PROPOSAL OF NEW CONCEPTS AND A NEW NOTATION IN FORMULATING ETYMOLOGICAL PROPOSITIONS AND DISCUSSIONS

Mário Eduardo VIARO[*]
Aldo Luiz BIZZOCCHI[**]

- ABSTRACT: This paper aims to present the work methodology and the main theoretical postulates that guide the preparation of *DELPo* (Etymological Dictionary of the Portuguese Language), under the responsibility of NEHiLP-USP (Center for Support of Research in Etymology and History of the Portuguese Language of the University of São Paulo), as it pertains to an innovative project both in its theoretical and empirical-operational aspects. It is intended here to present the conceptual, notational and terminological innovations, proposed by the authors, that support the preparation of *DELPo*. For this, we rely on theoretical references on neology (ALVES, 2007; BARBOSA, 1993, 1996), lexicogenesis (BIZZOCCHI, 1998), and etymology (VIARO, 2011) and propose an update on the symbology used in the formulation of etymological propositions, which can at the same time clear up ambiguities and inconsistencies of traditional notation and answer for the conceptual innovations introduced here. It is intended that both the described etymological processes and the symbols corresponding to them become standard in the research in etymology and etymological lexicography.

- KEYWORDS: Etymology. Lexicography. Etymological dictionaries. Etymological propositions. Etymological discussion. Etymological symbology. Linguistic notation.

Introduction: NEHiLP

The Center for Support of Research in Etymology and History of the Portuguese Language (NEHiLP, in its Portuguese acronym), linked to the Dean of Research at the University of São Paulo, aims to produce and disseminate scientific research on Historical Linguistics, Philology and Etymology. To this end, it brings together experts in various fields who are directly dedicated to such studies. Many researchers in this interdisciplinary center are experts in ancient and modern history, language structures

---

[*] USP - University of São Paulo - Department of Classical and Vernacular Letters, School of Philosophy, Letters and Human Sciences. São Paulo – SP – Brazil. 05508-900 – maeviaro@usp.br

[**] Post-doctoral researcher. USP - University of São Paulo - Department of Classical and Vernacular Letters, School of Philosophy, Letters and Human Sciences. São Paulo – SP – Brazil. 05508-900 – aldo@aldobizzocchi.com.br

and theories of linguistic reconstruction. The employed method is the research on past and current documents, with a view to the organization of linguistic information that generates quality data for the query of both experts in linguistics and other sectors of society who are interested in etymology (especially journalists and scientists from many areas). In addition to cataloging the first dating of words and their meanings, NEHiLP provides historical information, classified according to their sociolinguistic and stylistic characteristics, associated with information on frequency of use.

Among its projects are the *Antedating* project and *DELPo*.

The *Antedating* project

One of the etymologists' tasks is to collect contexts in certain works and associate them with the date of publication of the work, which must necessarily be a reliable edition. Through this dating, it is possible to establish the *terminus a quo* of the word and establish etymologies. Traditionally, the etymologist is thought of as a highly learned person who knows where to find the best sources and often knows details of the content of the works they consult/quote. Obviously, being learned and organized are always desirable qualities for the researcher, but they are not the main point. Ideally, fewer mistakes should be made. Errors that depend only on human propensities will always exist, but there are some kinds of errors that can be avoided, such as those motivated by fatigue and a lack of attention given the handling of a large amount of information. At this point, it is believed that a program that makes the automatic comparison of a given *terminus a quo* with the date of the analyzed work would aid a significant amount of research.

There would be several advantages of this program in a manual search:

(a) in a manual search, researchers do not have available in their erudition all *termini a quo* of all the words of a language, so that their findings are guided, in most cases, only by the "sensation" that the word "should not be in that text", since they assume (based solely on their experience as speakers and/or researchers) that use of this word is more recent than the date of the investigated text;

(b) to remedy the impossibility of an exhaustive investigation, investigators eventually specialize, either in the form (words with a particular suffix, for example), or the meaning (words of a certain semantic field), or a presumed origin (e.g., words of African origin). Thus, the text would need to be reviewed by an indefinite number of researchers so that all the information that may be of interest to etymological studies is extracted.

A program that ideally had a list of all the words of the Portuguese language with their respective *termini a quo* could theoretically provide the researcher with all instances in which the date of the rendered text is before the *terminus a quo*, without searching their erudition, and would solve the problems of item (a) above, at the same time that it made a full scan, which would cancel the unwanted bias of (b).

For that purpose, NEHiLP designed the computer software *Moedor* ("Grinder"), whose initial project was submitted to the Dean of Research at the University of São Paulo in 2012, along with the creation of the Center, and which was developed by NEHiLP in collaboration with the Institute of Mathematics and Statistics of the University of São Paulo between 2013 and 2014.

The *Dicionário Etimológico da Língua Portuguesa (DELPo)* project

The *Antedating* project aims at something broader than data collection. Unlike what happens with most of the European languages (English, French, Spanish, and Italian), the etymological information present in dictionaries of the Portuguese language is very flawed with regard to their etymological data. In them:

- suffixal/prefixal derivation is confused with etymology;
- etymon of the word and its remote origin are confused;
- there is not enough care taken with etyma of unwritten languages;
- much of the Arab influence is unknown;
- there is complete arbitrariness in relation to etyma of native American and African origin;
- there are plenty of fancy etyma that mischaracterize the etymological study as a scientific work.

One should add also that:

- a few etymological dictionaries have reliable dates for their *termini a quo*;
- there is not, to this day, a methodology for the work of the *terminus ad quem*;
- the Portuguese language is far from having etymological hypotheses and *termini a quo* for meanings, because what exists are mostly dates of lemmas.

All researchers devoted to historical aspects and the diachrony of Portuguese feel the lack of specialized material comparable to the *Oxford* etymological dictionary for the English language, *Le Robert* for French, Cortellazzo & Zolli for Italian or Corominas for Spanish. The most complete work we have in Portuguese are the various publications by Antônio Geraldo da Cunha, the dictionary by Houaiss & Villar and the dictionary by José Pedro Machado. Studies of the 17th, 18th and 20th centuries are absent or very flawed.

Therefore, etymological propositions are poorly developed in Portuguese, as the most elaborated and integrated ones in a Romance vision date back only to the late 19th century to the mid-twenties of the 20th century. Since then, there is an abundance of *ad hoc* etymologies and unknown etyma.

