

EL LÉXICO LATINO PATRIMONIAL, CRITERIO PARA LA CLASIFICACIÓN DE LAS LENGUAS ROMÁNICAS

DAN MUNTEANU COLÁN

In memoriam Mioara Avram

Resumen. El léxico latino de principios del s. I d.C. fue, en líneas generales, el utilizado en la expansión de la civilización, cultura y lengua de Roma y, por tanto, es el origen de las correspondencias y semejanzas léxicas entre las lenguas romances. Si bien, cuantitativamente las voces patrimoniales latinas representan un porcentaje relativamente reducido en los léxicos románicos, cualitativamente su importancia es trascendental, al tratarse de palabras que denominan realidades de la vida cotidiana, la civilización material y los elementos esenciales de la vida espiritual, o de términos indispensables en la comunicación, que sirven para estructurar las frases. Por tanto, el léxico latino patrimonial puede ser considerado un auténtico núcleo léxico de cada lengua y, consecuentemente, un elemento fundamental en la comparación entre las distintas lenguas y modalidades romances. Partiendo de estas premisas, proponemos una posible clasificación de las lenguas románicas basada en la comparación de los *corpora* de lexías latinas patrimoniales seleccionadas en los vocabularios representativos de los idiomas romances.

Uno de los auténticos retos para los especialistas, desde que la lingüística románica existe como ciencia *in se* y *per se*, ha sido y sigue siendo la clasificación de las lenguas románicas. A lo largo de la historia, incluso en la „etapa precientífica” (Munteanu Colán 2005: 41-62), a partir del siglo XIV, grandes personalidades de la vida cultural y, más tarde, destacados especialistas propusieron posibles y variadas clasificaciones, basadas en distintos criterios; o, en todo caso, se preocuparon por el tema, tomando parte directa o indirectamente en los debates al respecto, ofreciendo sus puntos de vista a las preguntas sobre el parentesco y la clasificación de las modalidades románicas, proponiendo criterios y principios metodológicos y soluciones diversas al problema, „de acuerdo con el espíritu de la época, el estado y las tendencias de la investigación científica general en el respectivo período y/o país, así como con las opciones teóricas y metodológicas personales” (Munteanu Colán 2005: 123).

Desde las primeras propuestas de clasificación, la lingüística románica ha registrado tres grandes modalidades de clasificar y estudiar comparativamente las lenguas romances: a) el comparativismo histórico o la lingüística genética, que analiza las similitudes y las reglas de correspondencia entre las modalidades comparadas teniendo en cuenta la dimensión temporal, es decir, la génesis común;

RRL, LII, 3, p. 245–253, București, 2007

b) la lingüística areal o espacial, cuyo fundamento es la comparación de las semejanzas entre las lenguas románicas desde la dimensión espacial, geográfica, partiendo de la idea de *Sprachbund*, es decir, la existencia de grupos de lenguas formados en función de la contigüidad espacial, idea esbozada ya por Schuchardt, y, más tarde, por Spitzer, Trubetzkoy, Jakobson, Pisani y otros; y c) la tipología (sincrónica o diacrónica) basada en el isomorfismo, sin considerar las dimensiones temporal y espacial, como destacaba tan acertadamente, en 1936, Jakobson (1963: 71): „L’isomorphisme peut unir différents états d’une même langue ou deux états (simultanés ou séparés dans le temps) de deux langues différentes, qu’elles soient contiguës ou éloignées, parentes ou non”¹.

A pesar de todos los intentos y loables esfuerzos, en el día de hoy no existe todavía una clasificación siquiera unánimemente aceptada. La explicación es relativamente fácil: la aceptación o el rechazo de las clasificaciones propuestas están condicionadas por las propias circunstancias extralingüísticas que marcaron la génesis de esas clasificaciones. A lo largo del tiempo, algunas clasificaciones han sido objeto de críticas más o menos justificadas, de actualizaciones y precisiones ulteriores debidas a investigaciones más recientes, y consecuentemente, han gozado de mayor aceptabilidad. Otras han quedado registradas en la historia de la lingüística románica como mero testimonio de una orientación o idea.

