

## SOBRE LA INFLUENCIA DE LAS LENGUAS INDÍGENAS EN EL LÉXICO DEL ESPAÑOL HABLADO EN MÉXICO

DESDE que Rodolfo Lenz afirmó, en 1893, que “el español de Chile es, principalmente, español con sonidos araucanos”<sup>1</sup> hasta nuestros días, mucho es lo que ha cambiado el parecer de los lingüistas respecto de la fuerza e influencia de los sustratos. Según Malmberg, por ejemplo, sólo debe pensarse en la acción del sustrato cuando fallan las explicaciones internas, sistemáticas, y ello siempre que las particulares condiciones socioculturales de los idiomas en contacto lo permitan.<sup>2</sup> En lo que respecta al español de América, pocos son los fenómenos —fonéticos o gramaticales— que podrían explicarse seriamente como resultado de la influencia de los sustratos prehispánicos. Pensando, concretamente, en las peculiaridades del español *normal* de México, creo que ni una sola de ellas sería consecuencia de la influencia náhuatl.<sup>3</sup>

Sin embargo, en los dominios de la entonación y del léxico, la situación parece ser muy diferente. Aquí, las huellas de las lenguas precolombinas *parecen* evidentes y numerosísimas. Sobre todo en el terreno léxico: el vocabulario del español americano se muestra esmaltado, a los ojos de casi todos los investigadores, por abundantes indigenismos.

Ahora bien: que la influencia léxica no sea, propiamente hablando, un fenómeno estricto de sustrato parece ser cosa generalmente admitida; la teoría de los préstamos explica satisfactoriamente las transferencias léxicas que se producen entre dos lenguas en contacto.<sup>4</sup>

---

<sup>1</sup> R. Lenz, “Beiträge zur Kenntnis des Amerikanospanischen”, *Zeitschrift für Romanische Philologie*, XVII (1893), 188-214. Traducción española de A. Alonso y R. Lida, “Para el conocimiento del español de América”, *Biblioteca de Dialectología Hispano-Americana*, VI (1940), 209-258; v. p. 249.

<sup>2</sup> Cf. B. Malmberg, “L’extension du castillan et le problème des substrats”, *Actes du Colloque International de Civilisations, Littératures et Langues Romanes*, Bucarest, 1959, 249-260; en especial, p. 258.

<sup>3</sup> No considero aquí, naturalmente, las hablas regionales del interior del país, algunas de las cuales muestran, sin duda, cierta influencia de la lengua de sustrato. Es bien sabido, por ejemplo, que la peculiar articulación de las oclusivas que distingue al español de Yucatán puede deberse a la influencia maya (lengua de adstrato, más que de sustrato).

<sup>4</sup> De acuerdo con el concepto de sustrato lingüístico expresado por B. Malmberg, los préstamos léxicos no representan, de ningún modo, el resultado de la acción del

Con todo, no deja de señalarse una y otra vez cuánto han contribuido las palabras de origen amerindio a colorear y a diferenciar el español hablado en cada uno de los países de América. Voluminosos diccionarios de indigenismos son prueba de ello. En el de voces chilenas recopilado por Lenz figuran unas 2.500 formas;<sup>5</sup> en el de indigenismos venezolanos de Lisandro Alvarado, unas 1,700;<sup>6</sup> en el de aztequismos publicado por Robelo, no menos de 1.500 formas de origen náhuatl, a las que habría que añadir los centenares de voces precortesianas de distinta procedencia (maya, zapoteca, etc.). Ante estas cifras, no puede extrañarnos que Darío Rubio haya escrito lo siguiente: “Si desaparecieran del lenguaje español que hablamos los mexicanos, todas las voces en dicho lenguaje incluidas y que tienen su origen en el idioma náhuatl. . . se produciría un caos verdaderamente horrible por la situación en que tal desaparición hubiera de colocarnos”.<sup>7</sup>

