

## CONTENT AND FUNCTION WORDS IN ORAL DISCOURSE: A COMPARISON OF A FORM OF THE AUXILIARY VERB *PODER* AND THE NOUN *PÓ* IN BRAZILIAN PORTUGUESE

SILVA, Luis Filipe Lima<sup>1\*</sup>

<sup>1</sup>PhD in Linguistic Studies, Federal University of Minas Gerais – ORCID: <https://orcid.org/0000-0003-0188-2861>

---

**Abstract:** *The goal of this paper is to verify whether there is an acoustic difference between a reduced form of the verb poder (po') 'to be able to' and the noun pó 'powder' in Brazilian Portuguese oral discourse. The reduction of the verb occurs in contexts in which it is an auxiliary in periphrases formed with the infinitive. This process results in a form segmentally homonymous with the noun pó [ˈpɔ]. As the literature states that function words are shorter in duration than content words, homonymy is expected to occur at the segmental, but not at the suprasegmental level. To verify this hypothesis, we analyzed data from Brazilian Portuguese spontaneous speech, selecting the parameters of normalized duration (in smoothed z-score), mean intensity, mean f0 of both [ˈpɔ] and its contiguous words, and the mean frequency of the formants (F1 and F2) of the vowel of both the verb and the noun. Statistically significant results were found only for duration, confirming that the auxiliary verb (a function word) is shorter than the noun (a content word). It is suggested that the suprasegmental level, namely duration, is an important correlate in differentiation between two segmentally homonymous words but from different categories.*

**Keywords:** syntax; function and content words; auxiliary verb; phonetic reduction; suprasegmental features.

---



---

\* Corresponding author: [luisfl397@gmail.com](mailto:luisfl397@gmail.com)

## 1 Introduction

In modern linguistics, the systematic study of the distinction of words into content (or lexical) and function (or grammatical) categories has been explored for more than a century within the scope of research in linguistic change<sup>1</sup>, specifically in grammaticalization, with the phenomenon initially conceived as the transition of items from the lexicon to grammar (cf. Meillet 1912; Kuryłowicz 1965; Heine, Claudi & Hünnemeyer 1991; Hopper & Traugott 2003; Narrog & Heine 2011; 2021; Lehmann 2015; 2024). Furthermore, the consequences of the proposal regarding the division of words into content and function categories have been increasingly explored in studies that investigate (i) speech pathologies, such as agrammatism, a characteristic of non-fluent aphasia in which individuals basically use content words, that is, there is a level of absence of function words in their speech (cf. Goodglass & Mayer 1958; Kean 1985; Bastiaanse & Thompson 2012; Wang et al. 2019; Fedorenko, Ryskin & Gibson 2023), as well as in studies that investigate (ii) the acoustic correlates of words from content and function categories, especially the effects of duration between them (cf. Swanson & Leonard 1994; Turner & Tjaden 2000; Bell et al. 2009; Lee, Kong & Wang 2014; Brunelle, Chow & Nguyễn 2015), and (iii) other studies that also investigate the differences between content and function words under different methodologies and populations (Mohr, Pulvermüller & Zaidel 1994; Diaz & McCarthy 2009; Liu et al. 2010; Baese-Berk et al. 2019).

This paper's main topic is the acoustic analysis of the phonetic reduction of the grammaticalized auxiliary verb *poder* 'to be able to' in Brazilian Portuguese. It focuses on a quantitative perspective of the phenomenon with data gathered from a spontaneous speech corpus. The analysis will be based on verifying the phonetic reduction of the verb *poder* to its inflected third person singular (3SG) form of the present indicative<sup>2</sup> *po'* ['pɔ] (a function word) in comparison with the noun *pó* ['pɔ] 'powder'<sup>3</sup> (a content word). The basic feature of the definition that oppose content words to function words is that the former have conceptual meaning, like the noun *pó*, and the latter have procedural meaning, like the auxiliary *po'*. The question to be answered is whether the phonetic reduction that the 3SG form of the verb *poder* undergoes makes it in fact homonymous with the noun *pó*. Based on the hypothesis that function words are shorter than content words (van Bergem 1993; Shi et al. 2005), it is expected that homonymy occurs at the segmental, but not at the suprasegmental level. Before exploring this hypothesis, however, it is important to understand the origins of the verb *poder* and how one of its forms became reduced.

## 2 The verb *poder*: Diachrony and phonetic reduction

### 2.1 Origins of the verb *poder*

The verb *poder* comes from the Latin *possum*, which, in turn, is the reanalysis of the adjective *potis*, *-e* 'able to' and a form of the verb *sum* 'be' (Fruyt 2011: 833). Fruyt (2011) states that the final consonant of *potis* tended to disappear, giving rise to the forms *potest* (3SG) and *potes* (2SG)<sup>4</sup>.

---

<sup>1</sup> It is worth mentioning that other older traditions of linguistic studies have already identified such a distinction (cf. Boye 2023: 72, for more information).

<sup>2</sup> From now on only the person/number feature will be mentioned when referring to the verb form.

<sup>3</sup> Throughout this paper, the reduced form of the verb *poder* will be transcribed as *po'*, while the noun meaning 'powder' will be transcribed orthographically as *pó*.

<sup>4</sup> This issue is also explained by Benveniste (2016: 63): "In Latin an extensive word family is organized around the word \**potis* either as a free form or in composition. Apart from *hospes* it forms the adjectives *impos*, *compos* 'who is not...' or 'who is a master of himself, of his senses' and the verb \**potire*, the perfect of which, *potui*, survives incorporated into the conjugation of the verb meaning 'be able', *possum*, which itself is formed from the adjective *potis* in a predicative use: *potis sum*, *pote est*, an expression which is simplified to *possum*, *potest*".

Some of the morphemes that originated from the inflected form of the verb are a root in complementary distribution (A1) *pot-* (before a vowel) and (A2) *pos-* (before *s*), and (B) a bound form of the verb *sum* (*pos-sum*, *pot-es*, *pot-est*, etc.). Still according to Fruyt (2011), the meaning of the adjective from which the verb *possum* originated is a case of grammaticalization, since it is an evolution of the Indo-European root *\*pot-* / *\*pet-*, which denoted possession and property to a word expressing ability (see also Benveniste 1954). The modal meaning of the verb *possum* started with *pot est* and *potest* in sentences whose subject was an infinitive or a subordinate clause, as evidenced by the examples in (1).

