

Clausura del  
IV Curso de Lexicografía Hispánica

José G. Moreno de Alba  
Director de la Academia Mexicana de la Lengua

## Unidad y diversidad del español: el léxico

### 1 *Introducción*

Es interesante constatar que la frecuencia de aparición de tal o cual lexema en la lengua hablada o escrita es inversamente proporcional a la riqueza de la clase a la que pertenece: evidentemente las más frecuentes son las palabras *gramaticales*;<sup>1</sup> por lo contrario, las voces más *raras*, más especializadas, las que aparecen muy de vez en cuando en un texto oral o escrito, pertenecen a las clases más ricas (los verbos, los adjetivos, y, sobre todo, los sustantivos). La función del léxico tiene que ver con un proceso de *designación*, que se aplica a entidades de la experiencia, mientras que la gramática trata con modalidades *agregadas* o *conferidas* a esas entidades. Los nombres sustantivos son sin duda las palabras léxicas (en cuanto no gramaticales) más representativas. La clase de los sustantivos es tan extensa porque a ella pertenecen no sólo esos pocos nombres que todos conocemos y empleamos cotidianamente, sino también los muchos millares de palabras *raras*. Son los sustantivos, sobre todo lo llamados *comunes*, los que nos permiten organizar el mundo en grupos, en clases, en taxonomías, de conformidad con las necesidades que va señalando la experiencia.

Ya sea por el frecuente recurso de la derivación y composición, ya sea mediante la combinación de raíces griegas y latinas, ya sea por la penetración de extranjerismos provenientes de otras lenguas modernas, continuamente llegan a los léxicos generales de

---

<sup>1</sup> En una investigación estadística de léxico en el español mexicano, en la lista de las 100 palabras más frecuentes de un corpus de casi dos millones de apariciones, las primeras diez fueron las siguientes: *la, el de, y, que, en, a, se, no, ser*. Como se ve, las nueve primeras forman parte de clases *cerradas*. El primer sustantivo que aparece en la lista (se trata de la palabra *vez*) ocupa el lugar 51 (cf. HAM CHANDE 1979).

cada idioma vivo palabras nuevas, que se incorporan a esos acervos no por casualidad o por capricho de tal o cual hablante, sino por imperiosas necesidades de designación. En un mundo cambiante, donde cada día hay novedades en todos los terrenos, también los nombres que sirven para designarlas nacen exactamente con la misma velocidad. Cualquier objeto o concepto nuevo requiere de inmediato ser nombrado; son los nombres, los sustantivos *neológicos* los que cumplen, siempre a tiempo, esta función designadora.<sup>2</sup>

Ciertamente la lengua española parece tender hoy más a la unidad que a la diversidad. Es innegable empero que en un idioma extendido por tan extensos territorios y con tantos millones de hablantes debe haber diferencias internas más o menos acusadas. En los profundos niveles lingüísticos correspondientes a la fonología y a la gramática estas divergencias no son particularmente graves; en la fonética y, sobre todo, en el vocabulario son fácilmente perceptibles para cualquiera. Hace años, ROSENBLAT (1971:11 y *ss.*) explicaba que son muy diferentes las opiniones que sobre la diversidad del español pueden dar los turistas, los puristas o los filólogos. A un turista español que llegue a México —escribía— le sorprenderá que en el desayuno le ofrezcan *bolillos* (cierto tipo de pan blanco), que en la calle pueda abordar un *camión* (autobús) o pedir un *aventón* (*autostop*), que sea un *bolero* el que le *bole* (asee) los zapatos, que en el teléfono le contesten ¡*bueno!*, que ahí las casas se *renten* (se alquilen), que pueda acompañar sus bebidas con *botanas* (tapas), que haya establecimientos como las *loncherías* (cierto tipo de cafeterías modestas), *tlapalerías* (ferreterías donde se venden también pinturas), *misceláneas* (pequeñas tiendas de comestibles), *rosterías* (asadores), etc. Algo semejante le pasará a cualquier otro turista que llegue a las demás ciudades del extenso mundo hispanohablante porque hay ahí necesariamente diferencias léxicas, aunque sin duda es una tarea pendiente determinar si éstas afectan o no a la esencia misma de la lengua o si impiden de manera notable la comunicación entre los diversos grupos.

---

<sup>2</sup> En la lengua, en su léxico, se dan simultáneamente dos condiciones: la de un sistema conservador (“nuestras lenguas son instrumentos anacrónicos [...] al hablar somos humildes rehenes del pasado”, escribió Ortega y Gasset) e innovador. Hace ver ROSENBLAT (1967) que palabras como *radar*, *neutrón*, *fanatismo*, *litro*, *burócrata*, son “increíblemente jóvenes”.

Aceptando que existen diferencias considerables en el léxico de las diversas regiones del extenso mundo que habla español, a las que además deben sumarse las que se originan por razones sociales, es asimismo indudable que hay también un enorme caudal léxico común a todos los hispanohablantes, observable por ejemplo en la literatura. Esto no quiere decir de ninguna manera que estén desapareciendo las diferencias regionales. Lo que habría que determinar es si tenía o no razón el mismo ROSENBLAT (1971:32) cuando escribía aquello de que las designaciones básicas (parentesco, partes del cuerpo, objetos más comunes, fórmulas de la vida social...) son comunes a todos:

Al pan lo seguimos llamando pan, y al vino, vino. Por encima de ese fondo común las divergencias son sólo pequeñas ondas en la superficie de un océano inmenso

## 2 *Abundantes diferencias en el léxico regional y rural*

En el prólogo al poema *Nastasio* de Soto y Calvo, el gran filólogo colombiano Rufino José Cuervo, a fines del siglo XIX, escribía lo siguiente:

Hoy sin dificultad y con deleite leemos las obras de los escritores americanos sobre historia, literatura, filosofía; pero en llegando a lo familiar o local, necesitamos glosarios.