Con el fin de averiguar hasta qué punto es importante la contribución léxica de las lenguas indígenas al habla común de la ciudad de México, desde hace dos años hemos orientado las labores del Seminario de Dialectología de El Colegio de México hacia la investigación de esa influencia léxica. Los resultados a que hasta este momento hemos llegado no dejan de ser sorprendentes, si bien sean todavía provisionales. La impresión que esos voluminosos diccionarios de indigenismos pueden producir, se aparta totalmente de las conclusiones a que en el Seminario estamos llegando. La influencia léxica de las lenguas indígenas sobre el español hablado en la ciudad de México es —numérica y proporcionalmente al menos— muy pequeña. Y ello, sobre todo, porque el vocabulario de origen indígena tiene un campo vital sumamente reducido. La gran mayoría de las voces que se consignan en los diccionarios de indigenismos es enteramente desconocida para el hablante medio. La opinión de Morínigo a este respecto me parece por completo atinada: la erudita acumulación de palabras pre-

---

sustrato. (Cf. su artículo “Encore une fois le substrat”, *Studia Linguistica*, XVII, 1963, 40-46; en especial, pp. 41-42). — Cf. también Frederick H. Jungemann, *La teoría del sustrato y los dialectos hispano-romances y gascones*, Madrid, 1955; p. 17.

<sup>5</sup> Correspondientes a más de 1.600 voces; cf. A. Alonso, “Substratum y superstratum”, *Revista de Filología Hispánica*, III (1941), p. 216, n. 3.

<sup>6</sup> Cf. Ángel Rosenblat, *El castellano de Venezuela: la influencia indígena*, Caracas, 1958, p. 9 (Sobretiro del *Boletín Indigenista Venezolano*, vols. III-V).

<sup>7</sup> Cf. Darío Rubio, *Refranes, proverbios y dichos y dicharachos mexicanos*, 2ª ed., Méjico, 1940; t. I, pp. XXII-XXIII.

hispánicas en tales diccionarios no responde a la realidad hablada.<sup>8</sup> Volvamos, un momento, sobre el *Diccionario de aztequismos* de Cecilio A. Robelo: del millar y medio de formas que reúne, sólo unas 160 eran conocidas —y reconocidas como de uso común— por los investigadores, mexicanos, del Colegio de México; esto es, aproximadamente el 10% de los artículos consignados en el diccionario. Y eran poco más de 250, *en total*, las formas que unos u otros conocían con mayor o menor precisión. Las 1.200 voces restantes les eran enteramente desconocidas. Pero hay que tener en cuenta, además, que varios de esos 160 indigenismos mexicanos de uso general —o casi general en la ciudad— son palabras que pertenecen ya al acervo común de la lengua española (*chocolate, tomate, jícara, chicle, petate*, etc.), y que por ello no particularizan —no distinguen dialectalmente— al español mexicano.<sup>9</sup> Ciertamente que el habla urbana no es campo fértil para el arraigo de los indigenismos; suelen éstos emplearse para designar realidades de la flora particular de cada región, realidades que prácticamente desconoce el hablante urbano.<sup>10</sup> A ello se ha referido también Ángel Rosenblat, al estudiar la influencia indígena sobre el español venezolano,<sup>11</sup> advirtiendo que los indigenismos de uso común en Venezuela son sólo 17. Las 1.700 voces que recoge Lisandro Alvarado resultan, en su casi totalidad, desconocidas para el hablante medio venezolano.<sup>12</sup>

En lo que a la ciudad de México se refiere, la situación que vamos descubriendo en el Seminario de Dialectología no es esencialmente distinta,

<sup>8</sup> Cf. Marcos A. Morínigo, "La penetración de los indigenismos americanos en el español", *Presente y futuro de la lengua española*, Madrid, OFINES, 1963; t. II, p. 226.