(1) Latin *possum* (Danckaert 2017: 222)

a. *ut omni-s contio audi-re pos-s-et*  
 so.that whole-NOM gathering.NOM hear-PRS.INF be.able-IPFV.SBJV-3SG  
 ‘so that the entire gathering may be able to hear’ (Cicero, *Pro Cluentio* 134)

b. *ut nemo [...] uoc-em acerb-am atque inimic-am*  
 so.that nobody.NOM [...] word-ACC bitter-ACC and hostile-ACC  
*bon-is pos-se-t audi-re*  
 good-DAT.M.PL be.able-IPFV.SBJV-3SG hear-PRS.INF  
 ‘so that nobody could hear any bitter or hostile word uttered against the good citizens’ (Cicero, *Post reditum in senatu* 26)

It can be seen, therefore, that Latin *possum* was already an auxiliary verb with a modal meaning. Furthermore, it can be noted that in Latin there were two possibilities for expressing verbal periphrasis, namely, [non-finite verb + auxiliary] and [auxiliary + non-finite verb]; the second order was the one that was fixed in Romance languages (cf. Danckaert 2016). An important point to be mentioned is that the verb *possum* could be used intransitively in Latin, as in *omnia possum* ‘I can (do) everything’ (Almeida 2000: 234). This is a lexical use of the verb, which also has instances in Portuguese as can be seen in the data below.

(2) Lexical *poder* (Fernandes 1993: 465)<sup>5</sup>

a. *Quem é esse inimigo tão formidável que tanto pode?*  
 who be.3SG.PRS.IND this.M enemy.M so formidable that so.much can.3SG.PRS.IND  
 ‘Who is this formidable enemy that wields such power?’

b. *Cuidaram que Cristo podia menos dormindo que acordado.*  
 care.3PL.PST.IND that Christ can.3SG.PST.IND less sleep.GER that awake.PST.PTCP.M.SG  
 ‘They thought Christ could do less while asleep than when awake’

c. *Quanto mais pode a fé que a força humana!*  
 how.much more can.3SG.PRS.IND the.F faith that the.F strength human.F  
 ‘How much more can faith achieve than human strength!’

d. *Amontoando razões que nada puderam com Álvaro.*  
 heap.GER reason.PL that nothing can.3PL.PST.IND with A.

<sup>5</sup> Fernandes (1993) warns that many authors consider the verb *fazer* ‘to do’ to be implied in some cases (see also Coelho 2013). Additionally, Almeida (2000) puts the verb *fazer* in parentheses in the translation of the Latin *omnia possum* example. This information is important, although the status of *poder* as a lexical verb is undoubted. Other examples of lexical *poder* can be found in Borba (1991).

‘Accumulating reasons that could do nothing with Álvaro’

However, it must be said that probably most occurrences have always been the grammatical instance of the verb, considering that the grammaticalized form is generally superior in frequency of use to the lexical form (Bybee, Perkins & Pagliuca 1994; Bybee 2003). This is especially true in the case of an auxiliary verb, a very common part-of-speech in Portuguese (Pontes 1973; Lobato 1975; Coelho 2006), although diachronic research is necessary to confirm this hypothesis. As shown above, the grammatical use was already latent in Latin, and it is possible to observe it in the oldest Portuguese documents, such as in a medieval song from the 13<sup>th</sup> century (Lopes & Ferreira 2011):

(3) Grammatical *poder* [13<sup>th</sup> century medieval poem]

*De parecer e de falar*

of seem.INF and of speak.INF

*e de boas manhas haver,*

and of good.F.PL slyness.PL there.is.INF

*ela, nōn'a pode vencer (...)* [João Garcia de Guilhade]

she not her can.3SG.PRS.IND win.INF

‘In her appearance and speech, and in her charming abilities, she cannot be surpassed’

Note the modal meaning of the auxiliary verb in the periphrasis *pode vencer* (lit. ‘can win’). According to Fleischman (1982: 51), this is the only meaning inherited by the verb *possum* in Romance languages; in Latin, it could also be used to form an infinitive future (*posse facere for facturum esse*). This study, however, specifically concerns the phonetic reduction that the 3SG form inflected in the present indicative of this verb undergoes in contexts in which it is a modal auxiliary. This is the topic of the next section.

## 2.2 Phonetic reduction of the verb *poder*

An important study conducted by Coelho & Paula (2011) on the phonetic reduction of the verb *poder* argues that this phenomenon is the result of a grammaticalization process. Among other factors, the study highlights that there is no reduction when the verb is in its simple form, that is, in its lexical use, contrary to the contexts in which it occurs in a periphrasis, as illustrated by the data below.

(4) Lexical *poder* (Coelho & Paula 2011: 106)

a. *Quem pode o muito pode o pouco.*

who can.3SG.PRS.IND the.M much.M can.3SG.PRS.IND the.M little.M

‘He who can do much can do little’

b. \**Quem po' o muito po' o pouco.*

(5) Lexical (a)/Grammatical *poder* (b) (Coelho & Paula 2011: 89)

a. \**Ninguém po' com esse menino levado.*

nobody can.3SG.PRS.IND with this.M boy.M naughty.M

‘Nobody can handle this mischievous boy’

b. *Ninguém po' dize(r) que ele não estava avisado.*

nobody can.3SG.PRS.IND say.INF that he not be.3SG.PST.IND warn.PST.PTCP.M.SG

‘Nobody can say he wasn’t warned’

(6) Grammatical *poder* (Coelho & Paula 2011: 106)

- a. **Pode**                    **deixar**    *em cima da*    *minha mesa*.  
can.3SG.PRS.IND leave.INF in top of.F my.F table  
'You can leave it on my desk'
- b. **Po'** **de(i)xa(r)** *em cima da minha mesa*.

Additionally, the authors state that the syntagmatic variability of the auxiliary verb is low, which would confirm the degree of grammaticalization proportional to the fixity of the form. This change, according to the authors, is not random, but governed by (i) syntactic criteria, such as the fact that phonetic reduction only occurs in cases in which the verb *poder* is linked to the infinitive to form a verbal periphrasis, that is, when it is a modal auxiliary, and by (ii) phonetic criteria, since they argue that reduction does not occur when the verb is followed by an infinitive that starts with a vowel and before a one-syllable verb, such as *ser* 'to be' and *ir* 'to go', although the authors recognize that reduction can occur with the one-syllable verbs *ter* 'to have' and *dar* 'to give', for example.

According to Coelho & Paula (2011), the verb *poder* undergoes phonetic reduction when inflected in the 1SG of the present indicative; however, this is only an elision of the vowel following the sibilant, which does not constitute a loss of the syllable. In this case, syllabic status is maintained, and this is justified by non-voicing of the sibilant despite the following sound environment, as in (7b). On the other hand, voicing of the sibilant occurs in the prefix *pós* 'post' in (7a) due to the influence of the following sound context. Such maintenance illustrated in cases as (7b) would ensure the syllabic status of the post-tonic, differently from what occurs with the inflected 3SG of the present indicative of the verb *poder*, in which there is an absolute erosion of the final syllable.

