

SCHWINDT, Luiz Carlos; DE BONA, Camila. Lexical frequency effects on reduction of final nasal diphthongs in Brazilian Portuguese. *ReVEL*, edição especial n. 14, 2017. [www.revel.inf.br].

LEXICAL FREQUENCY EFFECTS ON REDUCTION OF FINAL NASAL DIPHTHONGS IN BRAZILIAN PORTUGUESE

EFEITOS LEXICAIS SOBRE A REDUÇÃO DE DITONGOS NASAIS FINAIS NO
PORTUGUÊS BRASILEIRO

Luiz Carlos Schwindt¹

Camila De Bona²

schwindt@ufrgs.br

camilabona@ifsul.edu.br

ABSTRACT: In this article, we discuss the role of lexical frequency on the phenomenon of nasal reduction in final unstressed diphthongs [ẽjⁿ] and [ãwⁿ] in spoken Portuguese data from southern Brazil (e.g. homem ~ homi 'man', ontem ~ onti 'yesterday', chegaram ~ chegaru they arrived) departing from the statistical reanalysis of Schwindt and Bopp da Silva's data (2010). Our study showed a role of frequency in non-verbs. However, regarding nouns ending in *gem* (e.g. viagem 'travel', pilotagem 'piloting'), in the context of a palatal consonant before the nasal diphthong, lexical frequency appears to have no determining role, this sequence being a derivational suffix or part of a root. Concerning verbs, lexical frequency showed no role in the application of nasal reduction as well regardless of the verbal tense or the preceding context. These results contribute to the hypothesis defended by Schwindt (2015; 2016), according to which the output of nasal reduction is related to two processes, distinguished by access or not to morphological information.
KEYWORDS: lexical frequency; phonological variation; nasal reduction.

¹ Doctor; Federal University of Rio Grande do Sul (UFRGS), National Council for Scientific and Technological Development.

² Doctoral student; Federal University of Rio Grande do Sul (UFRGS); Coordination for the Improvement of Higher Level Personnel (CAPES); Sul-rio-grandense Federal Institute of Education, Science and Technology (IFSul), campus Novo Hamburgo.

RESUMO: Neste artigo, discutimos o papel da frequência lexical envolvido no fenômeno de redução de nasalidade dos ditongos finais átonos [ẽ̃ⁿ] e [ãw̃ⁿ] no português falado no sul do Brasil (ex. homem ~ homi; ontem ~ onti; chegaram ~ chegaru), a partir da reanálise estatística de dados de Schwindt e Bopp da Silva (2010). Nosso estudo evidenciou algum papel de frequência em não verbos. No que diz respeito aos nomes terminados em *gem* (ex. viagem, pilotagem), contudo, frequência lexical parece não ter papel determinante, sendo essa sequência um sufixo derivacional ou parte da raiz. Em relação aos verbos, frequência lexical também não mostrou papel significativo na aplicação do fenômeno de redução da nasalidade, independentemente do tempo verbal ou do contexto precedente. Esses resultados contribuem para a hipótese defendida por Schwindt (2015, 2016), de acordo com a qual o output da redução da nasalidade está relacionado a dois processos, diferenciados por acesso ou não a informação morfológica.

PALAVRAS-CHAVE: frequência lexical; variação fonológica; redução da nasalidade.

INTRODUCTION

Studies on lexical frequency of words have redimensioned the analysis of many variable phonological phenomena. From the perspective of lexical diffusion (Wang, 1969) or from the perspective of usage-based theories (Bybee 2001, 2002, 2007; Pierrehumbert 2001, 2003), it seems to be the consensus that variation can be a product not only of rules in the Labovian sense but can be also influenced by lexical frequency. Clements (2009) points out that frequency is revealing of what is or is not universal and what is or is not marked in the languages of the world or in a particular language. Processes such as reduction, assimilation, and the regularization of irregular forms can be enhanced by the high frequency of certain words.

In this paper, in addition to the studies conducted by Votre (1978); Guy (1981); Battisti (2002); Bopp da Silva (2005); Schwindt and Bopp da Silva (2010); Schwindt, Bopp da Silva, and Quadros (2012); Schwindt (2012); Cristófaró Silva, Fonseca, and Cantoni (2012, 2013); and Chaves (2017), among others, we intend to contribute to the discussion about the role of lexical frequency on the reduction of final nasal diphthongs (NR) in Brazilian Portuguese (BP).

- (1) a. hom[ẽjⁿ]³ ~ hom[ɪ] ‘man’
 b. órg[ãw̃^o] ~ órg[ɔ] ‘organ’
 c. corr[ẽjⁿ] ~ corr[ɪ] ‘(they) run’
 d. compr[ãw̃^o] ~ compr[ɔ] ‘(they) buy’

NR is a process that targets non-verbs and verbs and usually implies diphthong reduction⁴.

