

# **La posición de los vocablos en el cálculo del índice de disponibilidad léxica: procesos de reentrada en las listas del léxico disponible de la ciudad de Málaga**

Antonio M. Ávila Muñoz  
José María Sánchez Sáez

## **Resumen**

En este trabajo reflexionamos sobre el significado de la posición de los vocablos en las listas de léxico disponible y, por extensión, sobre su importancia como factor determinante del índice de disponibilidad léxica. Nuestra hipótesis parte de que existen procesos de reentrada léxica en el experimento que representa la elaboración de las listas de disponibilidad que, como es sabido, utiliza como base metodológica estrategias de asociación de palabras. Esta hipótesis considera que, a medida que los informantes perciben que se alejan del núcleo prototípico que representa el estímulo que se les facilita para que comiencen a elaborar sus listados (centro de interés), ponen en marcha estrategias que les permiten volver a introducirse o, al menos, acercarse al prototipo colectivo que representa dicho estímulo. De esta manera, los listados de los sujetos pueden presentarse gráficamente como continuos vaivenes que tienen su referencia en el núcleo de máxima prototipicidad de cada centro de interés. En definitiva, y teniendo en cuenta la esencia misma del concepto de ‘disponibilidad léxica’, es probable que la posición en la que aparecen los vocablos en los listados de disponibilidad merezca una discusión que repercuta en el propio concepto<sup>1</sup>.

## **0. Introducción**

La elaboración de listas de disponibilidad léxica cuenta con una sólida tradición en el contexto de las investigaciones léxico-estadísticas, sobre todo, en el ámbito hispano<sup>2</sup>. Las razones de ese arraigamiento hay que buscarlas, sobre todo, en las múltiples aplicaciones que los investigadores suelen encontrar para las listas de disponibilidad: disciplinas como la sociolingüística, la etnolingüística, la dialectología, la psicolingüística o la pedagogía, entre otras, han sabido aprovecharlas de manera diversa y eficiente.

Es sabido que los estudios sobre léxico disponible surgieron a mediados de la década de los cincuenta del siglo XX en el seno de un programa lingüístico promovido

---

<sup>1</sup> El presente artículo forma parte de las investigaciones realizadas a lo largo del desarrollo del Proyecto de Investigación de Excelencia de la Junta de Andalucía sobre el *Condicionamiento Social del Léxico Disponible* (CONSOLEX, Proyecto HUM-315) y se ha beneficiado de las ayudas que lo financiaron.

<sup>2</sup> Este trabajo se basa en el estudio previo de Ávila Muñoz y Sánchez Sáez (2010). De hecho, muchos de los datos presentados ahora pueden considerarse una revisión ampliada de los manejados entonces con objeto de demostrar algunas de las hipótesis en las que se basaba aquel estudio.

por la UNESCO con objeto de facilitar el aprendizaje de la lengua francesa a los habitantes de las colonias y a los inmigrantes no nativos desplazados a Francia. Se pretendía facilitar la integración social de estos ciudadanos a partir del dominio de una «lengua base» simplificada desde el punto de vista léxico y gramatical<sup>3</sup>. Este programa educativo se centraba, esencialmente, en el vocabulario, por lo que se insistió pronto en la importancia de seleccionar los vocablos que debían integrar el soporte de este modelo de enseñanza.

## 1. Disponibilidad léxica

### 1.1. *Concepto*

Los léxicos disponibles son considerados el reflejo del caudal léxico que los hablantes usamos en determinadas situaciones comunicativas. De hecho, este tipo de listado explota el concepto de ‘situación frecuente’. En realidad, el léxico de disponibilidad encuentra su sentido en la evidencia de que ciertas palabras muy usadas en una lengua determinada están estrechamente relacionadas con la aparición o no de determinados temas a lo largo de las interacciones comunicativas: palabras como *escenario*, *proscenio*, *actor* o *platea* suelen aparecer solo en intercambios comunicativos donde se traten asuntos relacionados con el teatro.

Para la recolección del vocabulario disponible, aunque no necesariamente frecuente, se parte de la elaboración de unas pruebas asociativas que giran en torno a unos estímulos o centros de interés<sup>4</sup>. Alrededor de estos núcleos temáticos surge determinado vocabulario relacionado que, se supone, es el léxico potencial que pertenece al acervo activo de los sujetos y del que estos hacen uso si una conversación

---

<sup>3</sup> En realidad, desde los orígenes de la República el sistema educativo francés manifestó una evidente desigualdad social que se ha intentado compensar en diferentes momentos de la historia. *Vid.* Balibar y Laporte (1974).

<sup>4</sup> En los distintos estudios sobre disponibilidad léxica, los centros de interés han sido diversos: el cuerpo humano, el correo, la prensa, la agricultura, los sindicatos, las finanzas, la meteorología, el vestuario, los transportes, el circo, las profesiones, la ganadería, etcétera. En realidad, la metodología de recogida de materiales aprovecha una tradición bien consolidada en el ámbito de la psicología experimental. De hecho, el asociacionismo de palabras está ligado en sus orígenes a las teorías de Freud. Aunque la psicología posterior modificó los planteamientos de base, el método ha seguido empleándose hasta nuestros días para la investigación en la construcción de redes léxicas (Aitchison, 1987; Nelson, Bennet y Leibert, 1997; Nelson, McEvoy y Denis, 2000). Para un análisis crítico de la prueba usada en la producción del léxico disponible, sus mecanismos y su evolución en la psicología experimental, así como sus repercusiones en la lingüística, véase Hernández Muñoz (2006: 85-125).

en un momento dado discurre por los cauces necesarios<sup>5</sup>. El número total de vocablos que componen estos listados de léxico disponible está normalmente limitado, bien por el tiempo de reacción de los hablantes, bien por el cierre de las listas tras una cantidad determinada de palabras. Se supone que son más disponibles las palabras que acuden antes a la memoria, es decir, aquellas que ocupan los primeros lugares de las listas. Así pues, el léxico disponible forma parte del lexicón mental de los hablantes, pero no suele actualizarse en los intercambios lingüísticos cotidianos, a menos que exista una especialización temática.

## 1.2. Tradición

El trabajo que sirvió de base, modelo y punto de partida para marcar la pauta de los estudios posteriores sobre el léxico disponible fue *L'élaboration du Français Fondamental*, realizado por Gougenheim, Michéa, Rivenc y Sauvageot en 1964. Las posibilidades de este tipo de estudios fue rápidamente entendida por los estudiosos que fueron realizando investigaciones sobre disponibilidad léxica de diferentes lenguas.

En el ámbito hispano, los trabajos que más han ayudado al asentamiento definitivo de la tradición de la disponibilidad léxica han sido los elaborados y amparados por Humberto López Morales<sup>6</sup>. A ellos debemos la puesta en marcha del Proyecto Panhispánico de estudio sobre la Disponibilidad Léxica (PPHDL), cuyo fin último es obtener un diccionario del léxico disponible del español a partir de la elaboración de listas de disponibilidad léxica en diferentes lugares hispanohablantes repartidos por todo el mundo.

---

<sup>5</sup> Para una crítica del concepto de 'léxico disponible', véase Hernández Muñoz (2006: 29-36). Allí se pone en duda que el léxico disponible sea realmente el caudal de léxico utilizable en una situación comunicativa dada, pues, tal y como señala la autora, no podemos olvidar que la tarea de producir el léxico disponible constituye un método artificial de activación y enumeración de palabras. Siguiendo el esquema teórico de Marconi (2000), Hernández Muñoz acepta que los hablantes manejamos dos tipos de destrezas diferentes a la hora de producir léxico: las inferenciales y las referenciales. Para la producción del léxico disponible se aplican las destrezas inferenciales basadas en el conocimiento de la red de conexiones establecidas entre una palabra y el resto de los elementos del mismo sistema lingüístico. Ello provoca que no sepamos con seguridad si los hablantes son capaces de utilizar con eficiencia esas palabras en el discurso cotidiano, ya que para incluirlas en las listas solo requieren el conocimiento básico de que pertenecen a cierta clase de elementos, pero no la información necesaria para su aplicación referencial, necesaria para actualizar las palabras en situaciones comunicativas reales. Cf. Hernández Muñoz (2006: 46).

<sup>6</sup> López Morales (1973, 1978, 1979 y, sobre todo, 1999).

La importancia de este macroproyecto coordinado reside, por encima de cualquier otra consideración, en la propuesta de homogeneización y unificación de los criterios de diseño metodológico comunes, que permiten el intercambio de datos y el desarrollo de estudios comparados entre los resultados de los diferentes proyectos locales.

Los distintos grupos adscritos al macroproyecto asumen unas directrices metodológicas que pueden sintetizarse en los siguientes puntos:

1. Se trabaja con estudiantes preuniversitarios. De esta manera, se evita la contaminación «técnica» o específica de un área profesional determinada, a la vez que el estudio se centra en sujetos que se suponen suficientemente «maduros» desde el punto de vista léxico<sup>7</sup>.
2. Los materiales se obtienen mediante pruebas asociativas a partir de 16 centros de interés o núcleos temáticos sobre los que el informante debe aportar los elementos léxicos que considere relacionados<sup>8</sup>. Los centros de interés propuestos son los mismos que se utilizaron en los trabajos pioneros de Gougenheim: 01. El cuerpo humano; 02. La ropa; 03. Partes de la casa (sin los muebles); 04. Los muebles de la casa; 05. Alimentos; 06. Objetos colocados en la mesa para la comida; 07. La cocina y sus utensilios; 08. La escuela: muebles y materiales; 09. Calefacción, iluminación y medios de airear un recinto; 10. La ciudad; 11. El campo; 12. Medios de transporte; 13. Trabajos del campo y del jardín; 14. Los animales; 15. Juegos y distracciones; 16. Profesiones y oficios.

---

<sup>7</sup> Benítez (1992: 73), López Morales (1999: 28).

<sup>8</sup> No obstante, como hemos señalado, siempre han existido discrepancias sobre el número y la clase de los centros de interés apropiados para la tarea (Butrón, 1987). Quizás el planteamiento más crítico y realista pueda leerse en Hernández Muñoz (2006: 127-188) quien fundamenta su exposición en: 1) no todos los centros de interés responden a la concepción tradicional de las categorías naturales organizadas en torno a un prototipo (organización interna de los centros de interés); 2) los centros de interés pueden ser de carácter inclusivo: los vocablos son elementos del nivel de base o subordinado, o de carácter relacional: los vocablos no establecen una relación de pertenencia directa (niveles de inclusividad). Como hay ocasiones en las que esta identificación se dificulta, los sujetos pueden desarrollar estrategias de respuesta que manipulen los resultados. Todo ello lleva a Hernández Muñoz (2006: 185-186) a afirmar que ni cualitativa ni cuantitativamente existe unidad entre los centros de interés seleccionados en los estudios de léxico disponible, y, lo más importante, hacen que la disponibilidad de los vocablos pueda deberse a diferentes factores. A partir de aquí, es fácil deducir que la disponibilidad no es una propiedad fija de las palabras (probablemente ni siquiera en su naturaleza), sino que depende, en gran medida, del marco categorial. *Vid.* Hernández Muñoz (2006: 186).

