido por formas gramaticales (artículos, preposiciones, conjunciones, etc.) en casi un 50%. El resto de las palabras son, en orden decreciente, verbos, adjetivos y sustantivos. En cambio, el segundo grupo está integrado casi exclusivamente por sustantivos. La razón de estas diferencias, insiste Michéa (1953:342), es que frecuencia y disponibilidad son de naturaleza distinta. Entre una y otra hay una fuerza de selección y de adaptación que lleva a emplear solo las palabras que son útiles en una situación específica.

La oposición entre las palabras disponibles y las frecuentes resalta de manera especial en campos léxicos estrechamente organizados, como los nombres de los meses. Puede suceder, por ejemplo, que en un corpus gigante el término *diciembre* tenga una frecuencia mayor que *mayo*. Pero cuando el hablante dice *diciembre*, el término *mayo* y los nombres de los otros meses están igualmente presentes en su mente porque los 12 forman una misma serie asociativa. Si elige uno en particular es porque se adapta a la situación concreta del momento. La suma de ambos conjuntos de vocabulario usual, el léxico frecuente (o estable) y el léxico disponible (o inestable), constituye el *léxico fundamental* de la lengua. Según esto, se puede decir que los estudios sobre disponibilidad léxica y los índices de frecuencia son mutuamente complementarios.

3.2 LOS ESTUDIOS DE DISPONIBILIDAD EN EL MUNDO HISPÁNICO

Las investigaciones sobre disponibilidad léxica en el mundo hispánico han experimentado un auge notable durante las últimas décadas. Ese florecimiento ha tenido sus raíces en la labor desarrollada por Humberto López Morales, quien, con el ejemplo de su incesante trabajo y con la coordinación de un amplio proyecto panhispánico, ha servido de inspiración a muchos investigadores de ambos lados del Atlántico. Y en esta ocasión no podía ser de otra manera: ha sido precisamente el léxico de Puerto Rico el que primero fue objeto de los análisis de disponibilidad. En un trabajo pionero, López Morales examina las diferencias y las semejanzas léxicas entre los hablantes de diversos estratos socioeconómicos de San Juan de Puerto Rico.¹⁰ Dos décadas después, con un intermedio de distintos artículos que publica acerca del tema, aparece su obra *Léxico disponible de Puerto Rico* (1999a). Otros investigadores incursionan también en el campo de la disponibilidad léxica puertorriqueña. Lorán (1983,

¹⁰El artículo *Disponibilidad léxica y estratificación socioeconómica* aparece recogido en su libro *Dialectología y Sociolingüística. Temas puertorriqueños*. 1979.

1987) escribió dos informes técnicos sobre el tema en la Universidad de Puerto Rico, Mayagüez; Román (1985) escribe *Disponibilidad léxica en escolares de Dorado*, tesina de maestría para la Universidad de Puerto Rico; y Butrón presenta su tesis de doctorado para la misma universidad en 1987: *El léxico disponible: índices de disponibilidad*.

De la disponibilidad léxica en el español de México se han ocupado varios investigadores, pero de manera especial López Chávez, quien, junto a J. Strassburger, discute la adecuación de las fórmulas utilizadas en los análisis de disponibilidad y propone una nueva en el estudio *Otro cálculo del índice de disponibilidad léxica*. A este trabajo, realizado en 1987, ha seguido la serie de publicaciones en las que, en colaboración con otros autores, López Chávez estudia el léxico disponible en niños de diferentes grados, desde el preescolar hasta el sexto grado de primaria. Son siete volúmenes publicados en 1993 por la Editorial Alhambra Mexicana. Más recientemente, el autor publica *¿Qué te viene a la memoria? La disponibilidad léxica: teoría, métodos y aplicaciones* (2003). También han incursionado en el tema H. Justo Hernández (1986), que trabaja el centro de interés de los colores, Ruiz Basto (1987), Meza Canales (1987), Núñez Alonso (1991), Arjona Iglesias (1999), entre otros.

