

UNDER COPYRIGHT

The publisher should be contacted for permission to re-use or reprint the material.

**Glide insertion to break a hiatus across words in European
Portuguese: the role of prosodic, geographic and
sociolinguistic factors**

Running head: Glide insertion to break a hiatus

Pedro Oliveira, Marisa Cruz, Nuno Paulino & Marina Vigário

(University of Lisbon, Center of Linguistics, Portugal)

Abstract

Glide insertion to break hiatus has long been reported in the literature to occur in the Northern and Central varieties of Portuguese but this has not been systematically investigated before. From a larger *corpus* of speech material collected in several locations in the North and Center of Portugal, we selected for analysis a sample coming from six locations in the Northwest of Portugal and one location in the Center. The effects of several linguistic and extralinguistic factors were evaluated, namely, prosodic domain, level of prominence in the second vowel, prosodic status of first word (W1), region, task, and age. A binary logistic regression model showed that all factors considered are good predictors of the presence of the glide, some considered in isolation (prosodic status of W1 and task), and

some in interaction with others (prosodic domain*region, prominence level*task, prominence level*age).

Key words: prosodic variation, epenthesis, hiatus resolution, prosodic constraints, extralinguistic factors, European Portuguese

1. Introduction

In this paper we investigate glide insertion to break a hiatus (GI) in varieties of European Portuguese (EP). It has long been noticed that in at least certain areas of the North and Center of Portugal, a [j] is inserted between two adjacent <a a> across words, when the second vowel (V2) bears word-stress (e.g. *a água* ‘the water’, [eá]>[ejá]) (Lopo, 1895; Pereira, 1908; Santos, 1897; Vasconcellos, 1901). By contrast, and like in the standard variety of EP (SEP), in unstressed environments the two non-high-vowels usually undergo vowel merger, and the resulting vowel is low (e.g. *amizade* ‘the friendship’, [eæ]>[a]).¹

¹ A more superficial, phonetic phenomenon operates across EP varieties consisting of the emergence of a transitional glide when V2 is stressed. This glide has the same value for the back feature as the adjacent stressed vowel (as in *a [j] água*, *a [j] Ema*; *a [w] hora*, *a [w] ostra*). We will not be concerned with this phenomenon here, which we believe to be phonologically different and possibly less variable across EP dialects. This transitional glide appears also in the standard variety of EP, and is reported to occur in other regions in very recent work by Brissos (2015), who further notes the existence in specific locations of [j] insertion between non-central vowels or with stressless V2. In the data we collected no such cases were found.

Only very recently, GI across words in varieties of EP has started to be systematically investigated (Oliveira, Paulino, Cruz, & Vigário, 2014; Oliveira, 2016), aiming at answering the following questions: (i) what is the prosodic domain of the phenomenon? (ii) what is its exact geographical distribution and the features of possible intra-dialectal variation?, and (iii) to what extent factors such as speaker age or speech style also condition its occurrence? This paper is a contribution towards the answer to these questions.

The paper is structured as follows. In section 2 we give some background information on dialectal and cross-linguistic variation in segmental phrasal phonology, as well as on the prosodic conditions that have been shown to play a role in segmental phrasal phonology in Portuguese and in other languages. Details on the corpus analyzed here and on data collection and analysis are given in section 3. The results are presented and discussed in sections 4 and 5, respectively. In section 6 we highlight the major conclusions reached and raise a number of questions that remain open for future work.

2. Background

Research on phonological dialectal variation in EP has traditionally focused on word-level segmental variation (Boléo & Silva, 1961; Cintra,

1971; Vasconcellos, [1901]1987). Phrase-level segmental variation, by contrast, has barely been reported until very recently (Aguiar, 2008; Barros, 2014; Cruz, 2013; Cruz, Oliveira, Palma, Neto, & Frota, this volume; Florêncio, 2001; Lopo, 1895; Maia, 1975; Martins & Vitorino, 1989; Paulino & Frota, 2015; Rodrigues, 2003; Segura, 2013; Vasconcellos, 1890-92, 1896, 1901/1987, pp. 87).

