

## DIVERSIDAD LÉXICA Y UNIFORMIDAD GRAMATICAL. EN TORNO AL PORVENIR DE LA LENGUA ESPAÑOLA

JUAN M. LOPE BLANCH  
Universidad Nacional Autónoma.  
El Colegio de México.

Hace no muchos años, el profesor Juan López Chávez hacía referencia a una investigación realizada por él mismo en colaboración con la profesora Leonilda Rodríguez Fonseca, en la que habían llegado a la preocupante —por no calificarla de trágica— conclusión de que «al comparar entre sí algunos de los inventarios lexicales que se han realizado en el mundo hispánico, hemos encontrado un bajo grado de compatibilidad [...] Las comparaciones que 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»<sup>1</sup>.

Inquietante, ya que no trágica, la situación en que se halla la lengua española actualmente, caso de ser acertada la estadística que he recordado; acertada o, más propiamente, sintomática, sin negar por ello la certeza de tal estadística en su propio y particular «universo».

La cual coincide en lo fundamental con la que comentó el profesor Humberto López Morales, en un coloquio celebrado recientemente en España, con base precisamente en las también recientes investigaciones del profesor López Chávez a que acabo de hacer referencia. Según las cuales la compatibilidad de los *léxicos básicos* de tres centros lingüístico hispánicos —Ma-

---

<sup>1</sup> Cf. J. López Chávez, «Léxico fundamental panhispánico: realidad o utopía», en *Actas del IV Congreso Internacional de «El español de América»*, Santiago de Chile, Pontificia Universidad Católica de Chile, vol. II, 1995, págs. 1006-1014 (la cita corresponde a la pág. 1006).

drid, Puerto Rico y la República Dominicana— resulta ser angustiosamente reducida. En efecto, la coincidencia entre el léxico básico español y el dominicano no llega siquiera al 50% (es del orden del 47,9%); algo superior es la coincidencia del habla de Madrid con la de Puerto Rico (56,63%); pero lo más sorprendente —y desconcertante— es que sea menor que ella la compatibilidad de las hablas puertorriqueña y dominicana entre sí (50,59%), tratándose de dialectos geográficamente próximos e históricamente afines.

Pero la discordancia en los vocabularios de las ciudades hispánicas comparados es todavía muchísimo mayor en el caso de los *léxicos disponibles*. En efecto, la comparación de los 6.267 vocablos reunidos en el habla de Madrid, con los 8.810 de Las Palmas, los 6.059 de Puerto Rico y los 5.143 de Santo Domingo permitía al profesor López Chávez advertir que la coincidencia léxica entre Madrid y Puerto Rico llegaba apenas al 14,8%, y que entre Puerto Rico y la República Dominicana —donde se alcanzaba el máximo de convergencia<sup>2</sup>— ascendía sólo al 24,18%. El promedio total, tomando en cuenta los porcentajes de las seis posibles confrontaciones<sup>3</sup>, sería del orden del 20,15%.

¿Permitirían estos precisos datos estadísticos suponer que cualquier diálogo que tratasen de mantener un madrileño y un puertorriqueño, o un dominicano y un castellano —o cualquier otra combinación de las seis anotadas— habría de tropezar con constantes detenciones, necesarias para explicar el significado del 80% de sustantivos diferenciales de cada dialecto? Evidentemente que no es éste el caso. Deducir de las estadísticas obtenidas tal conclusión sería un verdadero desatino, totalmente alejado de la realidad lingüística. Pero las divergencias léxicas están ahí precisa y exactamente consignadas<sup>4</sup>.

Por estas mismas o parecidas fechas, hacía el profesor Raúl Ávila diversas indagaciones también de carácter lexicográfico, sirviéndose de un método diferente, de acuerdo con el cual los resultados estadísticos se referían al total de palabras y de vocablos hallados en cada *corpus* o fuente de estudio, atendiendo así a todas las clases de palabras —no sólo nombres y verbos— que aparecen en el discurso español contemporáneo. Y llegaba a resultados absolutamente distintos de los considerados hasta ahora. En el primero de sus trabajos, relativos al español general de México<sup>5</sup>, hallaba el

<sup>2</sup> Aunque en el caso del léxico básico tal convergencia era, sorprendentemente, la mínima, según acabo de señalar.

<sup>3</sup> A saber: Madrid-Las Palmas = 22,02%; Madrid-Sto. Domingo = 17,21%; Las Palmas-Sto. Domingo = 23,7%, además de las dos ya señaladas en el texto.

<sup>4</sup> Véanse las listas estadísticas de las páginas 1008 a 1012 del ensayo de López Chávez.

<sup>5</sup> «Las palabras de todos y las de cada uno: Un análisis estadístico del español hablado en México», en V. Demonte y B. Garza (eds.), *Estudios de lingüística de España y México*, Universidad Nacional Autónoma y El Colegio de México, 1990, págs. 335-350.

profesor Ávila que, en un total de 430.000 palabras consideradas, aparecían 9.309 vocablos, el 92% de los cuales pertenecía al español común, al vocabulario hispánico general. La proporción crecía todavía un poco en el habla culta (95,6%), pero no se alejaba mucho de ella la propia del habla popular (93,2%).

