

History of the spanish lexicon and the World Wide Web: some examples

Historia del léxico español y la red global: algunos ejemplos



Soledad Chavez Fajardo. Professor in the Department of Linguistics of the University of Chile. Teacher in the area of Historical Linguistics (Romance Linguistics, History of the Medieval and Modern Spanish Language, History of Latin American and Chilean Spanish, Judeo-Spanish) and of Spanish Lexicography. Her research focuses on linguistic historiography and historical lexicology.

University of Chile, Santiago de Chile, Chile
schavez@uchile.cl
ORCID: 0000-0001-8173-8979

Received: 21/07/2020 - Accepted: 24/10/2020

Abstract:

In this essay, we would like to underline the relevance of international digital media as a corpus of research that contributes to our understanding of the history of the Spanish lexicon. We present tools such as corpora and word banks, especially elaborated for the Hispanic tradition. For example, we would like to examine some research projects in Romance or Hispanic linguistics that are in the Anglo-Saxon, French or German university networks or accessible in online libraries. Particularly, in this essay, we present some cases where research in historical lexicology was complemented with data provided by the whole digital ecosystem. For this research, we will only focus on two areas of Hispanic lexicology: that of Romance etymology and that of the history of the Spanish-American lexicon.

Keywords:

Historical lexicology; etymology; Spanish language studies; Romance language studies; dictionaries.

1. Introduction

Any investigation of a lexicon, especially regarding its history and its etymology, will always require a search for data and the collation of information, whether on paper or online. Dictionaries, monographs, studies on the subject, as well as pertinent or related bibliography, in addition to the information found on the Internet, are all fundamental sources for

Recibido: 21/07/2020 - Aceptado: 24/10/2020

Resumen:

Queremos mostrar en este ensayo la relevancia de los medios digitales internacionales como corpus de investigación para hacer historia del léxico español. Queremos presentar algunas herramientas como corpus y bancos de palabras, especialmente elaborados para la tradición hispánica. Por ejemplo, queremos trabajar con proyectos de investigación en romanística o hispanística que están en la red de universidades anglosajonas, francesas o germánicas. O, además, bibliotecas en línea. Justamente, en este ensayo queremos presentar algunos casos en donde la investigación en lexicología histórica se complementó con los datos que aportó el ecosistema digital total. Para esta investigación nos centraremos en dos espacios de la lexicología hispánica, solamente: el de la etimología romance y el de la historia del léxico español americano.

Palabras clave:

Lexicología histórica; etimología; hispanística; romanística; diccionarios.

How to cite this article:

Chavez Fajardo, S. (2020). History of the Spanish lexicon and the World Wide Web: some examples. *Doxa Comunicación*, 31, pp. 403-419.
<https://doi.org/10.31921/doxacom.n31a20>

historical lexicology. However, many times these works, or the kinds of information provided by the Hispanic lexicology network of what we could call “classical” investigation, will be insufficient. The same happens with the official word banks and corpora, such as those provided by the Royal Spanish Academy [i.e., CORDE, CREA, CORPES XXI or the online database for the *Nuevo Diccionario Histórico* (‘New Historical Dictionary’)] or the National Library of Spain (and its digital newspaper library, for example). Such resources can open up access to other spaces, which can help to complement or verify what is being investigated. What we want to account for in this essay is precisely the relevance of international digital media as a research corpus that contributes to the historical analysis of the Spanish lexicon. Indeed, the global digital ecosystem constitutes a fundamental linguistic space and provides us with data sometimes not contemplated by the official literature (especially the Royal Academy of the Spanish Language). Likewise, since the filter of Hispanic normative correction is not preponderant, many times the information in the global network provides extremely relevant information which might not appear within “codified” spaces. In other cases, we have corpora, and word banks specially elaborated for the Hispanic tradition. Often, these are projects by Anglo-Saxon, French or German universities and if we are not up to date on research regarding Romance or Hispanic linguistics, we can miss out on many free-access tools. Other times, as we will see below, in addition to projects carried out by philologists and linguists from universities, there is a diverse array of projects by online libraries whose features might be outright surprising. In this essay, we aim to present some cases where research in historical lexicology was complemented by data contributed by the whole digital ecosystem. In this study, in summary, we would like to account for the relevance of the use of the Internet as a linguistic corpus, not the one provided by the whole worldwide web, but by the world’s universities for the study of the Spanish lexicon. We will only focus on two areas of the Hispanic lexicology: Romance etymology and the history of the Spanish American lexicon.

2. Methodology

As this is a work of historical lexicology, its methodology is philological, which implies a review and collation of every lexical item studied here in various bibliographic sources, whether on paper or online. We will start with a probe of the meaning of every term in question, especially regarding its etymology, its validity, and its extension. This may involve the questioning of etymology or origin as attributed in previous studies. It can also be an alert to a state of doubt regarding a certain word: a philologist or etymologist comments about an aspect of the structure of the word signaling that further study is needed. It may also be the revision (which must be constant) of the hypothetical etymon of a word, under the advances of etymology and lexicology. Likewise, it can be the critical review of the diatopic labels of a word in an official repertoire (such as the Royal Spanish Academy’s dictionary) which may relativise or corroborate the information in question. In these cases, the work of the historical lexicologist usually requires, more than anything, expertise in dictionaries and the collation of dictionaries. These are usually Spanish-language dictionaries and most of the time, depending on the synchronic cut of the term in question, one may work with dictionaries that are contemporary or typologically fitting to the term (for example: 19th-century Spanish dictionaries, etymological dictionaries, Hispano-American dictionaries, Chilean dictionaries, dictionaries from the Southern Cone, among others). Additionally, we make use of official, Spanish-language word banks, which are those provided by the Royal Spanish Academy or by the *Hemeroteca digital* (‘digital

newspaper archives'). In cases where there is no conclusive or defining information, or there is a gap that needs to be verified, sources provided by international Romance or Hispanic studies, among others, will be utilised. These resources consist of webpages offered by international universities or other projects financed by research centres, whose data are available to both researchers and the general public. Due to the *modus operandi* of historical lexicology, reviewed and cited bibliography documents are often not current. However, this does not imply that the research itself is out of date. The fact that historical lexicology largely requires exhaustive bibliographic examinations regarding the use and competence of a given word makes resorting to digital resources absolutely necessary.

2.1. *Etymology: a case of interference*

An interesting case within etymological procedures occurs when, in the conflicting hotchpotch that is popular etymology, a key element is found. It may be a similarity in the form or content (or form and content) of the words studied. This similarity produces what is known in historical semantics as phonetic associative interference (cf. Kany 1962: 203-204), i.e., an exchange of sounds and meanings generated between two words. In general, etymologists blame the speakers' lack of attention or their ignorance of word meanings (especially foreign, dialectal, archaic words, or those that may seem strange to them) for this.