Three examples of well established phrase-level segmental phenomena under variation in the European Portuguese space are the realization of fricative consonants in sandhi contexts and two phenomena of insertion of an epenthetic vowel at the right-edge of oxytone words, briefly described below.

Work by Aguiar (2008) and Cruz (2013) has shown that there is variation in the point of articulation of fricatives in resyllabification contexts across words, i.e. when word final fricatives are followed by words starting with a vowel, within the intonational phrase (IP) domain (Frota, 2000). Whereas in SEP the fricative is systematically realized as a [z] across words and phonological phrases (PhP) within IP domain, e.g. ((a[z] *aluna*[z])_{PhP} (*ofereceram caneta*[z])_{PhP} (*ao*[z] *amigos*)_{PhP})_{IP}, in the Northeast of Portugal it surfaces as [ʒ] and there is variation between [z] and [ʒ] in Southern regions ([ʒ] being considerably more frequent in Algarve than in Alentejo).²

² The post-alveolar realization of the fricative in these contexts is also found in other regions, such as Braga, in the Northwest of Portugal (Rodrigues, 2003).

The second example of a phrasal segmental phenomenon is schwa-vowel epenthesis. It has been noticed that a schwa often emerges at the end of oxytone words when they appear at the right edge of an IP, as in (*Quanto ao MiGUEl*[i̯])_{IP}... – Frota, 2000; Vigário, 2003).³ Because this only seems to happen when there is complex tonal material that must be realized within a single tone-bearing unit, i.e. under tone crowding conditions, schwa epenthesis has been considered to be a strategy of text-to-tune accommodation. Importantly, it is well known that there are other strategies to deal with tone crowding in other languages, namely truncation (or deletion) of the tonal material (Grice, 1995; Roseano, Vanrell, & Prieto, 2015). For example, in Bari Italian the final low edge-tones of yes-no questions are deleted when sentences end in a stressed syllable to which L+H* is associated. In recent work, Frota et al. (2015), Frota, Oliveira, Cruz, and Vigário (2015) and Frota et al. (2016) show that there is variation as to whether insertion of text material is preferred, as in Standard EP (SEP) and in the Southern regions of Portugal, or the deletion of tonal material is favored instead, as in Brazil and in some Northern regions of Portugal. For example, in a transition region in the Center of Portugal, Castelo Branco, where both strategies are available, a yes-no question with L*+H associated with the stressed syllable of *angeliCAL* may show either truncation of the final L% or [i̯]-epenthesis (i.e. *angelical*[i̯]), with the realization of the L% on the new syllabic material.

³ Here and elsewhere in this paper, caps signal the syllable that bears word-level stress.

The last example of a phrasal segmental phenomenon that has long been reported to exist only in some Central-Southern regions of Portugal (Alentejo, in particular) is [i]-epenthesis at the end of oxytone words: [pɛ̃i] *pé* ‘foot’, [tʃuminɛ̃i] *chaminé* ‘chimney’ (Boléo & Silva, 1961: p. 101; Florêncio, 2001; Maia, 1975; Vasconcellos, 1890-92, 1896, 1987: p. 87). While there are some similarities in the context of this process of vowel epenthesis and the one mentioned in the previous paragraph, for instance this vowel insertion process also requires that the oxytone word is IP final, Cruz (2013) shows that it does not operate under tone crowding – for example, unlike in SEP it operates at the end of neutral declarative sentences, which are tonally similar to SEP neutral declaratives, bearing the low/falling basic melody (H+)L* L%.

These examples illustrate recent advances in the study of variation in segmental phrasal phonology in Portuguese. In this paper we investigate another segmental phenomenon that has also been noticed in the literature since long, which is GI to break a hiatus. From older reports and from more recent investigation by L. Segura, it seems that [j]-insertion operates not only in the North of Portugal, but also in some Central regions (Segura, 2013). It is nevertheless unknown how prevalent the process is today and if all Northern regions still maintain it equally active, namely across age groups and speech styles.