As the final product of NEHiLP research, *DELPo* (*Etymological Dictionary of the Portuguese Language*, in its Portuguese acronym) will be developed initially online, but with possible print versions. In addition to the desirable full-*DELPo*, it is thought that partial publications can also be produced for specific areas of expertise.

For the preparation of *DELPo* there are specialized groups of researchers in:

(a) Latin-based words or those associated with old substrates (old Hellenisms, Celtic, Iberian words, etc.) and the Germanic superstrate;
(b) words of Amerindian origin;
(c) words of African origin;
(d) words of Arabic origin;
(e) words of Indian, Chinese and Japanese origin;
(f) words of unknown origin;
(g) European international words that emerged from the 17th century (from various sources, mainly Italian, Spanish, French, and English).

Conceptual and notational innovations introduced by *DELPo*

In addition to the innovative aspect of the methodology used for the constitution of the corpus on which the dictionary will be developed, conceptual innovations have also been introduced with regard to the description of evolutionary phenomena that affect words throughout their history, with repercussions in the formulation of the so-called etymological propositions, which entailed the parallel development of a new descriptive notation of these phenomena.

Many of the concepts presented here are currently used in etymology; others had already been previously proposed (BIZZOCCHI, 1998, 2013; VIARO, 2011), while some others are being shown for the first time. At the time, these studies allied etymological research with quantitative lexicometry techniques (computerized statistical lexicon analysis), and consisted of a stage of literature review and reworking of the available theories and etymological knowledge and a stage of application of the new theoretical model thus developed to a corpus of academic and journalistic texts in six European languages, namely Portuguese, Spanish, French, Italian, English, and German.

The engine of this research was the realization that, when making translations of the same text to several European languages (for example, an instruction manual drawn up in several languages), it is noted that, where French and English use the Latin word *instruction* (from Latin *instructionem*), in which only the ending has been adapted, Portuguese replaces the suffix *-tionem* with the vernacular *-ção* (*instrução*) and suppresses c̲ from the stem, whereas Italian eliminates n̲ and c̲ and adapts the spelling (*istruzione*), and German creates a calque from Latin (*Anweisung*). That is, given the need to borrow a Latin word, each language incorporates it according to a different process.

With respect to Portuguese, it was found that the available etymological dictionaries do not make a clear distinction between lexical items inherited from Latin in prehistoric times (hereditary lexicon), those created from vernacular material in historical epochs (intralinguistic creations) and those resulting from loan, most of the time merely pointing out the etymon and presenting accreditations. Several of them classify learned words, especially technical terms (*termômetro*, *helicóptero*, etc.) as a result of composition between Greek and/or Latin elements as if this process had occurred simultaneously in Portuguese and other languages, that is, as if such terms had not been created in a particular language (almost never Portuguese) and only then lent to others.

In addition, many works of Etymology, Philology and Historical Linguistics divide the lexicon into learned and popular words, disregarding the existence of half-learned or hybrid words, as well as those of problematic classification; other works recognize half-learned words, but classify as learned words some that, strictly speaking, should be taken as half-learned; still others treat calques of foreign terms as vernacular words, ignoring also that a calque can be total or partial, syntagmatic or semantic. Finally, some authors, such as Bechara (1998), confuse popular and learned words with hereditary and loan ones, as if all the popular words were inherited (disregarding, therefore, modern compositions and derivations of these words) and all loans were learned, that is, from Greek or Latin provenance without any metaplasm (disregarding loans such as *futebol* [from English *football*], which is not learned at all, or half-learned words such as *artigo* and *cabido*).

Another aspect totally ignored by the literature in question is that, when creating and renewing their lexicon, languages make choices between the various possibilities that are given – intralinguistic creation, loan, calque, etc. – and that these choices, being made systematically, even becoming injunctions in certain cases, reveal something of the speakers' worldview regarding their own language. Moreover, as these choices change over time, the historical evolution of this worldview can be reconstructed by the synchronic analysis of the etymological spectrum of the lexicon, either at present or in past synchronies.

The DELPo project also incorporates and extends the notational innovations (VIARO, 2011) largely to account for the conceptual enrichment itself brought about by the project, as well as to eliminate or at least minimize inconsistencies that traditional etymological notation entails.

As in every lexicographical work, *DELPo* will be made from the filling in of records, fully computerized and online, in which the etymological propositions and etymological discussion are fields to be filled in by the researchers. Notations and concepts presented herein are defined as the standard to be followed by these researchers and are contained in a proper manual, available at the Center's website (www.nehilp.org).

The formulation of etymological propositions according to the methodology adopted in *DELPo*

As a first notational innovation, *DELPo* distinguishes between the symbols * (reconstructed datum) and ★ (non-existent datum), prefixed to the linguistic form in analysis. The latter symbol is used, according to Viaro (2011), to indicate inexistent or impossible forms instead of the Chomskyan asterisk. But the symbol * is strictly reserved in its older Schleicherian interpretation, as "reconstructed form" (thus supposedly existent, at least in theory). The reason is that the contradictory use of the asterisk creates aporias in the diachronic study.

An etymology can be succinctly expressed by one or more *etymological propositions*, which are statements about the transformations of data from the same synchrony or two contiguous synchronies.

In these transformations, we have the *original datum* (x) chronologically earlier than the *derivative datum* (x') separated by a symbol between them.

Both can be *attested data* (x), *inexistent* data (★x) or *reconstructed data* (*x). Attested data are found in works, non-existent and reconstructed data are not in any work; however, reconstructed data follow a predictable sequence of diachronic changes, while inexistent data are mere conjectural examples of the application of analogical regularities that have no frequency of use in a particular synchrony and a particular linguistic system.

We can also speak of *disused data* (†x), when, in a particular synchrony, the frequency of use is low or zero, compared to previous synchronies. The concept of disuse is connected to the study of the *terminus ad quem*.

Data, therefore, belong to linguistic *systems*, which, in turn, are linked to *languages* (defined also politically). So, before the datum, the name of the language that it comes from should appear.