La homogeneidad léxica de la lengua española aparecía aún más reforzada en la segunda de las investigaciones de Raúl Ávila, relativa al léxico empleado en esos dos poderosísimos medios de irradiación y de nivelación lingüística que son la radio y la televisión<sup>6</sup>. Contrariamente a lo que suele pensarse, el habla de los locutores de ambos medios de información está muy próxima a la norma hispánica general, al menos en lo que al léxico respecta y a la ciudad de México se refiere. El análisis de los 1.797 vocablos, encontrados en un *corpus* total de 11.357 palabras, permite advertir que el vocabulario empleado por locutores de una estación radiofónica mexicana coincide en un 98,4% con el léxico hispánico general; sólo el 1,6% de las voces reunidas parece ser propio —aunque quizá no exclusivo— de México. La proporción de coincidencia aumenta todavía un poco en el caso de los locutores informativos de la televisión: el 99% de los 1.974 vocablos hallados en el *corpus* total de 11.325 palabras pertenece al léxico hispánico general.

La enorme divergencia de los resultados obtenidos mediante las investigaciones aquí resumidas obedece, lógicamente, a la gran diferencia de los métodos empleados en cada una de ellas. Diferencia que depende de los diversos objetivos de cada investigación. No pretendo juzgar aquí la validez o pertinencia de dichos métodos, ni me parece siquiera asunto de particular interés. Como ya he dicho en ocasiones anteriores, me interesa mucho más la *lengua* española —y su porvenir— que la lingüística, sin que esto signifique, por supuesto, que me parezca intranscendente la metodología que haya de emplearse en cada investigación sobre nuestro idioma. Métodos diversos pueden ser igualmente válidos según sean los objetivos que con ellos se persigan. Y como mi objetivo principal es el de determinar el grado de cohesión —o de diversificación— de una lengua hablada en tantos países como la española, no deja de preocuparme advertir que la «compatibilidad de los léxicos básicos» de los países hispánicos pueda oscilar en torno al 50% solamente, aunque no deja de tranquilizarme el saber que el léxico total del habla mexicana de nuestros días coincide en un 90% con el léxico hispánico general.

Para determinar el grado de *comprensibilidad* léxica del habla hispánica espontánea, se me ha ocurrido hacer una prueba muy sencilla y metodológi-

<sup>6</sup> Cf. «El lenguaje de la radio y la televisión: Primeras noticias», en el *II Encuentro de Lingüistas y Filólogos de España y México*, Salamanca, Ediciones Universidad, 1994, págs. 101-117.

camente simplicísima: determinar en qué medida el léxico usado actualmente en el habla culta madrileña resulta extraño el caudal léxico de los hablantes mexicanos. Para ello, me he servido del volumen de encuestas hechas en Madrid bajo la dirección de Manuel Alvar y Antonio Quilis, y publicadas por M. Esgueva y M. Cantarero<sup>7</sup>. Con la colaboración de mi ayudante en el Sistema Nacional de Investigadores, Julieta López Olalde —quien leyó con todo cuidado dicho volumen para detectar el vocabulario madrileño ajeno o, al menos, extraño al habla mexicana— he obtenido la siguiente información.

He calculado que en ese volumen de encuestas se reúne un total de poco más de 133.000 palabras *aproximadamente*<sup>8</sup>. Figuran entre ellas 135 voces, correspondientes a 91 vocablos, que resultan «extrañas» —por sí mismas o por el significado con que se emplean— para el habla mexicana. Son ellas, en simple orden alfabético, las siguientes: *alférez* (apareció en 2 ocasiones), *americana*, *apetecer*, *asignatura* (en 2 lugares), *autocar* (también 2 veces), *bachiller*, *barracones*, *bedel*, *bomba* (pasarlo...), *botijero* y *botijo*, *breva*, *camilla*, *caqui*, *carnet* (2 veces), *cascar(se)*, *castañas*, *castizo*, *cisco* (2), *consigna*, *constipado*, *convalidación*, *catarro*, *cotilleo*, *crios*, *chalado* y *chaladura*, *chalet*, *charlar*, *chatitos*, *chaval(es)*, *chavalín* (2), *chuleta*, *chusquero* (2), *deberes*, *deje*, *doncella*, *económica*, *faena*, *fallé (sic)*, *fiambres*, *follón*, *fonendo*, *frigorífico*, *guateada*, *jaleo*, *jersey*, *lacito*, *largar*, *librería*, *maja* (2), *marcharse*, *matrícula*, *milicia* (4 veces) y *mili*, *millonetis*, *ministerio* (2), *mogollón*, *mus*, *narices* (de...), *pancho*, *pandereta*, *parador* (de turismo), *patata*, *pegas*, *perras*, *peso*, *pilarista*, *pillar* (4), *pisto*, *plantificar*, *plaza*, *portal* (2), *portante*, *rebujado*, *repollo*, *reválida* (2), *salado*, *salón*, *serial* (de televisión), *suspender*, *taco* (3), *telediario* (2), *tertulia*, *tiesto*, *tozudo*, *tresillo* (2), *tapas*, *veranear* (3), *veraneo* (4) y *zapatillas*.

Desde luego que la «extrañeza» de estos vocablos dentro del léxico mexicano actual no es siempre de la misma clase ni alcanza el mismo grado de intensidad. En algunos casos —como en el de *alférez*, por ejemplo— la voz es ajena al sistema léxico mexicano, en el que existe un vocablo equivalente: *subteniente*. En otras ocasiones, ni la palabra ni el concepto son conocidos en México: tal podría ser el caso del juego de naipes denominado *mus*, que no se practica en México. En otros casos, aunque la voz española se conoce —y a veces se identifica precisamente como española— no se usa en el habla mexicana, sino que se emplea otra con igual significado: así

<sup>7</sup> Cf. *El habla de la ciudad de Madrid. Materiales para su estudio*, Madrid, Consejo Superior de Investigaciones Científicas, 1981.