To the best of our knowledge, the exact phonological conditions for GI have also not been systematically investigated. The presence of two central

vowels where the second one bears word stress seems to be necessary conditions (Segura, 2013). Nevertheless, there are several other phrasal properties that may play a role in segmental prosodic phenomena, which have not been investigated before for this specific phenomenon. We review some of these in the following paragraphs.

Vowel adjacency resolution processes across words have been extensively investigated in SEP, spoken in the region of Lisbon (Frota, 1995, 2000). It has been found that depending on the vowels involved, there are different strategies for the resolution of hiatus, which are usually further conditioned by stress clash configurations. Because of its relevance for the understanding of the rest of this paper, before proceeding we will make a very short digression on prosodic phrasing and prominence in EP, essentially following Frota (2000, 2014) and Vigário (2003, 2010), developed within Nespor and Vogel's (1986/2007) approach to prosodic phonology.

Like in other Romance languages, intonational phrases include all adjacent phonological phrases that are part of a root sentence, while syntactic material not structurally attached, like parentheticals, form their own IP, as in (*As alunas,*)_{IP} (*até onde sabemos,*)_{IP} (*obtiveram boas avaliações*). Besides observing syntactic conditions, the formation of IP also depends on phonological constraints: long phrases tend to be broken into smaller IP, and there is preference for balanced IP or the longest IP is the rightmost. For example, while Subject-Verb-Object (SVO) may form a

single IP, if the subject is long, it will form an IP on its own, as in (*O boliviano mulherengo*)_{IP} (*gravava uma melodia*)_{IP} (Elordieta, Frota, & Vigário 2005). Similarly to other languages, IP formation is also sensitive to speech rate.

Phonological phrases (PhP) contain a lexical head (Lex) and the functional elements to its left within the maximal projection of Lex (Lex^{max}).⁴ When the conditions are such that a single Lex (Lex₁) would form a PhP by itself, Lex₁ is grouped with the PhP that contains the preceding Lex (Lex₀), if in addition Lex₁ is a complement or a modifier of Lex₀. Thus, for example, the noun phrase *os meninos lindos* forms a single PhP, despite the fact that it contains two lexical heads. Adverbs and adjectives at the non-recursive side of the Lex they modify, within Lex^{max}, are included in the same PhP as Lex. Therefore, for instance, the noun phrase (*um muito belo dia*)_{PhP} is a PhP despite including two lexical heads.

Lexical syntactic heads form a Prosodic Word Group (PWG).⁵ The difference between the PWG and the prosodic word becomes apparent when the syntactic terminal node includes an element that is formed of more than one prosodic word, as with certain types of compounds, like *porta-óculos* ‘eyeglass-holder’, or abbreviations, like KLM ‘id’.

⁴ Lexical heads are syntactic heads that belong to open class categories (i.e. Nouns, Verbs, Adjectives and Adverbs, like *carro, comeu, lindo, apenas*). Examples of lexical heads forming a PhP with the preceding functional elements within their maximal projection are the following: (*a tua mãe*)_{PhP}, (*uma destas tardes*)_{PhP}, (*muitas crianças*)_{PhP} (*tinham dito*)_{PhP}.

⁵ The PWG is a constituent of the same level as Nespor and Vogel’s clitic group, but is crucially claimed in Vigário (2010) not to play a special role in clitic attachment.

The prosodic word (PW) usually includes a stem plus any following suffix. Suffixes that constitute an independent domain for word-level stress form their own PW. Each PW bears one and only one word-level stress. Words that do not bear lexical stress are phonologically clitic. Clitics either incorporate into the preceding PW (essentially, postverbal weak pronouns) or they adjoin to the following PW (preverbal weak pronouns and other clitic words). There is phonological evidence of different sorts for all these prosodic domains, including segmental and prominence-related phenomena. Prominence at each level, from the PWG up, is rightmost. Word-stress is lexically assigned and located at the right edge of the PW within a three-syllable window.

Let us now return to the review of the processes involving hiatus in EP. In the case of adjacent sequences of central vowels, in SEP vowels usually undergo vowel merger when unstressed, and the resulting vowel is low – e.g. in *caNEta **anTIga*** ‘old pen’, the mid central vowels marked in bold are usually realized as [ã]. However, vowel merger is blocked if one of the two vowels bears word stress, like in *a caNEta **ÂMbar*** ‘the amber pen’, as experimentally shown in Frota (2000).