En Costa Rica, Sánchez Corrales y Murillo Rojas (1993) estudian la disponibilidad léxica de alumnos de nivel preescolar; Murillo Rojas (1993, 1994) también estudia el léxico disponible de preescolares en varios centros de interés.

Por su parte, en Chile cabe destacar la labor de Echeverría (1991, 2000) desde la Universidad de Concepción, y los trabajos de Valencia (1994, 1997, 1998). Unos años más tarde, aparece *Disponibilidad léxica en estudiantes chilenos* (1999), una obra de Alba Valencia y Max Echeverría, que ofrece los datos de más de 2,000 sujetos representativos de todo el país. Otros trabajos sobre disponibilidad léxica en Chile son la tesina de maestría de Mena Osorio (1986), sobre la disponibilidad léxica infantil en tres niveles de enseñanza básica; la de Urrutia (2001), que estudia la relación entre la edad y el acceso léxico en tres grupos de diferentes edades de informantes de nivel sociocultural bajo; y la de Urzúa (2005) sobre el vocabulario de la carrera de Ingeniería Civil.

En la República Dominicana, se cuenta con los trabajos de Alba (1995a, 1995b, 1996, 1998, 1999). Estos estudios incluyen el diccionario (*El léxico disponible de la República Dominicana*) y análisis sobre la presencia de anglicismos, aspectos sociolingüísticos de la disponibilidad y la comparación del léxico de la República Dominicana con el de otras variedades hispánicas.

Ha sido, sin embargo, en España donde las investigaciones sobre disponibilidad han florecido con mayor abundancia durante las últimas décadas. Con diferente nivel de desarrollo, en todas las comunidades autónomas del país se

han realizado ya o se llevan a cabo actualmente diversas investigaciones relativas al léxico disponible.¹¹

En Castilla y León se destaca el equipo de la Universidad de Salamanca que, además de realizar trabajos particulares, ha creado y mantiene la página Dispalex (<http://www.dispalex.com/>), que es coordinada por Bartol y Hernández Muñoz. Se han publicado las obras de Galloso (2003), que recoge el léxico disponible de Ávila, Salamanca y Zamora; y de Bartol (2004), sobre el léxico de Soria. En la provincia de Burgos, el estudio correspondiente ha sido hecho por Fernández Juncal (2008), y el de Segovia, por Santiago Guervós (2008). Otro investigador destacado del equipo de Salamanca que ha estudiado diversos asuntos relacionados con el tema de la disponibilidad léxica es Borrego Nieto (2002), que también, con la colaboración de Fernández Juncal (2002, 2003), ha analizado la aplicación de la metodología y de los datos de la disponibilidad a los estudios dialectales.

Las Islas Canarias son, sin duda, otra comunidad lingüística donde han florecido copiosamente los estudios sobre el léxico disponible, debido especialmente al trabajo de José Antonio Samper Padilla y su equipo de colaboradores. Entre las múltiples aportaciones de Samper Padilla se encuentran sus trabajos de 1998, 1999, 2006. Los materiales de Gran Canaria fueron recogidos en la década de los 90; luego se hizo lo mismo con los de Tenerife y, más recientemente, Samper Hernández encabezó un amplio proyecto que permitió realizar las encuestas en las demás islas. A la misma investigadora se debe un extenso análisis sobre la evolución del léxico disponible en estudiantes de distintos niveles educativos de Gran Canaria (Samper Hernández 2009). Entre los numerosos estudios llevados a cabo con los materiales ya reunidos, se incluyen el análisis de los dialectalismos (Hernández Cabrera y Samper Hernández 2002, 2004; Hernández Cabrera y Samper Padilla 2003), y la comparación del léxico de Canarias con el de otras comunidades hispánicas (Samper Padilla 1999, Samper Padilla, Hernández Cabrera y Bellón Fernández 2003).

En la Comunidad de Madrid, Paredes y Gómez Sánchez y Guerra han completado la realización de las encuestas a más de 600 informantes. Paredes ha realizado varios estudios particulares con respecto a la presencia de anglicismos y problemas de tipo ortográfico aprovechando sus materiales de Alcalá de Henares (Paredes 2001, 2005).