<sup>9</sup> El afán de aumentar al máximo el número de indigenismos a que se refiere Morínigo, podría explicar el que incluya Robelo en su *Diccionario* algunos vocablos de indudable ascendencia hispánica, procurándoles caprichosas etimologías náhoas. Así registra como aztequismos, entre otras, las palabras *cochino, apachurrar* y *nana*; e inclusive dedica una larga nota a refutar el origen hispánico de *cogote* —que él deriva del náhuatl *cocotl* 'esófago'— aunque la palabra figura ya en el *Universal vocabulario* de Alonso de Palencia, publicado en Sevilla en 1490.

<sup>10</sup> Cf. Gerhard Rohlfs, "Influence des éléments autochtones sur les langues romanes", *Actes du Colloque de... Langues Romanes*, Bucarest, 1959, p. 244.

<sup>11</sup> A. Rosenblat, *El castellano de Venezuela*, p. 12.

<sup>12</sup> Naturalmente que no pretendo decir que los diccionarios de indigenismos deberían limitarse a consignar únicamente las voces de uso general. Pero sí considero que sería necesario que los autores de esos léxicos declarasen siempre cuál es el alcance —la vitalidad— de cada uno de los términos recogidos. La delimitación geográfica y social de los indigenismos no debería faltar en ningún diccionario que pretenda ser riguroso y útil para la dialectología.

aunque el número de indigenismos usuales que hemos registrado no sea tan reducido como en Venezuela. Nuestro método de trabajo ha sido el siguiente: durante dos años, los miembros del seminario —18 en total<sup>13</sup>— han realizado 332 encuestas entre hablantes de todas las clases socioculturales que forman parte del heterogéneo conglomerado humano que es la ciudad de México. Todas las entrevistas quedaban siempre grabadas en cintas magnetofónicas, aunque se procuraba que la presencia de la grabadora portátil no coartara al informante en su manera de hablar. Cada entrevista duraba un mínimo de media hora, pero muchas veces —casi la mitad— se extendían hasta completar los 60 minutos y aun algunas se acercan a las dos horas.<sup>14</sup> En ellas se han abordado multitud de temas de muy distinta naturaleza. Acabada la encuesta, el investigador escuchaba atentamente la cinta y tomaba nota de todos los indigenismos que fuesen apareciendo. Hacía después un cálculo del número total de palabras pronunciadas a lo largo de la grabación, con el fin de determinar el porcentaje correspondiente a la aparición de indigenismos. Finalmente estudiábamos las peculiaridades de las voces indígenas: sus acepciones, su vitalidad, dominio semántico general a que pertenecían, correspondencia con determinada clase sociocultural, concurrencia con alguna voz hispánica equivalente, etc.

Hemos distinguido tres grupos o clases generales de indigenismos: topónimos (y patronímicos), gentilicios, y voces comunes (sustantivos, adjetivos y verbos). A estas últimas hemos dedicado particularmente nuestra atención, ya que son las verdaderamente reveladoras. Los topónimos —y sus derivados, los gentilicios— no tienen gran interés, pues son formas que están al margen de la estructura íntima de la lengua. El valor funcional de cualquier topónimo (como *Cuernavaca*, por ejemplo) es prácticamente siempre el mismo, tan ajeno, por así decir, al sistema gramatical como cualquier otro nombre de lugar, sea de origen ibérico, árabe o eslavo.<sup>15</sup>

<sup>13</sup> Son ellos: Raúl Ávila, Flora Botton, Miguel Capistrán, Elena Carrero, Julia Corona, Luz E. Díaz, Luz Fernández, Charles Frisbie, Beatriz Garza, Carmen Garza, Carmen Guardiola, Teresa Guzmán, Ivette Jiménez, Carlos Magis, Jaime del Palacio, José Polo, Teresa Piñeros y Gloria Ruiz de Bravo.