<sup>8</sup> Para llegar a tal número, he calculado la cantidad *promedio* de palabras que aparecen en cada una de las páginas del libro, y la he multiplicado por el total de páginas que lo integran: 417, descontando ya las páginas en blanco o incompletas que figuran en su numeración impresa (de 449).

sucede en el caso de *asignatura*, sustituida en el habla mexicana por *materia*. Otras veces la voz española significa en el habla mexicana cosa muy distinta que en Madrid: *pancho*, por ejemplo, es sólo el gentilicio de Francisco, pero no adjetivo calificativo sinónimo de ‘tranquilo, indiferente’ («quedarse tan *pancho*»). En algunas ocasiones, el habla mexicana no ha dado entrada a acepciones nuevas —generalmente familiares o aun populares— que sí ha aceptado el habla madrileña: tal el caso de *breva* ‘especie de higo’, pero no, como en España, «provecho logrado sin sacrificio... ventaja inesperada» (DRAE), cuyo equivalente mexicano podría ser tal vez *hueso* («agarrar un buen hueso», por ejemplo).

Aunque no era ése el propósito de estas páginas, presentaré esquemáticamente una posible clasificación de las «extrañezas» que esa lista de madrileñismos origina en el habla mexicana.

Casos en que hay coincidencia conceptual en ambos dialectos, pero con denominaciones diferentes: *alférez* (en México, *subteniente*), *autocar* (*autobús* o *camión* en México), *barracón* (*galerón*), *carné(t)* (*credencial*, documento de identidad; no se usa la voz francesa), *críos* (*chamacos*), *chalado* y *chaladura* (*chiflado* y *chifladura*), *chalé* (simplemente *casa*, frente a *departamento*), *chaval* (*chavo*, *chamaco*), *deje* (*sonsonete*, *tonito*), *fallé*<sup>9</sup> (creo que a partir del fr. *foyer*, *vestíbulo* o el anglicismo *lobby*), *follón* (*mitote*), *frigorífico* (*refrigerador*), *guateada* (*acolchada*), *jaleo* (*relajo*), *jersey* (*suéter*), *mogollón* (*montón*), *pegas* (sustantivo: *inconvenientes*, *dificultades*), *perros* (*monedas*, *fierros*, *morralla*), *pisto* (darse... *importancia*, *ínfulas*), *portal* (*zaguán* o *puerta*), *rebujado* (*arrugado*, *enredado*), *repollo* (*col*), *serial* (*serie* de T.V.) *telediario* (*noticiero*), *tiesto* (*maceta*), *tozudo* (*necio*), *tresillo* (*juego de sala*), *tapas* (*botanas*).

Con notable frecuencia la palabra española es también conocida en el habla mexicana, pero posee en ésta un significado distinto —muchas veces, el original— del que ha adquirido en Madrid. Así sucede en los casos de *americana* (la prenda de vestir, en México *saco*), *constipado* (*acatarrado*; *constipado* tiene en el habla mexicana el significado de ‘estreñido’), *chuleta* (en el habla escolar mexicana, *acordeón*), *deberes* (*tareas*, en el mismo sector lingüístico), *doncella* (no en el sentido primario de ‘virgen’, sino de *recamarera*), *económica* (una persona, por *ahorradora* o buena administradora), *faena* (no en el sentido de ‘trabajo’ sino de ‘mala pasada’ —DRAE—, lo cual sería en México *trastada*), *fiambres* (*carnes frías*), *laticito* (*moñito*), *librería* (impropiamente, en vez de *librero*<sup>10</sup>), *marcharse* (*ir-*

<sup>9</sup> Cf. pág. 256: «Y era muy bonita la salida del teatro Real por la noche. Y en la ópera pues había un *fallé* (*sic*) grandísimo... Un ‘fallé’, que es un salón muy grande».

<sup>10</sup> Cf. pág. 119: «Tengo una mesa de centro muy mona... después tengo dos *librerías*, y me falta la mesa de comedor».

*se*; *marchar* se reserva en México para el ejército o grupos de manifestantes), *matrícula* (*inscripción*, en un centro docente), *milicia(s)* o *mili*<sup>11</sup> (*servicio militar*), *Ministerio* de Educación, etc. (*Secretaría*), *peso* (por *balanza* o *báscula*: pág. 120), *pillar* (por *sorprender*, *agarrar*, *quedar*<sup>12</sup>), *plaza* (en el sentido de *lugar*, *espacio*<sup>13</sup>), *salado* (*gracioso*), *salón* (*sala*, en la casa), *suspender* (en algún examen: *reprobar*), *taco* (*lío*, *confusión*; o también *palabrota*, *grosería*; y aun *fajo* de billetes o de papel, etc.<sup>14</sup>), *zapatillas* (*pantufilas*; *zapatilla*, en México, es la de ballet o el calzado fino de mujer, hecho con tiras delgadas de piel).

Se da también el caso de que la palabra española se conozca —débilmente— en el habla mexicana, en la cual se prefiere otro término en algún modo equivalente. Lo que en Madrid es «no me *apetece* ir al cine» (pág. 125), aunque se entiende, sería preferible en México «no se me *antoja*», o *apetecer* algún alimento sería *gustar*. Cosa semejante sucede en los casos de *asignatura* (*materia*), *convalidación* (*revalidación*), *cotilleo* (*cotorreo* o *chismorreo*), *charlar* (*platicar*), *pandereta* (*pandero*) y *patatas* (*papas*).