En efecto, no es el vocabulario culto, bastante estandarizado, sino el popular y, sobre todo, el rural, el regional, el que hace ver a la lengua española (igual que a cualquier otra) como un inmenso mosaico constituido por infinidad de vocablos y acepciones de muy reducida extensión geográfica pero de hondo arraigo entre los hablantes de tal o cual pueblo o región. Fuentes confiables de este tipo de diferencias son sin duda los atlas lingüísticos, que no sólo organizan conceptos en campos semánticos sino también muestran la distribución geográfica de los vocablos en una zona determinada. Imposible en una comunicación como ésta detenerme a ejemplificar con detalle la enorme diversidad léxica del español regional y rural en el mundo. Basten

algunos pocos conceptos y vocablos tomados de cuatro atlas: el de Andalucía (ALEA), el de Canarias (ALEICan), el de Colombia (ALEC) y el de México (ALM).

Un concepto (y sus designaciones) tomado del mundo de la fauna: a la luciérnaga, en Andalucía se le llama, predominantemente, *bichito de luz* pero también: *candilico*, *gusanito de luz*, *linterna*, *bichuchico*, *reluzangana*...; en Canarias sobresale la designación *cucuyo* y, secundariamente: *luciérnaga*, *salvaja*, *fuego fatuo*, *fuego salvaje*, *miñoca*, *yuyú*...; en Colombia predomina el vocablo *candelilla* y también se emplean: *luciérnaga*, *alumbraculo*, *cocuy*, *cocuyo*, *lagaña de perro*, *manzana*...; la voz preferida en México es *luciérnaga*, que alterna con: *cocay*, *cocuyo*, *lucerna*, *alumbrador*, *linterna*, *copeche*, *chupiro*, *cucayo*, *chilasca*, *churrupitente*, *santamarta*, *tagüinche*, *gusanito*, *moscaque*, *cacusagui*, *pilil*, *pilín*, *arlomo*, *caballero*, *chupamecha*, *chuperete*, *currupete*, *lucecita*, *marticuil*, *martita*, *prendedor*, *prendeyapaga*, *santiaguillo*, *tachinole*, *trencito*, *viejita*... En sólo cuatro regiones del mundo hispánico se registran, para este concepto, más de 50 designaciones.

Suele pensarse que las partes del cuerpo humano se designan de manera uniforme porque se trata de las llamadas designaciones básicas. Fácilmente se encuentran contraejemplos. Véase el caso del concepto 'dedo meñique'. En Andalucía se le llama, preferentemente, *meñique*, aunque también, como secundarias, se dan las siguientes designaciones: *chico*, *margarito*, *chiquitín*, *miche*, *títete*, *tete*, *chiqueniño*, *chipilín*, *garite*... En Canarias se prefiere *margarito*, que alterna con: *márgaro*, *meñique*, *mangariño*, *pequeño*, *meringuiño*... En Colombia parecen prevalecer los vocablos *meñique* y *chiquito*, aunque también es frecuente *tientagallinas*. En México, finalmente, como en Andalucía y Colombia hay predominio de *meñique* y, como designaciones secundarias: *chiquito*, *niño chiquito* (y *bonito*), *chocoyote* (de origen náhuatl), *menor*, *pequeño*, *quinto*, *tup* (o *tupito*), de origen maya, *de la gallina*, *desconsolado*... Más de 20 designaciones para un concepto de designación básica.

Por último, en lo que respecta a léxico regional y rural, considérese un concepto cuyo referente no pertenece ni a la flora, ni a la fauna, ni al cuerpo humano. Aludo a 'diablo'. Llama la atención que, mientras en Andalucía y Canarias sólo se documentan

dos designaciones (*diablo* y *demonio*), en México y, sobre todo, en Colombia, además de esas dos, se registren muchas más. Designaciones secundarias en México son: *satanás*, *lucifer*, *chamuco* (de origen náhuatl), *pingo*, *chango*, *luzbel*, *satán*, *el enemigo*, *camándula*, *judío*... Sorprende que, en el mapa del ALEC correspondiente a este concepto, se documenten más de 100 formas de llamar al diablo en Colombia, algunas de las cuales son: *barrabás*, *busiraco*, *cachón*, *cachudo*, *caifás*, *capataz*, *compadre*, *cuco*, *chiras*, *diánchiro*, *diantre*, *enemigo malo*, *gato negro*, *guainás*, *judas*, *judío (errante)*, *luzbel*, *mahoma*, *mandingas*, *malingno*, *mañoco*, *perro negro*, *tentación*, *trampas*, *uñón*, *viruñas*...

### 3 Pocas diferencias en el léxico fundamental

El corpus que se viene utilizando para el importante proyecto del «Diccionario del español de México»<sup>3</sup> consta de 1 891 045 vocablos;<sup>4</sup> de él se obtuvo un vocabulario constituido por 64 183 palabras. Ahora bien, cuando se preparó el *DFEM* (1982), «al acumular los vocablos que constituyen el 75 % de todas las emisiones lingüísticas cultas en el español de México, se comprobó que están formadas por 1451 palabras. Esta acumulación es el vocabulario fundamental» (Introducción del *DFEM* 1982:11). Ésas son las entradas del *DFEM* (1982). Es posible que, a primera vista, parezcan pocas palabras. Sin embargo, como se señala en esa misma «Introducción» (p. 11),

«una constante entre muchas lenguas del mundo parece ser que su vocabulario fundamental se aproxime a esa cifra y no pase de dos mil palabras. Así lo han demostrado las investigaciones hechas sobre lenguas tan diferentes como el chino, el japonés, el francés o el inglés»

---

<sup>3</sup> Noticias metodológicas pueden verse en LARA, HAM Y GARCÍA (1979). Resultados no por parciales menos importantes son los tres diccionarios hasta ahora publicados, todos dirigidos por Luis Fernando Lara: *Diccionario fundamental del español de México* (DFEM 1982), *Diccionario básico del español de México* (DBEM 1986) y *Diccionario del español usual en México* (DEUM 1996).

<sup>4</sup> Los mil textos (orales y escritos), cada uno constituido por cerca de dos mil palabras, se clasifican así (de mayor a menor): lengua culta, lengua no estándar, lengua sub-culta, ciencias, periodismo, literatura, textos dialectales, técnicas, literatura popular, habla culta, documentos antropológicos, habla media, habla popular, lírica popular, jergas, discursos políticos, religión (cf. HAM CHANDE 1979:76).