2.1.1. It is what we find, in fact, in *baza*, understood as the game of cards with respect to the set of cards held by the winner of the game ("Conjunto de cartas que en ciertos juegos de naipes recoge quien gana la mano", *Diccionario de Lengua Española*, DLE); or what occurs when a player wins a hand and takes all the other players' cards ("ganar un jugador una mano y llevarse todas las cartas de los jugadores", *Diccionario de Uso del Español*, DUE); or the fact of having the card with the most value in the round ("tener la carta con más valor en la jugada", *Diccionario del Español Actual*, DEA) or the entire set of cards in a round ("conjunto de naipes que se utilizan en cada jugada", *Diccionario de Uso del Español Actual*, CLAVE). Therefore, due to semantic transition, the term variably refers to an entire well-calculated operation (DEA); or something that allows an advantage or profit (DLE, DUE); or a situation leading to either winning or losing the game (DEA).

2.1.2. In the *Diccionario de Chilenismos y de otras voces y locuciones viciosas* ('Dictionary of Chileanisms and Other Non-Standard Words and Phrases'), published between 1901 and 1918 by Manuel Antonio Román, a Catholic diocesan priest, we find the following question regarding the adverbial phrase *sentada esta baza* o *sentada la baza*: 'How come the Royal Spanish Academy's dictionary confers [a spelling with a *z*] to this term? Is not the word [*baza*] screaming to be written with a "z", as with an "s" it means 'beginning and foundation of a thing'?' The questioning attitude of Román with the RAE helped to develop, through examination, an interesting case of etymological expertise

2.1.3. Precisely, when investigating the history of the dictionarisation of *baza*, we see that more than an apparent academic error, it contains an interesting case of a possible homonymic blend. Indeed, Covarrubias (2006 [1611]) in the *base* article added, as a second meaning:

Basas, en el juego, son las cartas ganadas, las cuales van haciendo fundamento sobre la primera, de do tomaron el nombre. Cuando uno se lo habla todo, sin que otro alguno de los circunstantes pueda decir su razón, comúnmente se dice del tal,

que no dejó hacer *baça* a los demás aludiendo al juego de las *baças*. (s.v. *basa*). ('Basas, in the game, are the cards won, which gradually become the base of the first [card], thence they took the name. When someone speaks out, without anybody around being able to speak their mind, it is commonly said that he did not make *baça* to the rest, thus alluding to the game of *baças*)

Note that the indistinctness of the grapheme occurs in the article itself: it begins with *basa* and ends with *baça*, a surprising swap, due to its late appearance. Indeed, Alonso (1947) claimed that in the sixteenth century, and even earlier, there are some vacillations of the s-z type. He exemplifies this with cases found by Menéndez Pidal in *La leyenda de los infantes de Lara* ('The Legend of the Infants of Lara'), in the writing of a scribe from Toledo as well as some examples taken from Cuervo in the same rhyme in Castilian authors such as Juan Álvarez Gato from Madrid in the 15th century. On the other hand, Valdés, in his *Diálogos*, also gave an account of these exchanges (1969 [1535]: § 89), especially from some speakers of Castile. Alonso postulates that this type of confusion, in a moment when the paradigm continues to present phonemic pairs, implies exchanges between them, that is, between apicoalveolar, palatal and dental sibilants, always in the cases of isolated words (cf. 1947: 8). This is precisely the clue that we cannot ignore in this case: the possibility of finding sonority in the sibilant of this word.

2.1.4. Let's continue with the collation in dictionaries from the Covarrubias era: Rosal (1992 [1611]) lemmatised *basa* and defined: "*Basa* en el naype, y fuera de él, es asiento o peana y grada en edificio, tomado del gr. *basis*; que tal forma representan las *basas* en el juego". ("*Basa* in card games, and out of them, is a seat, or a step and plinth in a building, taken from the Greek *basis*; thus, this form is represented by the *basas* in the game").

Therefore, we find both in Covarrubias and Rosal, the *basal/baza* graphic alternation, where *baza*, a word that refers to the cards, can be written as *basa*. The authors treat both terms as derived from the same sign, *basa*, with a semantic transition: the first card is the *basa* (*base*) of the rest in the game, from there it derives, in semantic transition, the name of the set of cards.

2.1.5. The multilingual lexicographical tradition also offers interesting clues. Most of these dictionaries lemmatise the word referring to the card game in the same lexicographic article of the "basis" or "foundation", that is, in *basa*, like Covarrubias and Rosal. Thus, they continue to treat, rather than with a homonymy, with a polysemy. In other documents, we find that the word has variants: Palet (1604), in his bilingual French-Spanish dictionary, posits *basa* as well as *baça* with the same meaning. Oudin (1607), in his bilingual French-Spanish dictionary, like Palet has both *baça* and *basa* with the same meaning, while he also refers to *vaza*. Vittori (1609), in his trilingual French, Spanish and Italian dictionary, has *basa* and *vaza* too, like Oudin. Minsheu (1617), in his trilingual English, Spanish and Latin dictionary includes *basa*. Franciosini (1620), in his Spanish-Italian dictionary, out of *basa*, in another article lemmatises the two variants: *baça* and *vaça*. Mez de Braidenbach (1670), in his Spanish-German dictionary, lemmatises the two variants together: *basa* and *baça*; and Stevens (1706), in his Spanish-English dictionary has *basa*.

2.1.6. This multilingual panorama led us to think about a possible indistinction, especially supported by *basas-baças* in Covarrubias and *basa* in Rosal, something that continues with the multilingual lexicographic tradition, beginning little by little to opt for the use of the interdental sibilant *baza*, as we see in the later authors of bilingual dictionaries, such as Sobrino (1705), who in his Spanish-French dictionary only lemmatises the word in *baça*.

2.1.7. In the monolingual lexicographical tradition, in the second edition of *Autoridades* (1770) only *baza* was incorporated with the meaning of “en el juego de naipes es el número de cartas que recoge el que gana la mano” (‘in the card game, it is the number of cards picked by the winner of a hand’). However, it is relevant that variants continue to be presented, as in Domínguez’s *Suplemento* (1869), where one of the meanings of *baza* refers to *basa*. We pay attention to this because Román’s criticism focuses precisely on one of the cases where the word enters phraseology, specifically in the phrase *sentada estalla baza*, with the meaning of “establishing this principle”, a phrase that appears for the first time in the “Diccionario de la lengua castellana” edition of 1783. Román attributes a possible error to the academic dictionary and we think that, rather than an error, what we have is a phonetic associative interference.

2.1.8. In the first place, we have the word *baza*, which has a controversial etymology attributed to it until this day. It was thought that it could come from Arabic, as claimed by a long lexicographical tradition: *Autoridades* (1990 [1726]) indicates that it comes from the Arabic “to conquer, subdue, dominate”, which becomes “appropriate a thing” in the edition of 1899 and “something won in a dispute” in the edition of 1956. The etymological lexicographic tradition of the 19th century insists on the Arabic origins: Echegaray (1887) and Barcia (1880) suggest that it comes from the Arabic “[to] win, because the one who makes the trick wins”. In the etymological lexicographic tradition of the 20th century, the DCECH (1980) relativises this thesis, since the term does not appear in the most relevant Arabic-language lexicographic repertoires.