3. Las listas son abiertas en el sentido de que no se limita el número de palabras aportadas, sino que el informante debe escribir todas las posibles relacionadas con el asunto propuesto.
4. La única condición en este sentido se refiere al tiempo de que disponen los individuos para confeccionar sus listas, pues este se controla para que no sobrepase los dos minutos por centro de interés.
5. Una vez obtenidas las listas léxicas se editan los materiales mediante unos criterios consensuados de lematización y las bases de datos se someten a un proceso de análisis estadístico que nos proporciona, en última instancia, las listas de disponibilidad léxica. Para ello, los investigadores cuentan con un *software* de distribución gratuita desarrollado por la Universidad de Alcalá y aprobado por el Instituto Cervantes: *LexiDisp* es una aplicación para Windows© que puede calcular la disponibilidad léxica a partir de la fórmula matemática desarrollada por López Chávez y Strassburger en 1987<sup>9</sup>.

Al disponer de información referida a las características sociales y personales de los sujetos se pueden establecer correlaciones sociolingüísticas particulares que, posteriormente, son comparadas con las obtenidas en diferentes sintopías, con lo que los estudios contrastivos enriquecen considerablemente las posibilidades de partida.

### 1.3. *Muestra*

---

<sup>9</sup> Tanto el programa como las instrucciones de manejo pueden obtenerse en el siguiente enlace: <<http://www.linguas.net/Proyectos/LexiDisp/tabid/73/Default.aspx>> [última consulta: 16 de abril de 2010]. Además, el PPHDL cuenta con un lugar de encuentro electrónico desde el cual se informa de las últimas novedades a todas las personas que investigan la disponibilidad léxica o se interesan por el estado de la cuestión: <http://www.dispolex.com/> [última consulta: 16 de abril de 2010]. En realidad, Dispolex es algo más que un simple medio de contacto, pues gracias a su interfaz, el usuario, debidamente registrado, puede incorporar sus materiales a un banco de datos general. Según sus creadores (Bartol, Hernández Muñoz, Universidad de Salamanca) con ello se contribuye a configurar un almacén panhispánico de amplia capacidad, configurado de modo que se adecue a las características de las diferentes investigaciones particulares. Por otra parte, el investigador registrado puede acceder de manera gratuita a las herramientas necesarias para realizar los cálculos más habituales de disponibilidad léxica: índice de disponibilidad, frecuencia, porcentaje de aparición, número de palabras, número de vocablos, promedios por informantes, índice de cohesión y comparaciones entre proyectos.

Los datos que sirven de base al presente trabajo fueron obtenidos y analizados a lo largo del desarrollo del Proyecto de Investigación sobre el Condicionamiento Social del Léxico Disponible (Proyecto CONSOLEX)<sup>10</sup>.

El objetivo general de la investigación consistió en la elaboración del diccionario de disponibilidad y en el estudio paralelo de la variación social del léxico disponible de los hablantes de la ciudad de Málaga a partir de una muestra de afijación uniforme<sup>11</sup>. Las variables de preestratificación fueron la edad, el sexo y la instrucción. La hipótesis de partida asumía que existe un condicionamiento estratificacional, reticular e individual del vocabulario a disposición de los hablantes, que los diferencia en el número, la adecuación y la precisión de las palabras que pueden evocar en ciertos contextos conceptuales y semántico-situacionales. Mediante el análisis cuantitativo y cualitativo de la relación entre la variación léxica y las variables de hablante consideradas tratamos de confirmar esta hipótesis general. La idea era que la comprensión completa del comportamiento lingüístico solo se obtendría mediante el análisis correlacional (estratificación), interactivo (red social) e individual/interpretativo (vida social)<sup>12</sup>.

El objetivo general se articuló en los siguientes objetivos específicos:

(1) Elaboración del diccionario de disponibilidad léxica de la muestra de hablantes de la ciudad de Málaga según los cánones del mencionado PPHDL.

(2) Análisis de la variación social del léxico disponible: estudio multifactorial del condicionamiento del léxico disponible.

(3) Comparación de los resultados de la disponibilidad léxica en la ciudad con los obtenidos de los grupos de adolescentes (tanto malagueños como de otras localidades).

La investigación se benefició de los contactos, recursos y procedimientos del PPHDL (red de investigadores, trabajo de campo, recursos analíticos: *LexiDisp*,

---

<sup>10</sup> *Vid.* nota 1.

<sup>11</sup> En realidad, el objetivo resultó ser novedoso en sí mismo pues, como acabamos de señalar, los estudios de disponibilidad léxica adscrito al macroproyecto PPHDL estudian la disponibilidad léxica en alumnos preuniversitarios.

<sup>12</sup> Villena Ponsoda (1995).

*Dispolex*), con los que el equipo estaba familiarizado<sup>13</sup>. La ejecución del Proyecto supuso cumplir el circuito metodológico completo (hipótesis, pruebas previas, cuestionario, trabajo de campo, codificación, análisis, resultados, interpretación, comparación) y se basó en la experiencia previa en investigaciones sociolingüísticas y de análisis cuantitativo.

En el desarrollo del Proyecto CONSOLEX se utilizó una muestra representativa con afijación uniforme de hablantes estratificada por la edad, el sexo y el nivel de instrucción ( $n = 72$ ), sobre una población censada en Málaga de 557.770 habitantes en 2005, lo que significa una relación muestra/población de  $1/7746$ <sup>14</sup>. En la Tabla 1 se ofrece la distribución de la muestra.

Tabla 1. Distribución de la muestra de hablantes en el estudio CONSOLEX

| Hombre     |    |    |    |       | Mujer      |    |    |    |       |
|------------|----|----|----|-------|------------|----|----|----|-------|
| Educación  | 0  | 1  | 2  | Total | Educación  | 0  | 1  | 2  | Total |
| <hr/> Edad |    |    |    |       | <hr/> Edad |    |    |    |       |
| 20-34      | 4  | 4  | 4  | 12    | 20-34      | 4  | 4  | 4  | 12    |
| 35-54      | 4  | 4  | 4  | 12    | 35-54      | 4  | 4  | 4  | 12    |
| > 55       | 4  | 4  | 4  | 12    | > 55       | 4  | 4  | 4  | 12    |
| Total      | 12 | 12 | 12 | 36    | Total      | 12 | 12 | 12 | 36    |
| <hr/>      |    |    |    |       | <hr/>      |    |    |    |       |

Se tuvieron en cuenta, además, numerosas variables de postestratificación. La hipótesis era que algunas de ellas podían tener valor explicativo de la variación léxica. La información sociológica se recogió por medio de un cuestionario específico que los sujetos rellenaban tras la prueba léxica. En él, los individuos proporcionaron

<sup>13</sup> Ávila Muñoz (2006).

<sup>14</sup> En realidad, los informantes fueron seleccionados a partir de los criterios establecidos en el marco del Proyecto PRESEEA. Los principios teóricos y metodológicos de este proyecto panhispánico pueden consultarse en el siguiente enlace electrónico: <http://www.linguas.net/portaIpreseea/Metodolog%C3%ADa/tabid/474/Default.aspx>. [consulta 26 de abril de 2010].

información referida a variables sociales, psicosociales y reticulares<sup>15</sup>.

Los datos lingüísticos, por otro lado, se recogieron según los procedimientos usuales en el ámbito del PPHDL<sup>16</sup>. En nuestro caso, a los 16 centros de interés habituales, sumamos otros cuatro específicos: 17. Dinero y economía, 18. Internet y ordenadores, 19. Los colores y 20. El mar.

## 2. El índice de disponibilidad léxica

### 2.1. Frecuencia y orden de aparición

El índice de disponibilidad es, a grandes rasgos, un referente numérico que trata de poner en relación criterios de frecuencia y orden con objeto de obtener unos resultados compatibles con el concepto expresado en la sección anterior (1.1). El cálculo matemático en el que se basa pretende ponderar de manera adecuada la frecuencia con la que se actualiza un vocablo en un centro de interés con su posición en las diversas listas. El punto de partida es una matriz similar a la que se muestra en la Tabla 2.

Tabla 2. Matriz de aparición de vocablos por sujeto.

Fuente: adaptada de Butrón (1987: 23-35)

| sujeto 1       | sujeto 2       | sujeto 3       |
|----------------|----------------|----------------|
| v <sup>1</sup> | v <sup>2</sup> | v <sup>1</sup> |
| v <sup>3</sup> | v <sup>4</sup> | v <sup>2</sup> |
| v <sup>2</sup> | v <sup>1</sup> | v <sup>4</sup> |
| v <sup>5</sup> | v <sup>3</sup> |                |
|                | v <sup>5</sup> |                |

<sup>15</sup> Los cuestionarios empezaron a pasarse en 2005. Casi todas las pruebas se realizaron en una situación de entrevista cara a cara entre el investigador y el sujeto informante. En Ávila Muñoz y Villena Ponsoda (eds.) (2010, Apéndice 1) puede consultarse el cuestionario sociorreticular completo.

<sup>16</sup> Vid. 1.2.

Los sujetos 1, 2 y 3 confeccionan listas de diferente tamaño y, como es lógico, no siempre coinciden los mismos vocablos en las mismas listas individuales y, cuando lo hacen, no siempre se encuentran en la misma posición. A partir de la matriz anterior se puede construir un vector de posiciones para cada uno de los vocablos, tal y como se refleja en la Tabla 3.

Tabla 3. Vector de posiciones de vocablos.

Fuente: adaptada de Butrón (1987: 23-35)

|       |       |
|-------|-------|
| $v^1$ | 20100 |
| $v^2$ | 11100 |
| $v^3$ | 01010 |
| $v^4$ | 01100 |
| $v^5$ | 00010 |

---

De aquí se deduce que el vocablo 1 ( $v^1$ ) ha sido escrito por el conjunto de la población estudiada dos veces en primera posición, ninguna en segunda posición, una vez en tercer lugar y no se ha escrito ni en cuarta ni en quinta posición; el vocablo 2 ( $v^2$ ) se ha escrito una vez en las tres primeras posiciones y nunca en los dos últimos lugares, etcétera.