Other processes, like vowel deletion, are further conditioned by stress clashing configurations that would arise from the loss of a vowel intervening between the two word stresses. For example, in *o bailaRIno **Ama*** ‘the dancer loves’, V1 cannot be deleted if *ama* ‘loves’ is the head of its phonological phrase, because deletion would result in two adjacent

vowels bearing phonological phrase prominence. Importantly, there is variation as to which stress clashing configurations are tolerated and which are not, both across Portuguese varieties and across Romance languages (e.g. Cabré & Prieto, 2005; Tenani, 2002).

Besides being sensitive to prominence patterns, all these processes are also bound to occur within specific prosodic domains (Nespor & Vogel, 1986/2007; Selkirk, 1984, 1986), usually either spanning the IP or applying at the right edge of the PW within the IP or the PWG – see Frota (1995, 2000) and Vigário (2003, 2010).

GI as a strategy for breaking a hiatus across words is not found in SEP and to the best of our knowledge the prosodic domain within which GI operates, as well as the level of prominence in V2 that actually matters only now started to be investigated.

It is well established that sequences of similar adjacent phonological objects, for example two adjacent vowels, tend to be avoided in the languages of the world. While many languages tend to avoid vowel hiatus, the means for achieving this end vary greatly cross-linguistically. Common processes, across languages, are: deletion ($V_1V_2 > _V_2$ or $V_1 _$, which operates for example in Etsako, a Niger-Congo language) or semivocalization of one of the vowels ($.V_1.V_2. > .GV_2.$, as attested in Igede language, spoken in Nigeria), and vowel merger ($V_1V_2 > V_3$, which is found for example in European Portuguese). Other ways of breaking a hiatus are epenthesis ($V_1V_2 > V_1GV_2$, a process that applies for instance in Washo, a

Native American language) and diphthongization (.V₁.V₂. > .V₁V₂.) (Bergman, 1971; Cabré & Prieto, 2005; Casali, 1997, 2011; Elimelech, 1976; Frota, 2000; Hall, 2011, 2013; Kutsch Lojenga, 1994; Payne, 1981; Paulino & Frota, 2015; Tenani, 2002).

In the Romance area, as far as we know, insertion of a glide to break a hiatus across words is a rare phenomenon, but it has been reported to occur in Galician. In Galician, however, it may be restricted to particular morpho-phonological contexts involving clitics (CL) (Álvarez & Xove, 2002, p. 38; Colina, 1997). CL, and more in general highly frequent (combinations of) words, are known to often trigger specific phonology, in EP and other languages. They are frequently involved in processes of lexicalization of more or less idiosyncratic surfacing realizations (Bybee, 2001; Napoli & Nevis, 1987; Vigário, 2003: 273-324, a.o.). Besides this, word frequency and major word classes (i.e. open vs closed class) have also been found to correlate with the frequency of hiatus resolution in other languages (Alba, 2006; Chitoran & Hualde, 2007). This means that the prosodic category of the words involved, in particular of the first of two words displaying the relevant segmental and prominence requirements, may also play a role in GI.⁶

Bearing this in mind, in this paper we aim at adding to the characterization of the prosodic conditions under which GI takes place, by

⁶ In our research, only the first word may be either a PW or a clitic. The second word always starts with a vowel that bears word stress, and hence the second word is always a PW.

inspecting the possible effects on GI of (i) the prosodic domain within which the central vowels appear, (ii) the prominence level of V2, and (iii) the type of word (W1=CL or W1=PW) the first central vowel belongs to. In addition, external factors that may favor or inhibit the process are also considered, namely the geographic area, speakers' age and speech styles.