(7) *pós*-/*posso* comparison (Coelho & Paula 2011: 108)

- a. *Sua vida* **pós-bebê** [pɔz be'be] *não é*                    *mais a mesma*.  
your.F life post-baby                    not be.3SG.PRS.IND more the.F same.F  
'Your post-baby life is no longer the same'
- b. *Eu não* **posso**                    **beber** ['pɔs be'be] *hoje*  
I not can.1SG.PRS.IND drink.INF                    today  
*porque estou*                    *tomando antibiótico*.  
because be.1SG.PRS.IND taking.GER antibiotics  
'I can't drink today because I'm taking antibiotics'

Coelho & Paula (2011) conclude by saying that 3SG favors the loss of the syllable, while 1SG only favors the loss of the final phoneme. Additionally, they state that the reduction occurring in 1SG does not result from grammaticalization, since it can be seen in other non-grammaticalized verbs, as shown in (8).

(8) Verb *passar* (Coelho & Paula 2011: 108)

- Não* **passa**    **dentro** ['pas 'dẽtro] *desse buraco que você (se) machuca*.  
not pass.IMP inside                    this.M hole that you (REFL) hurt  
'Don't go through that hole because you get hurt'

Notwithstanding, although Coelho & Paula (2011) classify the phonetic reduction that occurs with the verb *poder* as a consequence of the grammaticalization process, we believe that

this phenomenon is more general. In other words, we do not believe that the phonetic reduction of the verb *poder* is carried out by a grammaticalization process. In this sense, we agree with the idea of Giomi (2023: 103) according to which “(...) there also are no processes of phonological adaptation specific to grammaticalization but all follow the phonological rules that regularly operate at the relevant layer of morphosyntactic organization, irrespective of the lexical or grammatical nature of the formatives involved”<sup>6</sup>. In this regard, note that the verb *estar* ‘to be’, for example, is reduced both in contexts where it is a lexical (9a) and in contexts where it is a grammatical (9b) verb. Therefore, it would be not possible to explain that the reduction of (9b) arises from grammaticalization, as in this example the verb is an auxiliary, and (9a) arises from specific phonological rules, since the result is the same for both cases, that is, it would be an *ad hoc* explanation.

(9)

a. *Ele tá em casa.*  
he be.3SG.PRS.IND in house  
‘He’s at home’

b. *Ele tá estudan(d)o agora.*  
he be.3SG.PRS.IND study.GER now  
‘He’s studying right now’

Both cases of phonetic reduction, but especially the one in (9b), would not result from a grammaticalization process, but probably from a phonetic context that favors reduction correlated with a high frequency of use of the verb. This idea is strongly associated with the work of J. Bybee, who states that “(...) use, which is measured by token frequency, provides the locus of change, without which sound change cannot occur. It would follow, then, that frequency of use is in a way part of the phonetic context” (Bybee 2017: 295)<sup>7</sup>.

The grammaticality judgment presented in the study by Coelho & Paula (2011) for lexical *poder* data, wherein sentences in which the verb was phonetically reduced were judged ungrammatical, may result from an introspection bias towards occurrences of very low frequency words, as it seems to be the case with lexical *poder*. In any case, although the phonetic reduction of lexical *poder* is possibly ungrammatical, this is probably due more to its low frequency of occurrence than to an issue involving grammaticalization. That is, low frequency would block phonetic reduction. The lexical or grammatical status of the verb, therefore, would have nothing to do with the phonetic reduction process that operates in the cases illustrated in (9), as can be seen from Giomi’s (2023) proposal. In its history, the verb *poder* went through a process of grammaticalization, as we showed in the previous section, since it is possible to observe lexical

---

<sup>6</sup> This idea was already latent in the work of Bybee, Perkins & Pagliuca (1994: 20): “There is a link between frequency of use and phonetic bulk such that more frequently used material, whether grammatical or lexical, tends to be shorter (phonetically reduced) relative to less often used material”. Although the terms content and function (or lexical and grammatical) words are not explicitly stated, what was proposed by Bybee, Perkins & Pagliuca (1994), in turn, can be recovered, in some way, from Zipf’s work in another context: “(...) the length of a word, far from being a random matter, is closely related to the frequency of its usage – the greater the frequency, the shorter the word” (Zipf 1968[1935]: xi).

<sup>7</sup> The germ of this idea was proposed in the work of Zipf (1929: 4): “*Principle of Frequency*. The accent, or degree of conspicuousness, of any word, syllable, or sound, is inversely proportionate to the relative frequency of that word, syllable, or sound, among its fellow words, syllables, or sounds, in the stream of spoken language. As usage becomes more frequent, form becomes less accented, or more easily pronounceable, and *vice versa*”.

and grammatical instances of its use. However, the reduction to the form *po'* would not necessarily result from grammaticalization for the reasons already mentioned.

It remains to be seen whether the process that leads to the phonetic reduction of the verb acts to make it entirely homonymous with the noun *pó*. This is what this study aims to answer. However, before delving into that, we will explain the methodology employed.

### 3 Materials and methods

The data for this study were gathered from the reference corpus of informal spoken Brazilian Portuguese C-ORAL-BRASIL I (Raso & Mello 2012). As expected, the frequency of the auxiliary verb in both its full form (*pode*) and in its reduced form (*po'*) is much higher than the frequency of the noun *pó*. Thus, for comparison purposes, we sought to equate the number of occurrences of *po'* with that of *pó*, randomly selecting fifteen occurrences of the verb *po'*, as this is the number of occurrences capable of acoustic analysis of the noun *pó* in the corpus consulted<sup>8</sup>.

To assess the measures related to the degree of reduction of the verb *po'* and compare it with the noun *pó*, the following variables were selected<sup>9</sup>:

- (1) Normalized duration
  - a. Comparison between the duration of the verb and the noun;
  - b. Comparison between the duration of the verb and noun and their respective contiguous words<sup>10</sup>;
  - c. Comparison between the sum of the duration of the verb and the noun and their respective contiguous words (duration of the verb + duration of its contiguous words with the duration of the noun + its contiguous words duration);
  - d. Comparison between subtracting the difference between the verb and the noun and their respective contiguous words ((verb duration - contiguous words duration) - (noun duration - contiguous words duration)).
- (2) Mean intensity: the same calculations as (1);
- (3) Mean f0: the same calculations as (1);
- (4) Comparison of mean frequencies of the first (F1) and second (F2) formants of the vowel [ɔ] of the verb and the noun.

All measures were extracted using Praat software (Boersma & Weenink 2024). For duration measures specifically, the SGDetector script (Barbosa 2006; 2013)<sup>11</sup> was used. This script normalizes the duration of syllables using a z-score transformation technique followed by smoothing (cf. Barbosa, Camargo & Madureira 2016: 165; Barbosa 2022: 115). The objective of

---

<sup>8</sup> One occurrence of the noun was excluded from the analysis because it was uttered in a context of speech overlap.