Como se observa en este desarrollo, en la determinación del índice de disponibilidad léxica solo se considera relevante la posición de los vocablos y la frecuencia de ocurrencia en esa posición. De hecho, desde los trabajos de Lorán (1983) y Lorán y López Morales (1983) se han diseñado estrategias de ponderación de la frecuencia que, en mayor o menor medida, incluían la posición de los vocablos en las listas como factor determinante del índice de disponibilidad. Tras diversas revisiones, López Chávez y Strassburger (1987) proponen una formulación definitiva para dicho cálculo que parte de una matriz de vectores en la que participan:

- a) la frecuencia absoluta que obtiene la unidad en cada posición de las listas;

- b) la frecuencia absoluta de la unidad; esto es, la suma de las diferentes frecuencias alcanzadas en cada posición;
- c) el número de individuos considerados;
- d) el número de posiciones alcanzadas en el centro de interés analizado;
- e) las posiciones en que fue escrita la unidad.

La aplicación del factor e), elevado a un exponente complejo, es la responsable de la auténtica ponderación de la posición y lo que permite combinar convenientemente la frecuencia y la posición de cada unidad, pues ofrece una ponderación variable entre 0.1 y 1, con independencia del número de participantes, de la extensión de cada listado individual y de la frecuencia de aparición de cada elemento léxico.

La forma final de la fórmula propuesta queda expresada así:

$$d(p) = \sum_{i=1}^n \exp \left( C - \frac{i-1}{n1} \right) \frac{f_{pi}}{N_i},$$

donde  $n$  es la máxima posición alcanzada,  $f_{pi}$  es la frecuencia absoluta de la palabra en la posición  $i$  y  $N_i$  es el número de hablantes investigados.

## 2.2. El índice de disponibilidad léxica y el concepto de 'prototipicidad'

Partimos de un modelo que considera el índice de disponibilidad como un excelente indicador del grado de prototipicidad que los vocablos poseen dentro de cada uno de los centros de interés<sup>17</sup>. Según nuestra propuesta, la disponibilidad de un vocablo en un centro de interés responde, esencialmente, al concepto de 'accesibilidad'. De partida,

---

<sup>17</sup> Tal y como se expuso más arriba (nota 8), conviene no asociar el índice de disponibilidad a la simple naturaleza de los vocablos, pues, en gran medida, ese índice está relacionado, como se verá enseguida, con la categorización conceptual colectiva.

asumimos que cada centro de interés gira en torno a un prototipo creado a partir del concepto que determina ese centro de interés<sup>18</sup>.

A partir de las listas individuales, el índice de disponibilidad puede permitirnos estimar cuantitativamente la estructura de la accesibilidad al léxico para una población en un centro de interés dado. La cuantificación de esta accesibilidad es la medida del concepto de disponibilidad de cada término en ese centro de interés, una vez que se ha integrado la información proporcionada por todos y cada uno de los individuos que forman la población.

La asociación entre las palabras y la accesibilidad que presenta cada una de ellas determina la representación de la estructura del léxico en un centro de interés. En esa representación los vocablos más próximos al prototipo del núcleo temático mostrarán mayor valor de accesibilidad (o lo que es lo mismo, mayor índice de disponibilidad).

Según nuestro planteamiento, es necesario establecer un grado de ‘compatibilidad’ entre, por un lado, cada uno de los vocablos aportados por los hablantes y, por otro, el concepto que para la comunidad estudiada representa cada uno de los centros de interés propuestos. Para establecer dicho grado de compatibilidad, nos basamos en el concepto matemático de ‘conjunto difuso’<sup>19</sup>. A grandes rasgos, los conjuntos difusos son una generalización de la teoría de conjuntos en la que, en lugar de la pertenencia de los elementos, se considera la compatibilidad de estos con el concepto representado por el conjunto. De este modo, pueden establecerse diferentes niveles de compatibilidad entre un elemento y el conjunto con respecto al que se mide esa compatibilidad, en lugar de la dicotomía pertenencia-no pertenencia propia de la teoría clásica de conjuntos. Este concepto de ‘compatibilidad’ se corresponde, en nuestro caso, con el concepto de ‘accesibilidad’<sup>20</sup>. Una de las herramientas que pone a nuestro alcance la teoría de los conjuntos difusos es la determinación del valor de compatibilidad

---

<sup>18</sup> En realidad, la propuesta parte del desarrollo de la teoría de los prototipos (*Prototypes theory*: Wittgenstein, 1953; Rosch, 1978; Lakoff, 1987).

<sup>19</sup> Para una exposición de los principios elementales de la teoría de los conjuntos difusos, véase L. A. Zadeh (1965: 338-353). Además, en H. J. Zimmermann (2001) se propone un estudio exhaustivo sobre el estado de la cuestión de la teoría de los conjuntos difusos, así como de las operaciones sobre ellos basadas en t-normas y otras aplicaciones varias, como el concepto de ‘negación’.

<sup>20</sup> Como se observa, este planteamiento se estructura a partir de un marco matemático previamente establecido y contrastado que nos permite emplear las herramientas disponibles en ese marco. Ello nos facilita la resolución de otros problemas a partir del ya planteado de la simple ordenación de los términos que hasta ahora parece haber ocupado un lugar preferente en los estudios sobre el léxico disponible. Para una revisión crítica de este asunto, *vid.* Ávila Muñoz y Sánchez Sáez (2010).

característico del conjunto difuso, FEV (*Fuzzy Expected Value*) o su variación WFEV (*Weighted Fuzzy Expected Value*). Con su empleo podemos establecer, por un lado, un límite de caracterización de los valores de pertenencia y, por otro, parámetros para identificar elementos «muy característicos» o «poco característicos». Se trata, en definitiva, de proponer una marca objetiva de corte en los niveles superiores e inferiores del conjunto difuso que no dependa directamente de la percepción subjetiva del investigador; tal marca pondría en relación directa el grado de compatibilidad de cada elemento con una valoración del conjunto de los elementos seleccionados, factor este último que permite parametrizar el proceso diferenciador.

Para determinar el espectro de un centro de interés partimos de la construcción de un conjunto difuso que represente la actuación lingüística llevada a cabo por cada individuo en las listas de palabras que ha producido. Para ello se recurre a una valoración de cada término ( $t$ ) en función de dos parámetros: la posición que ocupa en la lista ( $n$ ) y una constante para cada problema ( $k$ ) que se expondrá a continuación<sup>21</sup>:

$$t = \frac{k}{n}$$

Una vez que se han determinado los espectros individuales, se procede a la construcción del modelo colectivo. Para ello, se toma como referencia el valor de nula disponibilidad para todos los términos. De esta manera, incorporamos el valor proporcionado en el espectro individual para cada término, acumulándolo al modelo colectivo siguiendo una ley que premie su representatividad si el término aparece en la actuación de ese hablante con altos índices de disponibilidad. Una ley que opera de esta forma es la ley probabilística:

$$P(t) = \frac{t}{\sum t}$$

---

<sup>21</sup> Esta propuesta mantiene un modelo cuantitativo similar al ofrecido por la ley de Zipf, empleado ampliamente en los trabajos cuantitativos en el ámbito de la lingüística.

Es decir, si la disponibilidad de un término, en un momento dado del proceso, fuera de 0.3, y para un hablante ese término tuviera una disponibilidad personal de 0.2, el nuevo valor del término sería:

$$\left| \frac{0.3 + 0.2}{2} \right|$$

Si, por el contrario, el término no aparece en el nuevo hablante, no mejoraría, al sumar cero.

El mayor inconveniente de este tipo de valoración es que nunca decrece: o no varía (si se le añade un valor nulo) o crece, aunque sea poco (si el valor que se añade es no nulo). De este modo, es fácil que los valores se acerquen tanto a 1 que se hagan indistinguibles.

El parámetro  $k$  que se tomó en el paso inicial tiene como objetivo conseguir que las valoraciones de los hablantes sean tales que se puedan acumular según la ley probabilística y que las valoraciones cuantitativas sean fácilmente interpretables. Se trata de un elemento de utilidad experimental que no condiciona en absoluto el modelo sobre el que trabajamos; tan solo ralentiza el crecimiento de los valores, con el objetivo de que estos permanezcan en unos márgenes que faciliten su comparación. La determinación de este valor  $k$  ha de responder a la pregunta: ¿en qué posición la ocurrencia de un término en las listas se puede considerar que está en el núcleo del centro de interés<sup>22</sup>?

Esto permite establecer un valor de  $k$  que, si bien en este trabajo, como hemos señalado, se ha obtenido de forma experimental, plantea una situación que puede justificarse de forma relativamente poco subjetiva, determinándose de forma unívoca su interpretación. El proceso de determinación de este valor o de cualquier otro modelo de ralentización debería obtenerse a partir de una discusión en el seno de la comunidad investigadora interesada en el asunto, a partir de los presupuestos teóricos de acceso al

---

<sup>22</sup> Se entiende por núcleo de un centro de interés el conjunto de vocablos que tengan una muy alta compatibilidad con el propio centro de interés, y, en consecuencia, está formado por los vocablos a los que se supone que todos los hablantes tienen acceso.

léxico. En nuestro caso se ha optado, dentro de las posibilidades a nuestra disposición, por la solución más simple para mantener el control sobre los valores.

En consecuencia, se ha construido un conjunto difuso con valores de pertenencia que se identifican con el concepto de disponibilidad léxica, pero estableciendo un marco de trabajo en el que se pueden aplicar herramientas que no estaban a nuestra disposición en etapas anteriores de este tipo de estudios. Si bien los resultados serán similares, a partir de ahora pueden justificarse desde el punto de vista de la construcción de un modelo compatible con la génesis de los conceptos expuestos y tratados.

Una de las aplicaciones más evidentes de este planteamiento es la determinación del léxico relevante para un centro de interés. Gracias a la herramienta *Fuzzy Expected Value* (FEV) podemos determinar el valor característico de pertenencia en un conjunto difuso con respecto a una determinada medida, que en nuestro caso será el tamaño relativo de los conjuntos de términos que superan un nivel de compatibilidad  $k$  (conjuntos de corte). El FEV es capaz de proponer un valor de pertenencia que establece un equilibrio entre el número de términos que lo supera y el propio valor. Además, si ponderamos en diferente grado la relevancia del tamaño del conjunto de corte, se pueden establecer grados de restricción en su interior que nos permitirían construir conceptos como «muy representativo» o «poco representativo».