Con relación a Andalucía, en la actualidad se han recogido ya los datos del léxico disponible en todas las provincias de la región. Han sido publicados los

¹¹En su trabajo sobre los estudios de disponibilidad léxica y sus contribuciones, José Antonio Samper Padilla y Marta Samper Hernández (2006) hacen un detallado recuento de las investigaciones realizadas en España. El apretado resumen que se presenta a continuación aprovecha el valioso material del que han hecho acopio los investigadores grancanarios.

léxicos de Huelva (Prado y Galloso 2005), de Málaga (Ávila Muñoz 2006), de Jaén (Ahumada 2006). Ávila Muñoz y Villena Ponsoda (eds. 2010) han realizado también un minucioso estudio sobre la variación social del léxico disponible de la ciudad de Málaga.

Además de la publicación de los diccionarios del léxico disponible de cada zona andaluza, varios investigadores han realizado estudios sobre diferentes aspectos de la disponibilidad. En este sentido, se destacan los trabajos de Prado y Galloso, de la Universidad de Huelva, quienes además de sus estudios sobre los materiales del proyecto, han reunido también el léxico disponible de estudiantes del sexto curso de primaria y del cuarto de la secundaria; y han incursionado en el campo de la enseñanza del vocabulario en español como lengua extranjera y del español como lengua materna (Galloso 2004, Galloso y Prado 2004, Prado 2005). En la provincia de Córdoba ha recopilado los materiales Bellón (2003), que además ha trabajado el tema de los anglicismos en el léxico disponible cordobés (Bellón 2004, 2005). En lo que concierne a Cádiz, cabe mencionar los trabajos de González y Orellana (2006) sobre los anglicismos en los distintos centros de interés.

En cuanto a la Comunidad de Extremadura, se debe también a las investigadoras andaluzas Prado y Galloso la recolección de los datos del léxico disponible de dicha región. Y en Murcia ha sido Ortolano la encargada de la realización de las encuestas.

El diccionario del léxico disponible de Valencia ha sido publicado por Gómez Molina y Gómez Devís (2004). Esta obra no se limita a la presentación de los materiales, sino que también incorpora una serie de pruebas sobre análisis factorial simple, análisis de regresión múltiple, coeficientes de correlación múltiple, entre otros. Sobre el léxico disponible de la provincia de Castellón, Blas y Casanova (2001-2003, 2003) han publicado varios artículos que estudian diferentes aspectos del léxico disponible de esa provincia.

En Aragón, un equipo coordinado por Arnal ha realizado y editado el léxico de la región (Arnal, Castañer, Enguita, Lagüéns y Moliné 2004). Han sido realizados diversos trabajos sobre los dialectalismos (Arnal), los extranjerismos (Castañer), la ortografía (Saura), entre otros. Enguita ha realizado también una comparación del léxico urbano y el rural.

En Galicia, López Meirama y Rojo se han encargado de la coordinación de los trabajos de recolección de los datos. En el País Vasco, Etxebarria ha realizado las encuestas en las tres provincias de la región. En Cataluña, por su parte, Serrano (2003, 2004, 2006) ha reunido los materiales correspondientes a la provincia de Lérida, y un equipo formado por Blasco Mateo, Torres, Forment y Pérez Edo (2003) ha recogido los datos de Barcelona.

Los trabajos de recolección de los materiales de Cantabria fueron iniciados por Carrera de la Red. Fernández Juncal se ha ocupado de la realización de las últimas encuestas y de la edición de los datos. A Hernández Muñoz se debe la realización de los trabajos de recogida de los materiales de las cinco provincias de Castilla-La Mancha. En su tesis doctoral de 2005 en la Universidad de Salamanca, utilizando los materiales de la región, la misma investigadora analiza la relación de los estudios sobre el léxico disponible con la psicolingüística. Con respecto a La Rioja, un equipo coordinado por Balmaseda ha llevado a término la realización de las encuestas. Por su parte, en Navarra, ha hecho lo mismo el equipo coordinado por Tabernerero en el que participan Martínez, Areta y Mangado.