Nuestra intención de proponer un nuevo criterio para la clasificación de las lenguas románicas no debe ser considerada, por tanto, un descabellado afán de protagonismo y/o un irracional deseo de sumarse a la galería de los grandes romanistas que se dedicaron a estudiar la cuestión, sino, simplemente, un intento de averiguar, convalidar o invalidar propuestas anteriores, desde una perspectiva nueva: la comparación del léxico latino heredado o elemento latino patrimonial (ELP) de una lengua románica con el de otra/s, con el fin de establecer su parentesco partiendo de los elementos léxicos comunes que caracterizaban las distintas zonas de latinidad en las postrimerías del Imperio romano.

Es una tradición en la lingüística románica el estudio comparativo del léxico. Porque si el parentesco genealógico de las lenguas romances se refleja en las semejanzas existentes entre los sistemas y las estructuras, los vocabularios y los morfemas de las lenguas románicas, es, quizás, el léxico el que mejor refleja esta semejanza, por razones obvias. Sin embargo, como destaca Sala (1988: 11), las investigaciones de tipo comparativo del léxico románico, en su mayoría, „s-au situat cu precădere în perspectivă istorică, confruntând componența etimologică a lexicului limbilor romanice. Pe de altă parte, în aceste lucrări au fost puse pe același plan cuvinte cu o situație diferită în ansamblul vocabularului limbilor

¹ Un panorama reciente de las clasificaciones de las lenguas románicas presentan Maria Iliescu, „La typologie de langues romanes. État de la question”, en Fernando Sánchez Miret (ed.), *Actas del XXIII Congreso Internacional de Lingüística y Filología Románica, Salamanca 2001*, Tübingen, Niemeyer, 2003, vol. I, p. 61–81, y Dan Munteanu Colán, *Breve historia de la lingüística románica*, Madrid, Arco/Libros, 2005, p. 123-150.

respective: alături de cuvinte cu o largă răspândire, care se caracterizează printr-o mare bogăție semantică și familii numeroase, au fost luate în discuție cuvinte păstrate numai în anumite variante diacronice, diatopice sau diastratice ale unei anumite limbi romanice (cuvinte învechite, răspândite numai dialectal sau păstrate numai în unele terminologii sau stiluri funcționale)”.

Como hemos afirmado (*Vid. supra*), nuestra propuesta se fundamenta también en la comparación de los léxicos románicos, pero desde una perspectiva diferente. En primer lugar, consideramos que la comparación se puede realizar en el plano sincrónico, partiendo de la situación actual de las lenguas románicas, punto de vista que aclararemos más adelante. En segundo lugar, se trataría de comparar sólo los *corpora* constituidos por los elementos léxicos heredados del latín por las lenguas románicas (palabras latinas patrimoniales). Y, en tercer lugar, opinamos que se debería limitar la comparación a las palabras latinas patrimoniales seleccionadas en un determinado tipo de vocabulario, elaborado mediante la aplicación de criterios de selección idénticos al conjunto del léxico de cada una de las lenguas románicas, como, por ejemplo, el llamado por Sala (1988: 11-16) „vocabulario representativo”.

Es un hecho conocido que, desde hace tiempo, los estudiosos se dieron cuenta, aunque sin poder demostrarlo siempre de manera convincente, que en el conjunto del léxico de una lengua existe cierta jerarquía, que algunas palabras son más „importantes” que otras, que se puede identificar, aunque no delimitar con precisión, una especie de núcleo que ocupa una posición peculiar con respecto a la masa del vocabulario, como acertadamente destacaba A. Ernout: „Le vocabulaire d’une langue est composite: à côté d’un fond ancien de termes généraux dont la fixité n’est pas, du reste, immuable, il comporte une grande part d’éléments spéciaux et changeants, de toute provenance, créés à mesure qu’il faut exprimer des concepts ou des objets nouveaux” (Ernout, Meillet 1967: IX).