Si aplicamos este planteamiento a los resultados que obtuvimos al confeccionar las listas de disponibilidad léxica en la ciudad de Málaga, podremos observar de manera práctica cómo se manifiesta este planteamiento teórico. La Tabla 4 muestra los vocablos encontrados en el centro de interés 03. Partes de la casa (sin los muebles) con índice de disponibilidad (ID)  $> 0,1$  y empleados por el 30 % de la muestra.

Tabla 4. Vocablos con índice de disponibilidad (ID) > 0,1 y empleados por el 30 % de la muestra en la ciudad de Málaga. Fuente: Ávila y Villena (eds.) (2010)

03. Partes de la casa (sin los muebles)

| Vocablo           | Disponibilidad | % Aparición |
|-------------------|----------------|-------------|
| cocina            | 0,79100        | 95,833      |
| (cuarto de) baño  | 0,67669        | 91,667      |
| salón             | 0,67631        | 80,556      |
| dormitorio        | 0,58151        | 76,389      |
| terraza           | 0,51210        | 86,111      |
| comedor           | 0,39799        | 54,167      |
| salita            | 0,29979        | 43,056      |
| lavadero          | 0,26359        | 45,833      |
| pasillo           | 0,25338        | 47,222      |
| (cuarto de) aseo  | 0,22899        | 33,333      |
| (cuarto) trastero | 0,20348        | 40,278      |
| entrada           | 0,20138        | 31,944      |
| recibidor         | 0,16123        | 23,611      |
| ventana           | 0,15553        | 29,944      |
| jardín            | 0,15080        | 27,778      |
| escalera          | 0,14765        | 29,556      |
| patio             | 0,13784        | 27,778      |
| vestidor          | 0,13736        | 27,778      |
| sala de estar     | 0,12755        | 20,833      |
| garaje            | 0,12447        | 25,000      |
| habitación        | 0,12011        | 19,444      |
| hall              | 0,11664        | 18,056      |
| porche            | 0,11215        | 22,222      |

Como se observa, los doce vocablos empleados por más del 30% de la población estudiada son los que alcanzan mayor índice de disponibilidad. Ello nos lleva a confirmar que los vocablos con mayor índice de disponibilidad de cada centro de interés formarán la categorización conceptual colectiva del estímulo en cuestión, en tanto que a

medida que el índice de disponibilidad de los vocablos disminuye y, por tanto, es empleado por menor número de hablantes, los elementos léxicos se van alejando del núcleo prototípico del centro de interés. En este sentido, la Tabla 5 nos presenta los veinte vocablos que en el mismo centro de interés logran menor índice de disponibilidad y, según nuestra hipótesis, están más alejados del núcleo de categorización colectiva del centro de interés.

Tabla 5. Vocablos con menor ID en el léxico disponible en la ciudad de Málaga. Fuente: Ávila y Villena (eds.) (2010)

03. Partes de la casa (sin los muebles)

| Vocablo             | Disponibilidad | % Aparición |
|---------------------|----------------|-------------|
| rodapié             | 0,00475        | 1,389       |
| cuarto de juguetes  | 0,00475        | 1,389       |
| parqué              | 0,00475        | 1,389       |
| aparcamiento        | 0,00475        | 1,389       |
| sala de costura     | 0,00475        | 1,389       |
| celosía             | 0,00475        | 1,389       |
| techo de escayola   | 0,00440        | 1,389       |
| cuarto de plancha   | 0,00440        | 1,389       |
| salón de fumar      | 0,00440        | 1,389       |
| viga                | 0,00440        | 1,389       |
| cierre              | 0,00440        | 1,389       |
| gallinero           | 0,00407        | 1,389       |
| cámara              | 0,00377        | 1,389       |
| bajante             | 0,00377        | 1,389       |
| pretil              | 0,00377        | 1,389       |
| parking             | 0,00349        | 1,389       |
| dormitorio infantil | 0,00349        | 1,389       |
| arriate             | 0,00349        | 1,389       |
| leñera              | 0,00300        | 1,389       |
| caballeriza         | 0,00278        | 1,389       |
| habitáculo          | 0,00257        | 1,389       |

El modelo que hemos construido asume que de los 107 vocablos que componen el núcleo temático, solo treinta y seis (33,6 % del total) pueden considerarse en el rango de «relevantes» para el conjunto al que representan<sup>23</sup>. Como se muestra en el Cuadro 1, de estos treinta y seis vocablos, once forman el núcleo de máxima prototipicidad: *salón, cocina, (cuarto de) baño, dormitorio, comedor, terraza, entrada, (cuarto de) aseo, salita, pasillo, hall*<sup>24</sup>. A partir de aquí aparecen otros niveles de corte diferentes que corresponden a las distintas zonas que, aun siendo muy compatibles con el estímulo de referencia, se van alejando progresivamente del núcleo central<sup>25</sup>. En cualquier caso, los vocablos que están representados en el Cuadro 1 resultan ser los más compatibles con el enunciado del centro de interés o, lo que es lo mismo, suponen la categorización conceptual colectiva del estímulo propuesto (conjunto de referencia). Con este procedimiento de selección léxica evitamos trabajar con fenómenos particulares del habla individual gracias al empleo de herramientas matemáticas fiables y objetivas que nos permiten centrarnos únicamente en hechos de norma compartida<sup>26</sup>.

---

<sup>23</sup> Como hemos señalado, el nivel de corte fue establecido de manera automática en un índice de compatibilidad de 0,34 por la herramienta *Fuzzy Expected Value* (FEV).

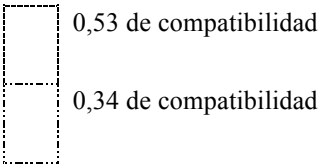
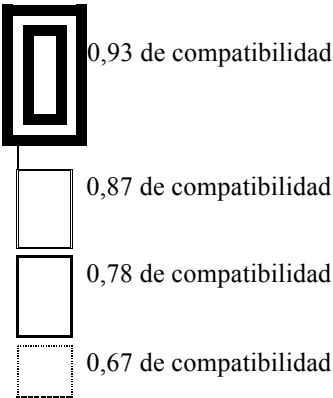
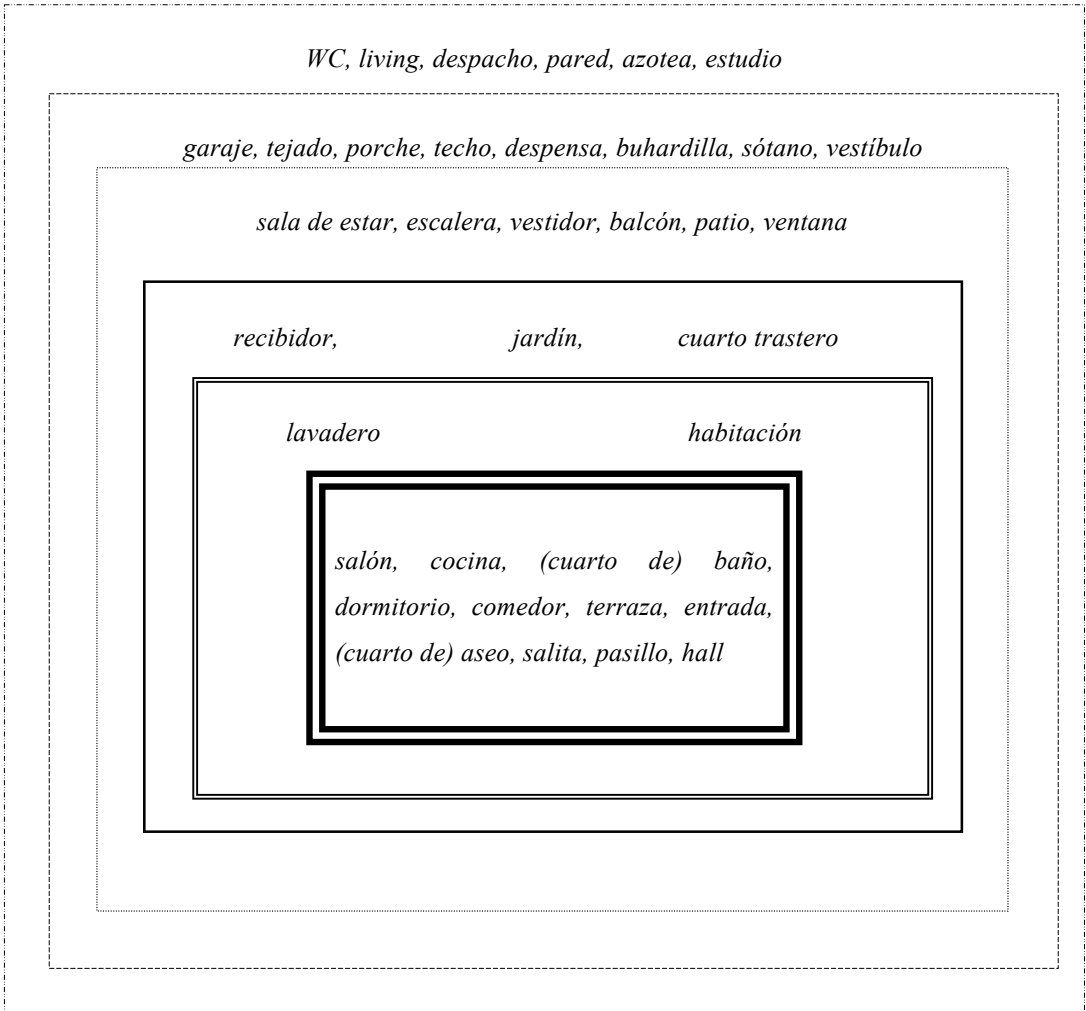
<sup>24</sup> Es decir, el 10,3 % del conjunto con un nivel de compatibilidad del 0,93.

<sup>25</sup> Como es lógico, el dibujo podría extenderse y abarcar los diferentes límites de seguridad hasta alcanzar la zona de incertidumbre límite del centro de interés.

<sup>26</sup> *Vid.* Samper Padilla (1999: 554).