Puede muy bien, por tanto, tomarse este léxico fundamental del español mexicano como ejemplo, pues muy probablemente la constitución de otros vocabularios fundamentales, correspondientes a otros dialectos (español europeo, argentino, puertorriqueño...) no variaría mucho, ni en el número ni en el tipo de vocablos. Si se revisan las entradas de este diccionario, si se revisa este léxico fundamental, con facilidad podrá comprobarse que casi no contiene mexicanismos.<sup>5</sup> La casi totalidad de entradas corresponde al español general. Véanse, sea por caso, las 47 entradas que comienzan por *b*: *b*, *bailar*, *baile*, *bajar*, *bajo*<sup>1</sup>, *bajo*<sup>2</sup>, *banca*<sup>1</sup>, *banca*<sup>2</sup>, *banco*<sup>1</sup>, *banco*<sup>2</sup>, *banco*<sup>3</sup>, *bañar*, *baño*, *barco*, *barrio*, *basar*, *base*, *básico*, *bastar*, *bastante*, *bastar*, *batir*, *beber*, *belleza*, *bello*, *beneficiar*, *beneficio*, *besar*, *beso*, *bien*<sup>1</sup>, *bien*<sup>2</sup>, *bilabial*, *blanco*, *boca*, *bolsa*<sup>1</sup>, *bolsa*<sup>2</sup>, *bonito*, *brazo*, *breve*, *buen*, *bueno*, *burgués*, *burguesía*, *burocracia*, *burócrata*, *buscar*, *búsqueda*. Ninguna de estas voces puede ser considerada mexicanismo; lo mismo puede decirse de las diversas acepciones de cada una de ellas. Algo similar puede afirmarse del resto de las entradas, aunque no faltan algunos pocos mexicanismos (*chile*, *tortilla* [de masa de maíz]...). En este *Diccionario* no tienen cabida numerosos mexicanismos que aparecerán en los otros dos, sobre todo en el *DEUM* (1996), como *atole*, *banqueta* ('acera'), *bolear*, *bolero*, *camote*, *comal*, *elote*, *nopal*, *petate*, *tamal*...

Es muy probable que, si revisáramos los vocabularios fundamentales de los demás dialectos del español, formados con técnicas estadísticas semejantes, encontraríamos lo mismo: la casi totalidad de vocablos y acepciones correspondería al español general y serían muy pocas las palabras y acepciones propias de cada uno de los dialectos. Más aún: quizá sea ésta una de las peculiaridades de los vocabularios fundamentales de los dialectos: las diferencias entre vocablos y acepciones entre unos vocabularios y otros es mínima. Evidentemente, «a partir del límite trazado para el vocabulario fundamental [75 % de todas las emisiones lingüísticas cultas] el léxico aumenta en relación con la enorme variedad de temas especializados...» (*DFEM* 1982:11). En esa misma proporción aumentarán también las diferencias en relación con el español general y con el español de cada uno de los demás dialectos.

---

<sup>5</sup> Entiendo por *mexicanismo* un vocablo o acepción propio de la mayor parte de los hablantes mexicanos y

#### 4 *Notables diferencias en el proyecto Varilex*<sup>6</sup>

Como se ve, no es interés de los trabajos sobre vocabularios fundamentales la búsqueda de diferencias léxicas entre las variedades lingüísticas. Hay empero investigaciones léxicas que ponen énfasis, precisamente, en las diferencias de vocabulario existentes entre los diversos dialectos. Los instrumentos de consulta, en este caso, están pensados no para encontrar el léxico fundamental de tal o cual dialecto ni tampoco los vocablos que son comunes a todos los hispanohablantes sino que les interesan, casi exclusivamente, aquellos conceptos que se manifiestan por diversas palabras y frases en los diferentes dialectos geográficos y sociales de la lengua española. Normalmente no es sino hasta que se concluyen las encuestas, cuando los investigadores se dan cuenta de la medida en que el cuestionario fue útil para sus fines (descubrir diferencias léxicas). A ello se debe que, en este tipo de investigaciones, se vayan depurando y perfeccionado sucesivamente los interrogatorios, hasta que se tiene la seguridad de que la casi totalidad de preguntas obtendrá una aceptable variedad de respuestas.

En VARILEX (1994: XVII-XIX), en la lista de «Las palabras encuestadas», en el campo semántico de la ropa, no aparecen, sea por caso, ‘pantalón’, ‘camisa’ o ‘zapatos’, conceptos que difícilmente serían redituables para encontrar diferencias; pero sí están los conceptos de los que se tiene cierta garantía de variación léxica, por ejemplo: ‘chaqueta’ (*americana, chaqueta, saco, vestón...*), ‘cremallera’ (*cremallera, cierre, cierre relámpago, zíper, cierre éclair...*) o ‘gafas’ (*gafas, anteojos, lengtes, espejuelos...*). De esta forma, la ‘lista 2’ (VARILEX 1994: XXIV-XXVIII) contiene las 919 palabras que se obtuvieron como respuestas a las 219 preguntas de ese cuestionario, correspondientes a 14 ciudades. Eso quiere decir que, en promedio, cada concepto se manifestó en cuatro vocablos diferentes. Unos pocos tienen sólo dos respuestas (‘abrigo’: *abrigo y tapado*)

---

ausente en la mayor parte de los hispanohablantes no mexicanos.

<sup>6</sup> Variación léxica del español del mundo. Un grupo de investigadores, coordinado desde Tokio por Hiroto Ueda y Toshiro Takagaki, han venido trabajando con cuestionarios léxicos en un buen número de ciudades hispanohablantes. Los resultados, aunque crecientes, todavía parciales circulan entre investigadores, en forma de cuadernos. Han aparecido seis. Llevan siempre la siguiente leyenda: “Esto es un informe intermedio del Proyecto “Español del Mundo”. No se trata de una publicación definitiva sino que sirve sólo para la circulación entre los miembros. Por lo tanto no puede ser objeto de reseña ni de crítica de las revistas públicas”.

pero hay otros que obtuvieron más de diez ('palomitas de maíz': *palomitas, cabritas de maíz, popcorn, cotufas, pop, pororó, copos de maíz, pachocle, canguil, cancha, rosita de maíz...*).