Somos conscientes que el análisis comparativo sincrónico ideal entre los vocabularios de las lenguas románicas debería llevarse a cabo a nivel de todo el léxico de cada una de las lenguas comparadas. No obstante, consideramos que ésta sería una tarea que, si bien realizable teóricamente, en realidad resultaría casi imposible de llevar a buen puerto, por no decir que sería una empresa utópica, al menos en la vida de una persona, debido a la riqueza del material.

Como es sabido, el vocabulario de una lengua, en su conjunto, es prácticamente una multitud infinita de unidades léxicas (Jordan, Manoliu 1989, II: 65). Y esta multitud se convierte en un número inabarcable de unidades cuando se trata de más lenguas. Por esa razón nos decantamos por una comparación basada en inventarios relativamente reducidos del léxico.

Por otra parte, nos damos cuenta perfectamente que no se pueden trazar límites rigurosos entre un núcleo fundamental o „representativo” y el resto del léxico de una lengua, si tenemos presente, como destacaba acertadamente Coteanu (1958), que „vocabularul ar putea fi asemuit cu o nebuloasă, a cărei materie se

rarifică în mod constant de la centru la margine. Densitatea din ce în ce mai scăzută neavând praguri, nu se micșorează dintr-odată”, es decir, que el paso de las unidades léxicas situadas en un escalón jerárquico superior, central, a otro inferior, periférico, se produce gradualmente.

Juillard defiende la misma idea cuando afirma que la división del léxico de una lengua en un vocabulario básico y el resto del vocabulario es arbitraria desde el punto de vista estadístico. Porque el paso de unas palabras con frecuencia elevada o gran difusión a palabras con frecuencia baja o escasa difusión se hace de manera gradual. Por tanto, no se puede elaborar y delimitar correcta y objetivamente un vocabulario fundamental o básico sólo sobre criterios estadísticos (Vintilă-Rădulescu 1974). Un vocabulario „representativo” debe ser, como destaca Sala (1988: 13), distinto de un simple inventario de frecuencia, debe reflejar la realidad léxica de la lengua y, no en último término, cumplir con las condiciones necesarias para convertirse en un „escenario básico abierto a la ampliación continua del *corpus* investigado desde los mismos planteamientos teóricos” (Munteanu 1995: 413).

Los intentos de jerarquizar el léxico, de elaborar vocabularios fundamentales, principales, selectivos, etc.², con denominaciones y finalidades varias, para diversas lenguas romances, basados en distintos criterios y diferentes metodologías, son numerosos, y no es éste el lugar apropiado para enumerarlos. La finalidad de estos intentos es también diferente, desde la puramente científica hasta la didáctica y práctica. Los resultados obtenidos son dispares en cuanto al número de unidades seleccionadas y las coincidencias entre los mismos. Incluso para una misma lengua se han obtenido resultados muy distintos acerca del número de unidades seleccionadas y la composición del vocabulario, de acuerdo con la metodología utilizada.

Entre los últimos estudios publicados dedicados a la jerarquización del léxico románico, merecen ser mencionados el de Reinheimer, Tasmowski (1997) y el de Vallés (2000).

Las primeras dos autoras proponen un diccionario básico de las lenguas románicas „mayoritarias”: portugués, español, francés, italiano y rumano, como fundamento para una gramática comparada práctica de las lenguas romances.

Según Reinheimer y Tasmowski, „Comme les langues romanes sont autant de formes prises par une espèce de latin, un sujet francophone, mettons, devrait déchiffrer une autre langue romane avec plus de facilité qu’un Anglais. C’est ce que nous avons voulu illustrer en prenant appui sur les acquis de la discipline” (Reinheimer, Tasmowski 1997: 9).