Cuadro 1. Representación de la categorización conceptual colectiva del prototipo

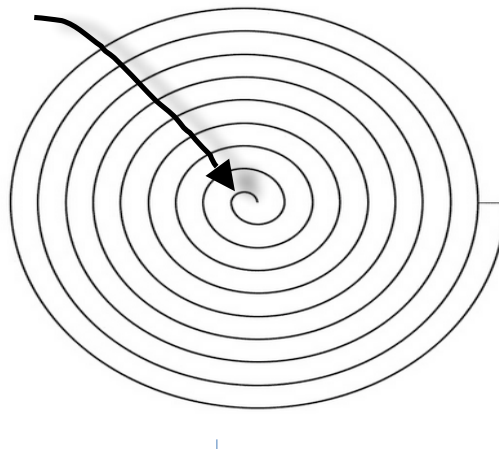
Centro de interés 03. Partes de la casa (sin los muebles)



### 2.3. Los procesos de reentrada

Cuando un individuo se somete al experimento de la disponibilidad léxica, este accede a su red léxica a partir del prototipo del centro de interés como puerta de entrada. Posteriormente, recorre la red léxica asociada, acción que lo irá alejando, progresivamente, del núcleo prototípico. Este proceso tiene la forma geométrica de una espiral en cuyo centro encontraríamos los vocablos más nucleares o prototípicos a los que recurren los sujetos cuando reciben el estímulo del centro de interés. A medida que se desarrolla la prueba, los individuos recorren esa espiral, de modo que los términos que van apareciendo sucesivamente pierden de forma paulatina esa condición de centralidad o nuclearidad (Figura 1).

Dibujo 1. Puerta de acceso a la red léxica: vocablos muy nucleares



Ocurre que, de vez en cuando, el individuo percibe que se ha alejado excesivamente del núcleo del centro de interés y busca entonces estrategias que le permitan reentrar en la red léxica original o, al menos, acercarse al núcleo prototípico que representa el centro de interés. Por ejemplo, en la lista del Ejemplo 1 se observa cómo el sujeto 019 utiliza la palabra *fruta* como puerta de acceso a su red léxica cuando recibe el estímulo ‘Alimentos’.

## Ejemplo 1. Acceso y desarrollo de una red léxica individual

Sujeto 019. Centro de interés 05. Alimentos

009 05 fruta, verdura, carne, **proteína, carbohidrato, mineral, vitamina**, *cocina*, aceite, legumbre, lenteja, huevo, pollo, **corral, campo**, *trigo*, pan, leche

Tras acceder a la red léxica mediante *fruta*, escribe *verdura* y *carne*, aunque esta última le lleva a alejarse del núcleo temático a partir de una asociación indirecta relacionada con lo que podríamos llamar «propiedades nutricionales» (marcadas en el ejemplo en negrita: *proteína, carbohidrato, mineral, vitamina*). Por alguna razón a la que no tenemos acceso, tras escribir *vitamina*, el informante decide reentrar en la red léxica propia del centro de interés, y para ello emplea una palabra «puente»: *cocina*<sup>27</sup>, que lo lleva a *aceite, legumbre, lenteja, huevo y pollo*, palabra esta última que, otra vez, lo aleja del núcleo prototípico propio del centro de interés y escribe, mediante otra asociación indirecta, *corral* y *campo*. Sin embargo, de nuevo decide entrar en la red léxica más próxima al estímulo de partida y vuelve a usar un elemento «puente» como *trigo*, que le conducirá, por último, a *pan* y *leche*<sup>28</sup>.

El ejemplo anterior evidencia que cuando los informantes se someten al experimento de la disponibilidad léxica, el estímulo que se les propone a modo de centro de interés activa una palabra muy nuclear y comienza un proceso de asociación de términos. Como señala Hernández Muñoz (2006: 87 y ss.), en el proceso de producción de las listas de léxico disponible se ponen de manifiesto dos de las tareas de activación de léxico más aceptadas en el ámbito de la psicología experimental: la fluencia categorial semántica (*category fluency*)<sup>29</sup> y la asociación libre de palabras (*free word association*)<sup>30</sup>. La fluencia semántica consiste en asociar palabras relacionadas con

---

<sup>27</sup> Uno de los mecanismos de asociación léxica observados en el trabajo de Galloso Camacho en las provincias de Ávila, Salamanca y Zamora consistía en encadenar palabras con un significante similar (Galloso Camacho, 2003: 140-142). Si bien es cierto que este procedimiento de producción léxica resultó reducido, en ocasiones se daba una «especie de rima entre los vocablos que van seguidos...» (Galloso Camacho, 2003: 141). Como advierte esta autora, la rima puede ser consonante, igual que se observa en el ejemplo que estamos comentando.

<sup>28</sup> Para una exposición detallada de los mecanismos de asociación léxica, véase Galloso Camacho (2003: 135-164).

<sup>29</sup> Warrington y McCarthy (1987); Damasio (1990); Saffran y Schwartz (1994); Garrard, Perry y Hodges (1997); Caramazza (1998).

<sup>30</sup> Deese (1962), Cramer (1968), Clark (1970).

una dada; la asociación libre consiste en actualizar las palabras que primero vengan a la mente después de escribir o escuchar otro término que actúa como estímulo. Hernández Muñoz (2006: 88) cree que la producción del léxico disponible, tal y como se concibe en la actualidad, es una prueba de fluencia semántica, aunque en el proceso subyacen algunos de los presupuestos de la asociación de palabras.

En cualquier caso, el recorrido léxico que los informantes ponen en marcha hace que, a medida que avanza la asociación de términos, el hablante pueda percibir que la cadena de asociaciones lo ha alejado en exceso del núcleo prototípico, situación que, en realidad, entra en conflicto con la esencia de la prueba. En ese momento, como hemos visto en el Ejemplo 1, pone en marcha estrategias que le permiten volver a introducirse en un punto cercano al prototipo.

Para identificar este proceso de reentrada hemos integrado en las muestras proporcionadas por cada usuario la compatibilidad de cada término con el centro de interés expuesto. Esto nos proporciona una visión de la evolución de la relevancia de cada término en el centro de interés según progresa la prueba, y permite observar la variación de las relevancias de los vocablos, así como los procesos de alejamiento del centro de interés y de reintroducción en la red léxica.

Consideremos, de nuevo, parte de la lista de palabras confeccionada por el sujeto 009 ante el estímulo 05. Los alimentos:

|                                                                                                |
|------------------------------------------------------------------------------------------------|
| 009 05 <i>fruta, verdura, carne, proteína, carbohidrato, mineral, vitamina, cocina, aceite</i> |
|------------------------------------------------------------------------------------------------|

Si añadimos a cada término de la secuencia su compatibilidad con el centro de interés, obtenemos los resultados de la Tabla 6.

Tabla 6. Compatibilidad de algunos vocablos con el centro de interés 05. Alimentos

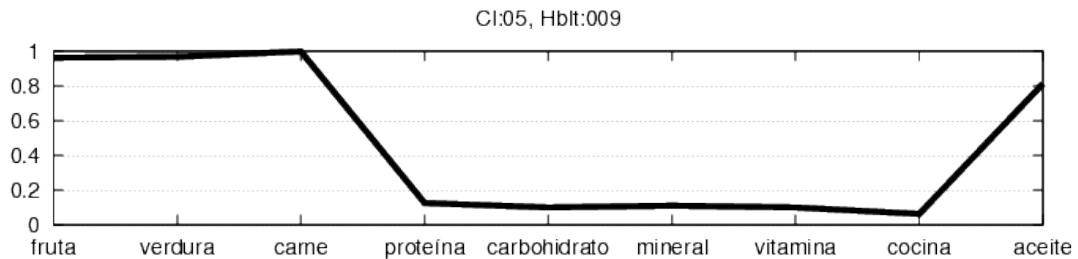
| Vocablo      | Compatibilidad |
|--------------|----------------|
| fruta        | 0.962964       |
| verdura      | 0.968115       |
| carne        | 0.999299       |
| proteína     | 0.125000       |
| carbohidrato | 0.1            |
| mineral      | 0.110294       |
| vitamina     | 0.100446       |
| cocina       | 0.062500       |
| aceite       | 0.812684       |

En la lista de base el hablante toma inicialmente elementos con una relevancia muy alta, cercana a la unidad (*fruta, verdura, carne*). Al escribir *proteína* se aleja del núcleo del prototipo, pero se reintroduce en él con *aceite*.

Podemos considerar, por tanto, que el término *aceite* ha representado una reintroducción en el léxico del centro de interés. La forma abrupta de transferencia entre ámbitos de compatibilidad de este ejemplo es poco frecuente, pero representa con nitidez la situación de evaluación numérica que estamos intentando analizar. La representación gráfica de esta situación es, si cabe, aún más evidente y se muestra en el Gráfico 1.

Gráfico 1. Transferencias consecutivas de compatibilidades.

Centro de interés 05. Los alimentos. Sujeto 009.



El elemento que consideramos clave de este análisis es el estudio de la variación en los grados de compatibilidad en la progresión de la prueba. Para el estudio de esta variabilidad basta con el cálculo de las diferencias de grado de compatibilidad en cada muestra (Tabla 7).

Tabla 7. Compatibilidad y variación de grados de compatibilidad de algunos vocablos en el centro de interés 05. Alimentos

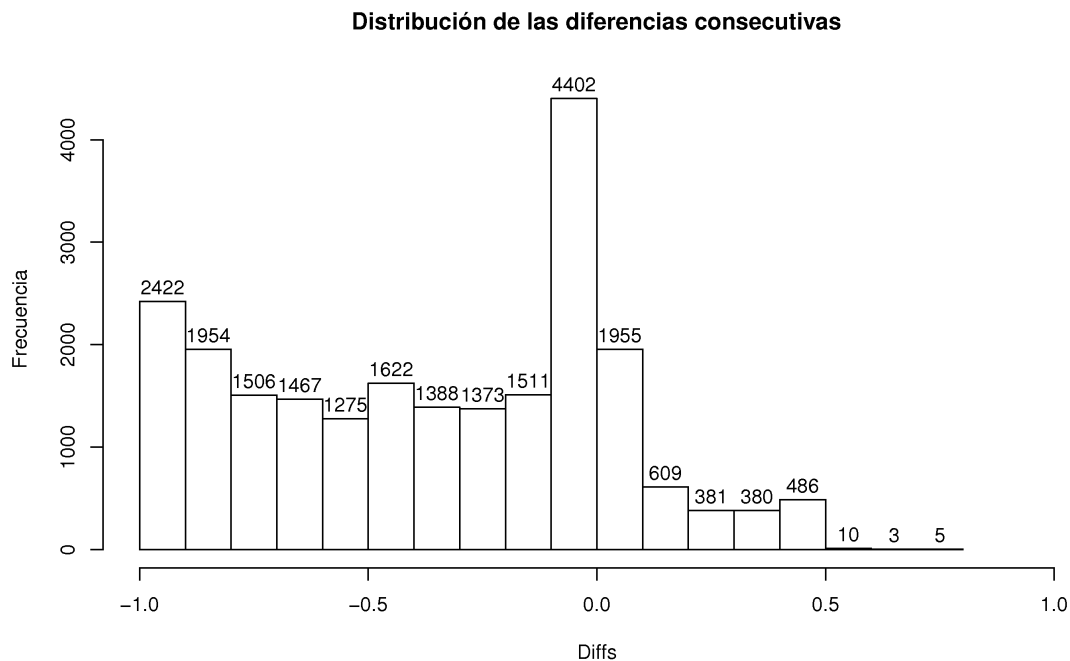
| Vocablo      | Relevancia | Variación |
|--------------|------------|-----------|
| fruta        | 0,962964   |           |
| verdura      | 0,968115   | 0,005151  |
| carne        | 0,999299   | 0,031184  |
| proteína     | 0,125000   | -0,874299 |
| carbohidrato | 0,100000   | -0,025000 |
| mineral      | 0,110294   | 0,010294  |
| vitamina     | 0,100446   | -0,009848 |
| cocina       | 0,062500   | -0,037946 |
| aceite       | 0,812684   | 0,750184  |

En el caso de que nuestra hipótesis fuese cierta, deberíamos detectar, sobre todo, movimientos descendentes, es decir, negativos, y, en menor grado, movimientos de relevancia ascendente, es decir, positivos, que implicarían reentrada en el sistema léxico.