Conforme va avanzando la investigación, en los años siguientes, se llevan a cabo dos importantes modificaciones de método. Por una parte, se van eliminando de la lista de conceptos aquellos que no se muestran suficientemente productivos de variantes léxicas. Por otra parte, se van añadiendo cada vez más ciudades y, por ende, mayor número de sujetos encuestados.<sup>7</sup> Estos dos ajustes llevan a resultados sorprendentes, en lo que toca a la productividad de vocabulario por concepto. En VARILEX (1999), por ejemplo, se ofrecen, para las mismas preguntas (Serie «A»), respuestas correspondientes no ya a 14 ciudades sino a 44. El cuestionario, por su parte, se redujo a sólo 150 preguntas. Esto quiere decir que se suprimieron 69 conceptos, que representan más de la tercera parte de los que aparecían en la lista de VARILEX (1994). El resultado de la aplicación de estas medidas, entre otras, es observable en el altísimo número de «palabras» o respuestas «diferentes» para un mismo concepto. Conté estas respuestas y obtuve como resultado un total de 2288. El promedio de variantes por concepto, pasa de cuatro, en la lista que aparece en uno de los artículos de VARILEX (1994: XXIV-XXVIII), a 15, en los resultados de la investigación de 1999.

Evidentemente que, si uno estudia con atención las respuestas a cada concepto, se encuentra con que no pocas de ellas son *incorrectas* por muy diversas razones (mala pregunta del encuestador, falta de comprensión o desconocimiento del concepto por parte del encuestado —a pesar de lo cual proporciona una respuesta inventada, equivocada o imprecisa— pregunta especialmente difícil o confusa, etcétera).<sup>8</sup> También habría que determinar hasta qué punto ciertas respuestas no son sino variantes poco significativas de otra: en la cuestión 8 (*jeans*), por ejemplo, que produjo 22 términos, se anotan, como respuestas *diferentes*: *blue jean* y *blue jeans, pantalones de mezclilla* y *pantalón de*

---

<sup>7</sup> Asimismo van apareciendo nuevos cuestionarios. Para los fines de esta presentación, me basta sin embargo ejemplificar con la serie "A", a la que me estoy refiriendo.

<sup>8</sup> Independientemente de lo que se quiso preguntar en la cuestión 2 ('cardigan, prenda de lana, con botones') me parece que en ningún dialecto es lo mismo *abrigo* que *chaleco, chaqueta* que *suéter, saco*

*mezclilla, tejero y tejeros, pantalón de jean y pantalón de jeans...* Pues bien, aún suprimiendo muchas de las respuestas, por estos u otros semejantes motivos, no cabe duda de que subsistirá la mayoría y desde luego, una cantidad muy superior a la obtenida en las primeras listas de 1994. Téngase en cuenta, además, que fueron consideradas, en las listas de VARILEX (1999), después de suprimir los menos rendidores, sólo 150 conceptos en apenas 44 ciudades. Imaginemos cuánto crecerá todavía el número de variantes si, por una parte, se encuestan más ciudades y si, por otra, se construye otro tipo de listas con preguntas igualmente ricas en respuestas diferentes.<sup>9</sup>

##### *5 No tan importantes diferencias en los léxicos disponibles de los diversos dialectos*

La lexicoestadística se ha centrado en el estudio de la frecuencia de las palabras. Sobre esa base, desde hace algunos años vienen determinándose, para diversos dialectos (geográficos y sociales) y para determinados conjuntos de hablantes (agrupados por edad, por sexo, por escolaridad, etc.) índices de disponibilidad léxica referidos a diversos campos nocionales o culturales. Se trata de fijar la probabilidad de aparición de las palabras que se asocian a un determinado centro de interés. Es una manera de relacionar la lexicoestadística con la búsqueda del léxico fundamental. Se han publicado ya numerosos trabajos de disponibilidad léxica para diversos dialectos y grupos de hablantes de lengua española. Para ejemplificar la manera en que puede observarse la unidad o diversidad del vocabulario español en estos índices de disponibilidad léxica me limitaré a valerme de los resultados de la reciente investigación de VALENCIA y ECHEVERRÍA (1999), que tuvo como sujetos a 2052 estudiantes chilenos (de todas las regiones del país) cuando asistían al curso terminal de su enseñanza media. Hubo equilibrio proporcional entre los dos sexos y los tres niveles socioeconómicos establecidos (bajo, medio y alto). Se distinguieron 18 centros de interés.<sup>10</sup> Las instrucciones fueron muy

---

*tejido que suéter abierto...* Pues bien todas éstas fueron respuestas consideradas buenas para esa pregunta. Algo semejante se observa en varias otras entradas.

<sup>9</sup> Esto último ya se viene haciendo. Se dispone ahora de cinco listas (de conceptos y de variantes léxicas). El número de ciudades también ha crecido considerablemente.

<sup>10</sup> 1: Las partes del cuerpo; 2: ropa y calzado; 3: la casa; 4: muebles; 5: alimentos; 6: la cocina; 7: medios de transporte; 8: trabajo de campo y jardín; 9: plantas y animales; 10: juegos y entretenimientos; 11: profesiones y oficios; 12: la política; 13: la actividad económica; 14: las artes; 15: ciencia y tecnología; 16: mundo espiritual; 17: procesos mentales; 18: problemas del ambiente.

sencillas. Por ejemplo, en relación con el primer centro de interés («Las partes del cuerpo»), a cada estudiante se le pedía simplemente que escribiera todas las palabras que pudiera sobre el cuerpo humano y sus partes. Así se procedió con los demás centros de interés. El corpus se presenta en cuatro columnas: vocablo, índice de disponibilidad léxica,<sup>11</sup> frecuencia absoluta y frecuencia relativa.