There are eight symbols used in the description of synchronies:

x → y or y ← x	the linguistic datum *y* is a morphological derivative of *x*
x ⇨ y or y ⇦ x	the linguistic datum *y* is affected by analogy with *x*
x ▶ y or y ◀ x	the loanword *x* becomes *y*
x ≡ y	*x* and *y* are homophones
x ≅ y	*x* is cognate to *y*
x ≈ y	*x* is a variant of *y*
x ~ y	*x* and *y* are inflections of the same paradigm
x + y → z	*z* is a composition of *x* and *y*

There are four symbols linked to diachronic description:

x > y or y < x	*x* is etymon/source of *y*
x >> y or y << x	the meaning *x* has become *y*
x ≥ y or y ≤ x	*y* is a calque of *x*
x ⇒ y or y ⇐ x	*x* has been replaced by *y*

The elements *x* or *y* will always have the "language *datum*" format, unless the language is Portuguese (in this case we would have just "*datum*"), except if it is not a word used throughout the Portuguese language; in this case it is necessary to specify: "European Portuguese", "Brazilian Portuguese", "Angolan Portuguese", "Portuguese of Ceará". However, the expression "Old Portuguese" (or "medieval Portuguese", "Renaissance Portuguese", etc.) is not used since each synchrony must be sufficiently clear by means of a superscript number prepended to the datum.

An etymological proposition has *authorship* (which is presented by means of a bibliographic indication) and also *degrees of certainty*.[1] The degree of certainty is defined by means of a number. To assign this number, it is necessary to confirm the regular application of phonetic laws in each synchrony. To check the regularity of the incidence of phonetic laws, NEHiLP has developed the *Metaplasmador* ("Metaplasmer") program, available for public use on its website (tab "Software/ Metaplasmer"). The degrees of certainty are:

- **[1] Certain:**
  when there is a regular application of phonetic laws on a not reconstructed datum, in addition to conservation of meaning.

- **[2] Probable:**
  when there is a regular application of phonetic laws on a reconstructed datum, and conservation of meaning;
  when there is irregularity in phonetic laws in only one *locus* of the etymology in a not reconstructed datum, and conservation of meaning.

- **[3] Possible:**
  when there is a regular application of phonetic laws on a reconstructed datum, but no conservation of meaning;
  when there is a regular application of phonetic laws on a reconstructed datum, but no conservation of meaning;
  when there is irregularity in phonetic laws in only one *locus* of the etymology in a reconstructed datum, but conservation of meaning.

- **[4] Improbable:**
  when there is irregularity in phonetic laws in only one *locus* of the etymology in a reconstructed datum or not, as well as no conservation of meaning;
  when there are irregularities in phonetic laws in more than one *locus* of the etymology in a reconstructed datum or not, with conservation of meaning.

- **[5] Impossible:**
  when there is irregularity in phonetic laws in more than one *locus* of the etymology in a reconstructed datum or not, and furthermore, there is no conservation of meaning.

---

[1] According to Jespersen (1954, p. 307, note 1): *It is of course, impossible to say how great a proportion of the etymologies given in dictionaries should strictly be classed under each of the following heads: (1) certain, (2) probable, (3) possible, (4) improbable, (5) impossible – but I am afraid the first two classes would be the least numerous.*

Briefly (with i = number of irregular *loci*):

Table 1 – Table 1 - Attribution of degrees of certainty
from indexes of the proposed etymology

Regularity of phonetic laws ($i \leq 1$)	yes		yes		yes		yes		no	no	no	no
Reconstruction	yes		yes		no		no		yes	yes	no	no
Conservation of meaning	yes		no		yes		no		yes	no	yes	no
Number of irregular *loci*	0	1	0	1	0	1	0	1	> 1	> 1	> 1	> 1
Degrees of certainty	[2]	[3]	[3]	[4]	[1]	[2]	[3]	[4]	[4]	[5]	[4]	[5]

Source: Author's elaboration.

This protocol is valid for inherited words. In cases of analogy, loans, replacements, and calques, other criteria should be adopted. The degrees of certainty could be increased by + when there has been an analogical performance. For example, if an etymological proposition of degree 4 does not follow the phonetic laws because it requires an explanation of analogical nature, its notation would be [4+]. Loanwords evidently undergo adaptation to the phonological system of the target language in the synchrony they occur; however, this system is not always known, because of the lack of sufficient data for its reconstruction and the possible nature of the loan. Issues such as the wealth of documentation necessarily influence the judgment. The same can be said of substitutions and calques.

The etymological discussion

The etymological discussion concentrates on arguing each etymological proposition, the variation of the form (diatopic, diaphasic, diastratic, diachronic) and meaning of the word and interlinguistic cognates (and possibly on their formal or semantic differences). Comments based on the topicality of a word or its rarity, its degree of specialization, its use or disuse are considered relevant.

Syntactical issues involving morphosyntactic aspects of agreement (such as gender, number or specific governments) and the participation of the word in lexemes are also relevant.

In this entry field, fully developed in a primary sense, one talks about the different senses that are subordinate to it.

For example, suppose that $estrela_1$, "celestial body", is an inherited word, but $estrela_2$, "famous actress" is calque from English *star* and not a direct derivation.

The solution to this is the main meaning, $estrela_1$, in whose *Etymology* field is the whole history of the word *estrela*.

Within this field, are numbered, first, the proposition "Latin *stellam* > *estrela* ⇦ Latin *astrum*" and all the etymological discussion plus its cognates and, next, the proposition "English *star* ≥ *estrela*", with its own comments and cognates. Thus, the comment of "English star ≥ *estrela*" will not be on the form "$estrela_2$", where there is only one reference to the main meaning, "$estrela_1$". Once the *Etymology* field is correctly filled, this discussion will reappear in the lemma.

In traditional discussions of Etymology, the words are classified as *learned* and *vulgar* or *popular*. These categories relate to the origin of the word and not to its usage: there are inherited words, such as *escorreito*, which only occur in ultraformal register, as well as learned words, such as *operário*, which are widely used in all registers, from ultraformal to informal and popular. It is important to distinguish, therefore, when studying a meaning, the learned or popular origin of its usage.

However, there are important gradations between these two extremes. We can say that there are *half-learned words* when there is hybridity of learned and popular elements, either by combination or vulgarization of a Greco-Latin form. There are also *half-popular words*, i.e. words that exist in the language from the very beginning, arising in the period called "Christian Latin", often linked to the Church, disseminated in popular speech, but with a low frequency of use, as well as restricted to certain genres in which the speech used to be more monitored. Only a detailed investigation of past synchronies (made with research and the help of tools like *Metaplasmer*) can tell exactly the times when Classical Latin acted with more or less force to play a role in the etymology of a word.

Irregular metaplasms, as in the case of *escola*, *cabido*, *cônego*, and *missa* point to the so-called *half-popular* words.