Hemos calculado todas las variaciones de relevancia de los vocablos consecutivos en todos los centros de interés y en todos los informantes. El estudio de estas variaciones se ha realizado utilizando los métodos estadísticos clásicos. En

particular, se ha desarrollado un estudio descriptivo, ya que es lo suficientemente evidente y simple como para considerar innecesaria cualquier otra comprobación. En el Gráfico 2, a continuación, se puede observar el histograma de las variaciones obtenidas.

Gráfico 2. Distribución de las diferencias consecutivas de grado de compatibilidad o relevancia en todas las listas de disponibilidad en la ciudad de Málaga (n = 1440)



Como se observa, por un lado, casi el 20% de las variaciones (4402) son pequeñas desviaciones descendentes, entre -0.1 y 0, y es, con diferencia, la variación más frecuente, casi doblando a cualquier otra. Esto fortalece nuestra hipótesis de la mecánica individual de elaboración de las listas de disponibilidad léxica consistente en el alejamiento progresivo del núcleo del centro de interés. Además, más del 80% de las variaciones son descendentes (entre 0 y -1), es decir, movimientos a términos menos relevantes, lo cual parece, de nuevo, corroborar nuestros planteamientos de partida.

Por otro lado, cerca del 30% (27.94%) de los movimientos son de baja intensidad, es decir, son movimientos que se mantienen a una distancia muy similar del núcleo del centro de interés, lo cual, junto con la consideración de la enorme importancia de los elementos con valoración descendente, incide de nuevo en el modelo similar a una espiral que se aleja del centro de interés.

Otro dato interesante podría ser el relativamente elevado nivel de descensos de relevancia brusco ( $> -0.5$ ), lo que puede indicar el fin del vocabulario relevante que posee el individuo, es decir, un alejamiento notable y pronunciado con respecto al núcleo del centro de interés. Este hecho favorece el razonamiento a través del cual las variaciones descendentes poco acentuadas reflejan un proceso continuo de alejamiento del núcleo prototípico representado por el centro de interés: fluencia categorial (*category fluency*), en tanto que las variaciones bruscas implicarían un proceso de asociación más libre de palabras (*free word association*).

La parte derecha del Gráfico 2 representa los puntos de reentrada o de acercamiento al núcleo prototípico, como es lógico mucho menos frecuentes que los que se observan en la parte izquierda del gráfico (alejamientos progresivos). No obstante, estos mecanismos son muy útiles para los hablantes, pues, en definitiva, les permiten reintroducirse en la secuencia lógica de asociaciones próximas al centro de interés cuando consideran que se han alejado demasiado del núcleo prototípico. En este sentido, si bien es cierto que la mayoría de acercamientos se observan en la secuencia entre 0 y +0.1 (1955), correspondiendo a la elección de términos de similar relevancia al mencionado inmediatamente antes, no deja de ser llamativo el elevado número de mecanismos de acercamiento que se registran a partir de este primer estadio: 609 entre 0.1 y 0.2, 381 entre 0.2 y 0.3, 380 entre 0.3 y 0.4, etcétera.

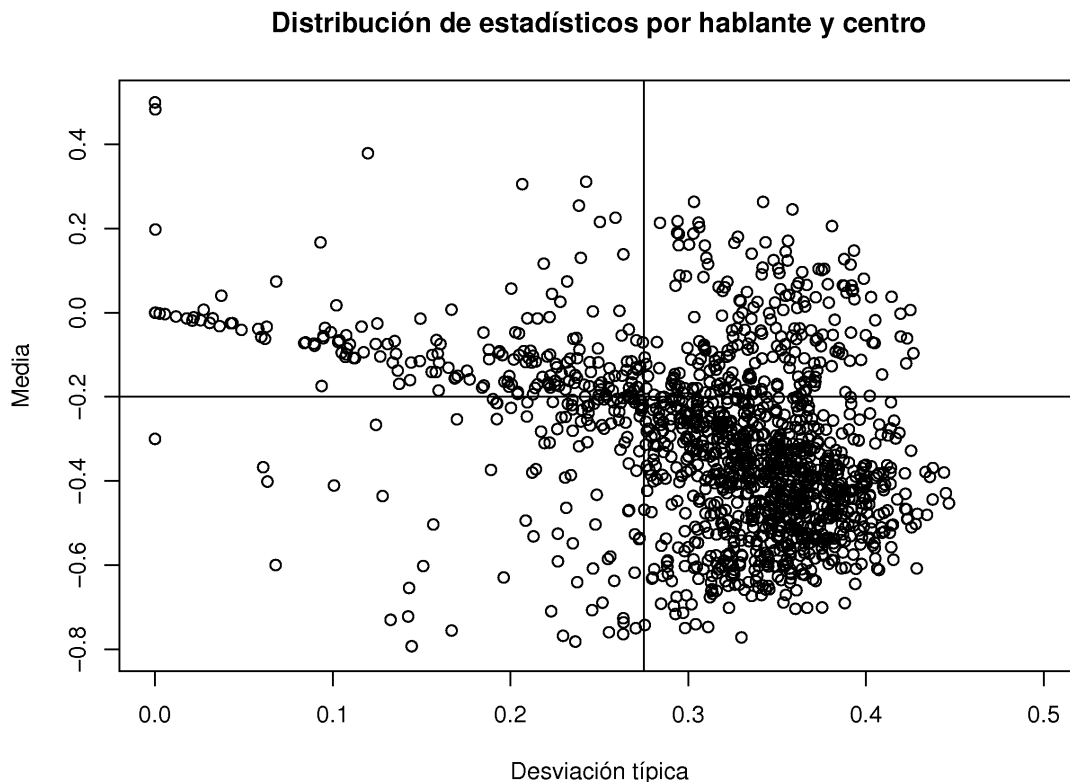
En consecuencia, parece razonable observar las series de palabras que representan las listas de léxico disponible como continuos procesos de alejamiento progresivo del núcleo prototípico, salpicado con estrategias de reentrada en momentos en que la distancia percibida por los informantes entre el núcleo del centro de interés y el lugar de la espiral al que les ha llevado el proceso de asociación de palabras entra en conflicto con la propia naturaleza de la prueba.

Este hecho queda demostrado en el Gráfico 3, generado a partir de la combinación de dos pruebas estadísticas:

1. La variación media en cada una de las listas individuales: esto es, una vez calculadas las variaciones de cada lista, la media de estas variaciones. Este dato es indicativo de la tendencia general de la prueba, en relación con el proceso de variación de la relevancia.
2. La desviación típica de la variación en cada lista: esto es, la desviación típica de las variaciones en la prueba. Este parámetro da una medida de la variabilidad de las

variaciones; esto es, si se producen de forma similar (y por tanto cercanas todas a la media) o, por el contrario, hay variaciones muy diferentes, lo que indicaría que se producen fenómenos no uniformes (en nuestro caso, procesos de reentrada).

Gráfico 3. Variación media-variación de la variación media



En el Gráfico 3 se observa una alta aglomeración de elementos con una tendencia media negativa y una variabilidad alta. En la nube de puntos, la densidad de las representaciones puede impedir la apreciación de la situación de los datos así que, como guía, hemos establecido dos límites, uno horizontal para la media en -0.2 y uno vertical para la desviación típica en 0.275 De esta manera establecemos una referencia para la discriminación de los comportamientos que nos permite realizar el recuento de los elementos que quedan en cada uno de los cuadrantes que se generan. Los datos obtenidos son los siguientes (Tabla 8):

Tabla 8. Referencia de media y desviación típica

|               | Desviación < 0.275 | Desviación >= 0.275 |
|---------------|--------------------|---------------------|
| Media >= -0.1 | 91                 | 133                 |
| Media < -0.1  | 204                | 1012                |

Aunque la división anterior solo tiene carácter didáctico, lo cierto es que, a partir de los datos expuestos, se observa una evidente aglomeración de muestras en un comportamiento claramente definido: el Gráfico 3 muestra una gran aglomeración de elementos con una tendencia media negativa y una variabilidad alta. Esto muestra: 1) que la mayor parte de las listas están compuestas a partir de una clara tendencia al descenso paulatino de la relevancia de cada término en cada uno de los centros de interés y 2) que estas variaciones tienen una elevada desviación típica. Esto último señala, a su vez, la existencia de comportamientos no uniformes a lo largo del proceso de elaboración de las listas (es decir, la tendencia del descenso progresivo de la relevancia de los términos en los centros de interés no se mantiene a lo largo del desarrollo completo de la prueba, sino que encontramos idas y venidas que refuerzan la hipótesis de los procesos de reentrada.).