En VALENCIA y ECHEVERRÍA (1999:258 y *ss.*) se proporcionan listas con los cien vocablos más frecuentes de cada centro de interés. Se trata, en total, de 1800 vocablos que, con un método diferente del explicado en el apartado anterior, pueden constituir el léxico fundamental de ese muy representativo grupo de hablantes chilenos. El vocabulario fundamental establecido sobre la base de disponibilidad léxica es ciertamente diferente del que se obtiene mediante la selección de los más frecuentes vocablos que aparecen en un corpus de textos orales y escritos.<sup>12</sup> Señalé antes que en el léxico fundamental de México, contenido en el *DFEM* (1982), casi no había mexicanismos. Por lo contrario, no son pocos los chilenismos que pueden detectarse entre las 1800 voces que se enlistan en la investigación chilena sobre disponibilidad léxica que estoy comentando. Revisé con cuidado esos 1800 vocablos y separé aquellos que, casi seguramente, no forman parte del vocabulario activo (y, muy probablemente, tampoco del pasivo) de los mexicanos. Si este ejercicio lo lleva a cabo un peruano, un español, un panameño, etc., probablemente separe si no precisamente las mismas palabras, sí muchas de ellas. Asimismo es posible que, si contáramos con listas de cada país, obtenidas con la misma estricta metodología, y nos pusiéramos a separar aquellas palabras que cada uno

---

<sup>11</sup> Que se expresa en decimales. Hay una relación directamente proporcional entre el índice y la frecuencia con que el vocablo fue mencionado. Así, en el centro de interés número 1 (“Las partes del cuerpo”), el vocablo que obtuvo el más alto índice de disponibilidad léxica fue *cabeza* (0.6230383149) y el que ocupó el último lugar fue *nervio-óptico* (0.0000255043).

<sup>12</sup> Una de las razones para esta diferencia está en el hecho de que en la lista léxica procedente del corpus de textos (mexicanos) entra todo tipo de palabras (tanto las pertenecientes a inventarios cerrados como las que forman parte de inventarios abiertos). Por lo contrario, en las listas procedentes de la disponibilidad léxica aparecen casi exclusivamente sustantivos. A ello se debe que entre las cien palabras más frecuentes del corpus mexicano (*cf.* HAM CHANDE 1979) casi no haya sustantivos. De hecho en esa lista sólo aparecen ocho: *vez* (lugar 51), *día* (62), *tiempo* (80), *hombre* (85), *vida* (95), *casa* (96), *forma* (99) y *trabajo* (100). Para un diccionario fundamental parece conveniente basarse en listas procedente de textos orales y escritos. Para una caracterización léxica de los dialectos creo que resultan más adecuadas las listas que se basan en disponibilidad léxica.

de nosotros o no empleamos o ni siquiera conocemos, llegaríamos, en cada caso, a proporciones semejantes.<sup>13</sup>

En algunos de los 18 centros de interés de la investigación chilena la casi totalidad de vocablos pertenecen, en mi opinión, al español general. No encontré allí chilenismos evidentes o, al menos, designaciones desconocidas para un hablante mexicano. Supongo que al mismo resultado habría llegado cualquier atento revisor de otras latitudes hispánicas. Esas listas corresponden a los siguientes centros de interés: 1: las partes del cuerpo; 13: la actividad económica; 14: las artes; 15: ciencia y tecnología; 16: mundo espiritual; 17: procesos mentales; 18: problemas del ambiente.<sup>14</sup> En estos 7 apartados (39 % del total de centros de interés) no llegan al 5 % de cada lista los vocablos que parecen no pertenecer al español general.

En 5 (28 %) de los apartados (el 3: la casa; el 8: trabajo de campo y jardín; el 9: plantas y animales; el 11: profesiones y oficios; y el 12: la política)<sup>15</sup> encontré vocablos no pertenecientes al español mexicano (y, probablemente, tampoco al general) que representan entre el 5 % y el 10 % de cada lista. Siguen algunos ejemplos. Del apartado 3: *living* (en México: *sala*), *velador* (buró, mesita de noche), *cálifon* (bóiler, calentador

---

<sup>13</sup> Un ejercicio semejante a éste, pero en relación con Madrid, Las Palmas de Gran Canaria, República Dominicana y Puerto Rico, puede verse en LÓPEZ (1992).

<sup>14</sup> Anoto en seguida, algunas posibles excepciones. Del centro de interés 1 (partes del cuerpo): el vocablo *mejilla* (lugar 86 de la lista), aunque se conoce en México, es mucho menos usual ahí que *cachete* (que si bien no aparece entre los 100 más frecuentes del corpus chileno, está entre las voces de baja frecuencia); algo semejante sucede con *canilla*, más frecuente en Chile que en México, donde se prefiere *espinilla*. Del centro de interés 16 (mundo espiritual), hay tres voces que no conozco: *trauco*, *caleuche* y *pincoya*; tampoco las encontré en los diccionarios que pude consultar. Lo mismo sucede con *agua-servida* y *atochamiento* (del centro de interés 18, problemas del ambiente). Del grupo de voces referentes a la ciencia y tecnología (centro de interés 15) me parece desconocida (o de uso esporádico) en México *multicarrier*. En el centro de interés 13 (la actividad económica) hay algunas abreviaturas y siglas que sólo parecen tener significado en el ámbito chileno (*uf*, *codelco*, *utm*...). Del centro de interés de las artes (14) sólo anoto, como poco usual en México, al menos en ese campo semántico, el sustantivo *greda*. Ninguno de los 100 vocablos que se anotan en el apartado 17 (procesos mentales) me parece ajeno al español general. Ahora bien, si se acude no a estas listas de los 100 vocablos más frecuentes sino a las listas completas, aun estos campos muestran ya, en voces que están después del lugar 100, algunas diferencias en relación con el español mexicano; es decir se pueden encontrar vocablos que o no se conocen en México o se emplean sólo muy esporádicamente. Por ejemplo, en el centro de interés número 1 (Las partes del cuerpo), están las siguientes denominaciones, entre otras, que, como hispanohablante mexicano, me parecen en alguna medida al menos, extrañas: *pera* (México: piocha; barba de chivo), *poto* (nalgas, sentaderas, glúteos...), *guata* (estómago), *mono* (?), *carrillo* (cachete, mejilla), *carretilla* (mentón), *piñén*, *piamadre*.