Finally, according to this distinction between learned and popular words, phonological neologisms (*poperô*, *tilim-tilim*),[2] derivations from acronyms (*ufologia*, *petista*, *aidético*) or from proper nouns (*amperímetro*, *kantiano*), and recompositions (*reprografia*, *informática*, *metrô*, *minissaia*, *showmício*) become unclassifiable.

Considering also that a word may originate in a language without any contribution, directly or indirectly, from another, or may have been formed from external morphic elements to the system or even with indigenous elements but according to a foreign structural model, two basic processes of lexical formation are recognized:

- autogenetic (heredity or intralinguistic creation);
- allogenetic (loan or creation from allogenic elements).

---

[2] Many of these formations are considered onomatopoeic (imitation of the sound of the meaning: *plim-plim*, *miau*, *zunzum*, *ziguezague*, *tique-taque*, *tilim*, and derivatives *miar*, *zunzunar*, *ziguezaguear*, *tilintar*). There is much discussion whether these words are actually *ex nihilo* neologisms (creation of a new signifier from nothing: *poperô*, *chinfrim*, *pirlimpimpim*, *zureta*), but there is still much to study about phonological expressiveness (VIARO; FERREIRA, GUIMARÃES-FILHO, 2013, p. 58-105).

It can be said that autogenetic words are those that do not contain any element from another linguistic system, either in the expression or on the content plane, and that allogenetic words are those containing at least one of these elements, be it morphological or semantic.

Therefore, the following terminology for the meanings will be part of the "Etymological Discussion" of the *Etymology* field (BIZZOCCHI, 1998, 2009, 2013):

1. Autogenetic processes (*autogeny*):
   a. *heritage* (vernacular or hereditary words);
   b. *phonological neology* (phonological neologism);
   c. *resemantization* (of autogenetic words);
   d. *composition* or *derivation* (from autogenetic words).

2. Allogenetic processes (*allogeny*):
   a. *loan* of a foreign word;
   b. *resemantization* (of an autogenetic word with borrowed meaning);
   c. *resemantization* (of an allogenetic word with vernacular meaning);
   d. *composition* or *derivation* (from allogenetic words).

The learned and half-learned words fall within the allogenetic category, as they result from elements borrowed from Greek and Latin. In the aforementioned example, *estrela*$_1$ is an autogenetic word, while *estrela*$_2$ is allogenetic.

One must also consider that the border between the two families of processes is blurred, as vernacular words can receive foreign meanings and be restored, that is, have their inherited signifier replaced by the Latin signifier that gave rise to it. For example:

mosteiro $\Rightarrow$ *monastério* ◄ Latin *monastērĭum* ◄ Greek μοναστήριον.

In this respect, the status of *eponyms*, proper nouns that become common, is arguable. Many of them pay homage to the creator of the object named by them. For example, *ampère*, *macadame*, *gilete* and *zepelim* are eponyms of surnames *Ampère*, *Gillette*, *MacAdam* and *Zeppelin*, respectively. In the case of *gilete*, the common noun has not derived directly from the name of the inventor of the razor blade, but from the product brand, originally manufactured by the inventor himself. Another case of a trademark that has become a common noun is *maisena* (< Maizena®). Among the compounds and derivatives of this kind we have: Portuguese *abreugrafia*,, French *voltmètre*, English *Newtonian*, German *Kantismus*, and, in general, all derivatives of names of people and countries, such as *hitlerista* and *zimbabuês*.

Abbreviations and acronyms may sometimes be lexicalized and come to have a syllabic pronunciation, while others remain spelled (BARBOSA, 1993, 1996). In any case, elements formed by *acronymy* as *USP*, *AIDS*, *PT*, and its derivatives *uspiano*, *aidético*, *anti-AIDS*, *petista* also form special groups to be studied.

The etymological processes involving the cases described above are detailed below.

Heritage

One classifies as *vernacular* or *hereditary*, words inherited directly from the linguistic system in an immediately preceding synchrony, without the intervention of any other system. In the case of Romance languages such as Portuguese, all that existed in Vulgar Latin and continued to exist is vernacular, so that, when they acquired the status of distinct languages from Latin, this collection of lexical and grammatical materials has become the patrimonial lexicon of these languages. The representation of an inheritance is $x > y$.[3] Latin noun forms, in this case, will come in the accusative (without apocope of -*m*).

It should be emphasized that, due to the exchange of words between Latin and Germanic during the last centuries of the Roman Empire and the early Middle Ages (1st to 7th centuries AD), there are also vernacular words in the Romance languages that are of Germanic origin (for example, Portuguese *guardar* < Vulgar Latin **guardare* ◄ Frankish *wardan*) and vernacular words in Germanic languages that are of Romance origin (e.g. English *dish* < Common Germanic **diskaz* ◄ Latin *discus*). Since these exchanges have taken place between Latin and a Germanic language, these words, depending on the synchrony, are either loanwords or heritage.

Unlike the traditional notation, there can be notations as $x > y < z$, meaning that the element y is a form with double origin. For example:

Latin *pro* > *por* < *por* ~ †*per* < Latin *per*.

Loan

The term *loan* is applied to all that is not created in synchrony nor is inherited, that is, a word from another system in a given synchrony. The term *foreign word* is often used as a synonym for loan, but it takes making some distinctions because there are the words considered "nationalized loanwords" and those "not nationalized". The term *foreign word* is therefore not a technical term adopted by NEHiLP in the section of "Etymological Discussion" of the *Etymology* field. One traditionally makes an exception to the words from Classical Greek and Latin. Save a few cases (*status*, *habitat*), Greek and Latin words enter already fully "nationalized", whereas some loanwords are either nationalized in spelling and pronunciation (*futebol*, *abacaxi*) or not (*pizza*, *bonbonnière*). Unlike traditional use, we do not make distinction between these two groups of words (i.e. those coming from Classical Greek and Latin and those coming

---

3   The symbol > is restricted to changes in signifier; therefore, it will be used by NEHiLP for changes not only of phonetic or phonological nature, but also graphic.

from other languages) and in both cases, representation is made as *x* ► *y* (the Latin noun forms, in this case, will come in the nominative). This applies to both direct and indirect learned words (when a Latin or Greek word, for example, enters Portuguese not directly from Latin or Greek, but through an intermediate modern language such as French, Italian, Spanish, or English).

Examples of etymological propositions:

Latin *temperātūra* ► French *température* ► *temperatura*;
Greek θέατρον ► Latin *theatrum* ► French *théâtre* ► *teatro*.