<sup>9</sup> These measures were selected to ensure that, under a specific speech rate condition, a reduction is still observed, as assessed by the z-score in relation to the context. In other words, without the context it is impossible to verify whether a word has been reduced.

<sup>10</sup> Two words were selected before and two words after each target word, that is, two words before and two words after each [‘pɔ] realization.

<sup>11</sup> SGDetector’s source code is available in the following GitHub repository: <https://github.com/pabarbosa/prosody-scripts/tree/master/SGDetector>

using the script is to minimize the impact of the intrinsic duration of the segments as well as irrelevant stress for some prosodic functions.

Normalization by z-score provides the quotient between the duration in milliseconds (ms) subtracted by the reference mean in ms of the phones, divided by their variance. In other words, for each measurement carried out, this value is subtracted from the mean and then divided by the standard deviation<sup>12</sup>. The z-score specifies, in standard deviation units, how far the measured value differs from the reference mean. Therefore, there will be negative results, which mean that the z-score value is lower than the reference mean, and positive results, indicating that the z-score value is higher than the reference mean.

Subsequently, there is a smoothing procedure for the z-score value calculated multiple times based on the mean and dispersion found by the script in certain intervals. The result provided by the script allows the viewing of the duration of the segments in milliseconds, in z-score and in smoothed z-score.

From each comparison conducted, the Shapiro-Wilk test was applied to verify the normality of the data, and subsequently, either the T-test for independent samples or the Mann-Whitney U test was used according to the distribution of each sample. All work was performed in the Python programming language with the help of the following libraries: Pandas (McKinney 2010), for organizing data into dataframes, NumPy (Harris et al. 2020) and SciPy (Virtanen et al. 2020), for statistical analysis, and Seaborn (Waksom 2021) and Matplotlib (Hunter 2007), for visualizations.

## 4 Results

Initially, we found that the number of occurrences of the 3SG form is much higher than the others. Around 80% of the data of the verb *poder* in the corpus corresponds morphologically to the 3SG form. This corroborates the hypothesis that the high frequency associated with automatization is an important factor in reducing items, as proposed by Bybee (2017) and Kapatsinski, Easterday & Bybee (2020). This means that the form *pode* is much more susceptible to reduction to *po'* than the other forms (around 71% of the 3SG corresponds to the present indicative form, cf. table 2), and this is what happens, given that there is no reduction from *posso* to *po'* or from *podemos* to *pô*, for example.

**Table 1:** Distribution of the verb *poder* in the corpus according to its morphological form

Morphological form of the verb	Occurrences	%
1PL	2	0.24
1SG	88	10.66
3PL	12	1.45
3SG	663	80.27
GER	1	0.12
INF	60	7.26
Total	826	100

Observing specifically the data from 3SG of present indicative, we found that the full form of the verb occurs most frequently in periphrases with an infinitive. Therefore, if it is hypothesized

<sup>12</sup> How the reference values were obtained, and more details of the calculations can be found in Barbosa (2006).

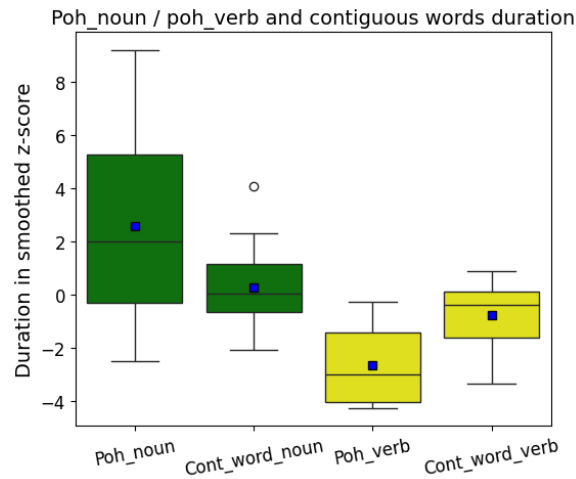
that there is a process of change underway, it is likely to be at its beginning, considering that the frequency of the reduced form is still lower compared to cases in which there is no reduction of the verb. In any case, it is important to note that around 14% of the data are in the reduced form, which shows that this phenomenon is not rare in oral discourse<sup>13</sup>.

**Table 2:** Distribution of 3SG of the present indicative form of the verb *poder* in the corpus

Constructions	Occurrences	%
[pode INF]	404	84.87
[pode ø]	2	0.42
[po' INF]	68	14.29
[po' ø]	2	0.42
Total	476	100

To check whether the reduced form of the verb *poder* (*po'*) is in fact a homonym with the noun *pó*, we proceeded to measure the acoustic variables, such as normalized duration, f0, intensity of [ˈpɔ] and its contiguous words, and the formants (F1 and F2) of the vowel [ɔ] in both the verb and the noun. Considering these four parameters analyzed, there was a statistically significant result only for duration. Regarding this variable, we can observe in the boxplot that the verb *po'* has a lower mean (-2.647) and median (-3.020) than the noun *pó* (mean = 2.565 and median = 1.980). This means that, in the data analyzed, the tendency was for *po'* to be uttered for a shorter period than *pó*. Additionally, there is greater dispersion in the data for the noun *pó* (std = 3.597) than for the verb *po'* (std = 1.361). Speakers therefore uttered the verb in a shorter time than the noun, with statistical significance confirmed in the T-test for independent samples (t = 5.24, df = 28.0, p-value < 0.001).

<sup>13</sup> The symbol ø in Table 2 means that no main verb was uttered, although it can be inferred from the context.



**Figure 1:** Boxplot of the duration in smoothed z-score

**Table 3:** Statistical summary of the results

	Construction	Poh_noun	Poh_verb
Construction_dur	sum	38.480	-39.700
	mean	2.565	-2.647
	median	1.980	-3.020
	std	3.597	1.361
	min	-2.500	-4.270
	max	9.200	-0.260
Cont_word_dur	sum	4.000	-11.370
	mean	0.267	-0.758
	median	0.020	-0.380
	std	1.628	1.231
	min	-2.070	-3.340
	max	4.080	0.880

There were also statistically significant results when comparing the duration between the verb *po'* and its respective contiguous words ( $t = -3.98$ ,  $df = 28.0$ ,  $p\text{-value} < 0.001$ ), as well as between the duration between the noun *pó* and its respective contiguous words ( $t = 2.25$ ,  $df = 28.0$ ,  $p\text{-value} = 0.03$ ). This means that the distribution of duration did not occur by chance, which in turn confirms, as can be seen in the boxplot, that the verb has a shorter duration than its respective contiguous words, just as the noun has a longer duration than its respective contiguous words.