Based on Schwindt (2015, 2016), our basic hypothesis is that lexical frequency can play a role on NR in non-verbs but not in verbs, considering that the latter are subject to the preservation of the morphological exponent (the final nasal is usually a suffix or part of a suffix indicating number-person). We sought to discover elements capable of differentiating between morphological conditioning and lexical frequency (diverging in part from what was disclosed in the neogrammarian hypothesis). Our postulate is that the first conditioner is accessed earlier in the grammar, while the second is more superficial, aligned with extralinguistic factors.

To examine lexical frequency effects on NR, we reanalyzed data from Schwindt and Bopp da Silva (2010) consisting of 144 informants distributed among the 3 states that form the southern region of Brazil, Rio Grande do Sul, Santa Catarina, and Paraná, stratified by age and educational level, totaling 5,332 contexts for the process in the portion of the sample adopted in this work.

The paper is organized as follows. In section 1, we present a brief contextualization of the role of lexical frequency in variable phonological phenomena, retaking the assumptions of lexical diffusion theory, as well as the main ideas of usage-based theories whenever possible in contrast to approaches of generative-based phonological theories. In section 2, we review the main results of previous analyses of NR. In sections 3 and 4, we describe the

³ We are assuming that nasal diphthongs in BP, after assimilation, preserve a nasal resonance (indicated by a superscript symbol). This is what, in our understanding, differentiates Portuguese from English and Spanish, where the nasal consonant is completely preserved in coda position (e.g. *bum*, *con with*), and from French, where it is completely deleted after assimilation (e.g. *avant before*).

⁴ Reduced forms resulting in oral diphthongs like hom[ej] and órg[aw] do not emerge in the language; forms like hom[ɪⁿ] and órg[õ^o], although they occur, are significantly less common than forms like hom[ɪ] and órg[ɔ]. According to Guy (1981: 215–216), this is probably motivated by the ordering of nasal reduction before vowel reduction, which creates the context for an ordinary process of final vowel reduction that only affects oral vowels (*hoje* → *hoj[ɪ]* ‘today’, *ovo* → *ov[ɔ]* ‘egg’).

methodology of our reanalysis of the phenomenon and present results followed by discussion, respectively. In section 5, we present our final remarks.

1. LEXICAL FREQUENCY EFFECTS ON SOUND CHANGE: AN OVERVIEW

For the neogrammarians—a German school of the late 19th century that opposed linguistic naturalism—a sound change affects all words that satisfy the structural conditions governing the implementation of change, that is, all words that show the environment or context expected in a given process. From this, it follows that sound changes are phonetically gradual and lexically abrupt. These sound changes are conditioned linguistically, configuring themselves as phonological rules—understood, therefore, as regular. Irregular changes, when they exist, are explained only on the basis of loan or analogy and have no access to grammatical properties.

Schuchardt (1885), in contrast to the neogrammarian ideas, pioneered the assumption that words with different frequency ranges tend to be affected differently in linguistic change: “rarely-used words drag behind; very frequently used ones hurry ahead. Exceptions to the sound laws are formed in both groups” (p. 58).

Wang (1969)—the main representative of the thesis known as lexical diffusion—proposed that the word is the main unit of change and, as a consequence, sound changes are understood as lexically gradual and phonetically abrupt since certain lexical items will gradually adhere to the changes. Under this approach, the idea is that changes are no longer conditioned by phonological rules but by lexical characteristics as frequency and predictability of occurrence. Factors such as the familiarity of the lexical item, for example, would be among the extra-grammatical aspects responsible for the greater or lesser vulnerability of a word to sound change. In this perspective, a sound change occurs initially in one word and is propagated to others that present a similar phonological structure. There may be words that are not subject themselves to lexical diffusion, just as there is also the possibility that the change reaches all potential words.

Irregularities in this scenario are justifiably expected since sound change need not be generalized.

More recently, another theoretical approach has made important advances in the discussion about the role of lexical frequency on change, the usage-based phonology, advocated, among others, by Bybee (2001, 2002). It is a word-storage model designed from some basic premises. The first one refers to the representation of the linguistic units of the mental lexicon: both linguistic and non-linguistic units are represented in the same way, and this representation is shaped by the experience of each individual.

The reason frequency or repetition plays a role in Grammar formation is that the mind is sensitive to repetition. This is a domain-general principle; that is, it does not apply just to language but to other cognitive domains as well. (Bybee 2007: 8)

According to this perspective, the most frequent words in the everyday vocabulary of certain people are more active and therefore will be accessed more easily. The mind can store words and whole sentences: it is not through phonemes or phonological features that individuals make, for example, morphological generalizations but through the phonetic and semantic association promoted between items stored in the mental lexicon.

The usage-based phonology aims to cover all the subsystems (i.e. phonology, morphology, syntax, semantics), although it does not consider them as representational primitives. Moreover, the model is not restricted to the study of linguistic structures since it takes into account the fact that the use of a language, by including all cognitive processing and social interactions, has a strong impact on its substance. Thus, the representation of a lexical item, basically constituted by phonetics and semantics, can be constantly shaped in relation to its experience of use in different contexts (cf. Bybee 2001).