[de agua]), *ampolleta* (foco, bombillo), *juguera* (licuadora), *frazada* (cobija, manta)... Del 8: *picota* (azadón), *huaso* (ranchero, campesino), *chuzo* (barra), *gualato*... Del 9: *chanchito* (cerdo, puerco), *gomero* (árbol, 'hule'), *oreja-de-oso*, *filodendro*, *malamadre* (árboles y plantas que o no existen en México o se les llama de otra forma)... Del 11: *carabinero* (policía), *gásfiter* (plomero, fontanero), *parvularia* (educadora, maestra de párvulos), *matrona* (partera), *dueña-de-casa* (ama de casa)...

Particularmente importantes, para este ejercicio, resultan los 6 apartados restantes (2: ropa y calzado; 4: muebles; 5: alimentos; 6: la cocina; 7: medios de transporte; y 10: juegos y entretenimientos), 33 % del total, pues en ellos los vocablos peculiares de Chile (en relación con el español mexicano al menos) oscilan entre un 11 % y un 23 %, porcentajes que me parecen altos. Sin embargo téngase en cuenta que, de acuerdo con este sencillo ejercicio, puede decirse que el grado de compatibilidad entre México y Chile oscilaría entre un 89 % y un 77 %; la media podría fijarse en un 83 %. En términos simples: si proyectamos estos resultados al léxico en general, podría pensarse que 17 de cada cien palabras que emplea un hablante chileno no las emplea y, a veces, no las comprende un hispanohablante mexicano. Algunos ejemplos: del centro de interés 2: *ambo* (combinación de saco o chaqueta y pantalón), *taco* (tacón [del zapato]), *soquete* (calcetín corto); diferentes tipos de calzado, como *chala* (sandalia, huarache), *zapatilla* (pantufla), *bototo* (botín), *ojota* (guarache, sandalia)...; suéteres y chaquetas de diversas clases y formas, como *polera* (playera), *parca* (makinof), *chomba* (suéter), *cuadros* (pantaletas, bragas), *beatle* (suéter con cuello de tortuga), *chaleca* ... Del 4: *velador* (buró, mesita de noche), *bifé* (buffet), *camarote* (litera), *mesón* (mostrador, mesa grande), *marquesa*, *bergier*, *rack*, *trinche*, *somier* (tambor), *mesa de living* (*mesa de centro*), *banqueta* (¿taburete?), *pizarrón*, *vanitorio*, *visor*, *arrimo*... Del 5: *poroto* (frijol, judías), *zapallo* (camote), *betarraga* (betabel), *arveja* (chícharo, guisante), *torta* (pastel), *choclo* (maíz, elote, mazorca), *palta* (aguacate), *frutilla* (fresa), *damasco* (albaricoque, chabacano), *queque* (panqueque), *manjar* (dulce de leche), *completo* (especie de hot dog, perro caliente), *ají* (chile)... Del 6: *juguera* (licuadora), *paila* (sartén), *fósforo* (cerillo),

---

<sup>15</sup> Estrictamente no debería haber incluido este último centro de interés (la política) porque las voces que un mexicano desconoce de la lista son, casi todas, siglas y abreviaturas de partidos y agrupaciones políticas chilenas (*udi*, *m*, *ppd*, *ps*, etc.).

*cálifon* (bóiler, calentador de agua), *cocina a gas* (estufa de gas), *sacajugo* (exprimidor), *pañó* (secador), *mesón* (mostrador, mesa), *cocina a leña* (estufa de leña), *uslero* (rodillo), *espumador*, *bol* (*tazón*)... Del 7: *bus* (camión, autobús), *colectivo* (taxi), *liebre* (taxi, pesero), *skate* (patineta), *citroneta*, *metrotrén* (metro), *bicimoto*, *metrobús*, *trolley*, *victoria* (calandria, carro tirado por caballos), *utilitario*... Del 10: *ludo* (parcasé), *dama* (damas), *luce* (avión), *flípper*, *ronda*, *bolitas* (canicas), *resbalín* (resbaladilla), *taca-taca* (fútbolito), *pillarse*, *la pinta*, *balancín* (sube y baja), *carioca*, *cordel* (cuerda), *papaya*, *tiña* (roña), *elástico*, *cacho* (cubilete), *emboque* (balero), *corre el anillo*...

LÓPEZ (1992) hace una comparación entre los léxicos disponibles de Madrid, Las Palmas de Gran Canaria, República Dominicana y Puerto Rico. Trabajó con índices de cercanía, compatibilidad y semejanza de cada uno de los dialectos. Los grados de compatibilidad a los que llega son sorprendentemente bajos. En trabajos anteriores (inéditos) también había llegado a conclusiones semejantes:

«Las comparaciones las hemos efectuado con diversos subconjuntos: con los primeros cien vocablos, con los quinientos, los mil y la totalidad; en todos los casos nos hemos encontrado con una compatibilidad mucho más baja de la que esperábamos. Los 100 primeros vocablos son compatibles por debajo del 70 % y ninguna comparación de los lexicones como totalidad alcanza un 56 % de compatibilidad (LÓPEZ 1992:29)»

El grado mayor de compatibilidad, en la investigación sobre los cinco dialectos mencionados, se da en el centro de interés 15 (Diversiones y deportes), entre Madrid y Las Palmas, y apenas fue de 49.68 %. El más bajo corresponde al centro 7 (La cocina y sus utensilios), entre República Dominicana y Puerto Rico: 8.13 % (!). En el mejor de los casos, ni siquiera la mitad de los vocablos registrados en uno de los dialectos se documentan en los otros. Como se ve, los resultados de esta investigación difieren notablemente de aquellos a los que yo llegué comprobando cuántas palabras chilenas suenan extrañas a un mexicano.