Cuando el concepto, acción u objeto, no es conocido en México, resulta natural que no exista en el habla mexicana el vocablo correspondiente. Así, han resultado ser voces desconocidas para mis informantes<sup>15</sup> *bedel*, *botijo*, y *botijeros*, *camilla*<sup>16</sup>, *caqui*<sup>17</sup>, *consigna* (en las estaciones del ferrocarril), *cotarro*, *chattitos*<sup>18</sup>, *chusquero* (oficial del ejército), *mus*, *parador* (de turismo, inexistente en México), *reválida* (al fin del bachillerato), *tertulias* y *veranear* o *veraneo*<sup>19</sup>.

Se desconocen también en México algunas frases o locuciones de carácter familiar usuales en España, como «pasárselo *bomba*» (pág. 188), que en

<sup>11</sup> Cf. «Estuve haciendo *milicia* en Lérida» (pág. 94), «al mismo tiempo que hacía *mili-cias*» (pág. 94), «haciendo las prácticas de la *milicia*» (pág. 152). La *milicia* es en México sinónimo, en cierta medida, de *ejército*.

<sup>12</sup> Cf. «Nos ha *pillado* una granizada endiablada» (pág. 239); «el colegio que más cerca *pillaba* de casa...» (pág. 323); «¿Sueles ver la televisión? —Pues, al rato que me *pillan*» (pág. 194). No se usa así el verbo en el habla mexicana.

<sup>13</sup> Cf. pág. 115: «Pero no había *plaza*» (en el hotel).

<sup>14</sup> Increíblemente, el *Diccionario* de la Real Academia —aunque muy generoso en la aceptación de mexicanismos en su última edición, de 1992— no da entrada en sus páginas al *taco* mexicano, el rollo de tortilla relleno de diversos alimentos, que es una de las formas más populares y comunes de la alimentación en el país.

<sup>15</sup> He consultado con un buen número de hablantes nativos mexicanos la validez y personalidad de cada uno de los vocablos diferenciadores.

<sup>16</sup> En el sentido de «*mesa-camilla*», no en el de «*cama portátil para transportar enfermos*».

<sup>17</sup> «*Vestirse de caqui*» (cf. pág. 105) en el sentido de enrolarse en el ejército. En México sí se usa *caqui* como nombre de la tela o de su color.

<sup>18</sup> Por (beber) *copitas* de vino, costumbre española no practicada en México.

<sup>19</sup> En México no es el *verano* la época en que la gente sale masivamente de vacaciones. Aquí «se va de vacaciones» —raramente se usa el verbo *vacacionar*— en diversas épocas, como Semana Santa, Navidad o en algún «puente», pero no precisamente en verano, que es la época *baja* del turismo.

México sería «pasarla *padre* o *de lujo* o *en grande*», «de *narices*»<sup>20</sup>, «que-darse o estar tan *pancho*», «tomar el *portante*», «ser muy *castañas*» (*latosos, exigentes?*), «*cascarse* algo muy pesado o aburrido» (pág. 112: *soplarse, aventarse*), «*madrileños castizos*», «*largar*(le) un sermón», *maja*, «caer o tener una *breva*», son expresiones que, por el contexto, pueden entenderse *grosso modo* en México, pero que no son aquí usadas.

Finalmente, algunos casos muy particulares: *bachiller* usado en Madrid (pág. 99) por *bachillerato*, *millonetis* (pág. 124), que en México sería *milloneta* (o *platudo*), *fonendo* (pág. 140), por *fonendoscopio* (que no es propiamente vocablo diferenciador de normas geográficas, sino voz especializada de la Medicina), y *pilaristas* (pág. 199), que acaso sea —imagino— escolar del Colegio del Pilar, en Madrid.

Como puede verse, una alta proporción de estos madrileñismos inusitados en México pertenece al habla familiar, popular y figurada, así como al habla estudiantil (*reválida, asignatura, chuletas, bachiller, suspender, etc.*). Y como puede también advertirse, no todas las voces consignadas presentan igual tipo ni equivalente grado de diferenciación. Pero las he considerado todas ellas e incluido en las estadísticas por igual, para que no se me pueda acusar de parcialidad en beneficio de la homogeneidad dialectal del idioma.

\* \* \*

Ahora bien, ¿qué conclusiones —provisionales, simplemente aproximativas, desde luego— se podrían deducir de esta confrontación, tan reducida y acaso simplista<sup>21</sup> del léxico madrileño con el vocabulario usual de los mexicanos? Creo que, en esencia, sólo lo siguiente:

Si dentro de un *corpus* léxico total integrado por algo más de 133.000 voces aparecen 135 palabras —correspondientes a 91 vocablos— usuales en el habla de Madrid pero no en la de México, ello significa que esas 135 palabras representan el 0,1% de apariciones «discordantes» en la *cadena hablada* o, lo que es lo mismo, que el 99,9% de las palabras utilizadas en el discurso por los hablantes madrileños son también conocidas y usadas por los mexicanos<sup>22</sup>.

<sup>20</sup> Cf. pág. 112: «Hacia un calor de *narices*».

<sup>21</sup> Aunque, por otra parte, tenga la virtud de estar hecha con base en el *habla* espontánea, real, es decir, con atención a la lengua viva, en funcionamiento, en su realidad comunicativa.