Frequency is one of the resources that the mental lexicon uses to categorize lexical items and can be divided into two different measures, namely, token and type frequency. The frequency of occurrence refers to how many times a given word or expression occurs in a given linguistic corpus. It is asked, for example, how many times the word *homem* 'man', a type, is registered in a particular database; the number of cases found is the token frequency of that word. Type

frequency may correspond to the frequency of a particular pattern or linguistic structure. We can look in a corpus for the number of words in which, for instance, the sequence [ẽⁿ] ('em') is present (in a noun like *homem* 'man' but also in a verb like *querem* 'they want'). As far as productivity is concerned, type frequency is extremely relevant since if certain patterns tend to be quite frequent, they will be applied to other items that fit in a similar structure.

High-frequency words and phrases grow strong with repetition and loom large, forming looser connections with other items, while low-frequency words and expressions are less prominent but gain stability by conforming to patterns used by other items. General patterns dominate networks where more specific patterns can be overpowered unless represented by high-frequency items. Words that have phonological similarities cluster together; constructions are connected if they have properties in common. Instances of constructions that grow to high-frequency slowly disengage from the more general pattern to become independent constructions. Thus the phonetic and semantic substance of language is ever being shaped by the effects of usage (Bybee 2007: 9).

Regarding the relation between lexical frequency and morphology, Bybee (1985: 118) defines **morphological relation** as the strongest sort of relation between phonological and semantic forms. Thus, in the scope of morphologically complex words, high-frequency items undergo less analysis and are less dependent on their related base words than low-frequency ones.

Philips (2001: 123–124) proposed what she called the Frequency Implementation Hypothesis, according to what “sound changes which require analysis—whether syntactic, morphological, or phonological—during their implementation affect the least frequent words first; others affect the most frequent words first.” Moreover, she recognized word class as an independent factor in sound change that takes precedence over word frequency: function words, since they characteristically receive low sentence stress, show higher frequencies, and usually change first in contexts of reductions or assimilations. In the domain of word class, sound changes that demand more detailed analysis (like neighborhood density effects, morphological make-up, and so on) target the least frequent words first.

One of the main goals of this paper is to discuss how lexical frequency and word class interact in NR, since all studies found different rates of application to verbs and non-verbs to this phenomenon.

As stated by Guy (2014), it is not trivial to reconcile rule-based and usage-based approaches to phonological variation. While the former seek for abstraction, regularity, and generalization, the latter seek for the lexically differentiated phonological phenomena, including variability, gradience, and probabilistic properties. The author points out that the most traditional approaches to phonology have been, in fact, rule-based approaches with the postulation of abstract mental representations of words, which are submitted to phonological operations capable of capturing the patterns of the most generalizable sounds of a given language.

According to Guy (2004), neogrammarian regularity is well instantiated in the historical record since probably most changes do not leave historical residues of unchanged segments in exceptional words. For the author, the rule-based model developed by the neogrammarians also presents a satisfactory explanatory adequacy considering that it predicts productivity; that is, it predicts the ability of the speaker to know how to pronounce neologisms and lexical loans—for which there are no previous models of pronunciation—besides being able to perform abstract phonological operations on certain classes of sounds throughout the lexicon. However, despite achieving greater explanatory adequacy, Guy reckons the rule-based model is not fully equipped to account for some types of phonological facts, especially those involving specific lexical items. In this sense, he stresses the contribution of the Usage-based Phonology (Bybee 2001, 2002) and the Exemplar Theory (Pierrehumbert 2001). We have already addressed some of the Usage-based Phonology. Exemplar Theory particularly states that speakers retain copies in memory of the words they have already heard with high phonetic detail. This alleged cloud of copies provides the speaker with information on the pronunciation of individual lexical items, the details of phonetic achievement, the apparent patterns of variation in the community, and the quantitative distribution of these facts.

Guy (2014), criticizing the fact that generalizations and abstractions are not fully captured by Exemplar Theory, suggests that the ideal alternative, which incorporates both generative productivity and quantitative precision, is still the variable rule (VR) model of Labovian sociolinguistics. The author admits, however, that lexical effects are not satisfactorily contemplated in the VR model,

which sometimes needs to postulate the existence of two or more underlying forms to account for what it considers **lexical exceptions**. He argues that phonological abstraction should not be abandoned simply because its traditional implementation does not include quantification: an adequate phonological theory needs both grammar and memory to account for all facts.

In this perspective, Walker (2012), following the debate proposed by Myers and Guy (1997), investigated formal, functional, and frequency effects on the process of t/d-deletion in the Canadian English spoken in Toronto. Although the initial results suggested a clear correlation between lexical frequency and deletion, only the factor groups concerning phonology and morphology were significant. By noting that formal constraints have the greatest effect on variation with phonological constraints as primary and morphological constraints as secondary, the author concludes that the usage-based hypothesis, which predicts that lexical frequency influences t/d-deletion, does not present consistent support after considering the interaction with morphological status and with a small group of lexical items. Walker argues we must recognize that frequency does not operate monotonically and has a more dynamic interaction with the lexicon.