### 2.3.1. Patrones dominantes de reentrada en las listas de disponibilidad léxica

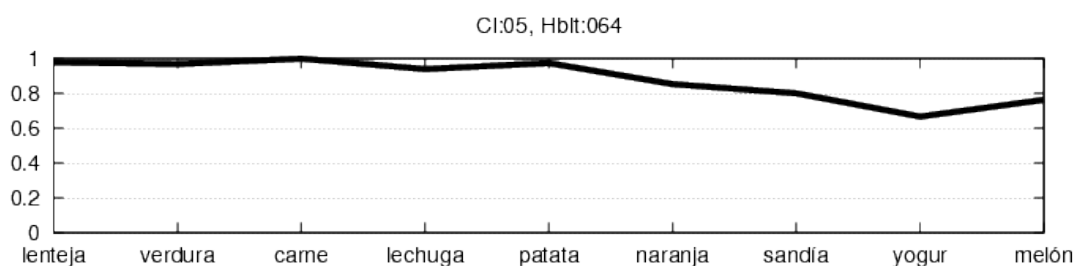
Una vez analizada la estructura de las listas y los procedimientos de aportación de léxico realizados por todos los informantes analizados en la ciudad de Málaga, puede establecerse una serie de patrones dominantes de elaboración de las listas de palabras en relación con nuestra propuesta de ‘compatibilidad’ o ‘centralidad’<sup>31</sup>.

- a. Patrón centralizado. Algunos informantes se mantienen en una línea de máxima centralidad y solo escriben palabras muy centrales, como por ejemplo, el caso representado en el Gráfico 4.

<sup>31</sup> Como es lógico, en la tipología puede haber patrones mixtos formados por la combinación de los que presentamos.

Gráfico 4. Muestra de compatibilidad de vocablos. Patrón centralizado

Sujeto 064. Centro de interés 05. Alimentos



b. Patrón de reentrada. Sin embargo, la posibilidad más frecuente es que tras la aportación de elementos muy centrales, se escriban de manera paulatina otros que lo son menos. En un momento determinado, el informante considera que se está alejando demasiado del centro del estímulo propuesto y realiza los procesos de reentrada que hemos comentado, tal y como aparece en los gráficos 5 y 6.

Gráfico 5. Muestra de compatibilidad de vocablos. Patrón de reentrada *a*

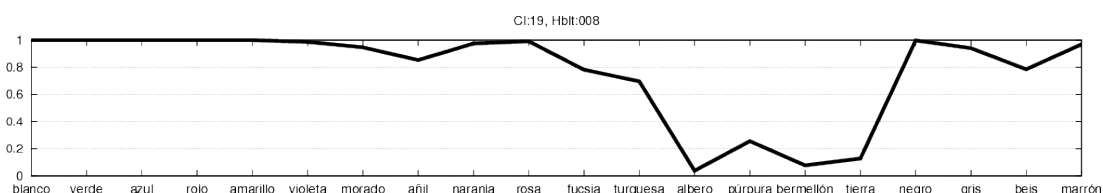
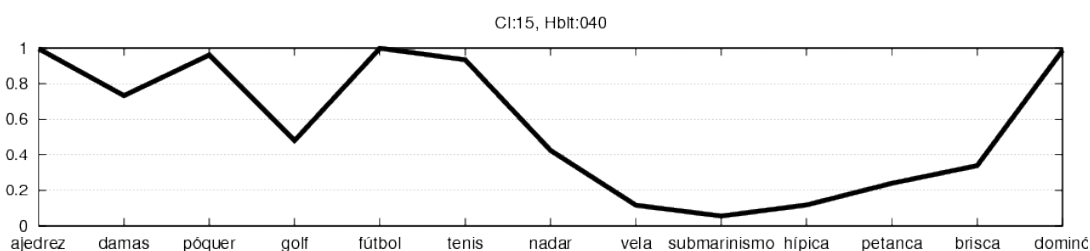


Gráfico 6. Muestra de compatibilidad de vocablos. Patrón de reentrada *b*



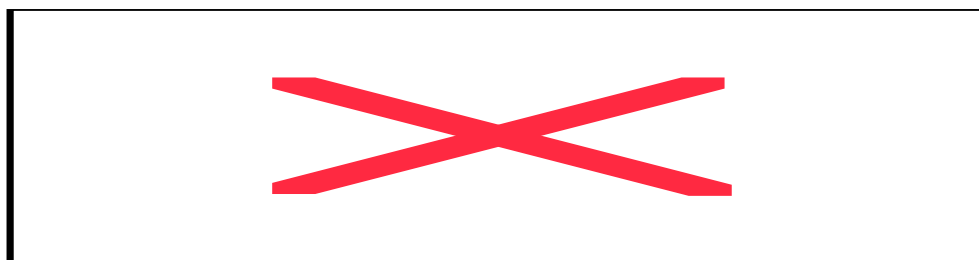
Los procesos de reentrada, como se observa en los gráficos anteriores, pueden darse en diferentes momentos de la realización de la lista de un mismo informante.

c. Patrón de desplazamiento y desviación. En ocasiones, cuando el hablante aporta un cierto número de palabras más o menos prototípicas suele comenzar



Gráfico 10. Muestra de compatibilidad de vocablos. Patrón de desplazamiento

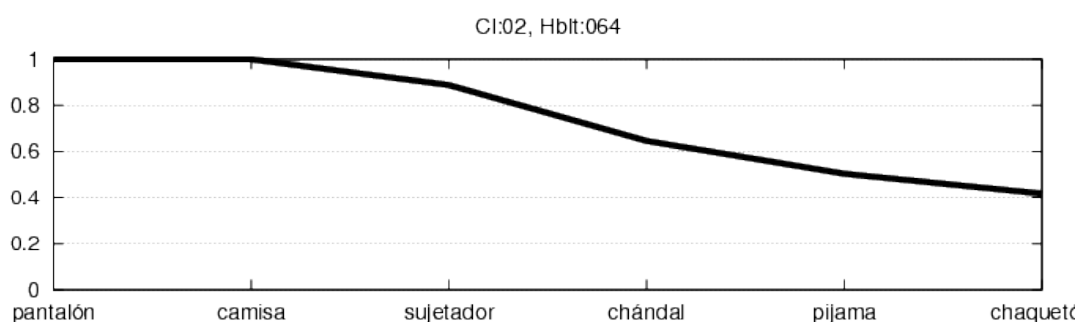
Sujeto 061. Centro de interés 04. Los muebles de la casa



d. Patrón de desviación sistemática. Por otro lado, quizás el proceso menos frecuente sea el que corresponde a los alejamientos continuos del centro prototípico sin reentradas (Gráfico 11). La poca frecuencia de este patrón confirma nuestra hipótesis de la importancia de los procesos de reentrada y, en consecuencia, el menor peso que habría que otorgar en la confección de las listas de disponibilidad individuales a la posición de las palabras. De hecho, los elementos que se ajustan a este patrón quedan acotados en la parte izquierda de la línea de corte en el Gráfico 3 (*supra*) que, como se observa, es la zona menos densa. De hecho, cuanto más hacia la izquierda de esa línea nos desplazamos (comportamientos más uniformes; menor desviación típica) menos ocurrencias encontramos.

Gráfico 11. Muestra de compatibilidad de vocablos. Patrón de desviación sistemática

Sujeto 064. Centro de interés 02. La ropa



En definitiva, la co-ocurrencia de estos patrones en los estudios generales de disponibilidad léxica evidencia que cada muestra individual no responde de forma

estricta a la forma del espectro léxico de cada sujeto. En realidad, esta hipótesis queda desvirtuada en el estudio de cada prueba individual por el proceso de reentrada. Sin embargo, tal y como ocurre en los sistemas probabilísticos, el mecanismo de muestreo que consiste en realizar múltiples pruebas a diferentes individuos, y la integración conjunta de la información obtenida, permite establecer un esquema del espectro del léxico compartido entre los individuos de una población estudiada<sup>32</sup>.

Al identificar los patrones dominantes en el Gráfico 3 (*supra*), es posible observar el comportamiento discriminatorio de esta representación. En el Gráfico 12 se representan algunos de los ejemplos usados en este trabajo<sup>33</sup>. Es posible observar que el binomio variación media-variación de la variación media, discrimina las situaciones descritas<sup>34</sup>:

---

<sup>32</sup> De hecho, si se aplicara esta metodología de múltiples pruebas a un único individuo, es probable que la naturaleza de la prueba afectase a la interpretación de su esquema léxico. Lo más posible es que se produjese un proceso de actualización reiterativo que se manifestaría como una modificación, a lo largo de las diferentes pruebas, del espectro del léxico del propio individuo.

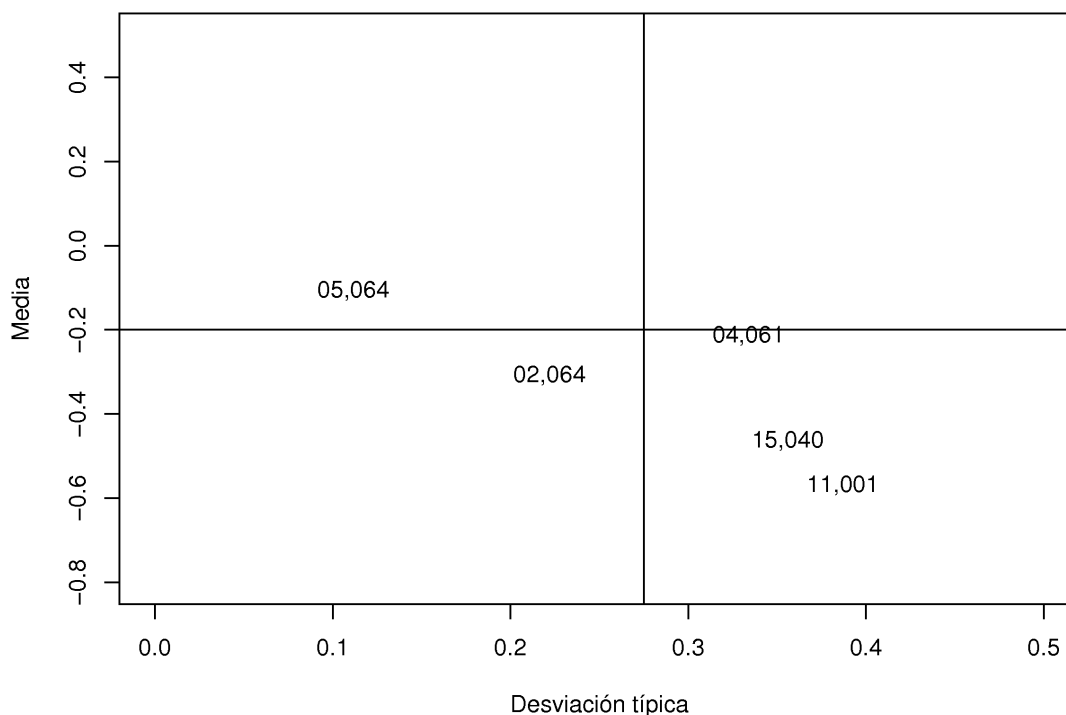
<sup>33</sup> Los elementos identificados en el Gráfico 12 deben leerse del siguiente modo: 05, 064 se refiere a la lista de palabras que el sujeto identificado como 064 escribe en el centro de interés 05. Alimentos (Gráfico 4); 02, 064 hace referencia a la lista del sujeto 064 en el centro de interés 02. La ropa (Gráfico 11), y así, sucesivamente. Para que se puedan identificar claramente los elementos del gráfico se ha optado por representar solo algunas de las muestras. El resto de ejemplos usados en este trabajo se agrupaban alrededor de los presentados en el Gráfico 12 de la siguiente forma: alrededor de 11,001 se encontraban las muestras 18,041 y 01,076; en la zona adyacente superior a 04,061 estaba la muestra 19,008.