This difference is one aspect that marks a differentiation between segmental homonyms at the categorial level, distinguishing [*po*] as a content word from [*pó*] as a function word. This also corroborates what is widely reported in the literature investigating the differences between content and function words. Such studies attest that function words are shorter than content words.

Additionally, it was noted that there was a difference in duration between the two groups when the target words were added up to their respective contiguous words. It was therefore verified that the mean duration of the verb and its respective contiguous words (-3.404) is lower than the mean duration of the noun and its respective contiguous words (2.832). This means that not only is the verb shorter than the noun, but also that the entire construction involving the environment in which the verb occurs is shorter than the corresponding noun ( $t = 5.36$ ,  $df = 28.0$ ,  $p\text{-value} < 0.001$ ).

This same statistical relevance was also attested when the difference in duration between pairs in constructions was subtracted. That is, the difference between the verb and its respective contiguous words was subtracted, as well as the difference between the noun and its respective contiguous words ( $t = 3.86$ ,  $df = 28.0$ ,  $p\text{-value} < 0.001$ ). It was found that the mean of the distribution after applying the subtraction of the verb group (-1.888) is lower than that of the noun group (2.298). It is suggested, therefore, that there is a greater difference between the verb and its contiguous words than between the noun and its respective contiguous words. All the results of the acoustic analysis attested by means of statistical tests indicate the clear difference, also evident in the boxplot, between [‘pɔ] as a verb and [‘pɔ] as a noun.

As mentioned previously, none of the other variables analyzed presented statistically significant results. Therefore, there was no statistically significant result for the same difference found in the duration parameter between verb and noun with regard to  $f_0$  ( $t = 0.97$ ,  $df = 28.0$ ,  $p\text{-value} = 0.33$ ) and intensity ( $t = 0.77$ ,  $df = 28.0$ ,  $p\text{-value} = 0.44$ ), nor between  $F_1$  ( $t = -0.15$ ,  $df = 28.0$ ,  $p\text{-value} = 0.87$ ) and  $F_2$  ( $t = -0.94$ ,  $df = 28.0$ ,  $p\text{-value} = 0.35$ ) of the vowel [ɔ] in the verb and the noun. These results mean that duration is the only pertinent variable to characterize the phenomenon acoustically.

## 5 Final remarks

This paper showed that the homonymy between *po*’ and *pó* occurs at the segmental, but not at the suprasegmental level. In this regard, the acoustic parameter of duration is an important correlate that distinguishes the two words. Additionally, it was found that the auxiliary verb, a function word, is uttered with a shorter duration, corroborating previous studies that attested that function words are shorter than content words.

It was also discussed that the reduction process responsible for transforming the 3SG form of the verb *poder* to the form *po*’ is not grammaticalization, considering that the same verb, such as *estar*, can be reduced both when it is lexical (a content word) and when it is an auxiliary verb (a function word). This process was attributed to the high frequency of occurrence of the 3SG form, as seen in the distribution of the data in the corpus. This phenomenon is explained extensively in the work of J. Bybee (see also Gahl 2008; Gahl, Yao & Johnson 2012). According to the author, the more a word is used, the greater the chance that it will be reduced in speech (Bybee 2002).

Phonetic reduction, in turn, is correlated to neuromotor processing, so that the repetition carried out by the high frequency of use mobilizes neuromotor commands and actions that lead the repeated sequences to become more efficient due to the reduction of the gestures involved in the process (Bybee 2007; 2010). Likewise, language efficiency has been precisely associated with phonetic reduction (cf. Zipf 1949; Gibson et al. 2019; Levshina 2022a).

Probably, the predictability factor also acts in the process of phonetic reduction of the verb *poder*, considering that such reduction occurs mostly in periphrasis contexts in which a verb follows in the infinitive, and the frequency factor does not act in isolation. That is, there are specific contexts that cause a word to be phonetically reduced (Bybee 2006; Levshina 2022b; Ryskin & Nieuwland 2023; Bybee & Brown 2024).

To our knowledge, the verb *poder* is one of the few verbs that reduces its entire person-number features and even part of its root (*pod-* to *po*’)<sup>14</sup>, so there should be a level of predictability

---

<sup>14</sup> Another verb known to exhibit such a reduction is *deixar* ‘to leave’ in the discourse marker *deixa eu* ‘let me’, as in *deixa eu ver* ‘let me see/think’. In this example, the entire discourse marker *deixa eu* can be uttered as [ʃo]. For more information, cf. Bybee, File-Muriel & Souza (2016), and Silva, Costa & Coelho (2022).

context to this reduction process to occur. Otherwise, the hearer couldn't retrieve some grammatical features of the verb in the communicative situation.

From a semantic-pragmatic point of view, we hypothesize the reduction could have to do both with (i) a weakening of the modal force provided by the verb in the pragmatic contexts of its use (cf. Silva-Corvalán 1985; Bybee, Perkins & Pagliuca 1994; Bybee & Fleischman 1995; Traugott & Dasher 2002; Narrog 2012, for the relation between semantic weakening and modality) and (ii) with some level of politeness and informality associated with the representation that the reduced form could suit in the contexts in which it is used (cf. Ofuka et al. 2000; Aalberse 2009; Caballero et al. 2018, for the relation between politeness and phonetic reduction).

Furthermore, in the theory of Z. Harris (cf. Harris 1976a; 1976b; 1978; 1988; 1991), the interplay of predictability, frequency and reduction could be explained by the level of information of the word:

The words that have highest likelihood, i.e., are expectable, in a given environment contribute little or no information when they enter there, in the information-theoretic sense (...). It is relevant that reduction takes place in several different high-likelihood and special status situations. This suggests that what determines reducibility is not simply high frequency but low information (...) (Harris 1988: 20).

Seyfarth (2014)'s work seems to attest this idea because the result of his study supports the fact that low-informativity<sup>15</sup> words exhibited shorter durations even when some variables such as local contextual predictability, frequency and speech rate were controlled for, although the author only included content words in his analysis. Considering that *po* ' is a low-information word, since it occurs mostly in one specific context, maybe this could be intertwined as well with the statements made in (i) and (ii) above. Future studies can investigate this relationship more in-depth.

The results found in this paper also corroborate the phonetic paradigm uniformity, according to which "durational targets of a target word's morphological relatives influence the realisation of that word" (Seyfarth et al. 2017: 45). As function words are shorter and the auxiliary verb is a function word, its phonetic realization as [*'pɔ*] would inherit an articulatory time distinct from that of the noun *pó*, a content word, especially because the reduction to the form *po* ' comes from the loss of the post-tonic syllable of the verb. And this is precisely what characterizes the differentiation between the function and the content word: the difference in duration, regardless of whether at the segmental level they are identical sequences.