2.1.9. However, we highlight another line within the 19th century etymological lexicographic tradition, which includes an Indo-European hypothesis, especially Germanic. Such is the case of Calandrelli (1881), who cites Diez’s thesis, which suggests that the word comes from the Middle High German *bazze*: “gain, benefit, advantage”, derived from the adjective *baz*: “good, profitable, useful”. This derived from the old High German *baz* or *paz*: “useful, profitable, good, excellent”, with an Indo-European base *bhad-*. Calandrelli proposes that from there derive both *baza* and *báciga* (‘old card game’), corresponding to the Italian term *bázzica* and the Catalan *basa*. Later, Meyer Lübke (1935 [1911–1920]) objected to Diez’s thesis, especially because of the competing sibilants in the word, which were not consistent with the Romance character of the term.

2.1.10. Corominas and Pascual (at DCECH 1980) proposed that the word could have been taken from the Italian *bazza* ‘bargain, gain’, of uncertain origin, which was already in Italian since the end of the 15th century, without adding further information. In order to justify more adequately this proposal and, therefore, to be able to unravel the relationship between *baza* and *basa* in the locution that Román criticises, we have been left without source or repertoire for this research. Hence, as paper repertoires become insufficient, we need an online search to obtain what Romance linguistics can provide us at length.

2.1.11. Indeed, we have found in current word banks, as in *Tesoro della lingua italiana delle Origini* (TLIO), a project directed by Pietro Beltrami, that the word was already recorded in Italian in the beginning of the 14th century (cf. TLIO: *Anonimo Genovese*, 1311). The TLIO proposes that the etymology comes from the Provençal *bauzejar* ‘frodare’ (‘to defraud’). Here again, some online Romance language lexicographic resources were helpful: the *Dictionnaire de L’Occitan Médiéval* (DOM) confirms that *bauzejar* is precisely ‘frauder’ and that it would come from a Germanic root **bauson*,

**bausjian* (so Diez and Calandrelli were not so far off), something that we collated in the *Französisches Etymologisches Wörterbuch* (FEW), also available online, but with the meaning of ‘defraud’.

2.1.12. Returning to the Iberian space, Corominas points out that the word in question went from Italian to Catalan *basa* and Portuguese *vasa*, both pronounced with the sonorous sibilant, as in Italian. This datum will be essential for the possible confirmation of this hypothesis: if the sibilant was voiced, as well as a loan, a late exchange could have occurred. This is a phenomenon we detected in the first lexicographical tradition of the word in Spanish.

2.1.13. Anyway, the fact that the homonym *basa*, *base* is presented, in turn, as “foundation or support” and that the formal disposition of the card game would constitute a *base* (something on which Covarrubias had already commented) will produce the blend with *baza*. This, plus the fact that the word, undoubtedly a loan, had a voiced sibilant helped to vacillate its articulation and, by extension, its spelling. Thus, this sort of “error” seen by Román is nothing more than the product of its lexical context.

2.1.14. The last part of Román’s lexicographical article provides reflections that escape from the reach of this essay. In fact, these refer to aspects linked to critical sociolinguistics or glottopolitics:

Sin duda los SS. Académicos oyeron la loc. de boca de algún andaluz, y como la oyeron así la escribieron, y así salió ella, [...]. Más acertados andamos los chilenos, que hemos inventado y usamos familiarmente la fr. fig. *hacer baza* en el sentido de ganar o prosperar en cualquier asunto o negocio. Ú. m. con negación. (1901-1908) (‘Without a doubt the Academics heard the locution from the mouth of some Andalusian, and as they heard it that is how they wrote it, and that is how it came out. We Chileans are much cleverer, as we have invented, and we use the figurative expression *hacer baza* with the sense of winning or prospering in any matter or business. Used more often with negation.’)

We cannot blame the Chilean priest for his discomfort around Andalusian Spanish: he did not have access to the etymological data that can unravel the tangle regarding this etymon.

3. Hypothetical etymology and its amendment

The amendment of etymons is a necessary, constant work within the field of etymology, due to the advances of the discipline or the discovery of new texts (especially in the cases of hypothetical etymons) or new study proposals. Indeed, this is a discipline that is constantly being remade, in pursuit of perfectibility. In the case of a non-etymological dictionary such as the Royal Spanish Academy’s *Diccionario de la Lengua Española* (DLE), the information that follows the headword also changes. At this point we must be careful and fair: the DLE is not a historical or etymological dictionary, but as it is a mixed work, it often includes the etymology or origin of the word. Many times this information may be rightly questioned or may be out of date; or it may not align with what is proposed in Corominas’ Etymological Dictionary, which as of today is the only Spanish language dictionary of this kind.

3.1. There is the case of *montaña*, a term derived from *monte* (“hill”), according to the 1947 edition of the Royal Spanish Academy’s dictionary. That is the information in the comment corresponding to the word’s etymology or its derivation. However, all editions of this dictionary (hereunder referred to as DRAE) from 1956 to the present –as well as Calandrelli

(1916) and Corominas (1981)– proposed a hypothetical **montanea*, neuter plural of **montaneum*, also present in Meyer-Lübke, (1935 [1911–1920]).

It is well known that, in the field of historical lexicology, the addition of an asterisk before a term marks it as a reconstructed word, that is, its existence has not been confirmed in any text. It may be convenient to avoid any analysis beyond this point and to recognise the value of this hypothetical etymon without further ado.

3.2. However, thanks to the etymological work carried out outside the Hispanic domain, there are a few possible new hypotheses. In effect, we would like to provide here information for the non-etymological community about some instruments available online, in languages other than Spanish. In 1995, in the context of the XXI Congress of Romance Linguistics, a round table debated critically about how to continue the line of the monumental *Französisches Etymologisches Wörterbuch* (FEW), probably the most complete etymological work in the field of Romance linguistics, written by Walther von Wartburg between 1910 and 1940. The project *Dictionnaire Étymologique Roman* (DÉRom) emerged in 2008 from the reflections, projections, and challenges discussed at that round table. Its purpose was to continue previous etymological work, with the FEW as the main source of reference, while expanding its reach towards the entire Romance language domain. New advances in linguistics and the digitalisation of a significant number of textual traditions, as well as the arrival of word banks, were fundamental for this process. The project's results are constantly updated and are available online (free access) so that anyone can obtain, with variable luck, new etymological information.