Not intending to exhaust this debate, which involves numerous developments, we assume for the purposes of our descriptive analysis the perspective suggested by Guy (2014) and Walker (2012), among others, according to which lexical frequency effects may be associated with the VR approach, reconciling quantification and abstraction.

2. STUDIES ON NASAL REDUCTION IN BRAZILIAN PORTUGUESE

In this section, we briefly summarize some of the main approaches to NR in Brazilian Portuguese, focusing on the role of word class on account of its possible relation to frequency.

The first studies on NR in BP were by Votre (1978) and Guy (1981).

Votre (1978) observed a preponderant role of linguistic variables over social variables in his analysis of the phenomenon. Moreover, considering the reduction extends to all categories present in each of the analyzed variables, he classified NR as a change in progress. In Votre's analysis, past tense verbs and monomorphemic nouns showed positive effects in preserving the final nasal segment, differently from non-past verbs. Nonetheless, since the author does not consider non-past as a natural class, there is no conclusion in his work on the role of morphology as a conditioner of NR, remaining dependent on more detailed investigations.

Guy (1981) studied the phenomenon in the speech of young and adult informants from the city of Rio de Janeiro in the literacy process. As far as the linguistic variables are concerned, the author noticed close rates of application of NR for verbs and non-verbs and a significant influence of palatal consonants in the previous context of nasal diphthongs. For this reason, words ending in *gem* also showed high rates of application in his study.⁵

Departing from Guy's findings, Battisti (2002) analyzed data from the three capitals of the southern Brazilian states, cities that make up the VARSUL database.⁶ Word class was selected as the main conditioner of NR in the statistical analysis. The author separated verbs, nouns in general, and nouns ending in *gem*. Words with this ending may be subject to denasalization for different reasons, including its instability in the history of the Iberian languages (e.g. *passaje*

⁵ Guy (1981) included in his work data of unequivocal nasal reduction in verbs, such as *compram* ~ *compru* 'they buy', and also cases in which the verbal form is replaced by another form of the paradigm, like *vieram* ~ *veio* 'they came'. This may explain why he did not find a difference between verbs and non-verbs in his study. We interpret, however, the latter case—common in BP among people with low schooling, such as Guy's informants—to be only partially related to the first since, despite also involving variable verbal agreement, it does not involve nasal reduction (cf. Schwindt 2015, 2016).

⁶ Project *Variacão Linguística na Região Sul do Brasil*: <http://www.varsul.org.br>.

‘passage’ and *viaje* ‘travel’ in Spanish but *passagem* and *viagem* in Portuguese), which resulted in double forms in the standard Portuguese, like *garagem/garage*. In Battisti’s study, in fact, nouns ending in *gem* led the application of NR followed respectively by non-verbs in general and verbs. She further observed a significant role of dorsal consonants in the previous context (where palatals are included).

Schwindt and Bopp da Silva (2010) analyzed all cities that form the VARSUL databank, including data already studied by Battisti (2002) and Bopp da Silva (2005) in addition to localities not yet analyzed.⁷ The data were normalized in such a way as to account for some new variables. The social constraints analyzed in this study were geographic location (i.e. city), age, and schooling. The linguistic ones were morphological category, preceding context, stress of the following word, and following context. Regarding the morphological context, the authors observed that NR is favored in non-verbs, especially those ending in *gem*; verbs, in principle, inhibit the process. Frequency effects were tested within the sample by removing the most frequent non-verbs from the statistical analysis (homem ‘man’, jovem ‘young’, viagem ‘travel’, origem ‘origin’). Nonetheless, the results remained proportionally unaltered. In regard to verbs, they were reclassified in their different tenses (i.e. indicative present, indicative perfect past, indicative imperfect past, subjunctive present, subjunctive past, subjunctive future, and inflected infinitive). The idea was to verify if forms that become identical to non-reduced forms after reduction (e.g. eles fal[ʊ]/eu falo ‘they speak/I speak’) would be more likely to be deleted. This rearrangement promoted the indicative simple past (a non-neutralizable tense) to a favorable context for NR, like non-verbs.

Schwindt, Bopp da Silva, and Quadros (2012) addressed NR specifically from the perspective of morphological conditioning. Their results corroborate those of previous studies: the phenomenon showed higher rates in non-verbal forms, and nouns ended in *gem* are presented as more subject to the process, regardless of whether *gem* is part of the root (e.g. vantagem ‘advantage’) or suffix

⁷ The cities studied by Schwindt and Bopp da Silva (2010) are Porto Alegre, São Borja, Panambi, and Flores da Cunha (RS); Florianópolis, Blumenau, Chapecó, and Lages (SC); and Curitiba, Londrina, Pato Branco, and Irati (PR).

(e.g. engrenagem ‘gear’). This is explained by Schwindt (2012), who noticed that palatal consonants in the previous context favor NR both in non-verbs and verbs (e.g. ga[n]am ~ ga[nʊ] ‘they win’, dei[j]em ~ dei[ji] ‘that they leave’). The analysis of verbs classified in verbal tenses confirmed the predominance of NR in the indicative perfect past. A possible explanation, according to the authors, is that this tense is the only one in which the nasal consonant does not carry the morphological information of number-person by itself (e.g. chega+∅+ram/chega+∅+m ‘they arrived/they arrive’), as suggested by Mateus and D’Andrade (2000). Under this analysis, the deletion is avoided in order to asseverate overt morphology.