What this study has shown, therefore, is that the suprasegmental level is an important correlate in distinguishing word classes in cases of segmental homophony together with frequency, the syntactic context, and the predictability involved in the realization of the forms. To conclude, this study opens perspectives to explore similar phenomena associated with linguistic change (Bybee 2003; Heine & Song 2011; Heine et al. 2021), as well as to explore the interplay between function words and rhythm in the speech flow (Barbosa 2007; Dilley & Pitt 2010).

---

<sup>15</sup> The definition of informativity considered by the author is the average predictability of a word in context. Word informativity is therefore defined and calculated mathematically (cf. Seyfarth 2014: 141, for more details).

## Acknowledgements

I gratefully thank Professor Sueli Maria Coelho for carefully reading and giving enriching suggestions for the improvement of this paper, as well as Professor Plínio Almeida Barbosa for his keen advice and support with the methodological procedures. My gratitude is extended to Professor Júlia Batista Castilho de Avellar, for her support and recommendation of references on Latin language. All remaining errors are my own responsibility. I also thank the two anonymous reviewers who greatly contributed to improving the text and posed important questions that can be addressed in future studies. Finally, I thank the organizers for the invitation to publish in this volume.

## Abbreviations

1	1st person
2	2nd person
3	3rd person
ACC	accusative
DAT	dative
F	feminine
GER	gerund
IND	indicative
INF	infinitive
IPFV	imperfective
M	masculine
NOM	nominative
PL	plural
PRCP	participle
PROG	progressive
PRS	present
PST	past
SBJV	subjunctive
SG	singular

## REFERENCES

1. Aalberse SP. Inflectional economy and politeness: Morphology-internal and morphology-external factors in the loss of second person marking in Dutch [dissertation]. Amsterdam: University of Amsterdam; 2009. 266p. Available from <http://dare.uva.nl/document/134127>
2. Almeida NM. Gramática latina: curso único e completo. 29th ed. São Paulo: Saraiva; 2000.
3. Baese-Berk MM, Dilley LC, Henry MJ, Vinke L, Banzina E. Not just a function of function words: Distal speech rate influences perception of prosodically weak syllables. *Atten Percept Psychophys*. 2019;81:571-589. <https://doi.org/10.3758/s13414-018-1626-4>
4. Barbosa PA. Incursões em torno do ritmo da fala. Campinas: FAPESP/Pontes Editores; 2006.
5. Barbosa PA. From syntax to acoustic duration: A dynamical model of speech rhythm production. *Speech Commun*. 2007;49(9):725-742. <https://doi.org/10.1016/j.specom.2007.04.013>
6. Barbosa PA. Semi-automatic and automatic tools for generating prosodic descriptors for prosody research. In Bigi B, Hirst D, editors. *Proceedings of Interspeech satellite event Tools and Resources for the Analysis of Speech Prosody*; 2013 Aug 30; Aix-en-Provence, France. Aix-en-Provence: TRASP; 2013. p. 86-89. Available from <http://www2.lpl-aix.fr/~trasp/Proceedings/19874-trasp2013.pdf>

7. Barbosa PA. Manual de prosódia experimental. Editora da Abralin; 2022. <https://doi.org/10.25189/9788568990230>
8. Barbosa PA, Camargo ZA, Madureira S. Scripts for the acoustic analysis of speech data. In Madureira S, editor. Sonorities: speech, singing and reciting expressivity. São Paulo: Pontificia Universidade Católica de São Paulo; 2016. p. 164-174. Available from <https://www5.pucsp.br/liaac/download/sonoridades-sonorities.pdf>
9. Bastiaanse R, Thompson CK, editors. Perspectives on agrammatism. London: Psychology Press; 2012.
10. Bell A, Brenier JM, Gregory M, Girand C, Jurafsky D. Predictability effects on durations of content and function words in conversational English. *J Mem Lang*. 2007;60(1):92-111. <https://doi.org/10.1016/j.jml.2008.06.003>
11. Benveniste É. Problèmes sémantiques de la reconstruction. *Word*. 1954;10(2-3):251-264. <https://doi.org/10.1080/00437956.1954.11659526>
12. Benveniste É. Dictionary of Indo-European concepts and society. Palmer E, translator. Chicago: Hau Books; 2016.
13. Boersma P, Weenink D. Praat: Doing phonetics by computer [Computer program]. Version 6.4.18. Available from <http://www.praat.org/>
14. Borba FS. Dicionário gramatical de verbos do português contemporâneo do Brasil. 2nd ed. São Paulo: Editora UNESP; 1991.
15. Boye K. Lexical vs grammatical words. In Van Lier E, editor. *The Oxford Handbook of Word Classes*. Oxford: Oxford University Press; 2023. p. 72-84. <https://doi.org/10.1093/oxfordhb/9780198852889.013.4>
16. Brunelle M, Chow D, Nguyễn TNU. Effects of lexical frequency and lexical category on the duration of Vietnamese syllables. In The Scottish Consortium for ICPhS 2015, editor. *Proceedings of 18th International Congress of Phonetic Sciences; 2015 Aug 10-14; Glasgow, UK*. p.1-5. Available from <https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0341.pdf>
17. Bybee J. Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Lang Var Change*. 2002;14(3):261-290. <https://doi.org/10.1017/S0954394502143018>
18. Bybee J. Mechanisms of change in grammaticization: The role of frequency. In Joseph BD, Janda RD, editors. *The Handbook of Historical Linguistics*. Malden: Blackwell; 2003. p. 602-623. <https://doi.org/10.1002/9781405166201.ch19>
19. Bybee J. From usage to grammar: The mind's response to repetition. *Language*. 2006;82(4):711-733. <https://dx.doi.org/10.1353/lan.2006.0186>
20. Bybee J. *Frequency of use and the organization of language*. Oxford: Oxford University Press; 2007. <https://doi.org/10.1093/acprof:oso/9780195301571.001.0001>
21. Bybee J. *Language, usage and cognition*. Cambridge: Cambridge University Press; 2010. <https://doi.org/10.1017/CBO9780511750526>
22. Bybee J. Grammatical and lexical factors in sound change: A usage-based approach. *Lang Var Change*. 2017;29:273-300. <https://doi.org/10.1017/S0954394517000199>
23. Bybee J, Brown EK. The role of constructions in understanding predictability measures and their correspondence to word duration. *Cogn Linguist*. 2024;35(3):377-406. <https://doi.org/10.1515/cog-2023-0077>
24. Bybee J, File-Muriel RJ, Souza RN. Special reduction: a usage-based approach. *Lang Cogn*. 2016;8(3):421-446. <https://doi.org/10.1017/langcog.2016.19>
25. Bybee J, Fleischman S, editors. *Modality in grammar and discourse*. Amsterdam: John Benjamins; 1995. <https://doi.org/10.1075/tsl.32>
26. Bybee J, Perkins R, Pagliuca W. *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago; London: The University of Chicago Press; 1994.
27. Caballero, JA, Vergis N, Jiang X, Pell MD. The sound of im/politeness. *Speech Commun*. 2018;102:39-53. <https://doi.org/10.1016/j.specom.2018.06.004>