En la elaboración del diccionario parten de una base francesa. Y como su finalidad es, según declaran, obtener „un vocabulaire de dimensions réduites et qui reflète avec plus ou moins de bonheur un usage moyen, ni trop littéraire ni trop

² A título de mera información, *Vid.* para el francés, G. Gougenheim, René Michéa, Paul Rivenc, Aurélien Sauvageot, *L’élaboration du français fondamental*, Paris, 1964, donde, en la primera parte, “Historique des vocabulaires simplifiés”, se presentan varios intentos de elaboración de tales vocabularios. Para el rumano existen también varios intentos de elaborar vocabularios selectivos, entre los que destacan los de Graur (1954, 1957), Coteanu (1957, 1965) y Iliescu (1977).

argotique” (Reinheimer, Tasmowski 1997: 87), eligen como punto de partida el vocabulario fundamental de Gougenheim (1958). Tras eliminar las palabras „qui sont moins employés de nos jours et les mots qui d’après notre intuition sont moins fréquents dans la langue parlée” por una parte, y añadir, por otra parte, palabras que denominan „des réalités que tout le monde connaît et dont tout le monde parle aujourd’hui”, las autoras presentan un inventario de 1800 unidades, con sus respectivas correspondencias en las otras lenguas incluidas en el diccionario.

El diccionario de Vallés tiene el español como lengua base y es el resultado de una selección llevada a cabo por el autor tras la consulta de siete diccionarios conceptualmente distintos (de frecuencias, básico, de aprendizaje, primero, visual, para extranjeros e imprescindible), todos ellos de contrastada base científica, homogeneizando gramaticalmente el vocabulario y completando las series o grupos conceptualmente equivalentes o antónimos³.

El resultado es un inventario de 1300 entradas, que el autor considera el vocabulario “fundamental” del español. El autor presenta las correspondencias de las entradas españolas en 21 modalidades románicas. Por tanto, no se puede hablar de un vocabulario representativo de las lenguas románicas, sino del español, con correspondencias en otras lenguas romances.

Para la clasificación de las lenguas románicas en base al elemento latino patrimonial consideramos que la más idónea propuesta de vocabularios jerarquizados es la del “vocabulario representativo” de Sala (1988) y sus colaboradores, por razones que exponemos a continuación.

Según sus autores, “el vocabulario representativo de las lenguas románicas” se propone establecer, sobre la base de una metodología común, el inventario de las unidades representativas de todo el vocabulario para cada lengua romance, primer intento de estas características: „Este prima cercetare care își propune să stabilească pentru fiecare limbă romanică în parte, pe baza unei metodologii comune, inventarul unităților reprezentative pentru întregul vocabular, selectate prin aplicarea unora și acelorași criterii. Aceasta este prima încercare de lingvistică romanică comparată de acest tip” (Sala 1988: 12).

Para llevar a cabo su propósito, los autores aplican tres criterios de selección – uso (U), riqueza semántica (S) y poder derivativo (D) –, adaptándolos a la situación específica de cada lengua. Así, para el catalán y el portugués se utilizan los criterios frecuencia (F), riqueza semántica (S) y poder derivativo (D), mientras para el sardo, el retorrománico y el occitano (debido a su situación particular, principalmente la ausencia de una variedad estándar, normada, y, consecuentemente, de unos diccionarios explicativos monolingües), se han aplicado sólo dos criterios de selección: riqueza semántica (S) y poder derivativo (D). Como se observa, el procedimiento fundamental es combinar los criterios lingüísticos – número de significados y derivados – utilizados normalmente en la

³ Comunicación personal del autor.

elaboración de los vocabularios fundamentales, con las listas de frecuencia. Por otra parte, los autores han renunciado a otros criterios corrientes para establecer vocabularios fundamentales, como la antigüedad – ya que se trata de una investigación estrictamente sincrónica – así como la importancia de la noción designada por la palabra.