Sometimes a word from system A is borrowed by system B and, some time later, returns to A. Often this return occurs when the word had already fallen into disuse in A or had changed formally and/or semantically. Other times, the newly coming form coexists as an allotrope of another, older one. This particular case is called *retroversion*. For example, the Portuguese word *fetiche* is a loan from French *fétiche*, in turn borrowed from Portuguese *feitiço* and resemantized. Symbolically:

fetiche ◄ French *fétiche* ("spell" >> "fetish") ◄ *feitiço*.

In this case, the allotropic forms are represented as cognates: *feitiço* ≅ *fetiche*. Retroversion of learned forms can also occur, as seen in the French †*parformer* ► English *perform* ► French *performer*. Lat. *humor*, "liquid", passed in this sense to medieval Fr. *humour* (current *humeur*) by translation of the suffix. The latter was lent to E. *humour*, which later acquired the sense of "humor, fun". This sense returned to French as *humour*, which now coexists with *humeur* (divergent forms). Similarly, Fr. *entrevue* went to E. *interview* by restitution[4] of the prefix and later returned to French as *interview*. Today, both forms coexist in French with different meanings: *entrevue* is an interview to discuss business (a job, for example) and *interview* is the journalistic interview.

Some learned words, being indirectly introduced, enter the language with the phonetic form of the language from which they were borrowed. This produces a divergence between the form taken by that learned word and that which would be expected if it came directly from Greek or Latin.

For example:

Greek φρένησις ► Latin *phrenēsis* ► French *frénésie* ► *frenesi*;
Greek Σειρήν ► Latin *Sīrēn* ► French *sirène* ► *sirene*;
Latin *domĭnō* ► French *domino* ► *dominó*.

---

4   For the definition of restitution, see item *Calque*, further on.

When a language borrows a half-learned word, the word usually keeps its half-learned nature. For example, Fr. *nécessaire* comes from Lat. *necessarĭus* (with adaptation of the ending), which went to Port. *nécessaire* in the sense of "female bag or case for toilet utensils". Similarly, Lat. *socĭĕtās* generated Fr. *société* and, next, this form passed to E. *society*.

Typically, learned words enter the vernaculars through writing and then come into speech. Therefore, borrowing languages usually retain the Greco-Latin spelling (except compulsory orthographic adaptations), even to the detriment of pronunciation. However, sometimes there may be oral transmission of an indirect learned word, that is, from a language other than Greek or Latin. In this case, the pronunciation of this language can be reproduced with adaptation of spelling. For example:

French *estime* ► English *esteem* [ɪs'ti:m] (and not ★*estime* [ɪs'tajm]).

As with popular words, learned words can also be borrowed with their endings, which incorporate onto the word stem in the new language. The French learned words *privé* and *habitué* have passed without change to Port. *privé* and *habitué*, respectively, thus maintaining the participle ending *-é* [e] and not ★[ɛ]. Being a French element rather than Latin, it gives the Portuguese terms a hybrid, i.e. half-learned character.

Calque

One traditionally calls *calque*, *loanword translation* or *clipping* the translation of a word (or the elements that compose it) from a system other than that studied by the equivalents of the system in question. Its representation is $x \geq y$.

Examples:

English *skyscraper* ≥ *arranha-céu*;
English *hot-dog* ≥ *cachorro-quente*;
French *chou-fleur* ≥ *couve-flor*.

NEHiLP expands the use of the concept *calque*, since it is also applied to the case of partial substitutions. Many learned words, in certain synchronies, made a sort of "translation" of suffixes, which are inexplicable phonetically. For example, the word *relação* was not inherited from Latin (so we could not display it, as it is usually done in books of Romance philology and historical linguistics, as Latin *relātĭō > relação*, much less as if it were an inherited word, such as Latin *relātĭōnem > relação* or Latin *relātĭōne- > relação*), nor was it a mere loan (and it would be inaccurate to display it as Latin *relātĭō ► relação*, because it is not a regular phonetic change, but systematic). The most appropriate way to indicate it would be Latin *relātĭō ≥ relação*. That is, the symbol ≥ used for calque will also apply to partial translation:

French †*estrang**ier** ≥ estrang**eiro**;
English *goal**keeper*** ≥ go**leiro**;
Spanish *cañ**ón*** ≥ canh**ão**.

Another particular case of calque is the so-called *semantic loan* (also known as *semantic extension* or *loanshift*), as in the aforementioned E. *star* ("celestial body" >> "famous actress") ≥ *star* (both senses) or E. *mouse* ("animal" >> "computer device") ≥ European Port. *rato* (both senses). The semantic loan also occurs with learned words:

Latin *nuclĕus* ("core" >> "nucleus") ≥ German *Kern* (both senses, cf. *Kernphysik*, "nuclear physics").

The so-called *restitution loans* are also included as calque subtypes (BIZZOCCHI, 1998, p. 124), in which a part of a foreign popular or half-learned word is translated by Greek or Latin elements. For example:

English *feed back* ≥ *retroalimentar*.

The same can occur with parts of a word as French suffix -*el*, which becomes -*al* in Portuguese in cases such as French *opérationnel* ≥ *operacional*.
It can be said that there has been a partial restitution when some half-learned or popular parts are translated by learned correspondents and others not, as in:

English *starship* ≥ French *astronef* ≥ *astronave*.

The reverse can also happen. When a calque occurs through the substitution of Greek or Latin elements by vernacular equivalents (*learned word translation*). In this case, it is the replacement of a part of a learned word for another of half-learned, half-popular or popular in nature, with the same meaning. This may occur in whole or in part. Examples:

Latin *super**pōnĕ**re* ≥ *sobre**por***;
Latin *inter**rumpĕ**re* ≥ *inter**romper***;
Latin *per**fect**us* ≥ *per**feit**o*;
Latin *commō**tĭō*** ≥ como**ção**.

The same may occur in half-learned or popular etyma:

French *dés**ordre*** ≥ des**ordem**.