<sup>14</sup> El tiempo total de grabaciones es de más de 238 horas. El número total de palabras grabadas en las cintas rebasa ampliamente los dos millones (2.147.000). Hablan en esas encuestas un total de 477 personas de todos los estratos sociales.

<sup>15</sup> Estructuralmente, lo mismo da decir que "Pasé las vacaciones en *Cuernavaca*", como decir que "las pasé en *Segovia*", o "en *Medina*", o "en *Varsovia*" o "en *X*", según se solía hacer durante el siglo pasado para dar una localización indeterminada a las novelas.

A ese carácter marginal —aunque étnicamente importantísimo— de los topónimos, se refirió ya Amado Alonso al distinguir entre sustrato racial y sustrato lingüístico.<sup>16</sup>

Los resultados a que hasta ahora hemos llegado —resultados quizá algo provisionales todavía, pues seguimos trabajando en el mismo terreno— son, en esencia, los siguientes: El número total de indigenismos ‘comunes’ registrados en las grabaciones asciende, excluyendo topónimos y gentilicios,<sup>17</sup> a 1.907. Estas 1.907 formas corresponden a sólo 193 artículos léxicos.<sup>18</sup> El porcentaje que estas cifras suponen en relación con el vocabulario castellano registrado en las encuestas es francamente pequeño: los 1.907 indigenismos consignados, frente a los dos millones de voces articuladas por nuestros informantes, no representan ni siquiera el 0.09% del total. Y aun incluyendo en nuestro recuento todos los topónimos, patronímicos y gentilicios americanos<sup>19</sup> a que se hace mención en las grabaciones, el porcentaje no llega sino al 0.42 %. Claro que estos porcentajes corresponden, más que a la proporción de raíces prehispánicas dentro del catálogo léxico del español mexicano, a la *incidencia* de tales indigenismos en la cadena hablada.<sup>20</sup> Pero ello nos permite sopesar, precisamente, la *vitalidad* real de dichas voces. Por ser términos que aluden,

<sup>16</sup> Cf. A. Alonso, “Substratum y superstratum”, *Revista de Filología Hispánica*, III, 1941, pp. 210-211.— Cf. también B. Malmberg, “Encore une fois le substrat”, p. 41.

<sup>17</sup> Pero sin excluir, ni siquiera, los mexicanismos comunes de uso más amplio, tal vez del español general, como *jicara*, *coyote* o *cacao*. (En todo lo que sigue, al hablar de los indigenismos mexicanos, excluyo —a no ser que advierta lo contrario— tanto los topónimos y patronímicos, cuanto los gentilicios de ellos derivados. Repito que no creo que ofrezcan verdadero interés lingüístico. Baste pensar, en efecto, que *México* y *mexicano* por sí solos aparecen en nuestras grabaciones tanto o más que el conjunto de todos los demás indigenismos registrados).

<sup>18</sup> Y, como algunos de esos artículos son palabras que derivan de una misma raíz indígena (por ejemplo *pulque* y *pulquería*, o *chile* y *enchilada*), el número de bases léxicas de los 193 artículos se reduce a 166 solamente. Además, algunas de ellas pertenecen al acervo común de la lengua española: *aguacate*, *cacahuate*, *cacao*, *coyote*, *chicle*, *chocolate*, *hule*, *jicara*, *nopal*, *petaca*, *petate* y *tomate*.

<sup>19</sup> Del total absoluto de indigenismos recogidos, los topónimos, patronímicos y gentilicios representan el 80%.