Tras aplicar los mencionados criterios al léxico actual de cada lengua romance, el vocabulario representativo propuesto por Sala (1988) y sus colaboradores ofrece unos *corpora* de entre 2300 y 2600 unidades léxicas (salvo algunas excepciones) que responden al menos a uno de los criterios de selección aplicados en la investigación. Este núcleo permite afirmar, junto con los autores, que las palabras seleccionadas tienen un peso particular en cada lengua y pueden integrar lo que los autores han llamado „vocabulario representativo”, aunque hay excepciones que, a pesar de notarse en todos, representan un porcentaje que afecta muy poco el carácter representativo del vocabulario elaborado (Sala 1988: 15). Teniendo en cuenta todas estas características, resulta evidente, a nuestro juicio, que si bien el material que se compara no abarca todo el caudal del vocabulario de una lengua, es lo suficientemente amplio y representativo para que los resultados de una investigación comparativa resulten fiables y objetivos.

Los especialistas son unánimes en reconocer que el léxico latino de principios del primer siglo de nuestra era fue, en líneas generales, el utilizado en la expansión de la civilización, cultura y lengua de Roma y, por tanto, es el origen de las correspondencias y semejanzas léxicas entre las lenguas romances (Reinheimer, Tasmowski 1997: 59). A pesar de que en la herencia latina surgieron tendencias diferentes, desarrollos originales, en cada una de las modalidades que se fueron desgajando de la lengua madre, debidos a factores varios, entre los que destaca por su propio peso el contacto lingüístico, más allá de las modificaciones y transformaciones sufridas por el elemento latino, éste sigue siendo fácilmente reconocible y confiere una innegable unidad al conjunto lingüístico románico.

Según el REW, el número de voces latinas heredadas directamente por cada lengua romance es de aproximadamente dos mil, sin los derivados y compuestos formados en la época románica. De las 6700 palabras latinas patrimoniales registradas por Meyer-Lübke, 1300 son panrománicas, es decir un 20%, menos de 4000, eso es, un 60%, se han conservado en algunas lenguas románicas, y algo más del 20% (unas 1500 palabras), en una sola lengua románica (Jordan, Manoliu: 1989, II: 69).

Cuantitativamente las voces patrimoniales representan un porcentaje relativamente reducido en los léxicos románicos, entre el 15 y el 20% (Reinheimer, Tasmowski 1997: 59), incluidas las palabras gramaticales (artículos, pronombres, preposiciones, conjunciones). No obstante, cualitativamente su importancia es trascendental. Porque se trata, por una parte, de palabras que denominan sobre todo realidades de la vida cotidiana, la civilización material y los elementos esenciales de la vida espiritual, afectiva. Y, por otra parte, de los términos más indispensables

en la comunicación, que sirven para estructurar las frases: „[...] dans ces 15 à 20% se trouvent justement les termes les plus indispensables à la communication, soit qu'ils servent à structurer les phrases, soit qu'ils aient trait à des aspects fondamentaux de la vie de l'homme, à sa personne, à ses besoins matériels et affectifs et à son environnement domestique [...]” (Reinheimer, Tasmowski, *ibid.*).

Teniendo en cuenta estas características, podemos considerar, junto con Reinheimer, Tasmowski (1997: 60), que las voces latinas transmitidas a las lenguas románicas presentan varios de los rasgos que definen un vocabulario fundamental: a) denominan realidades relacionadas directamente al hombre y a su entorno; b) se utilizan frecuentemente; c) son productivas, con poder derivativo relativamente grande, y entran en redes de derivación; y d) desarrollan campos semánticos amplios, que incluyen numerosos usos metafóricos. Lo que nos permite afirmar que el elemento latino patrimonial seleccionado en los vocabularios representativos románicos puede ser considerado un auténtico núcleo léxico de cada lengua y, consecuentemente, es un elemento fundamental en la comparación entre las distintas lenguas y modalidades.