The etymological research of loans and calques must be careful because there is no guarantee that there is an etymological connection in words that may have arisen

independently and were motivated by the concept itself and not by one another. For example, one cannot claim with certainty that *saca-rolhas* is a translation of Fr. *tire-bouchons* ("corkscrew"), nor *acendedor de cigarros* is a translation from E. *cigarette lighter*. As the signifier is strongly motivated by the meaning (as *laranjeira*, "orange tree", is motivated by *laranja*, "orange", which takes almost all languages to name the tree from the name of the fruit), it is very difficult to say that a language has influenced another in the choice of the designation. Nevertheless, neither can this hypothesis be discarded: when an artifact of foreign origin is introduced into a society, it is natural that, along with the thing, comes the name. So whoever coined *saca-rolhas* or *acendedor de cigarros* in Portuguese was certainly not unaware of the original names of these objects. These are therefore cases in which the exact explanation of the word etymology depends on rarely available empirical data, which can make these lexical items remain indefinitely defined as problematic or insoluble classification, decreasing, so to speak, the degree of certainty of the etymological proposition.

Derivation

The symbol $x \rightarrow y$ widely denotes a word y formed in synchrony within a system by various processes, such as prefixation, suffixation, parasynthesis, regression derivation or improper derivation (also known as *conversion*) either from a popular, or inherited, or foreign stem. Examples:

cabeça → cabecear;
mesa → mesário;
saudoso → saudosismo;
transação → transar.

Care must be taken, as many words are misclassified in etymological dictionaries as derivatives when they are actually translations. Example:

desafiar ≤ French *défier* (and not "*desafiar* ← French *défier*").

Another case where the symbol → is used, is the case known as *inflectional neologism* (ALVES, 2007). When Portuguese creates *soldada* from *soldado* ("soldier"), *professora* from *professor* ("teacher") and *presidenta* from *presidente* ("president"), it is a kind of derivation similar to that occurring in Spanish and Italian when they derive *banano* ("banana tree") from *banana* by analogy with *pero/pera* (< Lat. *pirum/pira*, from *pirus/pirum*), i.e. the derivation occurs by changing the thematic vowel (as there is no gender inflection in nouns). It is a very different case from simple inflection. A form such as *cantamos* is not derived from *cantar* or any other inflection because in the

system there is no hierarchy or, properly said, no inflection older than another. Cases like these are simply indicated as follows:

cantamos ~ cantar,

indicating that both inflectional forms are within the same paradigm.

There are derivations formed through learned stems and also learned affixes. In this case, these are not loans because the etymon does not belong to any synchrony and system at all. For instance, Port. *iniciativa* is a loan from Fr. *initiative*, which was created from a non-existent Lat. ★*initiativa*, derived from *initiare*, that is, Latin elements were used but the Latin word itself does not exist. The indication of this etymological proposition would, however, be as if the Latin word were in the same synchrony as French, for it was the source of inspiration of the learned neologism:

Latin *initiare* → French *initiative* ▶ *iniciativa*.

Other examples:

Latin *fractus* → English *fractal* ▶ *fractal*;
Greek γένος → German *Gen* ▶ English *gene* → English *genome* ▶ *genoma*.

We call *truncation* (ALVES, 2007, p. 68) the removal of a part, usually the end, of a lexical sequence. The result of this process is part of a composition element or the sequence formed by an element and part of another.[5] The truncated portion is represented between vertical bars. Examples in Portuguese:

preju|ízo| → *preju*;
micro|computador| → *micro*;
vice|-presidente| → *vice*;
ex|-marido| → *ex*;
French *métro|politain|* → French *métro* ▶ *metrô*;
French *photo|graphie|* → French *photo* ▶ *foto*;
French *cinéma|tographe|* → French *cinéma* ▶ *cinema*.

Sometimes derivation occurs from a learned stem with a half-learned or popular affix (Port. *agricultável*, *deseducar*) or from a half-learned stem with affixes of any kind (Fr. *désordre*). This category also includes the cases of the so-called *false derivation* (or *pseudo-derivation*). Often, when borrowing a Greco-Latin word, a language adds popular suffixes that do not alter the meaning of the word, while the form without

---

5   These elements are known as *prefixoids*, *suffixoids* or, more generally, *quasi-morphemes* and *fractomorphemes*, among other names. Compositions made with these elements are called *portmanteau words* and the process is sometimes known as *recomposition* (LINO, 1990, p. 30-31; ALVES, 2007, p. 69).

the suffix does not exist in the language (or does exist, but the suffixed form does not derive from it). It is in this case a *pseudosuffixation*, and the added element is actually a *pseudosuffix*, since there is no primitive word where the alleged derivation would originate. For example:

Latin *commodus* ► English *commodious* (there is no E. ★*commod*);
Fr. *photographe* ► English *photographer* (for E. *photograph* is a regressive derivation);
Latin *philosŏphus* ► English *philosopher* (there is no E. ★*philosoph*);
Latin *litterārĭus* ► German *literarisch*;
Latin *physĭcus* ► German *physikalisch*;
Latin *mūsĭcus* ► French *musicien*.

Composition

A compound word is formed by the composition of two or more stems (vernacular or foreign) in a given synchrony. Words only enter this category if they have actually been created by this process in the language that is being analyzed. Thus, *puxa-saco* ("who pulls a bag" >> "sycophant"), *pernilongo* ("with long legs" >> "mosquito"), and *cabisbaixo* ("head down" >> "crestfallen") are legitimately Portuguese popular compounds. Instead, *cachorro-quente*, mentioned in the previous item, is not the result of composition in Portuguese, but, as we saw, translation from E. *hot-dog*. The notation of the composition is $x + y \rightarrow z$.[6]

Many of the currently existing learned words did not exist in Greek or Latin, having been created in the very modern European languages, especially French and English and, secondarily, in Italian and German, since they are more influential languages culturally in certain synchronies. Since a large number of these new learned words are technical and scientific terms, it is natural that they arise in the languages of the countries where most of the scientific and technological innovations happen. These learned words are basically created in two ways: by composition between learned stems or by derivation from a learned stem with also learned affixes.

For example:

Greek ἕλιξ + Greek πτερόν ► French *hélicoptère* ► *helicóptero*;
French *social* + French *démocrate* → French *social-démocrate* ► *social-democrata*.

Composition may also occur between elements that have undergone truncation. Examples:

---

6   The name *hybrid composition* is sometimes given to that form constructed with a learned and a half-learned stem, with a learned and a popular stem, with a half-learned and a popular stem, or even two half-learned stems (for example, Port. *auriverde*, *rubro-negro*, *bafômetro*).

show + |co|*mício* → *showmício*;
brasi|leiro| + |*para*|*guaio* →*brasiguaio*;
portu|guês| + |*espa*|*nhol* → *portunhol*;
French *inform|ation|* + French |*auto*|*matique* → French *informatique* ► *informática*;
English *repro|duction|* + English |*photo*|*graphy* > English *reprography* ► *reprografia*;
English *auto|mobile|* + English *part* → English *autopart* ≥ *autopeça*;
English *mini|ature|* + English *skirt* → English *miniskirt* ≥ *minissaia*.