<sup>22</sup> Trabajo artesanalmente: no juzgo indispensable contar con toda exactitud los *vocablos* que aparecen en las encuestas madrileñas. He calculado por procedimientos no computacionales que pueden oscilar en torno a los 3.000 o poco menos. De ser así, los 91 vocablos madrileños no usuales en México representarían el 0,3% aproximadamente del total empleado por los hablantes de la capital de España. Inclusive si mis cálculos estuviesen equivocados, y las 133.000 palabras incluyeran sólo 2.000 vocablos —cosa que dudo mucho—, el porcenta-

Cierto es que en este tipo de recuentos se incluyen todas las palabras, cualquiera que sea su clase o categoría, y no sólo las voces semánticamente primarias o primordiales: sustantivos, verbos, calificativos. Con el procedimiento empleado por mí, se cuantifican igualmente preposiciones, conjunciones, determinantes (artículos, demostrativos, posesivos, etc.) y adverbios. Pero creo que, si se trata de determinar el grado de coherencia y de «intercomprensibilidad» de los dialectos de una misma lengua, eso es precisamente lo que debe hacerse. Porque esos vocablos que algunos llaman «gramaticales» son tan significativos como los «nocionales» y de más constante e importante empleo en la lengua hablada, en la lengua viva, no en la lengua codificada en casilleros o sectores gramaticales.

Creo, pues, acertado que en una de las primeras y más amplias —extensas, profundas— investigaciones sobre la variedad del léxico hispánico, realizada hace años en Puerto Rico<sup>23</sup>, se diera entrada en su recuento a todas las *palabras* que aparecieron en las muestras de la lengua española —escrita o hablada— que se espigaron léxicamente. Los resultados obtenidos en esa amplia investigación<sup>24</sup> son muy reveladores. Recordaré aquí algunos de ellos:

Entre las «Primeras 87 unidades léxicas más frecuentes» en español, figuran, de un lado, doce sustantivos<sup>25</sup> (=14,6%), once verbos<sup>26</sup> (=13,4%) y trece adjetivos, pero de éstos sólo tres son «conceptuales» y diez «gramaticales»<sup>27</sup>. Y, de otro lado, siete preposiciones y también siete conjunciones<sup>28</sup>, así como diecisiete formas pronominales y quince adverbiales<sup>29</sup>.

---

je de discordancia ascendería al 4,5%, lo cual sigue estando muy lejos del 80% calculado por López Chávez para el *léxico disponible no compartido* o del 50% correspondiente al *léxico básico* divergente. A este calificativo de *básico* habré de volver a referirme muy pronto.

<sup>23</sup> Ismael Rodríguez Bou (Director), *Recuento de vocabulario español*, Río Piedras, Universidad de Puerto Rico, 1952; 2 vols. (en realidad, tres volúmenes).

<sup>24</sup> «El total de palabras contadas alcanza la cifra de 7.066.637; el total de unidades léxicas es de 20.542; el total de formas de inflexión, 62.888 y el total de palabras diferentes llega a 83.430» (pág. 8). Naturalmente que estadísticas hechas con un registro de más de 7 millones de palabras serán más fidedignas y sintomáticas que las hechas con base en 133.000 palabras o en *corpora* mucho menores aún.

<sup>25</sup> Son, en orden alfabético, *agua, casa, día, escuela, flor, gato, libro, maestra(-o), mamá, niño, papá, y vez*. Más adelante haré un breve comentario sobre esta lista.

<sup>26</sup> Fueron *decir, estar, ir, gusta(r), haber, hacer, jugar, mira(r), ser, tener y ver*.

<sup>27</sup> *Bueno, grande y todo* (=3,7%). Los otros diez son *una(-o), dos, tres, el, la, este, ese y mi, tu, su* (=12,2%). El total de 13 representa el 15,8%.

<sup>28</sup> También en orden alfabético: *a, con, de, donde, en, para y por* (=8,5%), y *como, cuando, pero, porque, que, si e y* (=8,5%). Total de nexos = 17%.

<sup>29</sup> Son *ellas yo, me, mi, tú, te, él, ella, le, la, lo, se, nos, mío, tuyo, qué, cómo y otro* (=20%) de un lado, y *ahora, allí, aquí, así, bien, después, entonces, más, mucho* (no se especifica en el estudio proporción de empleo, si es como adverbio o como adjetivo), *muy, nada* (también aparecen indiscriminadamente reunidos sus diversos valores de adverbio o de pronombre indefinido), *no, sí, también y tan* (=18,3%).