Por otra parte, no podemos perder de vista otro aspecto que apoya totalmente nuestro punto de vista. El léxico de origen latino heredado por las lenguas romances demuestra claramente que numerosas palabras que en el latín imperial se diferenciaban sólo desde el punto de vista sociolingüístico, semántico-funcional o, incluso, morfológico, y se empleaban, muy probablemente, como sinónimos, se fueron difundiendo en las diferentes regiones del Imperio romano, según preferencias, en su mayoría, difíciles de explicar. Se acepta, casi unánimemente, la existencia de una zona central innovadora (Italia y Galia) y de unas áreas periféricas conservadoras (Iberia y Dacia): lat. *diurnum* > it. *giorno*, fr. *jour*; lat. *dies* > pg. *dia*, esp. *día*, rum. *zi*; lat. *caballa* > it. *cavalla*, lat. *jumentum* > fr. *jument*, lat. *equa* > pg. *égua*, esp. *yegua*, rum. *iapă*. Lo que no significa que en la periferia no se produjeron innovaciones: lat. *afflare* > pg. *achar*, esp. *hallar*, rum. *afla*; lat. *tropare* > it. *trovare*, fr. *trouver*. Por no hablar de las situaciones en que cada dominio emplea un término latino diferente para denominar la misma realidad concreta o abstracta: lat. *matutinum* > it. *mattino*, *mattina*, fr. *matin*; lat. **maneana* > pg. *manhã*, esp. *mañana*; lat. **de + mane + -itia* > rum. *dimineață*.

El léxico latino patrimonial representa la constante del vocabulario románico, una especie de fondo fundamental, muy resistente, generalmente, porque las nociones que denominan las palabras latinas patrimoniales no están afectadas por los cambios socio-económicos y culturales, y tiene una gran frecuencia de uso. Al mismo tiempo, el ELP puede ofrecer indicios muy claros sobre la pertenencia de una lengua románica a un grupo u otro o sobre el parentesco de la misma con otras lenguas vecinas, porque el léxico latino participa desde el principio en la génesis y la ulterior cristalización de la futura lengua. Es decir, ofrece datos claros y objetivos sobre el área de latinidad que originó la respectiva modalidad románica a partir de los primeros momentos en que se produce la fragmentación de la unidad latina.

Todo esto demuestra, en nuestra opinión, que el estudio comparado del léxico latino patrimonial seleccionado en los vocabularios representativos románicos puede arrojar más luz sobre las relaciones de parentesco entre las lenguas romances y contribuir, al menos desde este punto de vista, a aclarar algunas de las cuestiones sobre las que todavía los especialistas no han llegado a acuerdos definitivos, como la cuestión de la lengua rumana y sus dialectos, el estatus de unidad independiente o no del retorrománico, el grupo franco-provenzal, el gascón, la posición del catalán en el conjunto de la Romania, las áreas laterales de la romanidad, y las áreas laterales de Iberorromania.

La metodología de la investigación consistiría en establecer los inventarios del ELP de las lenguas que se pretende comparar. La siguiente etapa sería la comparación propiamente dicha de los lexemas seleccionados en los inventarios, estableciéndose las coincidencias entre dos o varias modalidades románicas, desde el punto de vista del ELP (étimo común), en base a los criterios según los cuales han sido seleccionadas las unidades léxicas. A continuación, se computarían estas coincidencias. El resultado indicaría, a nuestro juicio, por ejemplo, si las palabras rumanas y españolas que tienen el mismo étimo latino justificarían considerar que las dos áreas periféricas se parecen desde el punto de vista del ELP, es decir, si, aplicando el criterio que proponemos, se verificaría la teoría de las áreas laterales. Los resultados de nuestro estudio comparativo sobre el español y el rumano, limitado al reducido número de lexemas que responden al menos en una de las lenguas a los tres criterios de selección aplicados por los autores del vocabulario representativo de las lenguas románicas (uso, riqueza semántica, poder derivativo), invalidan la hipótesis de las similitudes entre el español y el rumano como áreas laterales de la romanidad. Pero, insistimos, únicamente desde la perspectiva propuesta, es decir de la comparación del léxico latino patrimonial seleccionado en los vocabularios representativos de las dos lenguas (Munteanu 1995).