Analogy

The phenomenon of analogy[7] is a transformation that occurs in synchrony and within the same system. In analogy, a word or group of words (a mold) acts on others forming a third element. Analogy itself is not in the same *dimension* as the diachronic phenomenon, but it should be represented *at right angles* to it. Since this notation is complex, the special arrow representing the analogical phenomenon (⇨) is placed in the opposite direction of *transmission* (by inheritance or loan). Thus, if *x* influences *y* such that *z* reflects this influence, the etymon is marked as *z* ⇨ *y* < *x* or as *x* > *y* ⇦ *z* (other symbols may occur in place of > depending on the case), whence we say *x* is the etymon, *y* is the analogical product and *z* is the analogical mold.
Example:

Latin *foresta > *floresta* ⇦ *flor* (and not ★*foresta*);
Latin *consecrare* ► French *consacrer* ⇦ French *sacrer* (and not French ★*consécrer*).

In some cases, a learned word can be borrowed and have the same learned elements, but combined in a manner inconsistent with the phonological system of Greek or Latin. For instance, Spanish has made the Latin words *dīminuĕre*, *immortālis*, and *commōtĭō* match the forms *disminuir*, *inmortal* and *conmoción*, that is, it has rehabilitated archaic forms of Latin, prior to assimilation (although *disminuĕre, *inmortālis, and *conmōtĭō are not attested, but deductible). Yet, the emergence of these forms in Spanish should not have taken place by a desire to reconstruct the primitive form of these words, but by the effect of the analogy with other words (e.g. *disponer*, *intenso*, and *contracción*). Other times, a learned word undergoes phonetic changes typical of popular words, although there are problems with respect to the involved synchronies. This phenomenon is often known as *metamorphism* (BIZZOCCHI, 1998, p. 104). Sometimes these changes are mandatory. The investigation of the synchronies in which such changes have occurred is an urgent factor for NEHiLP because characterizing such forms as results of an analogy can only be done by determining the vocabulary of the past synchronies in which they occurred. Examples:

---

[7]   The product of many analogies is sometimes known as *folk etymology*.

Latin *secta* ▶*se<u>it</u>a* ⇐ -IT-
Latin *doctor* ▶ *do<u>ut</u>or* ⇐ -UT-
Latin *conceptus* ▶*conce<u>it</u>o* ⇐ -IT-
Latin *statŭa* ▶*estátua* ⇐ E-

In this case, we can also include the *half-popular words*, which cannot be considered loanwords, but, at the same time, do not have exactly the same status as the inherited words, since their usage was initially restricted to monitored speech environments, such as the church and university, for example, which protected such words from undergoing all the regular metaplasms to which vernacular words were subject, resulting in hybrid forms, as Lat. *canonĭcus*, *clerĭcus*, *capitŭlum*, and *articŭlus*, which resulted in Portuguese the words *cônego*, *clérigo*, *cabido*, and *artigo*. The category of half-popular words is the one that is not subject to gaining new units, since the phenomenon that produced them occurred only once, in the passing of one historical language to another.

The analogical performance can generate popular, half-learned and, theoretically, even learned words. Examples:

Latin *dromedārĭus* > French *dromedier* ▶ German *Trampeltier* ⇐ German *trampeln* + German *Tier*;
Latin *corporalis* ▶ Italian *caporale* ⇐ Italian *capo*;
Latin *impedĭō* > *impeço* ⇐ *peço* < Latin **petĭō* ⇐ Latin *petō*;
Latin *impressa* > *imprensa* ⇐ *prensa* < Latin *prehēnsa*).

Substitution

The phenomenon of substitution typically involves two distinct transmissions that intersect and is represented by $x \Rightarrow y$.

This symbol is particularly useful for understanding the diachronic rearrangement that systems go through. Thus, a subset of the interrogative adverbs involves the meanings "where", "wherefrom", "whereby", and "whereto", that were expressed in Latin respectively by *ubi*, *unde*, *qua* and *quo*. We can say that Lat. *unde* > *onde*, but Lat. *qua* $\Rightarrow$ *por onde*.[8]

This is the case, for example, of the phenomenon (common in many synchronies) of *restoration* (also called *refection*) of a popular form (which falls into disuse) by another learned one, of the same origin (MAURER JR., 1951, p. 62), meaning something like $z > x \Rightarrow x'$ ◀ *z*. This substitution may be partial or total. For example:

---

[8]   This case is particularly rather complex, for we can say that in a given synchrony: *onde* (< Lat. *unde*) ~ *u* (< Lat. *ubi*) and then, with the disuse of *u*, the word *unde* > *onde* ("wherefrom" >> "where"), whereas, more recently, *a* + *onde* > *aonde* ("whereto" >> "where") as, longer before, Sp. *donde* ("wherefrom" >> "where") < *de* + *onde*.

Latin *silentĭum* > †*seenço* ⇒ *silêncio* ◄ Latin *silentĭum*;
Latin *flōrem* > †*chor* ⇒ *flor* ◄ Latin *flōs, -ris*;[9]
Latin *monastērĭum* > *mosteiro* ⇒ *monastério* ◄ Latin *monastērĭum*.

Sometimes restoration is partial:

Latin *inimīcum* > †*ẽemigo* ⇒ *inimigo* ≤ Latin *inimīcus* (and not ★*inimico*);
Latin *fĕlīcem* > †*fiiz* ⇒ *feliz* ≤ Latin *fĕlix, -īcis* (and not ★*felice*).

Restoration many times also occurs only in spelling, for example:

Latin *nīdum* > French *ni* ⇒ French *nid* ◄ Latin *nīdus*;
Latin *salūtem* > French *salu* ⇒ French *salut* ◄ Latin *salūs, -ūtis*.

In such cases, it only affects spelling, not pronunciation, but there are cases in which the pronunciation is changed because of the spelling, such as Latin *sub* > *so* ⇒ *sob* ⇐ Latin *sub* or Latin *nāscĕre* > *nacer* ⇒ *nascer* ◄ Latin *nāscĕre*, which in European Portuguese is pronounced [nɐʃˈseɾ].