Si interesa atender a la facilidad o dificultad de intercomunicación, de comprensión, entre unas hablas y otras, no cabe duda de que todos los elementos gramaticales del idioma —y no sólo los que tienen forma léxica autónoma (lexemática), sino también los morfemas— desempeñan un papel importantísimo. Es obvio que, si un sustantivo cualquiera —como *mesa* o *virtud*, por ejemplo— posee un significado propio, mejor o peor delimitado conceptualmente, también una preposición o una conjunción puede tener un significado, aunque quizá de distinta naturaleza, muy claro y preciso, e imprescindible en el acto de la comunicación, del habla. *Para*, en efecto, portará un significado de finalidad, o *porque* otro de causalidad, tan comprensible para los hablantes de español —tanto o más— que el de *virtud* o *indefinición*, pongamos por caso. No sólo eso: las palabras o los elementos *morfosintácticos* de la lengua son los más importantes e indispensables para el proceso de comunicación, que es —creo— la finalidad esencial de cualquier sistema lingüístico. Y digo que tan trascendental importancia no es exclusiva de las partículas o de las palabras determinativas del sustantivo o del verbo, sino que disfrutaban también de esa importancia los morfemas propios de las lenguas flexivas. Así, el morfema *-abas* de cualquier raíz verbal de la primera conjugación, siempre comportará la idea, el significado, de tiempo pasado, de segunda persona pronominal (*tú*) en su forma singular (frente a la plural de igual tiempo y persona *-abais*), en tanto que el morfema *-aré* en la misma clase de verbos siempre evocará un tiempo venidero, una primera persona gramatical (*yo*) en su singularidad. Y no dejan de ser elementos sumamente significativos los morfemas de género y número en el nombre, o los valorativos en su gran variedad: *-ito*, *-ejo*, *-ísimo*, etc. De tal modo que la diversidad léxica que pueda existir actualmente entre unos dialectos hispánicos y otros en el sector de los sustantivos y de los verbos —y aun de los adjetivos— será impedimento mucho menor para la fluida comunicación entre los hablantes, que el que originaría una hipotética diversidad de los elementos —palabras o morfemas— gramaticales<sup>30</sup>.

Y es el caso, además, de que esas «palabras de significado gramatical» son las que con mayor frecuencia se emplean en la lengua hablada, según

<sup>30</sup> Me permitiré construir, a manera de ejemplo, un imaginario enunciado hispánico en el que la determinación dependiera sólo de diversidades léxicas sustanciales, primarias: «Lo haré con mi *truslo* cuando llegue tu *ferco*», mensaje que para el interlocutor de diferente dialecto resultaría ser en gran medida incompleto, pero cuya oscuridad relativa podría ser superada fácilmente preguntando «¿Qué son *truslo* y *ferco*?»; y con la respuesta «Un *truslo* es aquí un *cuchillo* y *ferco* es otra manera de llamar al *sobrino*» quedaría solucionado el problema. En cambio el enunciado resultaría mucho más difícil de interpretar si los elementos divergentes no fueran los nominales *truslo* y *ferco*, sino los morfosintácticos: «*Cu hata pe li cuchillo, ro llemu sa sobrino*». Felizmente, las divergencias dialectales de la lengua española no han penetrado en su sistema morfosintáctico.

muestra la estadística presentada por el equipo puertorriqueño que dirigió el profesor Rodríguez Bou. En tanto que los sustantivos, adjetivos calificativos y verbos de mayor uso representaban en su conjunto sólo el 31,7%, las preposiciones, las conjunciones, las formas pronominales y las determinativas alcanzaban en total el 49,4%; el 18,3% restante correspondía a los adverbios *latu sensu*. Y aunque el estudio de Rodríguez Bou no lo precisa, supongo que el índice de aparición de cada una de esas palabras nexuales o complementarias era muy superior al de cada uno de los sustantivos o verbos reunidos en ese apartado de máximo empleo<sup>31</sup>. Por todo ello, me parece que lo más adecuado y justo sería dar el nombre de «léxico *básico*» precisamente al integrado por esas palabras gramaticales que aparecen constante y reiteradamente en todas las hablas hispánicas. Y es el caso, asimismo, de que esas palabras gramaticales son las que se han mantenido y viven prácticamente sin cambios, sin diferenciaciones, en todas las hablas hispánicas. En lo que respecta a las preposiciones y a las conjunciones, sólo una de ellas presenta diversidad de significado en uno de los grandes dialectos hispánicos: el mexicano. Donde la preposición *hasta*, así como el nexos conjuntivo *hasta que*, sin dejar de poder expresar, como en el español general, el límite *final* de una acción («Trabajé *hasta* la medianoche» o «Esperaré *hasta que* lo acabes»), han pasado a expresar el límite *inicial* de una acción durativa o, inclusive, puntual: «Lo entierran *hasta el domingo*» o «*Hasta que* se casó supo lo que era tener responsabilidades»<sup>32</sup>, lo cual proporciona una falsa información a los hablantes de los demás dialectos hispánicos.

Alguna mayor divergencia existe en el caso de ciertas formas pronominales: de un lado, la originada por el *voseo* usual en amplias zonas de América; del otro, por la casi total desaparición del pronombre de segunda persona plural, *vosotros*, en la casi totalidad de las hablas americanas, con la consiguiente decadencia del posesivo *vuestro*, en favor de *suyo* o *de ustedes*. Divergencias éstas, entre los dialectos españoles y los hispanoamericanos, que, sin embargo, no impiden la fácil comunicación entre los hablantes de unos y otros.

La homogeneidad esencial de los elementos gramaticales, morfosintácticos, del idioma español permite suponer que su unidad fundamental está

<sup>31</sup> Los artículos, por ejemplo, la conjunción *y*, así como muchas preposiciones hubieron de aparecer, a lo largo de las encuestas realizadas, con mayor frecuencia que el sustantivo *agua* o el verbo *gustar*. No considero que puedan ser muchas las oraciones gramaticales en que no haga acto de presencia algún artículo o alguna preposición, pero sí serán muchísimas, obviamente, las oraciones en que no aparezcan ni el sustantivo *agua* ni el verbo *gustar*.