Cristófaró Silva, Fonseca, and Cantoni (2012) studied the nasal diphthong [ãw]⁸ in verb-final position in a sample of 12 informants born and living in Belo Horizonte, southeast Brazil, as they read sentences. The data were acoustically analyzed and discussed from the perspective of what they named Multirepresentational Models (Bybee 1985, 1995, 2001, 2010; Johnson 1997; Pierrehumbert 2001). Lexical frequency was controlled by taking indexes from the ASPA databank, assumed as a reference corpus.⁹ The results showed similar rates of NR for the three tenses analyzed—indicative present, indicative past perfect, and indicative past imperfect. Type frequency was also similar in these three tenses. Token frequency, however, was considerably higher in the present, respectively followed by the past perfect and the past imperfect. Despite the regular distribution of the phenomenon in the studied tenses, among present and perfect past verbs, the most frequent verbs have proven to be more likely to undergo NR. In addition to admitting that NR is a process of phonetic weakening in word-final position, Cristófaró Silva, Fonseca, and Cantoni (2012) defended the hypothesis of analogical leveling between present and past imperfect forms, which converge to a centralized /a/, and between forms of the second and third person plural, which converge to a back vowel, /u/.

Chaves (2017) addressed the correlation between two variable phenomena in Brazilian Portuguese spoken data from the non-urban community in Costa da Lagoa (Florianópolis – Santa Catarina, Brazil), the reduction/denasalization of

⁸ Transcription assumed by the authors.

⁹ Project *Avaliação Sonora do Português Atual*: <http://www.projetoaspa.org>.

final unstressed nasal diphthongs and the explicit marking of verbal agreement for the third person plural. Her main objective was to discuss how these two phenomena were embedded in the linguistic and social concomitants matrix of the community under study. When analyzing the effect of lexical frequency on the explicit marking of verbal agreement in the third person plural, Chaves observed that less frequent verbs favored explicit agreement, while the most frequent ones showed an opposite tendency.

3. METHOD

Inspired by the discussion held by the works presented in the previous section, we propose in this paper an analysis of NR in southern Brazil considering lexical frequency effects with a particular focus on its relation to morphological conditioning. For this, we reanalyzed data of non-verbs and verbs closed by the diphthongs [ẽ̃ⁿ] and [ã̃ⁿ] from the sample studied by Schwindt and Bopp da Silva (2010).

Two core questions characterize the exercise we conducted here:

- (i) Are more frequent or less frequent words more likely to undergo NR?
- (ii) Does lexical frequency interact with morphological category?

The data were reclassified to include a group factor relative to three frequency levels, built from a logarithmic scale, taking indexes from the ASPA corpus as sources. The sample consisted of 283 types and 1,515 tokens for [ẽ̃ⁿ] and 694 types and 3,817 tokens for [ã̃ⁿ]. Since there was only one non-verb that ended in the diphthong [ã̃ⁿ], we restricted the analysis of this class to the diphthong [ẽ̃ⁿ]. For verbs, however, both diphthongs were analyzed. The statistical analyses were computed by GOLDVARB X (Sankoff, Tagliamonte, and Smith 2005).

4. RESULTS AND DISCUSSION

Like several other variable phonological phenomena in BP characterized as change from below (Labov 1966), NR has a moderate application in southern Brazil, 34% in our sample.

According to different studies, word class has a crucial role in the process, which distinguishes between non-verbs and verbs.

Word class	%	N
Nonverb ontem 'yesterday'	56	426/826
Verb pedem 'they ask'	32	2678/8487

Table 1: Word class on NR in southern Brazil (adapted from Schwindt and Bopp da Silva, 2010)

Following this section, we bring our analysis on the role of lexical frequency, focusing especially on the role of the word class. This analysis, as we mentioned, uses frequency from the ASPA project, considered as a reference corpus. 98% of the words of our sample have been found in the ASPA corpus. In addition, among non-verbs and verbs, 8 of the 20 most frequent words in ASPA are also among the 20 most frequent in our sample. In the scope of only non-verbs, this number rises to 12; among verbs, it falls to 8. This contributes to the belief that these two corpora are able to be compared.

Word class	Common words
non-verbs and verbs	foram 'they went', homem 'man', eram 'they were', estavam 'they were', viagem 'trip', querem 'they want', dizem 'they say', fizeram 'they did'
non-verbs	reportagem 'report', homem 'man', ordem 'order', viagem 'trip', origem 'origin', jovem 'young', vantagem 'advantage', passagem 'passage', homenagem 'tribute', margem 'margin', mensagem 'message', coragem 'courage'
verbs	foram 'they went', eram 'they were', estavam 'they were', querem 'they want', dizem 'they say', fizeram 'they did', tinham 'they had', ficam 'they stay'

Figure 1: Common items subject to NR in ASPA/VARSUL among the 20 most frequent words

Considering the analysis of the whole sample in the three levels of lexical frequency, there is no important difference between the levels as we can see in Table 2.