<sup>20</sup> El porcentaje de artículos o raíces léxicas de origen indígena dentro del vocabulario usual de la ciudad de México es más difícil de precisar. Habría que conocer con exactitud el número de palabras que integran el “diccionario” total del español mexicano. Sólo con el fin de proporcionar una idea general, aproximada, de ese porcentaje, calculemos que el léxico global de la ciudad comprenda unas cuarenta mil palabras; los trescientos indigenismos —como máximo, en vez de los 193 por nosotros reunidos hasta ahora— que en la ciudad puedan emplearse, no representa-

por lo general, a realidades muy particulares del medio mexicano, la frecuencia de su aparición en la frase hablada es pequeña, casi ocasional. Sólo una veintena de voces indígenas tiene relativa vitalidad en la conversación no especializada, y su aparición es, en cierta medida, frecuente.<sup>21</sup>

El estudio de los 193 artículos léxicos reunidos, así como la consideración de su vitalidad relativa, permite hacer las siguientes anotaciones:

Por lo que respecta a su origen, la inmensa mayoría de los indigenismos procede del náhuatl. Prácticamente ésta es la única lengua prehispánica de México que ha enriquecido el vocabulario usual en la capital. Los préstamos debidos a las demás lenguas indígenas son insignificantes. Del maya, lengua que sigue en importancia al náhuatl, pueden derivarse seis o siete voces (*cenote*, *henequén*, *maquech*, *papazul*, *pibil* y *salbute*). Al tarasco pertenecen *charal*, *huarache* y *uchepo*. Del otomí no hemos recogido más que *naco*. Del zapoteco, *guelaguetza*. Y *guare* suele considerarse palabra de origen cahita.

Desde el punto de vista gramatical, casi todas las voces recopiladas son sustantivos, mientras que los verbos y los adjetivos juntos no ascienden sino a 14. A las restantes categoría funcionales —adverbios, preposiciones y conjunciones— no corresponde ni un solo indigenismo. Por otra parte, su productividad es mínima, salvo alguna rara excepción, como *chile*, del que se derivan *enchilada*, *enchilarse* y *chilero*; o *pulque*, del que proceden *pulquería* y *pulquero*; o *petate*, de donde *petatearse*, *petateo*, *petateada*, *petatero* y *petatillo*. Pero la inmensa mayoría de los indigenismos no da origen a ningún derivado usual.

Los sustantivos son en su mayoría —casi la mitad del total de artículos— designaciones correspondientes a la flora y a la fauna peculiar del país.<sup>22</sup>

---

rían ni siquiera el 1% de ese vocabulario urbano total.

<sup>21</sup> Son, en orden decreciente: *chile* (163 veces), *pulque* (104), *mole* (102), *cuate* (81), *chamaco* (73), *elote* (53), *tequila* (39), *tamal* (38), *aguacate* (33), *barbacoa* (33), *milpa* (30), *guajolote* (27), *atole* (27), *metate* (27), *zapote* (25), *molcajete* (23), *escuincle* (22), *zacate* (21), *enchilada* (21) y *mezcal* (20), además de los mexicanismos de uso hispánico general *jitomate* (67) o *tomate* (18), *chocolate* (52), *cacahuate* (30), y *petate* (25).— (Para el origen de *barbacoa*, cf. Corominas, *Diccionario Crítico Etimológico de la Lengua Castellana*, s. v.)

<sup>22</sup> De un total de 91 artículos, 65 corresponden a la flora y 26 a la fauna. Algunos revelan cierta vitalidad (como *henequén*, *nopal*, *camote*, *cempasúchil*, *ocote*, *tule*, *mezquite* en lo relativo a la flora, o *guajolote*, *pípilo*, *coyote*, *chapulín*, *tuza*, *zopilote* y *cacomiscla* en la fauna), pero la incidencia de otros es mínima (*oyamel*, *huisache*, *mecuete*, o *ajolote*, *maquech* y *sacamiche*, documentados una o dos veces tan sólo en nuestras grabaciones).