3.3. As a way to corroborate, for example, whether the hypothetical form **montanea* has been updated, we decided to check the DÉRom, which includes this example of the proto-Romance term: **/mont-ani-a/*: “Ce lexème s’analyse, en synchronie protoromane, comme un dérivé en **/-ani-/*, suffixe (rare) à valeur collective” (cfr. DÉRom: s.v. **/mont-ani-a/*). Evidently, the problem is still there, as the reconstructed etymon persists, but with a new compositional element that can help understand further the structure of the word itself. This information can shed light onto the structure, distribution, and presence of vulgar Latin, which later led to this proto-Romance term as described in DÉRom.

4. *Hápax?*

There are cases in which a word has been defined only once in a lexicographical repertoire and there is no other reference to it. Can we speak of *phantom words* (Álvarez de Miranda 2007)? Or should we simply speak of a single recurrence in a dictionary? Could this be an example of a *hápax*? In general, the historiographic study of a dictionary requires an analysis of the lexical items it contains. This determines which elements are true contributions by distinguishing original work from mere transfers from previous publications. More often than expected, there are surprises, for instance, not finding a word in other lexicographic repertoires.

4.1 In the case of the Spanish term referring to the inhabitant of Betsamés, *betsamita*, we have detected that it has not appeared lemmatised within the collated body of work, except in the *Diccionario de chilenismos y de otras voces y locuciones viciosas* by the Chilean priest Manuel Antonio Román:

BETSAMITA, adj. y ú.t.c.s. com. Habitante de Betsamés, ciudad levítica de la tribu de Judá en la frontera con la de Dan, célebre por haber estado en ella el arca de la alianza y por el castigo que se llevaron sus moradores. Falta esta voz en el Dicc., en el cual figuran otras de esta misma naturaleza, tanto o menos conocida que esta (1901-1908).

4.2. From the same lexical family, we only have the case of the *Diccionario enciclopédico hispano-americano* (editors: Montaner and Simón), which includes the place name *Betsames*, but not its demonym (cf. Volume III s.v. *Betsames*).

4.3. Failing to locate *betsamita* on online dictionaries, official word banks, and other Internet sources, we have finally found it in the online dictionary of the *Watch Tower Bible and Tract Society of Pennsylvania*, part of the massive, worldwide Jehovah's Witnesses publishing machine. There, outside of Roman's *Diccionario*, we find the term *betsamita* in the following text (Spanish version): "En 1 Samuel 6:18 la Versión Torres Amat (1953) hace referencia a "la Piedra Grande llamada después Abel", y la nota al pie de la página lee: "Abel significa 'luto' o 'llanto': nombre que se cree dado a aquel lugar por causa de la gran mortandad de los betsamitas." (Watchtower Online Library).

4.4. This does not correlate, however, with the number of times that we find the adjective, either about *Josué, el betsamita* (the one with the highest number of occurrences) or about the biblical passages referring to the inhabitants of *Betsamés*. Further study of other versions of the Bible is necessary to register how this demonym is expressed.

5. An archaic word and its Andalusian origin

There are cases within the lexicographic tradition of the Royal Spanish Academy (which we call the *academic* lexicographical tradition in this essay, too) in which a word is labelled as belonging to an area in the Iberian Peninsula, although it has been detected in Latin America, at least partially or occasionally. Most of the time, it is not clear whether the word in question is archaic or not. In this context, the academic lexicographic tradition tends to treat certain words as peninsular provincialisms. It happens, above all, with provincial terms from Asturias or Andalusia, key sources of lexical transfer towards the New World, as described by the literature on American Spanish vocabulary. Due to the infrequent revision of certain lexicographical articles, this information is often not widely mentioned. Additionally, *mutatis mutandis*, the diatopic information appearing in the academic lexicographical tradition is about American Spanish, silencing provincial data from Spain. As it can be observed, this is a critical space for further work and research for historical lexicologists.

5.1. Let us take the case of a historical phrase: *hilo de acarreto*, which is defined as "thin hemp cord" and to which the DLE assigns the diatopic label of *Andalucía*.

5.2. Rivodó (1889: 174) verified the usage of this word in Latin America by claiming that it is of "common use" in Venezuela. The term is indeed widespread in Latin America and has been recorded there for a long time. In order to corroborate this, rather than CORDE or the BNE *Hemeroteca*, we rely on the *Léxico hispanoamericano*, a project by Peter Boyd-Bowman, a US Hispanist. It is important to mention this open-access website, as its contents are extremely useful for anyone who wants to study the etymology of Latin American Spanish. In 1967, Peter Boyd-Bowman began collecting data in a project called the Hispano-American Lexicon (LHA) and in 1971 he published materials for the study of 16th-century American Spanish. However, the amount of collected material was so vast that the possibility of publishing it all on paper was unfeasible. This

is why, in 1982, the *Hispanic Seminary of Medieval Studies* began publishing the LHA on microfiche. With advances in computing, in 1994, thanks to the National Endowment for the Humanities, the LHA was transformed into a database on CD-ROM. Finally, in 2015, the Graduate School of the University of Wisconsin-Madison published the online version of the LHA, which is today is freely accessible. In fact, on the LHA *hilo de acarreto* is recorded in 1512 for Puerto Rico, in 1544 for Puebla, in 1559 for Potosí, in 1683 for the Kingdom of New Granada and also in 1780 for the River Plate area, the most recent document within the corpus.

5.3 In Chile this lexical unit is known and used since colonial times, as demonstrated in the *Histórica relación* by P. Ovalle (l. I, c. IV): “Sacan también *el hilo* que llaman *de acarreto*, y otros géneros de cordeles que sirven para varios efectos” (2007 [1646]).

5.4. Within the wider scope of Spanish lexicography, the most remote mention found in a dictionary by the Royal Academy corresponds to the 1780 edition and includes the label *Andalucía*, something that we continue to verify with Alcalá Venceslada (1980 [1933]) in his *Vocabulario andaluz*. This is cited by Alvar Ezquerro (2000) in his *Tesoro de las hablas andaluzas*, who mentions Alcalá Venceslada’s *Vocabulario*. The term is also found in the Canary Islands, according to the *Diccionario Histórico del Español de Canarias* (DHECan), a project directed by Corrales and Corbella, who dates it to 1519, while citing Peter Boyd Bowman’s lexicon for reporting its presence in Latin America.

5.5. *Hilo de acarreto* has been registered in CORDE, in *Relación del origen y suceso de los Xarifes*, published in Seville in 1575 by Diego de Torres Bollo, an author from Zamora. The word only has a second occurrence in CORDE for Spain: in 1653, by Bernabé Cobo and his *Historia del Nuevo Mundo*. Later, it is only documented in Peru (1685, Jacinto Hevia Bustos, *Vejamen al doctor Antonio Correas*) and the Philippines (1754, *Historia general de Filipinas*). There are no references in CREA, so we think that the term, of Andalusian origin, quickly spread to the Canary Islands (where it remains) and to Latin America and the Philippines.