<sup>32</sup> Cf. mi artículo «Precisiones sobre el uso mexicano de la preposición *hasta*», *Anuario de Lingüística Hispánica* (Universidad de Valladolid), VI, 1990, págs. 293-321 (recogido también en mi libro de *Ensayos sobre el español de América*, México, Universidad Nacional Autónoma, 1993, págs. 157-189).

plenamente garantizada por mucho tiempo más, no obstante las limitadas diferencias léxicas que pueda haber entre unos y otros de sus dialectos. Diferencias menos importantes y menos dañinas para la comunicación de todos los hispanohablantes entre sí, que las derivadas de las desviaciones o peculiaridades fonéticas de las diversas hablas hispánicas. No creo que las dificultades de comunicación oral que puedan darse entre un andaluz y un mexicano, por ejemplo, se habrían de deber a diferencias de vocabulario, sino posiblemente, en mayor escala, a diferencias de carácter fonético.

\* \* \*

Una última consideración sobre los alcances de investigaciones en torno al léxico como las aquí consideradas. Creo que todas ellas son, sin duda, provechosas y válidas, pero con una validez sólo relativa, *sintomática*, y no, de ninguna manera, *matemática*. Solemos presentar las estadísticas resultado de nuestros recuentos numéricos con una aparente exactitud asombrosa. Los porcentajes que proporcionamos revelan hasta las centésimas de punto correspondientes al fenómeno de que se trate: 22,02%, por ejemplo, o 18,97% precisamente, y no, ya redondeando, 22% o 19%. Y claro está que esos porcentajes son absolutamente ciertos en lo que se refiere a la muestra analizada en cada caso, pero no, desde luego, al hecho lingüístico de que se trate. Cualquier cambio o ampliación de la muestra estudiada originaría cambios estadísticos de no sólo centenas o decenas de puntos, sino de uno o varios puntos porcentuales. Que la lingüística no es una ciencia exacta, dada la enorme variabilidad de las lenguas<sup>33</sup>.

La relativa validez sintomática de nuestras investigaciones depende también de otros diversos factores. No sólo, desde luego, del método de trabajo que se elija, sino además de circunstancias muy particulares, como por ejemplo del lugar, de la sede geográfica, en que se lleve a cabo la investigación. En el caso del estudio dirigido por Rodríguez Bou en Puerto Rico, influyó determinantemente esa localización geográfica, así como el propósito —acaso sería mejor decir los destinatarios principales— de la investigación. En efecto, no obstante que el equipo investigador analizó muestras, orales o escritas, de diversa procedencia regional hispánica, sus vecindades

---

<sup>33</sup> Atendiendo a este hecho indiscutible, al presentar los resultados estadísticos de los diversos fenómenos lingüísticos cuantificados en el *Atlas Lingüístico de México* (El Colegio de México y UNAM, 6 vols., 1990-2000), preferí referirme a los porcentajes encontrados, no con cifras exactas, sino con calificativos amplios, a saber: G = general (entre 95% y 100% *aproximadamente*), F = frecuente (alrededor del 80%, entre 90% y 70%), M = medio (entre 60% como máximo y un mínimo del 40%, también *aproximadamente*), P = poco (entre 30% y 10%) y E = esporádico (alrededor del 5%).

geográficas pesaron notoriamente en el trabajo... y en los resultados. Esto queda probado por el hecho de que el gentilicio *puertorriqueño*, en sus distintas variantes —femenina, plural, superlativa, etc.— apareció con mucha mayor frecuencia que *mexicano* o que *dominicano*, y muchísimo más que *panameño*, *boliviano* y *nicaragüense*, por ejemplo; sólo *español* superó ligeramente a los derivados de *Puerto Rico*, pero imagino que ese gentilicio haría referencia, en la mayor parte de los casos, a la «lengua *española*», no a los originarios de España. Paralelamente, *yanqui*, designación puertorriqueña de los estadounidenses, superó ampliamente a *gringo*, forma más usual en otros países hispanoamericanos. Y la finalidad docente del lexicon reunido por Rodríguez Bou y sus colaboradores determinó —me parece— que entre los doce sustantivos de más frecuente empleo en español, figurasen *escuela*, *libro*, *maestro*, *niño*, *mamá* y *papá*, que posiblemente no sean los de mayor uso en otros sectores temáticos.

A estas restricciones creo que habría que añadir la referente al «tamaño de la muestra». No pocas de las investigaciones de carácter lexicográfico se hacen con base en un *corpus* lingüístico reducido: quinientas o doscientas mil palabras, cuando no menos. Y ello me parece muy insuficiente si es que se pretende llegar a estadísticas porcentuales que estén muy próximas a la realidad. En Puerto Rico se trabajó con base en un *corpus* de algo más de siete millones de palabras y, no obstante tan elevado número, se advierte en los resultados limitaciones como las que acabo de señalar. Muy sintomático me parece, a este respecto, el recuento —y los resultados— obtenidos por mí mismo, con ayuda de un nutrido equipo de colaboradores, al estudiar el número y la proporción de uso de los indoamericanismos mexicanos empleados en el español de México en su actual totalidad<sup>34</sup>. Nuestra base de estudio estuvo integrada por un total de 4.600.000 palabras aproximadamente, espigadas tanto de la lengua hablada mexicana (más de 2.211.000 voces), cuanto de la escrita (2.393.000 palabras). En ese *corpus* total se reúnen muestras del habla propia de toda clase de informantes: hombres y mujeres, jóvenes y ancianos, cultos y analfabetos, etc., hasta un total de 490 hablantes; las muestras de la lengua escrita proceden, asimismo, de textos de muy diversa índole: novelas, cuentos, ensayos, poesía, obras de teatro, publicaciones periódicas —diarios y revistas de amplia difusión—, algunas cartas, etc. Nuestro único y simple propósito era el de determinar cuántos y cuáles eran los indoamericanismos mexicanos que se usaban en el español de México actualmente, y el grado de *vitalidad* de cada uno de ellos. Pues bien, lo que importa ahora y aquí señalar es que, ascendiendo a más de 4.600.000 palabras el *corpus* analizado, aparecieron

<sup>34</sup> Cf. *Léxico indígena en el español de México*, El Colegio de México, 2.<sup>a</sup> ed., 1979.