Sometimes the allegedly restored popular word has not been documented and must be reconstructed, for example:

Latin *sŏlum* > **soo* ⇒ *solo* ◄ Latin *sŏlum*;
Latin *altum* > **outo* ⇒ *alto* ◄ Latin *altus*;
Latin *clārum* > **charo* ⇒ *craro ~ claro* ◄ Latin *clārus*.

Some restorations occur by means of half-learned words. This phenomenon is quite common in some synchronies such as, for example, after Renaissance. Some (simplified) etymologies:

Latin *elĕphas, -antis* ► †*alifante* ⇒ *elefante* ◄ Latin *elĕphas, -antis*;
Latin *astronomĭa* ► †*astrolomia* ⇒ *astronomia* ◄ Latin *astronomĭa*;
Latin *informāre* ► †*enformar* ⇒ *informar* ◄ Latin *informare*.

This also includes cases referred to as *transcreation*, in which a learned word is created to replace a part of an existing learned word. For example:

Latin *signātārĭus* ► Italian †*segnatario* ⇒ Italian *firmatario* ← Italian *firma*;
Latin *jūdaismus* ► Italian *giudaismo* ⇒ Italian *ebraismo* ← Italian *ebreo*.

---

9   The restored form can sometimes undergo a new metamorphism; it is what happens to *flor* > †*frol*.

Conclusion

This paper aimed to briefly demonstrate how NEHiLP intends, through its innovative computer technology and large team of researchers from Brazil and abroad, to draw up an etymological dictionary of the Portuguese language that matches in size, quality and data reliability the best works of its kind available in the world for other languages. By its computing platform, *DELPo* will not only be a reference for experts and the general public, but also allow searches by theme (words of a given time, words of a certain origin, words with a certain prefix or stem, etc.), as well as the generation of charts, statistics, reports and many other resources that are likely to position Brazil and the Portuguese language as future references in the field of etymological science.

As we have seen, one of the innovative aspects of the *DELPo* project is the adoption of new concepts and categories for the classification of etymological phenomena, with the consequent introduction of a new symbolic notation that includes them. It is expected that such concepts and symbols will become progressively integrated into mainstream etymological and philological literature, which will be another Brazilian contribution to scientific advancement in the area.

As no scientific theory is ever fully completed, the very practice of research and development of the dictionary will reveal whether such innovations are sufficient or whether new concepts and symbols need to be introduced, or if the current ones need to be modified, revised or corrected. Suggestions and contributions are always welcome.

- *RESUMO: O presente artigo visa a apresentar a metodologia de trabalho e os principais postulados teóricos que norteiam a elaboração do DELPo (Dicionário Etimológico da Língua Portuguesa), a cargo do NEHiLP-USP (Núcleo de Apoio à Pesquisa em Etimologia e História da Língua Portuguesa da Universidade de São Paulo), visto tratar-se de projeto inovador tanto em seus aspectos teóricos quanto empírico-operacionais. Pretende-se aqui apresentar as inovações conceituais, notacionais e terminológicas propostas pelos autores, que embasam a elaboração do DELPo. Para tanto, baseamo-nos em referenciais teóricos sobre neologia (ALVES, 2007; BARBOSA, 1993, 1996), lexicogênese (BIZZOCCHI, 1998) e etimologia (VIARO, 2011), bem como propomos uma atualização na simbologia utilizada na formulação de proposições etimológicas, que ao mesmo tempo dirima ambiguidades e inconsistências da notação tradicional e dê conta das inovações conceituais aqui introduzidas. Pretende-se que tanto os processos etimológicos descritos quanto os símbolos a eles correspondentes se tornem padrão na pesquisa em etimologia e na lexicografia etimológica.*

- *PALAVRAS-CHAVE: Etimologia. Lexicografia. Dicionários etimológicos. Proposições etimológicas. Discussão etimológica. Simbologia etimológica. Notação linguística.*

REFERENCES

ALVES, I. M. **Neologismo**: criação lexical. 3.ed. São Paulo: Ática, 2007.

BARBOSA, M. A. **Léxico, produção e criatividade**: processos do neologismo. 3.ed. São Paulo: Plêiade, 1996.

BARBOSA, M. A. Acrograma e sigla: estatuto semântico-sintáxico e tratamento na obra lexicográfica. In: REUNIÃO ANUAL DA SBPC, 45., 1993. **Anais…** Recife: Sociedade Brasileira para o Progresso da Ciência, 1993. p. 477.

BECHARA, E. As palavras também escondem a idade. In: ELIA, S. et al. (Org.). **Na ponta da língua 1**. Rio de Janeiro: Liceu Literário Português; Lucerna, 1998.

BIZZOCCHI, A. **Léxico e ideologia na Europa ocidental**. São Paulo: Annablume, 1998.

BIZZOCCHI, A. A ideologia das raízes. **Língua Portuguesa**, São Paulo, v.40, p. 60-63, Feb. 2009.

BIZZOCCHI, A. Processos de formação lexical das línguas românicas e germânicas: uma nova perspectiva teórica. **Domínios de Lingu@gem: Revista Eletrônica de Linguística**, v.7, n.1, p. 9-39, July 2013.

JESPERSEN, O. **Language**: its nature, development and origin. London: Allen & Unwin, 1954.

LINO, M. T. R. da F. Observatório do português contemporâneo. In: COLÓQUIO DE LEXICOLOGIA E LEXICOGRAFIA, 1., 1990. **Actas**. Lisbon: Universidade Nova de Lisboa, 1990. p. 28-33.

MAURER JR., T. H. **A unidade da Romênia ocidental**. São Paulo: FFCL-USP, 1951.

UNIVERSIDADE DE SÃO PAULO. **Programa Metaplasmador**. São Paulo, [2016]. Disponível em: <www.usp.br/nehilp/infos/metaplasmador-gh-pages/index.php>. Acesso em: 20 jul. 2016.

UNIVERSIDADE DE SÃO PAULO. Núcleo de Apoio à Pesquisa em Etimologia e História da Língua Portuguesa. **Manual do NEHiLP**. São Paulo, 2015. Versão 2.1. Disponível em: <www.usp.br/nehilp/infos/manual.pdf>. Acesso em: 20 jul. 2016.

VIARO, M. E. **Etimologia**. São Paulo: Contexto, 2011.

VIARO, M. E.; FERREIRA, M. J.; GUIMARÃES-FILHO, Z. O. Derivação ou terminação: limites para a semântica, lexicologia e morfologia históricas. In: VIARO, M. E. (Org.). **Morfologia histórica**. São Paulo: Cortez, 2013. p. 58-105.