Finalmente, las coincidencias se analizarían y agruparían por categorías gramaticales y dominios onomasiológicos. Creemos que estos análisis finales contribuyen a definir el peso de las coincidencias no sólo desde el punto de vista cuantitativo, sino también cualitativo.

Otra investigación que estamos llevando a cabo es la comparación del ELP del catalán, español y occitano, para intentar aclarar el antiguo debate sobre la pertenencia del catalán al grupo galorrománico o iberorrománico. Los resultados de nuestra comparación son los siguientes: se han registrado 161 étimos comunes para el catalán y el occitano y 280 étimos comunes para el catalán y el español (el 4.08% respecto al 4.67% del total del inventario catalán). Es decir, una diferencia de 119 étimos a favor del español (casi el 10% del total del inventario catalán). Lo que pone de manifiesto inequívocamente, a nuestro parecer, que, desde el punto de vista del elemento latino heredado en el léxico del catalán, esta lengua presenta un mayor grado de parentesco con el español que con el occitano. Es decir, que, al menos desde el punto de vista del léxico latino patrimonial, el catalán forma parte indudablemente del grupo iberorrománico.

Nuestra propuesta metodológica se fundamenta en el léxico latino heredado en los vocabularios actuales de las lenguas románicas. Este componente es muy resistente, porque se trata, por lo general, de términos que denominan nociones que no están afectadas por los cambios socioeconómicos y culturales, y tiene una gran frecuencia de uso. Por esto, creemos poder afirmar sin riesgo de equivocarnos, que los resultados obtenidos al aplicar ese criterio en la clasificación de las lenguas románicas reflejaría una situación real, objetiva, ofreciendo datos irrefutables sobre la pertenencia de una modalidad lingüística romance a un determinado grupo genealógico y las relaciones de parentesco que se establecen entre las lenguas románicas.

REFERENCIAS BIBLIOGRÁFICAS

- Coteanu, I., 1957, *Cum dispare o limbă (istroromâna)*, București, Societatea de Științe Istorice și Filologice a Republicii Populare Române.
- Coteanu, I., 1965, *Premise pentru stabilirea vocabularului străromânei*, SCL, 16, 579–604; 772–788.
- Ernout, A., Meillet, A., 1967, *Dictionnaire étymologique de la langue latine. Histoire des mots*, 4^a ed., 2^a tirada, aumentada y corregida, Paris, Klincksieck.
- Gougenheim, G., 1958, *Dictionnaire fondamental*, Paris, Didier.
- Graur, Al., 1954, *Încercare asupra fondului principal lexical al limbii române*, București, Editura Academiei.
- Graur, Al., 1957, *Fondul principal al limbii române*, București, Editura Academiei.
- Iliescu, M., 1977, „Pentru un vocabular fundamental al limbii române”, *Limba română*, 26, 3, 239–263.
- Iordan, I., M. Manoliu, 1989, *Manual de lingüística románica*, revisión, reelaboración parcial y notas por Manuel Alvar, 2 vols., Madrid, Gredos (2^a reimpresión de la primera edición española, 1972, Madrid, Gredos).
- Jakobson, R., 1963, *Essais linguistiques*, traduit et préfacé par N. Ruwet, Paris, Éditions de Minuit.
- Munteanu, D., 1995, „Estudio comparativo de los vocabularios representativos del español y rumano”, REL, 25, 411–426.
- Reinheimer, S., L. Tasmowski, 1997, *Pratique des langues romanes. Espagnol, français, italien, portugais, roumain*, Paris, L'Harmattan.
- Sala, M. (coord.), 1988, *Vocabularul reprezentativ al limbilor romanice*, București, Editura Academiei.
- Vallés, M., 2000, *Vocabulario románico (Diccionario de las lenguas románicas)*, 2 vols., Barcelona, El autor-Fundació Bastir.
- Vintilă-Rădulescu, I., 1974, „Statistica și determinarea vocabularului de bază”, SCL, 25, 243–266.