312 vocablos (correspondientes a 237 lexemas) de origen amerindio mexicano, algunos de ellos enteramente desconocidos para la mayor parte de los hablantes mexicanos —como *pizote*, *guare* o *chomite*—, en tanto que no aparecieron —y esto es lo que me interesa destacar aquí— en nuestras encuestas, entre esos cuatro millones y medio de palabras reunidas, nada menos que 50 indigenismos (correspondientes a 38 lexemas) que son de uso amplio —algunos de conocimiento general— por parte de los hablantes mexicanos de nuestro tiempo, como sucede en los casos de *capulina*, *chípil*, *chipote*, *matatena*, *mitotero*, *petacón*, *tocayo*, *tololoche* entre otros<sup>35</sup>. Si partiendo de una base que rebasa los cuatro millones y medio de palabras, permanecen ocultos 50 vocablos, ¿qué seguridad puede proporcionar una estadística hecha con base en medio millón, o menos, de palabras? Para determinar con rigor y precisión el grado de homogeneidad o de divergencia léxica de la lengua española en los diversos territorios en que se habla actualmente, habría que hacer un registro total de los términos empleados en cada uno de esos territorios, medir su respectiva *vitalidad*, y compararlos entre sí cuidadosamente. Tarea sin duda irrealizable, ya que ni la más potente computadora eximiría a los investigadores de llevar a cabo los miles y miles de encuestas previas que permitirían acumular los millones y millones de palabras necesarias para hacer un recuento exhaustivo del léxico hispánico actual. Nuestras indagaciones no podrán, por ello, dejar de ser estudios muy parciales, simples *calas* superficiales, de alcance relativo, sólo *sintomático* o aproximativo, sin pretensiones de exactitud matemática.

El hecho de que la compatibilidad léxica de los dialectos hispánicos por él estudiados no llegara nunca ni siquiera al 56%, inclinó al profesor López Chávez a hacer, en el estudio citado en la nota 1 de estas páginas, «tres posibles reflexiones: o se trata de dialectos muy alejados léxicamente hablando, o las investigaciones recogieron muestras no representativas, o la técnica de recopilación de léxicos básicos no refleja con fidelidad el léxico de un dialecto» (pág. 1006). Dado que las hablas de Puerto Rico, la República Dominicana, México y aun España no representan dialectos muy alejados entre sí, habré de declararme enteramente de acuerdo con la tercera reflexión de López Chávez. Y añadir que, en mi opinión, no basta con recopilar léxicos básicos, sino que es preciso medir también el índice de *vitalidad* de cada uno de los vocablos recogidos, atendiendo con sumo cuidado a su frecuencia de aparición en la lengua hablada, en la lengua viva<sup>36</sup>.

<sup>35</sup> Cf. el libro citado en la nota anterior, pág. 32, n. 47.

<sup>36</sup> A este respecto, sería ingenuo y equívoco conceder la misma importancia —numérica y contrastiva dialectalmente— a indoamericanismos mexicanos como *chocolate* y *tololoche*, o a vocablos hispánicos como *hombre* y *mangosta*, o a verbos asimismo hispánicos como *vivir* o *desvirar*. Grave, para la homogeneidad y el futuro de la lengua española, sería que en vez

Creo, en conclusión, que el índice de comprensibilidad existente todavía hoy entre unos y otros dialectos hispánicos sigue siendo muy elevado, particularmente en el habla culta y aun en la estándar. Disminuye, sin duda, en las diversas jergas socioculturales o profesionales, sobre todo en el sector léxico, pero también en ellas las palabras gramaticales básicas siguen siendo las mismas, como lo sigue siendo la estructura morfosintáctica del idioma español. Todo lo cual me permite seguir viendo con pleno optimismo el futuro *histórico* de nuestra lengua, sin atreverme a hacer pronóstico alguno sobre lo que podría sucederle en lo que Dámaso Alonso denominó —en correlación con la *prehistoria*— la *posthistoria* de la lengua<sup>37</sup>. Cuya «hipotéticamente posible» fragmentación tampoco la veo como un hecho inevitable, al menos en todavía muchos siglos.

---

de *hombre* se dijera en algunos de sus dialectos *mon* o *petián*, pero mucho menos inquietante sería que la *mangosta* se denominara en alguno de ellos *peturia* o cualquier otra cosa divergente. No basta con calcular números y porcentajes, sino que es preciso añadir, a lo cuantitativo, lo cualitativo.

<sup>37</sup> Cf. «Para evitar la diversificación de nuestra lengua», en *Presente y Futuro de la Lengua Española*, Actas de la Asamblea de Filología del I Congreso de Instituciones Hispánicas, Madrid, Ediciones Cultura Hispánica, 1964, págs. 259-268 (véase en especial la pág. 259; publicado también en *Arbor*, LV, 1963, núms. 211-212, págs. 7-20).