5.6. The term *hilo de acarreto* must have been used frequently in Latin America, as the *Hemeroteca digital* presents numerous cases from Mexico since 1729. Interestingly, it has ceased to be used in Spain (it does not appear in Moliner, nor DEA 1999 nor CLAVE) while it remains in the Canary Islands. Likewise, the DRAE continues to label it as an Andalusian phrase. Nevertheless, the confirmation that the phrase settled early in Latin America was given by Boyd-Bowman’s LHA corpus.

6. Polygenesis?

Sometimes no matter how often the monogenetic thesis of a word origin is confirmed, data unveiled through philological research may reveal one or two surprises, as in the case of a word considered exclusive to Latin America, described below.

6.1. We will use as an example the Chilean word *amasandería* (‘bakery’), as well as *amasandero* (‘person who kneads the flour dough to make bread and other similar foods’), which are words so characteristic of Chilean Spanish that they are usually treated as strict *chilenismos*.

6.2. However, a comparison with other lexicographical corpora has demonstrated the term’s use outside of Chile, in Colombia, as verified by Cuervo in his *Apuntaciones*. In the editions of 1867-1872, 1876 and 1885, Cuervo adds *amasandería*

to the list of words that derive from Spanish roots “y no lo son ellas mismas” (in other words, they look like Spanish words but originated outside Spain). Indeed, in those editions, Cuervo reports that in Bogotá the “*panadería, tahona*” (‘bakery’) is called *amasandería*. In the 1907 and 1914 editions, in the chapter “Voces Nuevas”, Cuervo points out that nouns ending in *-ero* (referring to someone who exercises a trade or craft) derive from nouns with *-ía*, which refer either to the craft itself, the place where it is exercised, or where its products are sold. In the case of *amasandería*, he explains, there is no primitive form ending in *-ero*, but the noun referring to the place, with the suffix *-ería*, does exist. This can be seen in other cases as well. The information presented above by Cuervo is not valid for Chile, since *amasandero* is a long-standing word there, already lemmatised in the first dictionary of *chilenismos*, by Rodríguez (1875), who describes the *amasanderías* as ‘small bakeries, generally run by women’. Uribe (1887), for Colombia, also includes *amasandería* in his dictionary without further normative indications. In contrast, the Chilean lexicographer Ortúzar (1893) labels the term as incorrect and, unlike Rodríguez, he provides only ‘*tahona*’ or ‘*atahona*’ as definitions.

6.3. In the European lexicographic tradition, Zerolo (1895) is the first who includes *amasandería*, with two meanings: one for Colombia as ‘bakery, tavern’ and another one for Chile as ‘small or women-run bakery’, citing Rodríguez (1875) as a source. In Europe, Toro y Gómez (1901), Alemany (1917) and Rodríguez-Navas (1918) accept the information provided by Rodríguez and continue using his definition.

6.4. The Dictionary Manuals by the Royal Spanish academy from 1927 to 1989 list *amasandería* as a lexical unit used in Colombia and Chile, adding a diastratic mark of “vulgarism”. Lemmatised for the first time in the DRAE (2001) only for Chile, it refers to a *panadería*, ‘location, house or place where bread is sold’.

6.5. The lexicographic tradition of dictionaries of Americanisms suggests the usage of the term in Venezuela, because Malaret (1931) and Morínigo (1966) include Venezuela, along with Chile and Colombia, when labelling their corpora . However, this interpretation is doubtful since we do not find any information that verifies such dialectal distribution. In addition, Malaret’s 1942 supplement had only left the labels for Colombia and Chile.

6.6. Also, the fact that *amasandería* is a term still used only in Chile is corroborated by the latest edition of the dictionary of Chilean Spanish by the Chilean Academy of Language (DUECh 2010) and by the only case found in CREA. This also matches with Boyd-Bowman’s description for *amasandero* (1962). We have, however, an interesting anomaly: Orellana (1891 [1871]), from Barcelona, wrote: “[*Amasandería*] se trata del departamento destinado a amasar y cocer el pan en un hospital militar. Pues a eso se le llama en todas partes panadería y no amasandería. ¡Qué cosas tienen esos madrileños!” (“It is the unit destined to knead and bake bread in a military hospital. Well, that is called everywhere *panadería* and not *amasandería*. The things these Madrileños come up with!” (s.v. *amasandería*). We consider this a clue for necessary further research, since records that are currently available both online and on paper, either Hispanic or international, have proven insufficient.

7. A historical word that does not seem historical

There are cases where, within the academic lexicographic tradition, a word is typically labelled as exclusive to Latin America, or some of its territories. Frequently, it is not specified whether the term in question is historical. Given this context, in most cases, this lexicographic tradition continues to treat some words as simply being in common usage. In turn, *mutatis mutandis*, the diatopic information that appears in the academic lexicographic tradition is largely the Spanish-American one, rather than the Iberian one. This constitutes a critical space requiring further work and investigation in the field of historical lexicology.

7.1. This occurs in the case of *belduque*, defined as ‘cuchillo grande de hoja puntiaguda’ (‘a large, pointed knife’) (DLE). Román, in his *Diccionario de chilenismos y de otras voces y locuciones viciosas*, suggests the following definition:

Belduque, m. Cierta cuchillo de hoja puntiaguda y mango de madera y de una sola pieza, que estuvo aquí muy en uso treinta años atrás. Llamábase también **cuchillo de belduque** o **cuchillo belduque**, adjetivando esta última voz. Créela Cuervo venida de España y derivada de *Balduque*, que era como pronunciaban los españoles del siglo XVI el nombre de *Bois-le-Duc*, ciudad de Holanda, célebre en las guerras de los Países Bajos y en la cual hasta hoy florecen las fábricas de cuchillos. Tanto en Colombia como en Chile se ha dicho también **balduque**, lo que se acerca más a la casi cierta etimología. Entre otros usos, sirvió este cuchillo para matar y descuartizar reses: en este caso podría reemplazarse por el castizo *jifero*. (1901-1908)

(**Belduque**, m. A certain knife with a pointed blade and a wooden handle, in one piece, which was very much in use here thirty years ago. It was also called **cuchillo de belduque** or **cuchillo belduque**, this last word adjetivised. Cuervo believes it came from Spain and derived from *Balduque*, which was how the Spaniards of the 16th century pronounced the name of *Bois-le-Duc*, a city in Holland, famous in the wars in the Netherlands and for knife factories that still exist today. In both Colombia and Chile, **balduque** has also been used, which is closer to Cuervo’s etymology. Among other uses, this knife was utilised to kill and butcher cattle: in this case it could be replaced by the traditional *jifero*. (1901-1908))

7.2. The first cases of this word are found practically all across Latin America from an early stage. Once again, the *Léxico hispanoamericano* (LHA) helped us to corroborate the first dates: 1549 for Mexico (*Protocolos de Puebla de los Ángeles*, Boyd Bowman); 1559 for Bolivia (*Historia de la villa imperial de Potosí*, Arzáns de Orsúa y Vela, Boyd Bowman); 1581 for Guatemala (*Archivo documental centroamericano*, Boyd Bowman); 1626 for Chile (*Historia del tribunal del Santo Oficio de la Inquisición en Chile*, Medina, Boyd Bowman); and 1638 for Colombia (Juan Rodríguez Freyle, CORDE), all like *cuchillo de belduque*.