La explicación de esta diferencia es muy fácil. En LÓPEZ (1992) cualquier vocablo que sólo apareció en uno de los dialectos se considera «inexistente» en otro. En la «Comparación del centro de interés 16: ‘profesiones y oficios’» (61 y ss.), hay tres columnas: 1) vocablos comunes; 2) vocablos solo de Madrid; 3) vocablos sólo de Las Palmas.<sup>16</sup> Así, por ejemplo, *administrador, aduanero, alcalde, árbitro, astrólogo...* están en la columna de vocablos comunes (a Madrid y a Las Palmas). Sin embargo, la mayor parte de los vocablos que aparecen en las otras dos columnas, a pesar de no haber sido producidas en las listas de cada informante, es innegable que se emplean o, al menos, se conocen en el otro dialecto. Por ejemplo, entre los «vocablos solo de Madrid» están: *alguacil, arcipreste, arzobispo, bodeguero, cargador, churrero*, entre cientos más que, sin duda alguna, pertenecen también al dialecto de Las Palmas (y a muchísimos otros dialectos o, casi con seguridad, al español general). En la columna «vocablos solo de Las Palmas» aparecen: *afilador, agrónomo, ajedrecista, almirante, alpinista, apicultor, aprendiz...* y muchísimas más que, obviamente, o son del español general o al menos pertenecen a la mayor parte de sus dialectos, el de Madrid, desde luego, incluido. Hay por tanto una evidente confusión: los grados de compatibilidad léxica de diversos dialectos no pueden obtenerse mediante la simple comparación de los vocablos que libremente producen informantes de cada uno de esos dialectos. El que a ningún informante de Madrid se le haya ocurrido mencionar, entre las profesiones y oficios, al afilador o al agrónomo, de ninguna quiere decir que esas palabras no estén disponibles en el dialecto madrileño. Lo mismo puede asegurarse de muchas otras «omisiones». Ningún sujeto de Las Palmas mencionó al alguacil y al arzobispo: ¿quiere esto decir que ahí los alguaciles y los arzobispos se llaman de otra manera? En mi opinión, para saber cuáles voces de Las Palmas, entre las que se registraron con los informantes, no se emplean en Madrid, es necesario averiguarlo con hablantes madrileños (y viceversa). El porcentaje bajará muchísimo. Estoy seguro de que las diferencias serán aún menos numerosas que las que se pueden observar entre los léxicos chileno y mexicano.

##### 5 *El concepto de variación léxica es algo necesariamente relativo*

---

<sup>16</sup> Algunas páginas después cambian los títulos de las columnas: vocablos comunes, vocablos solo de

Cuando se dice que la lengua española tiende a la unidad, mejor que a la diversidad, en lo que atañe a la fonología, es un tipo de afirmación que puede demostrarse con relativa facilidad: aunque fuera difícil establecer el número de sistemas fonológicos del diasistema de la lengua, debido precisamente a que se trata de sistemas y, por tanto, de inventarios cerrados, de elementos contables, llegará un momento en que sabremos cuántos son —no serán desde luego muchos— y cuál es el inventario y distribución de los fonemas en cada uno de ellos, que tampoco será notablemente diferente entre unos y otros. Este tipo de certezas no son posibles en el nivel del léxico. El carácter verdaderamente abierto del vocabulario lo impide. A esto hay que añadir otro tipo de dificultades. Si resulta muy difícil establecer, así sea aproximadamente, el acervo léxico de un individuo, es no sólo difícil sino imposible hacerlo con respecto a un dialecto determinado, sea este geográfico o social, y ya no digamos si lo que se pretendiera hacer fuera enlistar el léxico actual de *toda* la lengua española. Ello no sólo por razones que podríamos llamar prácticas —no parece factible que alguien, ni persona ni institución, pueda enlistar todas las palabras y, sobre todo, todas las acepciones de todas las palabras, de todos los dialectos geográficos y de todos los niveles sociales de todos los hispanohablantes del mundo—, sino también por algunas razones que pueden designarse teóricas o, al menos, conceptuales. Por ejemplo: dificultades para distinguir el léxico pasivo frente al activo, para establecer si dos voces casi iguales en su forma son o no palabras diferentes, si dos o más significados muy semejantes son o no acepciones distintas de una misma palabra, cómo considerar dentro del vocabulario general de la lengua una voz que es empleada en un dialecto sólo en el ámbito rural y en otro en el urbano, o un vocablo que es común en tal o cual lugar o tipo de hablante y, en alguna manera, rara en otro, etcétera.

A todo esto hay que añadir las dificultades resultantes de la aplicación de criterios diversos para las diferentes investigaciones que suelen hacerse para adelantar en el conocimiento de la variedad léxica de una lengua. En esta nota he tratado de ejemplificar este tipo de inconvenientes con unos pocos proyectos de trabajo. Independientemente de que, en efecto, nos parezcan unos criterios mejores que otros, lo que interesa destacar es

---

Madrid, vocablos solo en Dominicana, y así sucesivamente, hasta agotar todas las combinaciones.

que son *diferentes* y que, por tanto, unos resultados no completan necesariamente a otros, sino que muchas veces resultan simplemente contradictorios. Ahora bien, aceptando la premisa, muy conocida, de que el español, como cualquier otra lengua, manifiesta mayor diversidad en el nivel léxico que en el fonológico, no queda otro remedio, para dar otro paso y aclarar en qué grado es importante esta diversidad, que reconocer, ante todo, que se trata de un concepto eminentemente relativo. Empleo la palabra *relativo* como antónima de *absoluto*.

Quien desee resaltar que esta diversidad no es tan importante como se piensa, puede, por ejemplo, echar mano de los léxicos fundamentales de varios dialectos, extraídos de las voces de mayor frecuencia en textos orales y escritos. Seguramente, con ese criterio, llegará a concluir —creo que con acierto— que no hay muchas diferencias en las listas de las 1500 palabras más frecuentemente presentes en corpus correspondientes a diversos dialectos. Podría decirse, por tanto, que *en relación con* los léxicos fundamentales, las diferencias léxicas son poco impresionantes; esto es, entonces, algo *relativo*. Habrá también quien quiera demostrar lo contrario: que el vocabulario de la lengua es enorme y que una manera de comprobarlo es viendo como un mismo referente o significado puede expresarse por muy diversas significantes o palabras. Hay muchas maneras de hacer que esta afirmación resulte aceptable. Bastará elegir con cuidado ciertos conceptos y ver en diferentes atlas lingüísticos cómo se expresan en las hablas rurales. También pueden citarse las listas de conceptos que, en las investigaciones orientadas a buscar la diversidad léxica, tengan numerosas designaciones. No será nada difícil juntar varios cientos de conceptos de este tipo, que tengan al menos un promedio de 10 variantes léxicas. Nadie podrá negar que se está así demostrando que el léxico de la lengua española tiende más a la diversidad que a la unidad. Sin embargo, aquí también, se tendrá que aclarar que esta diversidad se da en relación con cierto tipo de léxico correspondiente a las hablas rurales o de otra determinada clase de vocabulario, que corresponde a las hablas urbanas: es una diversidad, también, *relativa*.