28. Coelho SM. Estudo diacrônico do processo de expansão gramatical e lexical dos itens TER, HAVER, SER, ESTAR e IR na língua portuguesa [dissertation]. Belo Horizonte: Universidade Federal de Minas Gerais; 2006. 323p. Available from <https://repositorio.ufmg.br/handle/1843/ALDR-6PGGWS>
29. Coelho SM. Um estudo de redução fônica. In Ramos JM, Coelho SM, editors. Português brasileiro dialetal: Temas gramaticais. Campinas: Mercado de Letras; 2013. p. 43-56.
30. Coelho SM, Paula TF. Estudo do processo de gramaticalização do verbo poder no dialeto mineiro: análise comparativa entre o dialeto belo-horizontino e o dialeto ouro-pretano. *Rev Estud Ling*. 2011;19(1):87-113. <http://dx.doi.org/10.17851/2237-2083.19.1.87-113>
31. Danckaert L. Variation and change in Latin BE-periphrases: Empirical and methodological considerations. In Adams JN, Vincent N, editors. *Early and Late Latin: Continuity or change?* Cambridge: Cambridge University Press; 2016. p. 132-162. <https://doi.org/10.1017/CBO9781316450826.007>
32. Danckaert L. The origins of the Romance analytic passive: evidence from word order. In Mathieu E, Truswell R, editors. *Micro-change and macro-change in diachronic syntax*. Oxford: Oxford University Press; 2017. p. 216-235. <https://doi.org/10.1093/oso/9780198747840.003.0014>
33. Diaz M, McCarthy G. A comparison of brain activity evoked by single content and function words: An fMRI investigation of implicit word processing. *Brain Res*. 2009;1282:38-49. <https://doi.org/10.1016/j.brainres.2009.05.043>
34. Dilley LC, Pitt MA. Altering context speech rate can cause words to appear or disappear. *Psychol Sci*. 2010 ;21(11) :1664-1670. <https://doi.org/10.1177/0956797610384743>
35. Fedorenko E, Ryskin R, Gibson E. Agrammatic output in non-fluent, including Broca's, aphasia as a rational behavior. *Aphasiology*. 2023;37(12):1981-2000. <https://doi.org/10.1080/02687038.2022.2143233>
36. Fernandes F. Dicionário de verbos e regimes. 39th ed. São Paulo: Globo; 1993.
37. Fleischman S. *The future in thought and language: Diachronic evidence from Romance*. Cambridge: Cambridge University Press; 1982.
38. Fruyt M. Grammaticalization in Latin. In Baldi P, Cuzzolin, P, editors. *New perspectives on historical Latin syntax. Volume 4: Complex sentences, grammaticalization, typology*. Berlin: Walter de Gruyter; 2011. p. 661-864. <https://doi.org/10.1515/9783110253412.661>
39. Gahl S. Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language*. 2008;84(3):474-496. <https://doi.org/10.1353/lan.0.0035>
40. Gahl S, Yao Y, Johnson K. Why reduce? Phonological neighborhood density and phonetic reduction in spontaneous speech. *J Mem Lang*. 2012;66(4):789-806. <https://doi.org/10.1016/j.jml.2011.11.006>
41. Gibson E, Futrell R, Piantadosi SP, Dautriche I, Mahowald K, Bergen L, Levy R. How efficiency shapes human language. *Trends Cogn Sci*. 2019;23(5):289-407. <https://doi.org/10.1016/j.tics.2019.02.003>
42. Giomi R. *A functional discourse grammar theory of grammaticalization. Volume 2: Formal change*. Leiden: Brill; 2023. <https://doi.org/10.1163/9789004520585>
43. Goodglass H, Mayer J. Agrammatism in aphasia. *J Speech Hear Disord*. 1958;23(1):99-111. <https://doi.org/10.1044/jshd.2301.99>
44. Harris, CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, et al. Array programming with NumPy. *Nature*. 2020;585:357-362. <https://doi.org/10.1038/s41586-020-2649-2>
45. Harris Z. A theory of language structure. *Am Philos Q*. 1976a;13(4):237-255. Available from <https://www.jstor.org/stable/20009633>
46. Harris Z. On a theory of language. *J Philos*. 1976b;73(10):253-276. <https://doi.org/10.2307/2025530>
47. Harris Z. Grammar on mathematical principles. *J Linguist*. 1978;14(1):1-20. <https://doi.org/10.1017/S0022226700005636>
48. Harris Z. *Language and information*. New York: Columbia University Press; 1988.
49. Harris Z. *A theory of language and information: A mathematical approach*. Oxford: Clarendon Press; 1991. <https://doi.org/10.1093/oso/9780198242246.001.0001>
50. Heine B, Claudi U, Hünnemeyer F. *Grammaticalization: A conceptual framework*. Chicago: University of Chicago Press; 1991.