7.3. The word, as reported by Román, was usual in Chile until the 19th century. This is corroborated by Gormaz (1860), who amends *berduque* to *balduque*. Cuervo, from the first edition of his *Apuntaciones*, proposes the etymon of the word (1867-1872: §732) and confirms it in subsequent editions. Likewise, Cuervo gives an account of the extension of the word in Latin America: “from Mexico to Chile” (1907: §656). We found that Rodríguez (1875) for Chile, Uribe (1887) for Colombia, and Echeverría y Reyes (1900) for Chile are the only authors who cited the term in the Hispanic American lexicographic tradition. García Icazbalceta (1899), for Mexico, mentions fact that the word has fallen out of standard usage and that “today only the common people use this word.”

7.4. The term *belduque* was incorporated in DRAE (1925) with geographic labels (Colombia, Chile, and Mexico) precisely because it had appeared in some lexicographical repertoires from these countries. However, we have found that most of such authors indicate this word as out of use or as restricted to certain spaces. For example, from Chile, Yrarrázabal (1945) proposes the elimination of the term *belduque* from the DRAE, claiming that it is not used in his country, and Morales Pettorino (1984) labels it as unusual. The maintenance of diatopic labels is irregular within more contemporary lexicographical tradition. For example, Malaret, in the third edition of his dictionary (1946), discards Colombia and leaves only Mexico and Chile.

7.5. The 1956 edition of DRAE discards the Chile label. Its 1992 edition, Mexico is also put aside but reappears in the 2001 edition. The *Dictionary of Americanisms* (DA) gives an account of the word for Mexico and Colombia but regards both as being obsolete.

7.6. One last piece of information that we want to present and that invites us to continue investigating the destination and uses of *belduque* is what we found in Fernández (1900). He reports that a semantic transition occurred in Chile towards “débil, enclenque, apocado” (‘feeble, sickly, timid’) citing the authority of the writer Daniel Barros Grez (a 19th-century author of the *costumbrista* movement, one of the founders of Chile’s dramaturgy). This transition did not prosper, as it has not reappeared in any Chilean lexicographical repertoire and we have not found it in our search. This proves that a word, in this case, historical or unusual, requires constant philological research, especially when has a rather unique diachronic typology. As for now, it is relevant to emphasize that *Léxico hispanoamericano* (LHA), an online resource, has been greatly helpful to trace this diachrony.

8. An indigenous term in use in Chile

Another instance in which the *Léxico hispanoamericano* (LHA) has been of great help in determining the diatopic distribution of a lexical item is the case of the construction *al apa* (‘piggyback’). It can exemplify the function of the LHA, as an excellent source of the lexical wealth of non-Hispanic origin that constitutes and characterises Latin American Spanish. This particular case is an example of indigenous words that have passed into the Spanish language.

8.1. In the case of the word *apa*, Lafone Quevedo, in his *Tesoro de Catamarqueñismos* (1927 [1898]), comments that it comes from *apa*, which in Quechua means *to carry* and is the expression that children say to their wet nurses: *apa*, “take me or bring me”. In Chile, Román (1901-1908) adds that it is used to mean “piggyback”, among all kinds of people, in the adverbial phrase *al apa*.

8.2. Interestingly, *al apa* is an extensively used expression of Quechua origin that continues to be in widespread use in Chile. In Argentina, other than Lafone Quevedo’s account, the word has no record whatsoever. In Chile, within the lexicographic tradition of Chilean dictionaries, it is lemmatised for the first time in Rodríguez (1875), followed by Ortúzar (1893), Echeverría y Reyes (1900), Román (1901-1908) and Medina (1928). Lenz, in his *Diccionario etimológico de las voces chilenas derivadas de las lenguas indígenas americanas* (1979 [1904-1910]), indicates that the term is used frequently in Santiago.

8.3. In the European lexicographic tradition, *al apa* was incorporated for the first time in Alemany (1917) for Chile and the DRAE followed in 1970. We have no data or certainty as to why the *Diccionario Histórico de la Academia* (1996) added the use of *al apa* in Ecuador, as well as in Chile.

8.4. All references to *al apa* that we can find in the Hispano-American Lexicon (LHA) and CREA are, with this meaning, from Chile or referring to Chile.

9. Results

With this small sample, we intended to highlight the pertinence and importance of online word banks and dictionaries, their different levels and aspects. We focused, for example, on the phrase: *sentada esta/la baza*, by a critical observation of an early 20th-century dictionary. This observation, incidentally, had not been taken into account in the Hispanic lexicographic universe. Indeed, thanks to this observation and its subsequent breakdown and analysis, we detected an interesting case of phonetic associative interference, which could only be corroborated with the identification of the etymology of the word in question. In this case, the use of online lexicographic corpora from other linguistic domains, such as medieval Italian and Occitan and French was essential for testing the hypothesis.

On the other hand, due to the constant review of the etymons that both Corominas and the DLE provide, we wanted to corroborate to what extent the word *montaña* still has a hypothetical etymon. To accomplish this, we turn to DÉRom, the most up-to-date online etymological lexicographic project for Romance languages. Although the etymon continues to be hypothetical, thanks to the advances in the study of the proto-Romance the etymon has been reformulated and a novel morphological element has been highlighted as a result.

At the same time, the investigation of a word may take us to other spaces on the Internet, which are not related to university projects for databases or online dictionaries. An example of this is the digital library of the *Watch Tower Bible and Tract Society of Pennsylvania*, part of the massive Jehovah's Witness publishing system.

In other cases, we made use of the word bank directed by the American Hispanist Peter Boyd-Bowman: the *Léxico Hispanoamericano* (LHA), which is online and freely accessible. This database is the most important for analysing and collating documents regarding Spanish in Latin America. It helped us determine whether a historical word was used, and when and where it was used, as in the case of *belduque* or *hilo de acarreto*. In turn, thanks to the LHA we were able to verify that certain terms such as *amasandería* or *al apa* are still used only in Chile to this day.

10. Conclusions

As we mentioned in the introduction, any research about lexicon, especially from a historical and dialectological perspective, will always require data search and collating information. This is verified above, where the bibliographic references provide, in themselves, independent critical support. In effect, dictionaries, as well as the information found on the Internet, are key sources for historical lexicology. We also posit that the information provided by the bibliography from official or classical sources will be fundamental, but insufficient for a full investigation. For this reason, the Royal Spanish

Academy's word banks like CORDE and CREA, and the *Hemeroteca digital* of the BNE will be fundamental (but not the only) sources for our research. For this reason, it is essential that anyone who wishes to investigate the Hispanic lexicon should be prepared to embrace other linguistic traditions or other spaces that are not solely and exclusively Hispanic. Indeed, these websites can help to complement or verify what is being researched. In short, in this essay, we attempted to underline the relevance of international digital media as a research corpus that can help to elucidate the history of the Spanish lexicon. We insist on this: the global digital ecosystem is a fundamental linguistic space, which provides data often unavailable through literature from official or academic sources, via online dictionaries of medieval Italian, Occitan, or etymological dictionaries of French or other Romance languages.