De ellos, 28 son nombres de vegetales o frutas comestibles (como *tomate*, *zapote*, *aguacate*, *capulín*, *chayote*, *epazote*, *chile*, *elote*, etc.), y ocho nombres de animales también comestibles (como *guajolote*, *totol*, *charal*). Dado que 18 de los sustantivos comunes corresponden a nombres de comidas (*enchilada*, *mole*, *tamal*, *pozole*, etc.) y 6 a nombres de bebidas (*pulque*, *mezcal*, *tepache*), el número de voces relacionadas con la alimentación asciende a 60 artículos, lo cual representa casi la tercera parte de los indigenismos registrados. Y hay que tener en cuenta además que estos últimos vocablos son los que muestran mayor vitalidad en la lengua hablada: de las 25 voces que con mayor frecuencia aparecen en nuestras grabaciones (cf. nota 21), 17 se relacionan con la alimentación popular, y únicamente ocho pertenecen a los demás dominios semánticos.

Aparte de estos términos de la flora, la fauna y la alimentación, sólo otro aspecto del vocabulario mexicano aparece esmaltado por un número de indigenismos digno de tomarse en consideración: el de los enseres o utensilios domésticos. A él pertenecen nahuatlismos como *molcajete*, *comal*, *chiquihuite*, *jícara*, *metate*, *nixcómil*, *petate*, *tanate* o *tepalcate*, hasta un total de 18 términos. En cambio indigenismos de carácter afectivo hemos reunido muy pocos: *cuate*, *chamaco*, *escuíncl*e y *mitote* son los únicos que aparecen con alguna frecuencia.

Otra circunstancia que debe tenerse en cuenta para valorar la vitalidad de los indigenismos mexicanos es el hecho de que algunos de ellos alternan con otras voces sinónimas, ya sean de origen español, ya procedan de otras lenguas indígenas. Así, *cuate* está en concurrencia con *amigo*, *compañero*, *camarada*, *(her)mano* y *gemelo*; *petaca* con *maleta* o *velis* (fr. valise); *achichinar* con *quemar*; *mecate* con *reata*,  *cuerda* o *hilo*; *guajolote*, con *pavo* y, al mismo tiempo, con *totol* y *pípilo*; *chamaco* con *niño*, *muchacho*, y también con *escuíncl*e. Esto puede conducir o a la eliminación de la palabra indígena, sofocada por la castellana de uso general en otros países de lengua española e, inclusive, en otras regiones de México, o a su especialización (reducción semántica) frente a la voz hispánica; así, ante el general *mercado*, el indigenismo *tianguis* designa específicamente el mercado indígena, levantado al aire libre; y *escuíncl*e se carga de sentido peyorativo frente al común *niño* o *muchacho*.

También valdría la pena valorar con precisión el prestigio de que goza cada indigenismo, empezando por descubrir hasta qué punto se sirve de ellos la lengua escrita. A esta tarea se ha abocado durante el presente año el Seminario de Dialectología, mediante el estudio de diversas obras literarias —novelas, dramas, ensayos— y aun de revistas y periódicos de

gran difusión en todo el país. Esperamos que los resultados puedan tener algún interés.

En resumen, considero que la influencia léxica de las lenguas prehispánicas es mucho menor de lo que los diccionarios de indigenismos harían tal vez suponer. Limitada a ciertos aspectos del vocabulario (alimentación, flora y fauna), y menos productiva cada día dado el veloz retroceso de las lenguas indígenas en la actualidad,<sup>23</sup> esa influencia resulta casi insignificante frente el caudal del vocabulario hispánico, frente a los casos de constante creación léxica o de incesante traslación semántica que se producen dentro de la propia lengua española, y aun frente al torrente de anglicismos que afluye actualmente hacia el español.<sup>24</sup>

JUAN M. LOPE BLANCH.

*Universidad Nacional de México*

---

<sup>23</sup> Según los datos del último censo levantado en México (1960), la población que habla sólo alguna lengua indígena apenas si rebasa el millón. Los monolingües de habla náhuatl no llegan a trescientos mil (297.000), y los indígenas que hablan exclusivamente yaqui son sólo 545.

<sup>24</sup> El texto completo de esta comunicación se publicará en el *Anuario de Letras* de la Universidad Nacional de México, tomo V.