Aunque a mi ver con menos convencimiento, alguien, para hablar de esta tendencia a la diversidad de vocabulario, podrá acudir a la comparación de las listas de

palabras que, asociadas a ciertos campos de interés, producen libremente diversos sujetos de diferentes dialectos (disponibilidad léxica). Ningún sujeto producirá exactamente las mismas palabras, en el mismo orden, que otro. Habrá notables diferencias de inventario y de jerarquía. Ya señalé que con trabajos de esta naturaleza se corre el riesgo de interpretar que toda palabra que no esté presente en tal o cual lista, se dé por inexistente en este o en aquel dialecto. Esto es falso. Si se desea hacer uso de investigaciones sobre léxicos «disponibles», me parece más adecuado proceder de otra manera. Adecuadamente establecidas las listas de palabras de los diversos dialectos que se pretenden comparar, convendría que, con métodos de dialectología tradicional, se averigüe cuáles palabras de la lista de determinado dialecto realmente se desconocen por los hablantes de los otros dialectos. Reconozco la complejidad de este sistema, pero me parece que los resultados resultarían más confiables. Sin embargo ello no quita que también aquí, se trata de una diversidad *relativa*, debido, entre otras razones, a que se parte de un muy limitado universo léxico.

Ahora bien, el que cualquier tipo de investigación sobre variedad léxica conduzca a resultados relativos no quiere decir que se trate de datos deleznable o que éstos estén distorsionando o tergiversando la naturaleza del lenguaje. Todo lo contrario. Lo que este tipo de trabajos nos está demostrando —y ello es ya algo sumamente importante— es precisamente el hecho de que, por definición, podría decirse, la variedad léxica de una lengua siempre es algo relativo. Aclaro de inmediato que aquí lo relativo hace alusión a la comparación, en cierto sentido *interna*, que puede establecerse entre el grado de diversidad de los elementos correspondientes a los varios niveles lingüísticos (fonología, gramática, léxico) de los diferentes sistemas (geográficos o sociales) del diasistema llamado lengua. Se puede decir así que, en relación con grado de diversidad fonológica, la diversidad léxica de la lengua española es mucho mayor. Es grande en ese sentido, comparada con la variedad fonológica. Si contáramos con datos suficientes, podría también hacerse otro tipo de comparación, que podría llamarse *externa*: ¿tiene el español mayor o menor diversidad léxica que el inglés o el francés? Las casi imposibles respuestas a esta clase de preguntas serían, obviamente, doblemente *relativas*.

José G. Moreno de Alba  
Universidad Nacional Autónoma de México  
Madrid, 06 de junio de 2005

### BIBLIOGRAFÍA

ALEA: *Atlas Lingüístico y Etnográfico de Andalucía* (por M. Alvar López, con la colaboración de A. Llorente y G. Salvador), Universidad de Granada y CSIC, Madrid, 1964-1973, 6 vols.

ALEC: *Atlas lingüístico-etnográfico de Colombia*, Asesores: M. Alvar y T. Buesa Oliver, Investigadores: L. Flórez (y otros), Instituto Caro y Cuervo, Bogotá, 1981-1983, 6 vols.

ALEICan: Manuel Alvar, *Atlas Lingüístico y Etnográfico de las Islas Canarias*, Ed. del Excmo. Cabildo Insular de Gran Canaria, Madrid, 1975-1978, 3 vols.

ALM: *Atlas Lingüístico de México*, Juan M. Lope Blanch (director), Antonio Alcalá, Gustavo Cantero, Juan López Chávez, Antonio Millán y José G. Moreno de Alba (investigadores), El Colegio de México y UNAM, México, 1990-2001, 6 vols.

DBEM 1986: *Diccionario básico del español de México*, dir. por Luis Fernando Lara, El Colegio de México, México.

DFEM 1982: *Diccionario fundamental del español de México*, dir. por Luis Fernando Lara, Comisión Nacional para la Defensa del Idioma Español, El Colegio de México y Fondo de Cultura Económica, México.

DEUM 1996: *Diccionario del español usual en México*, dir. por Luis Fernando Lara, El Colegio de México, México.

HAM CHANDE 1979: Ham Chande, R., “Del 1 al 100 en lexicografía”, en LARA, HAM y GARCÍA, 41-86.

LARA, HAM y GARCÍA 1979: Lara, L.F., Ham Ch., R. y García H., M.I., *Investigaciones lingüísticas en lexicografía*, El Colegio de México, México.

LÓPEZ 1992: López Chávez, J., “Alcances panhispánicos del léxico disponible”, *Lingüística*, 4, 26-124.

ROSENBLAT 1967: Rosenblat, Ángel, “El futuro de la lengua”, *Revista de Occidente*, 56-57, 155-191. 1971: *Nuestra lengua en ambos mundos*, Salvat, Estella (Navarra).

VALENCIA y ECHEVERRÍA 1999: Valencia, Alba y Echeverría, Max, *Disponibilidad léxica en estudiantes chilenos*, Ediciones Universidad de Chile, Universidad de Concepción, Santiago de Chile.

VARILEX 1994: *Varilex 2, Variación léxica del español del mundo: índice de palabras*, coord. Hiroto Ueda y Toshiro Takagaki, Tokio.

VARILEX 1999: *Varilex 7, Variación léxica del español en el mundo, Serie E*, coord. Hiroto Ueda, Tokio.