In other cases, we unveiled the existence of databases and word banks especially elaborated for the Hispanic tradition, which are not amongst the academic or official banks, and thus not amongst the most commonly used. Often, these are projects by universities from the English-speaking world, and if we are not up to date on research regarding Romance or Hispanic linguistics, we can miss out on many tools that are freely accessible on the internet.

As we saw above, in addition to the university projects by philologists and linguists, the work displayed on other, perhaps non-academic online resources, such as the *Watch Tower Bible and Tract Society of Pennsylvania*, turn out to be surprisingly helpful.

In this essay, in summary, we wanted to present some cases where research in historical lexicology was complemented with the data mined from the total digital ecosystem. Thanks to a methodological philological approach, which involved the review and investigation of various bibliographic sources, starting with dictionaries, and followed by the use of official word banks (i.e., RAE or the *Hemeroteca digital*). Other resources were essential in aiding our understanding of cases where there was no standard conclusive or defining information, either from the Hispanic tradition or from Romance linguistics. Therefore, as researchers of the lexicon and its history, we owe a debt of gratitude to the work emanating from international universities and academic projects, and also to the work coming from non-typical centres of research. Together they have made a vital body of information freely accessible to researchers and to those who are curious about the topic.

11. Bibliographic references

- Academia Chilena de la Lengua. 2010. *DUECh. Diccionario de uso del español de Chile*. Santiago: MN.
- Alcalá Venceslada, Antonio. 1980 [1933]. *Vocabulario andaluz*. Madrid: Gredos.
- Aleman, José. 1917. *Diccionario de la lengua española*. Barcelona: Ramón Sopena.
- Alonso, Amado. 1947. "Trueques de sibilantes en el antiguo español". *Nueva Revista de Filología Hispánica*. First edition, nº1, pp. 1-12.
- Alvar Ezquerro, Manuel. 2000. *Tesoro léxico de las hablas andaluzas*. Madrid: Arco Libros.
- Álvarez de Miranda, Pedro. 2007. "Palabras y acepciones fantasma en los diccionarios de la Academia". Alicante: Biblioteca Virtual Miguel de Cervantes.
- Asociación de Academias de la Lengua Española (ASALE). 2010. *DA. Diccionario de americanismos*. Madrid: Santillana.

- Barcia, Roque. 1880. *Primer diccionario general etimológico de la lengua española*. Madrid: Establecimiento tipográfico de Álvarez Hermanos. First Volume.
- Beltrami, Pietro (dir.). 1998-. *Tesoro della lingua italiana delle origini* (TLIO). Florence, CNR/Accademia della Crusca. <http://tlio.ovl.cnr.it/TLIO/>.
- Boyd-Bowman, Peter. 2003. *Léxico hispanoamericano. 1493-1993*. Edited by Ray Harris-Northall and John J. Nitti, CD-ROM edition. New York: Hispanic Seminary of Medieval Studies.
- Calandrelli, Mathias. 1916. *Diccionario Filológico-Comparado de la lengua Castellana*. Buenos Aires: Imprenta de G. Kraft. Twelfth Volume.
- Calandrelli, Mathias. 1881. *Diccionario Filológico-Comparado de la lengua Castellana*. Buenos Aires: Imprenta de Biedma. Third Volume.
- Corominas, Joan and Pascual, José Antonio. 1981. *Diccionario crítico etimológico castellano e hispánico*. DCECH. Madrid: Gredos. Vol. iv.
- Corominas, Joan and Pascual, José Antonio. 1980. *Diccionario crítico etimológico castellano e hispánico*. DCECH. Madrid: Gredos. Vol. i.
- Corrales, Cristóbal and Corbella, Dolores. 2013. *Diccionario Histórico del Español de Canarias (DHECan)*. Tenerife: Instituto de Estudios Canarios.
- Covarrubias, Sebastián de. 2006 [1611]. *Tesoro de la lengua castellana o española*. Madrid: Iberoamericana.
- Cuervo, Rufino José. 1914. *Apuntaciones críticas sobre el lenguaje bogotano*. París: Macon, Protat frères, imprimeurs. 6th Edition.
- Cuervo, Rufino José. 1907. *Apuntaciones críticas sobre el lenguaje bogotano*. París: A. & R. Roger y F. Chernoviz, editores. 5th Edition.
- Cuervo, Rufino José. 1885. *Apuntaciones críticas sobre el lenguaje bogotano*. Chartres: Imprenta de Durand.
- Cuervo, Rufino José. 1876. *Apuntaciones críticas sobre el lenguaje bogotano*. Bogotá: Imprenta de Echeverría hermanos.
- Cuervo, Rufino José. 1867-1872. *Apuntaciones críticas sobre el lenguaje bogotano*. Bogotá: Arnulfo M. Guarín.
- DÉRom=Buchi, Éva and Schweickard, Wolfgang (dirs.). 2008-. *Dictionnaire Étymologique Roman (DÉRom)*. Nancy: ATILF. <http://www.atilf.fr/DERom>.
- Domínguez, Ramón Joaquín. 1869. *Nuevo suplemento al Diccionario Nacional o gran Diccionario Clásico de la lengua española*. Madrid: Imprenta y Librería Universal de los Sres. Crespo, Martín y Comp. Editors.
- Echegaray, Eduardo de. 1887. *Diccionario general etimológico de la lengua española*. Madrid: Álvarez Hermanos, Impresores. First Volume.
- Echeverría y Reyes, Aníbal. 1900. *Voces usadas en Chile*. Santiago: Imprenta Elzeviriana.
- Fernández, Abraham. 1900. *Nuevos chilenismos o Catálogo de las voces no registradas en los diccionarios de Rodríguez y Ortúzar*. Valparaíso: Talleres de San Vicente de Paul.
- Franciosini, Lorenzo. 1620. *Vocabolario español-italiano, ahora nuevamente sacado a luz*. Roma: Juan Pablo Profilio.