Word frequency	%	N
High viagem ‘travel’, estavam ‘they were’	32,5	719/2214
Medium garagem ‘garage’, mataram ‘they killed’	31,9	816/2554
Low nuvem ‘cloud’, ajoelharam ‘they knelt’	29,3	165/564

Table 2: Lexical frequency on NR – verbs and non-verbs

In the case of non-verbs, however, there is prevalence of NR in the context of high-frequency words.

Word frequency	%	N
High homem ‘man’, origem ‘origin’	64,3	182/283
Medium ordem ‘order’, chantagem ‘blackmail’	50,0	65/130
Low nuvem ‘cloud’, ladroagem ‘thievery’	43,8	14/32

Table 3: Lexical frequency on NR– non-verbs

The same pattern is not observed in verbs, where NR shows a relatively balanced role for the three frequency levels with some prominence in the context of medium-frequency words.

Word frequency	%	N
High eram ‘they were’, dizem ‘they say’	27,8	537/1931
Medium pensam ‘they think’, bebem ‘they drink’	31,0	751/2424
Low arrumaram ‘they arranged’, tecem ‘they weave’	28,4	151/532

Table 4: Lexical frequency on NR – verbs

So far, we can conclude that the role of lexical frequency in NR is limited to certain word classes. As we expected, the analysis suggests that non-verbs are more sensitive to frequency than verbs. Yet subgroups of these classes deserve to be separately analyzed for lexical frequency in order to check whether it plays any role within these word classes. This is the case of nouns closed by the suffix *-gem* versus nouns in which *gem* is part of the root and other non-suffixed words. Regarding verbs, different tenses can be examined separately.

As can be seen in Table 5, lexical frequency appears to have no determining role on suffixed non-verbs, suggesting that the process is, in principle, blind to the internal structure in this word class.

Morphological context	High	Medium	Low
-gem suffix ladroagem ‘thievery’, pilotagem ‘piloting’	61%	72%	65%
gem part of root viagem ‘travel’, origem ‘origin’	58%	65%	--
Other non-suffixed words homem ‘travel’, nuvem ‘cloud’	73%	21%	8%

Table 5: Lexical frequency on NR – suffixed and non-suffixed non-verbs

The question that remains, however, is why, although NR is blind to word structure in non-verbs, the words ending in the sequence *gem* are more prone to NR. According to Schwindt (2012), it is probably due to the phonetic context that precedes the nasal diphthong, particularly because of the effect of palatal consonants, as already noticed by Guy (1981).

Interestingly, lexical frequency is relevant to NR only in contexts other than those in which a palatal consonant precedes the diphthong. This is shown in Tables 6 and 7 below.

Word frequency	%	N
High ontem 'yesterday', homem 'man'	72,8	83/114
Medium ordem 'order', jovem 'young'	20,8	11/53
Low nuvem 'cloud'	8,3	1/12

Table 6: Lexical frequency on NR – preceding context in non-verbs (non-palatal)

Word frequency	%	N
High origem 'origin', linguagem 'language'	58,6	99/169
Medium chantagem 'blackmail', garagem 'garage'	70,1	54/77
Low ladroagem 'thievery', serragem 'sawdust'	65,0	13/20

Table 7: Lexical frequency on NR – palatal in preceding context in non-verbs

Among non-verbs, we only identified variation in the context of palatal consonants preceding the diphthong, since in other contexts we have specific types for each point of articulation (labio-dental: nuvem 'cloud'; nasal labial: homem 'man'; alveolar: ontem 'yesterday', ordem 'order'). In two of these words, NR is clearly influenced by high frequency: *homem*, a noun, and *ontem*, an adverb. In the context of palatals, although we find several different word types, all words end in the phonological sequence *gem* as a derivational suffix or a part of the root.

As pointed out by Guy (1981), the suffix *-gem* presents an irregular behavior in its evolution from Latin, resulting in words with and without a nasal diphthong in Portuguese. Diphthongized forms are always subject to variable reduction. This, in addition to phonetic conditioning, is probably responsible for the high application of the phenomenon, but it does not allow us to say with

certainty whether this application is due to a restructuring of the underlying form of the word or to variation itself.

Regarding the different verb tenses, there appears to be no remarkable role for lexical frequency on the application of NR. In the past indicative, the only tense that favors the process according to Schwindt, Bopp da Silva, and Quadros (2012), frequency rates are evenly distributed among the three levels. With lower rates of NR, the same pattern is found in the other tenses if we analyze them conjointly.

Verbal tense	High	Medium	Low
Past Perfect – Indicative Mood foram ‘they went’	37%	43%	45%
Other tenses comem ‘they eat’	24%	26%	24%

Table 8: Lexical frequency on NR – verb tenses

The results shown in Table 8 do not support the idea that memory is more responsible for the lower application of NR in verbs than the word structure. This is true even for the higher application in the sub-group of the past indicative forms. These results differ in some extent from those found in Cristófaró Silva, Fonseca, and Cantoni (2012) and in Chaves (2017).