51. Heine B, Kaltenböck G, Kuteva T, Long H. The rise of discourse markers. Cambridge: Cambridge University Press; 2021. <https://doi.org/10.1017/9781108982856>
52. Heine B, Song K-A. On the grammaticalization of personal pronouns. *J Linguist.* 2011;47(3):587-630. <https://doi.org/10.1017/S0022226711000016>
53. Hopper PJ, Traugott EC. Grammaticalization. 2nd ed. Cambridge: Cambridge University Press; 2003. <https://doi.org/10.1017/CBO9781139165525>
54. Hunter, JD. Matplotlib: A 2D Graphics Environment. *Comput Sci Eng.* 2007;9(3):90-95. <https://doi.org/10.1109/MCSE.2007.55>
55. Kapatsinski V, Easterday S, Bybee J. Vowel reduction: A usage-based perspective. *Ital J Linguist.* 2020;32(1):19-44. <https://doi.org/10.26346/1120-2726-146>
56. Kean M-L. Agrammatism. New York: Academic Press; 1985. <https://doi.org/10.1016/C2013-0-10949-X>
57. Kuryłowicz, J. The evolution of grammatical categories. *Diogenes.* 1965;51:55–71. <https://doi.org/10.1177/039219216501305105>
58. Lee T, Kong APH, Wang H. Duration of content and function words in oral discourse by speakers with fluent aphasia: Preliminary data. *Front Psychol.* 2014;64:39. <https://doi.org/10.3389/conf.fpsyg.2014.64.00039>
59. Lehmann C. Thoughts on grammaticalization. 3rd ed. Berlin: Lang Sci Press; 2015. <https://doi.org/10.5281/zenodo.7520796>
60. Lehmann C. Ten lectures on grammaticalization: An introduction. Leiden: Brill; 2024. <https://doi.org/10.1163/9789004692701>
61. Levshina N. Communicative efficiency: Language structure and use. Cambridge: Cambridge University Press; 2022a. <https://doi.org/10.1017/9781108887809>
62. Levshina N. Frequency, informativity and word length: Insights from typologically diverse corpora. *Entropy.* 2022b;24:280. <https://doi.org/10.3390/e24020280>
63. Liu B, Jin Z, Wang Z, Wu G. Chinese function words grammaticalized from content words: Evidence from ERPs. *J Neurolinguistics.* 2010;23(6):663-675. <https://doi.org/10.1016/j.jneuroling.2010.07.002>
64. Lobato LM. Os verbos auxiliares em português contemporâneo. In Lobato LM, editor. *Análises lingüísticas.* Petrópolis: Vozes; 1975.
65. Lopes GV, Ferreira MP et al. Cantigas medievais galego-portuguesas [online database]. Lisboa: Instituto de Estudos Medievais, FCSH/NOVA; 2011. Available from <https://cantigas.fcsch.unl.pt/>
66. McKinney W. Data structures for statistical computing in Python. In: *Python in Science Conference (SCIPY 2010)*; 2010; Austin, United States of America. p. 51-56. <https://doi.org/10.25080/Majora-92bf1922-00a>
67. Meillet A. L'évolution des formes grammaticales. *Scientia (Riv Scienza).* 1912;12(26-6):130-148.
68. Mohr B, Pulvermüller F, Zaidel E. Lexical decision after left, right and bilateral presentation of function words, content words and non-words: Evidence from interhemispheric interaction. *Neuropsychologia.* 1994;32(1):105-124. [https://doi.org/10.1016/0028-3932\(94\)90073-6](https://doi.org/10.1016/0028-3932(94)90073-6)
69. Narrog H. Modality, subjectivity, and semantic change: A cross-linguistic perspective. Oxford: Oxford University Press; 2012. <https://doi.org/10.1093/acprof:oso/9780199694372.001.0001>
70. Narrog H, Heine B, editors. *The Oxford Handbook of Grammaticalization.* Oxford: Oxford University Press; 2011. <https://doi.org/10.1093/oxfordhb/9780199586783.001.0001>
71. Narrog H, Heine B. *Grammaticalization.* Oxford: Oxford University Press; 2021.
72. Ofuka E, McKeown JD, Waterman MG, Roach PJ. Prosodic cues for rated politeness in Japanese speech. *Speech Commun.* 2000;32(3):199-217. [https://doi.org/10.1016/S0167-6393\(00\)00009-1](https://doi.org/10.1016/S0167-6393(00)00009-1)
73. Pontes E. *Verbos auxiliares em português.* Petrópolis: Vozes; 1973.
74. Raso T, Mello H, editors. *C-ORAL-BRASIL I: Corpus de referência do português brasileiro falado informal.* Belo Horizonte: Editora UFMG; 2012.

75. Ryskin R, Nieuwland MS. Prediction during language comprehension: what is next? *Trends Cogn Sci.* 2023;27(11):1032-1052. <https://doi.org/10.1016/j.tics.2023.08.003>
76. Seyfarth S. Word informativity influences acoustic duration: Effects of contextual predictability on lexical representation. *Cognition.* 2014;133(1):140-155. <https://doi.org/10.1016/j.cognition.2014.06.013>
77. Seyfarth S, Garellek M, Gillingham G, Ackerman F, Malouf R. Acoustic differences in morphologically-distinct homophones. *Lang Cogn Neurosci.* 2017;33(1):32-49. <https://doi.org/10.1080/23273798.2017.1359634>
78. Shi R, Gick B, Kanwischer D, Wilson I. Frequency and category factors in the reduction and assimilation of function words: EPG and acoustic measures. *J Psycholinguist Res.* 2005;34(4):341-364. <https://doi.org/10.1007/s10936-005-6138-4>
79. Silva LFL, Costa JC, Coelho SM. Um estudo de construções com o verbo deixar em fala espontânea do português brasileiro: aspectos gramaticais/discursivos e análise acústica. *Rev Estud Linguagem.* 2022;30(2):743-779. <http://dx.doi.org/10.17851/2237-2083.30.2.743-779>
80. Silva-Corvalán C. Modality and semantic change. In Fisiak J, editor. *Historical semantics – Historical word formation.* Berlin: Mouton; 1985. p. 547-572. <https://doi.org/10.1515/9783110850178.547>
81. Swanson LA, Leonard LB. Duration of function-word vowels in mothers' speech to young children. *J Speech Lang Hear Res.* 1994;37(6):1394-1405. <https://doi.org/10.1044/jslr.3706.1394>
82. Traugott EC, Dasher RB. *Regularity in semantic change.* Cambridge: Cambridge University Press; 2002. <https://doi.org/10.1017/CBO9780511486500>
83. Turner GS, Tjaden K. Acoustic differences between content and function words in amyotrophic lateral sclerosis. *J Speech Lang Hear Res.* 2000;43(3):769-781. <https://doi.org/10.1044/jslhr.4303.769>
84. Van Bergem, DR. Acoustic vowel reduction as a function of sentence accent, word stress, and word class. *Speech Commun.* 1993;12(1):1-23. [https://doi.org/10.1016/0167-6393\(93\)90015-D](https://doi.org/10.1016/0167-6393(93)90015-D)
85. Virtanen, P. et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat Methods.* 2020;17(3):261-272. <https://doi.org/10.1038/s41592-019-0686-2>
86. Waksom, ML. Seaborn: Statistical Data Visualization. *J Open Source Softw.* 2021;6(60):1-4. <https://doi.org/10.21105/joss.03021>
87. Wang H, Liu X, Wang X, Cui G, Xie X, Zhong X. Functional category production and degrees of severity: findings from Chinese agrammatism. *Aphasiology.* 2019;33(10):1227-1247. <https://doi.org/10.1080/02687038.2019.1626081>
88. Zipf GK. Frequency as a determinant of phonetic change. *Harv Stud Class Philol.* 1929;40:1-95. <https://doi.org/10.2307/310585>
89. Zipf GK. *Human behavior and the principle of least effort: An introduction to human ecology.* Cambridge: Addison-Wesley Press; 1949.
90. Zipf GK. *The psycho-biology of language: An introduction to dynamic philology.* Cambridge: The M.I.T. Press; 1968[1935].