<sup>34</sup> En realidad, una vez identificados los patrones referidos a los procesos de reentrada en la zona de máxima densidad de elementos (media:  $<-0.1$ ; desviación típica:  $\geq 0.275$ ) tenemos que recurrir a una ulterior discriminación de carácter cualitativo para distinguir los diferentes patrones expuestos.

Gráfico 12. Variación media-variación de la variación media.

Identificación de elementos

#### Distribución de estadísticos por hablante y centro. Algunos ejemplos



### 3. Conclusiones

A lo largo de este estudio hemos mostrado que el centro de interés es un punto de acceso a una red de elementos léxicos que el sujeto relaciona entre sí y con el concepto enunciado como centro de interés. Así concebido, su estructura es subjetiva e inherente al propio individuo, por lo que no es siempre extrapolable a otros sujetos. No obstante, por diversas razones, entre ellas sociales y culturales, las personas que pertenecemos a las mismas comunidades de habla poseemos estructuras léxicas similares<sup>35</sup>. Esa metaestructura sociocultural compartida es lo que consideramos el léxico de un centro de interés. Esto significa que los centros de interés no existen por sí mismos, sino como

<sup>35</sup> La lengua es, al tiempo, producto de  $n$  actos de habla efectivamente realizados y almacenados en la memoria de los hablantes, y sistema virtual que permite actualizar actos repetidos, pero también actos inauditos. Este equilibrio entre lo efectivo y lo virtual es la esencia de la propuesta de Saussure, como es bien sabido.

una realización del conjunto de hablantes. Si se cambia el conjunto de hablantes es posible, e incluso probable, que cambie la estructura del centro de interés.

Cuando un individuo se somete al experimento de la disponibilidad léxica y se le solicita que escriba palabras relacionadas con un determinado centro de interés, este suele actualizar, en primer lugar, los elementos léxicos más cercanos al núcleo del concepto propuesto. Posteriormente, recorre su red léxica hacia elementos más alejados de ese punto de entrada a medida que consume los elementos más accesibles. Como hemos mostrado, son frecuentes los procesos de reentrada en la red cuando el sujeto considera que se ha alejado demasiado, para lo cual utiliza diversas estrategias, aunque mucho menos frecuentes que los procesos de alejamiento progresivo del núcleo del centro de interés, probablemente por un proceso de asociación.

En consecuencia, el orden de aparición de las palabras en las listas de disponibilidad léxica es relevante con respecto a su accesibilidad, aunque pueden producirse puntos de reentrada (reinicialización de la prueba en un mismo proceso) que, en definitiva, parecen indicar que el orden de aparición de las palabras en las listas de disponibilidad léxica debe ser considerado en función de las circunstancias anteriores.

En definitiva, la demostración de la existencia de los procesos de reentrada nos indica que algunas palabras que aparecen en posiciones atrasadas en algunas listas individuales de disponibilidad (y las asociadas directamente a ellas) pueden resultar más representativas de un determinado centro de interés que otras cuya aparición se registre antes, siempre que su actualización sea frecuente en las diferentes listas consideradas.

#### **4. Bibliografía**

- Aitchison, J. (1987): *Words in the Mind: An introduction to the mental lexicon*. Oxford-Nueva York: Basil Blackwell.
- Auer, P., F. Hinskens y P. Kerswill (eds.) (2005): *Dialect Change: Convergence and Divergence in European Languages*, Cambridge: Cambridge University Press.
- Ávila Muñoz, A. M. (2006): *Léxico disponible de los estudiantes preuniversitarios malagueños*, Málaga: Universidad.
- Ávila Muñoz, A. M. y J. M. Sánchez Sáez (2010): *La disponibilidad léxica. Antecedentes y fundamentos*. En Ávila Muñoz, A. M. y J. A. Villena Ponsoda (eds.) (2010).

- y J. A. Villena Ponsoda (eds.) (2010): *Variación social del léxico disponible en la ciudad de Málaga. Diccionarios y análisis*. Málaga: Sarriá.
- Balibar, R. y D. Laporte (1974): *Le français national: Politique et pratiques de la langue nationale sous la Révolution française*. París: Hachette (trad. esp., Barcelona: Avance: 1976).
- Benítez, P. (1992): Disponibilidad léxica en la zona metropolitana de Madrid. *Boletín de la Academia Puertorriqueña de la Lengua Española (BAPLE)* 1/1: 71-102.
- Butrón, G. (1987): *El léxico disponible: índices de disponibilidad*. Río Piedras: Universidad de Puerto Rico (Tesis Doctoral inédita).
- Caramazza, A. (1998): The interpretation of semantic category-specific deficits: What do they reveal about the organization of conceptual knowledge in the brain? *Neurocase* 4: 265-272.
- Clark, H. H. (1970): Word associations and linguistic theory. En Lyons, J. (ed.) 1970: 271-286.
- Cramer, P. (1968): *Word Association*. Nueva York: Academic Press.
- Damasio, A. R. (1990): Category related recognition defects as a clue to the neural substrates of knowledge. *Trends in Neurosciences* 13/3: 95-98.
- Deese, J. (1962): On the structure of associative meaning. *Psychological Review* 69: 161-175.
- Galloso Camacho, M. V. (2003): *El léxico de los estudiantes preuniversitarios en el distrito universitario de Salamanca (Ávila, Salamanca y Zamora)*. Burgos: Fundación Instituto Castellano y Leonés de la Lengua.
- Garrard, P., R. Perry y R. Hodges (1997): Disorders of semantic memory. *Journal of Neurology, Neurosurgery and Psychiatry* 62: 431-435.
- Gougenheim, G., R. Michéa, P. Rivenc y A. Sauvageot (1956): *L'élaboration du français élémentaire. Étude sur l'établissement d'un vocabulaire et d'une grammaire de base*. París: Didier.
- , R. Michéa, P. Rivenc y A. Sauvageot (1964): *Français Fondamental (1<sup>er</sup> degré). Étude sur l'établissement d'un vocabulaire et d'une grammaire de base*. Nouvelle édition refondue et augmentée. París: Didier.
- Hernández Muñoz, N. (2006): *Hacia una teoría cognitiva integrada de la disponibilidad léxica: el léxico disponible de los estudiantes castellano-manchegos*. Salamanca: Universidad.
- Lakoff, G. (1987): *Women, Fire and Dangerous Things: What categories reveal about the mind*. Chicago-Londres: The University of Chicago Press.
- López Chávez, J. y C. Strassburger Frías (1987): Otro cálculo del índice de disponibilidad léxica. En *Presente, y perspectiva de la investigación computacional en México. Actas del IV Simposio de la Asociación Mexicana de Lingüística Aplicada*. México: UNAM.

- López Morales, H. (1973): *Disponibilidad léxica en escolares de San Juan*. San Juan de Puerto Rico: ms.
- (1978): Frecuencia léxica, disponibilidad y programación curricular. *Boletín de la Academia Puertorriqueña de la Lengua Española* 6/1: 73-86.
- (1979): Disponibilidad léxica y estratificación socioeconómica. *Dialectología y sociolingüística. Temas puertorriqueños*. Madrid: Hispanova de Ediciones: 173-181.
- (1999): *Léxico disponible de Puerto Rico*. Madrid: Arco/Libros.
- Lorán, R. (1983): *Un índice de disponibilidad léxica*. Mayagüez: Departamento de Matemáticas. Universidad de Puerto Rico.
- y H. López Morales (1983): Nouveau calcul de l'indice de disponibilité. Puerto Rico: ms.
- Lyons, J. (1970) : *New Horizons in Linguistics*, Londres : Penguin.
- Marconi, D. (2000): *La competencia léxica*. Madrid: Antonio Machado Libros.
- Morales, A., E. Forastieri, J. Cardona y H. López Morales (eds.) (1999): *Estudios de lingüística hispánica. Homenaje a María Vaquero*. San Juan de Puerto Rico: Universidad.
- Nelson, D. L., C. L. McEvoy y S. Denis (2000): What is free association and what does it measure? *Memory and Cognition* 28: 887-899.
- , D. J. Bennett y T. W. Leibert (1997): One step is not enough: making better use of association norms to predict cued recall. *Memory and Cognition* 25: 785-706.
- Rosch, E. (1978): Principles of categorization. En Rosch, E. y B. Lloyd (eds.) 1978: 27-48.
- y B. Lloyd (eds.) (1978): *Cognition and Categorization*, Hillsdale: Lawrence Erlbaum.
- Saffran E. M. y M. F. Schwartz (1994): Of cabbages and things: Semantic Memory from a neuropsychological perspective. A Tutorial Review. En Umiltà, C. y M. Moscovitch (eds.) 1994.
- Samper Padilla, J. A. (1999): Léxico disponible y variación dialectal: datos de Puerto Rico y Gran Canaria. En Morales, A., E. Forastieri, J. Cardona y H. López Morales (eds.) 1999: 550-573.
- Umiltà, C. y M. Moscovitch (eds.) (1994): *Attention and Performance XV*. Cambridge-Londres: MIT Press.
- Villena Ponsoda, J. A. (2005): How similar are people who speak alike? An interpretive way of using social networks in social dialectology research, en P. Auer, F. Hinskens y P. Kerswill (eds.) 2005: 303-334.
- Warrington E. K. y R. A. McCarthy (1987): Categories of knowledge. *Brain* 110: 1273-1296.

- Wittgenstein, L. (1953): *Philosophical Investigations*. Nueva York: McMillan.
- Zadeh, L. A. (1965): Fuzzy sets. *Information and Control* 8: 338-353.
- Zimmermann, H. J. (2001): *Fuzzy Set Theory and its Applications*. Boston: Kluwer Academic Publishers.