- García Icazbalceta, Joaquín. 1899. *Vocabulario de mexicanismos comprobado con ejemplos y comparado con los de otros países hispano-americanos. Obra póstuma publicada por su hijo Luis García Pimentel*. México: Tipografía y litografía “La europea”.
- Gormaz, Valentín. 1860. *Correcciones lexicográficas*. Valparaíso: Imprenta del Comercio.
- Kany, Charles. 1962. *Semántica hispanoamericana*. Madrid: Aguilar.
- Lafone Quevedo, Samuel. 1927 [1898]. *Tesoro de catamarqueñismos. Nombres de lugares y apellidos indios con etimologías y eslabones aislados de la lengua cacana*. Buenos Aires: Imprenta y casa editora Coni.
- Lenz, Rodolfo. 1979 [1904-1910]. *Diccionario etimológico de las voces chilenas derivadas de las lenguas indígenas americanas*. Edition directed by Mario Ferreccio. Santiago: Universidad de Chile, Seminario de Filología Hispánica.
- Malaret, Augusto. 1946. *Diccionario de americanismos*. Buenos Aires: Emecé editores. Third Edition.
- Malaret, Augusto. 1942. *Diccionario de americanismos: suplemento*. Buenos Aires: Academia Argentina de Letras. Tomo primero, letras A-E. 2nd Edition.
- Malaret, Augusto. 1931. *Diccionario de americanismos*. San Juan: Imprenta Venezuela. 1st Edition.
- Maldonado, Concepción (dra.). 1996. CLAVE. *Diccionario de uso del español actual*. Madrid: SM.
- Medina, José Toribio. 1928. *Chilenismos. Apuntes lexicográficos*. Santiago de Chile: Imprenta universo/Sociedad Imprenta/Literaria Universo.
- Meyer-Lübke, Wilhelm. 1935 [1911–1920]. *REW. Romanisches Etymologisches Wörterbuch*. Heidelberg: Winter.
- Mez de Braidenbach, Nicolás. 1670. *Diccionario muy copioso de la lengua española y alemana hasta agora nunca visto, sacado de diferentes autores*. Viena: Juan Diego Kürner.
- Minsheu, John. 1617. *Vocabularium hispanicolatinum et anglicum copiosissimum*. Londres: Joannum Browne.
- Moliner, María. 1966-1967. *DUE. Diccionario de uso del español*. Madrid: Gredos.
- Montaner y Simón editores. 1887-1910. *Diccionario enciclopédico Hispano-Americano*. Barcelona: Montaner y Simón, editores.
- Morales Pettorino, Félix (director). 1984. *DECh Diccionario ejemplificado de chilenismos y de otros usos diferenciales del español de Chile*. Santiago: Editorial Universitaria. Volume 1.
- Orellana, Francisco. 1891 [1871]. *Cizaña del lenguaje. Vocabulario de disparates, extranjerismos, barbarismos y demás corruptelas, pedanterías y desatinos introducidos en la lengua castellana*. Barcelona: Librería de Antonio J. Bastinos.
- Ortúzar, Camilo. 1893. *Diccionario manual de locuciones viciosas y de correcciones del lenguaje. Con indicación de algunas palabras y ciertas nociones gramaticales*. San Benigno Canavese: Imprenta salesiana.
- Oudin, César. 1607. *Tesoro de las dos lenguas francesa y española. Thresor de deux langues françoise et espagnolle*. Paris: Marc Orry.
- Ovalle, Alonso de. 2007 [1646]. *Histórica Relación del Reyno de Chile*. Cervantes Virtual.
- Palet, Juan. 1604. *Diccionario muy copioso de la lengua española y francesa*. París: Matthieu Gillemot.
- Real Academia Española. 2014. *Diccionario de la lengua española (DLE)*. Madrid: Espasa.

- Real Academia Española. 2001. *Diccionario de la lengua española*. Madrid: Espasa.
- Real Academia Española. 1996. *Diccionario histórico de la lengua española*. Madrid: Imprenta Aguirre. Volume III, Parts 22 & 23.
- Real Academia Española. 1992. *Diccionario de la lengua española*. Madrid: Espasa Calpe.
- Real Academia Española. 1990 [1726]. *Diccionario de Autoridades*. Madrid: Gredos. Letras A y B.
- Real Academia Española. 1970. *Suplemento al Diccionario de la lengua española*. Madrid: Espasa-Calpe.
- Real Academia Española. 1956. *Diccionario de la lengua española*. Madrid: Espasa-Calpe.
- Real Academia Española. 1899. *Diccionario de la lengua castellana por la Real Academia Española*. Madrid: Imprenta de los señores Hernando y compañía.
- Rivodó, Baldomero. 1889. *Voces nuevas en la lengua castellana*. París: Librería española de Garnier y hermanos.
- Rodríguez, Zorobabel. 1875. *Diccionario de chilenismos*. Santiago: Imprenta del Intendente.
- Rodríguez-Navas y Carrasco, Manuel. 1918. *Diccionario general y técnico Hispanoamericano*. Madrid: Cultura Hispanoamericana.
- Román, Manuel Antonio. 1901-1908. *Diccionario de chilenismos y de otras voces y locuciones viciosas. Tomo I: A, B, C y suplemento á estas tres letras*. Santiago: Imprenta de la Revista católica.
- Rosal, Francisco del. 1992 [1611]. *Diccionario etimológico*. Madrid: Consejo Superior de investigaciones Científicas.
- Seco, Manuel, Andrés, Olimpia and Ramos, Gabino. 1999. *DEA. Diccionario del español actual*. Madrid: Aguilar.
- Selig, María (dra.). 2012-. *Dictionnaire de l'occitan médiéval (DOM)*. Institut de Philologie romane del' Université de Munich (LMU). <http://dom-en-ligne.de>.
- Sobrino, Francisco. 1705. *Diccionario nuevo de las lenguas española y francesa*. Bruselas: Francisco Froppens.
- Stevens, John. 1706. *A new spanish and english dictionary*. London: George Sawbridge.
- Toro y Gómez, Miguel de. 1901. *Nuevo diccionario enciclopédico ilustrado de la lengua castellana*. París/Madrid: Librería Armand Colin.
- Uribe, Rafael. 1887. *Diccionario abreviado de galicismos, provincialismos y correcciones de lenguaje con trescientas notas explicativas*. Medellín: Imprenta del Departamento.
- Valdés, Juan de. 1969 [1535]. *Diálogo de la lengua*. Edición de Juan M. Lope Blanch. Madrid: Clásicos Castalia.
- Vittori, Girolamo. 1609. *Tesoro de las tres lenguas francesa, italiana y española. Thresor des trois langues françoise, italliene et espagnolle*. Ginebra: Philippe Albert and Alexandre Pernet.
- Wartburg, Walther Von. 1910-1940. *Französisches Etymologisches Wörterbuch (FEW)*. Analyse et Traitement Informatique de la Langue Française (ATILF) y Université de Lorraine. <http://atilf.fr/lecteur FEW>.
- Watch Tower Bible and Tract Society of Pennsylvania. 2018. *Biblioteca en línea Watchtower*. [<https://wol.jw.org/es/wol/h/r4/lp-s>]
- Yrarrázabal, José Miguel. 1945. *Chilenismos*. Santiago de Chile: Imprenta Cultura.
- Zerolo, Elías. 1895. *Diccionario enciclopédico de la lengua castellana*. París: Garnier hermanos.