As mentioned, Schwindt (2012) found that palatal consonants in a context preceding nasal diphthongs favor NR as in non-verbs. This favoring, however, does not appear to be strongly related to lexical frequency, which is shown in Table 9 below.

Preceding context	High	Medium	Low
Palatal beijam ‘they kiss’, tinham ‘they had’	25%	31%	14%
Other preceding contexts chamam ‘they call’	28%	31%	29%

Table 9: Lexical frequency on NR – preceding context in verbs

With regard to social factors, Schwindt and Bopp da Silva (2010) showed that NR is conditioned by age, with a slight preference for application among younger and less-educated informants.

In the domain of verbs, no correlation between social factors and lexical frequency has been observed. Concerning non-verbs, on the other hand, the conjoined analysis of age and schooling in contrast to lexical frequency pointed to two remarks: (i) older/lower schooling informants show reduced rates of NR in low-frequency words, signaling for the preservation of the diphthong in rarer forms in the language; (ii) younger/higher schooling informants seem to ignore frequency with respect to NR, revealing a typical variable phonological process, which is applied across the board. A more accurate analysis of the relationship between social variables and lexical frequency requires, in our understanding, to examine the individual behavior of the informants, which exceeds the limits of this paper.

5. FINAL REMARKS

In this paper, we proposed an analysis of NR in southern Brazil considering lexical frequency effects with a particular focus on its relation to morphological conditioning. For this, we reanalyzed data of non-verbs and verbs closed by the diphthongs [êjⁿ] and [ãwⁿ] from the sample studied by Schwindt and Bopp da Silva (2010). Two core questions characterized the exercise we conducted here:

- (i) Are more frequent or less frequent words more likely to undergo NR?
- (ii) Does lexical frequency interact with morphological category?

Considering the analysis of lexical frequency in combination with verbs and non-verbs, there is no significant difference between the three levels investigated. Nonetheless, if we look separately at non-verbs, there is prevalence of NR in the context of high-frequency words. The same pattern, however, is not observed in verbs, where NR shows a relatively balanced role for the three frequency levels.

The analysis of subgroups within the class of non-verbs shows that lexical frequency is relevant to NR only in contexts other than those in which a palatal

consonant precedes the diphthong, though the restrictiveness of types in each phonetic context suggest a need for more accurate research. With regard to nouns ended in *gem*, this sequence being a derivational suffix or being part of the root, the balanced rates for the three frequency levels show that NR is, in principle, blind to the internal structure in this word class. Considering the variability between diphthongized and non-diphthongized forms (*gem/ge*) in diachrony and in the current Portuguese lexicon, further investigation is needed to verify whether the high application rates of NR in these words is due to a restructuring of their underlying forms or to variation itself.

Regarding the different verb tenses, there appears to be no remarkable role for lexical frequency on the application of NR even in the past indicative, the only sub-group of verbs reported to be favorable to the process. The preceding context, including palatal consonants, also does not interact with lexical frequency in verbs.

In the domain of social factors and lexical frequency, no correlation has been observed in verbs. Concerning non-verbs, on the other hand, the conjoined analysis of age and schooling in contrast to lexical frequency showed that this may be a typical variable phonological process, which is applied across the board. A more accurate analysis of the relationship between social variables and lexical frequency requires the analysis of individual behavior.

We can affirm that lexical frequency contributes to the distinction between verbs and non-verbs with respect to NR, although the examination of subclasses of non-verbs deserves a more accurate analysis. This finding contributes to the hypothesis defended by Schwindt (2015, 2016), according to which NR corresponds to two distinct processes that converge to the same output. The first process, which affects verbs, interacts with morphology, preserving person-number marks, and is not conditioned by lexical frequency or social factors; the second, which reaches predominantly non-verbs, is phonetic and is subject to extralinguistic conditioning.

REFERENCES

- BATTISTI, Elisa. A redução dos ditongos nasais átonos. In: BISOL, Leda; BRESCANCINI, Cláudia. *Fonologia e Variação: recortes do português brasileiro*. Porto Alegre: EDIPUCRS, 2002.
- BYBEE, Joan. *Morphology: a study of the relation between meaning and form*. Philadelphia: Benjamins, 1985.
- BYBEE, Joan. Regular morphology and the lexicon. *Language and Cognitive Processes*, 10, 1995.
- BYBEE, Joan. *Phonology and Language Use*. Cambridge: Cambridge University Press, 2001.
- BYBEE, Joan. Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change*, 14, 2002.
- BYBEE, Joan. *Frequency of use and the organization of language*. New York: Oxford University Press, 2007.
- BYBEE, Joan. *Language Use and Cognition*. Cambridge: Cambridge University Press, 2010.
- BOPP DA SILVA, Taís. *A redução da nasalidade em ditongos de sílaba átona em final de vocábulo entre falantes bilíngües e monolíngües do Rio Grande do Sul*. Porto Alegre: Universidade Federal do Rio Grande do Sul (master's dissertation), 2005.
- CHAVES, Raquel. *A redução/desnasalização de ditongos nasais átonos finais e a marcação explícita de CVP6: um estudo de correlação*. Florianópolis: Universidade Federal de Santa Catarina (doctoral dissertation), 2017.
- CLEMENTS, George Nick. The role of features in phonological inventories. In: RAIMY, Eric; CAIRNS, Charles. *Contemporary Views on Architecture and Representations in Phonology*. Cambridge, Ma.: The MIT Press, 2009.
- CRISTÓFARO SILVA, Thaís. FONSECA, Marco Silva; CANTONI, Maria. A redução do ditongo [ãw] postônico na morfologia verbal do português brasileiro: uma abordagem baseada no uso. *Letras de Hoje*, v. 47, 2012.
- CRISTÓFARO SILVA, Thaís. FONSECA, Marco Silva; CANTONI, Maria. Conflict in patterns of diffusion in diphthong reduction in Brazilian Portuguese. *Phonological studies*, v. 16. Tokyo: Kaitakusha Publicshing Co. Ltd., 2013.
- GUY, Gregory Riordan. *Linguistic variation in Brazilian Portuguese: aspects of the phonology, syntax and language history*. University of Pennsylvania (doctoral dissertation), 1981.
- GUY, Gregory Riordan. Linking usage and grammar: Generative phonology, exemplar theory and variable rules. *Lingua*, 142, 2014.

- JOHNSON, Keith. Speech perception without speaker normalization: An exemplar model. In: JOHNSON, Keith; MULLENIX, John. *Talker Variability in Speech Processing*. San Diego: Academic Press, 1997.
- LABOV, William. *The social stratification of English in New York City*. Washington DC: Center for Applied Linguistics, 1966.
- MATEUS, Maria Helena; D'Andrade, Ernesto. *The phonology of Portuguese*. Oxford: Oxford University Press, 2000.
- MYERS, James; GUY, Gregory Riordan. Frequency effects in Variable Lexical Phonology. *University of Pennsylvania Working Papers in Linguistics*, 4, 1997.
- PHILLIPS, Betty. Lexical diffusion, lexical frequency, and lexical analysis. In: BYBEE, Joan; HOPPER, Paul. *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins, 2001.
- PIERREHUMBERT, Janet. Exemplar dynamics: Word frequency, lenition, and contrast. In: BYBEE, Joan; HOPPER, Paul. *Frequency and the Emergence of Linguistic Structure*. Amsterdam: John Benjamins, 2001.
- PIERREHUMBERT, Janet. Phonetic diversity, statistical learning, and acquisition of phonology, *Language and Speech*, 46(2-3), 2003.
- SANKOFF, David; TAGLIAMONTE, Sali A.; SMITH, Eric. Goldvarb X. Department of Linguistics, University of Toronto, Toronto, Canada. http://individual.utoronto.ca/tagliamonte/Goldvarb/GV_index.htm, 2005.
- SCHUCHARDT, Hugo. On sound laws: against the Neogrammarians. In: VENNEMANN, Theo; WILBUR, Terence. *The Neogrammarians and the Transformational Theory of Phonological Change*. Frankfurt: Athenäum, 1885 [1972].
- SCHWINDT, Luiz Carlos.; BOPP DA SILVA, Taís. Panorama da redução da nasalidade em ditongos átonos finais no português do sul do Brasil. In: BISOL, Leda; COLLISCHONN, Gisela. *Português do Brasil: variação fonológica*. Porto Alegre: EDIPUCRS, 2010.
- SCHWINDT, Luiz Carlos; BOPP DA SILVA, Taís; QUADROS, Emanuel Souza de. O papel da morfologia na redução da nasalidade em ditongos finais no português do sul do Brasil. In: LEE, Seung Hwa. *Vogais além de Belo Horizonte*. Belo Horizonte: FALE/UFMG, 2012.
- SCHWINDT, Luiz Carlos. Condicionamento morfológico em fenômenos fonológicos variáveis no português brasileiro. *Letras & Letras*. v. 28, n.1, 2012.
- SCHWINDT, Luiz Carlos. Um output, dois processos. *Revista da ABRALIN*, v. 14, n. 1, 2015.
- SCHWINDT, Luiz Carlos. Sobre a preservação de expoentes morfológicos na fonologia variável do português brasileiro. *Domínios da Linguagem*. v. 10, n. 2. Uberlândia: Editora da Universidade Federal de Uberlândia, 2016.
- VOTRE, Sebastião Josue. *Aspectos da variação fonológica na fala do Rio de Janeiro*. Rio de Janeiro: Pontifícia Universidade Católica do Rio de Janeiro (doctoral thesis), 1978.

WALKER, James. Form, function and frequency in phonological variation. *Language Variation and Change*, 24, 2012.

WANG, William S-Y. Competing changes as a cause of residue. *Language*, v. 45, n.1, 1969.