3. Method and materials

The materials analyzed here are part of the data collected in the project *Interactive Atlas of the Prosody of Portuguese* (<http://labfon.letras.ulisboa.pt/InAPoP/>, Frota (coord.) 2012-2015). Within this project, data were collected in 40 urban and rural locations within Portugal, including the islands. For the present research, we have selected a small sample of the data, coming from 7 locations, in particular 3 adjacent areas in the Northwest of Portugal and 1 area in the Center of Portugal. The three Northwestern areas included an urban location (U) and a rural location (R): 1. Arcos de Valdevez (ArV, U); Castro Laboreiro (CtL, R); 2. Ermesinde (Erm, U); Gião (Gia, R); 3. Braga (Bra, U); Fiscal (Fis, R). In the Center, we have only selected one rural location, Nisa (Nis, R). The selection of these data points, depicted in Fig. 1, enables us to investigate variation across a contiguous space within an area where the process is expected to operate, including possible differences between rural and urban

areas. Aiming at contributing to the delimitation of the area where GI insertion operates currently, we have considered the Southern area where glide has been reported to occur (Segura, 2013). By selecting this point close to the limits of the geographic area where GI operates, we also expect to be able to see if there are signs of loss of the process, and if so, if some of the linguistic and sociolinguistic variables contribute in any particular way to the pattern of loss observed.



Figure 1. Geographic distribution of the locations where the speech materials analyzed here were collected: Arcos de Valdevez (ArV), Castro Laboreiro (CtL), Ermesinde (Erm), Gião (Gia), Braga (Bra), Fiscal (Fis), Nisa (Nis).

In most major areas, six female speakers were recorded, three between 20 and 45 years-old (20-45) and three above 59 years-old (60+). Speech was collected *in loco*, and was obtained in two different speech styles: controlled (in a reading task), and (semi-)spontaneous (in an interview and in a map task – Anderson et al., 1991, here adapted in order to obtain specific segmental phenomena in various prosodic contexts). Read sentences were produced twice by each speaker. For various reasons, deviations to this general plan may have occurred, for example, the subject was too tired and did not complete all tasks, canceled her participation on the spot or was incorrectly reported in terms of age during the recruitment phase. Table 1⁷ shows the number of subjects that produced the data analyzed here, per location and age group.

Age group	ArV	CtL	Bra	Fis	Erm	Gia	Nis	Total
20-45	3	4	3	3	5	3	2	22
60+	3	2	2	3	2	3	3	18
Total	6	6	5	6	7	6	5	40

Table 1. Number of subjects that produced the data under analysis in each location and age group

Reading task, including repetitions, yielded a total of 1431 potential contexts for insertion. Sentences read with disfluencies were eliminated. In

⁷ Note that there is a gap in the number of speakers that performed all tasks: in Castro Laboreiro and in Ermesinde in the older subjects (60+), only one speaker was engaged in the reading task and did not complete it.

the interview and in the map task, 39 and 91 potential contexts for insertion were obtained, respectively.

The read corpus contains 24 sentences, controlling for prosodic constituency (see examples 1-5), level of prominence in V2 (see examples 6-8), and type of W1 – PW or CL (see examples 9-10).

(1) *V1 V2 inside PW*

Nunca tinha ouvido falar da região de Simaári Cura, ...
never (I-)had listened talk about-the region of Simaári Cura
'I had never heard about Simaári Cura's region...'

(2) *V1 V2 across PW within PWG*

O João era mesmo ganacioso, ultra-ávido de vencer!
the John was really greedy, overeager to win
'John was really greedy, overeager to win'

(3) *V1 V2 across PWG within PhP*

Tive um amigo que montava asas de aviões ultra-leves.
(I-)had a friend who set up wings of airplanes ultra-light
'I had a friend who used to set up wings for ultra-light airplanes.'

(4) *V1 V2 across PhP*

Um amigo meu importava aves raras do Brasil.
a friend of-mine imported birds rare from-the Brazil
'A friend of mine imported rare birds from Brazil'

(5) *V1 V2 across IP*

Quanto à Maria, aulas às oito da manhã nunca lhe agradaram.
as for-the Mary, classes at-the eight of-the morning never her pleased
'As for Mary, classes at eight in the morning never pleased her.'

(6) *V2 nonhead of PWG*

Os políticos deram agora à J(ota).A.E. a tarefa de vigiar a
the politicians gave now to JAE the role of supervising the

sinalização nas auto-estradas Portuguesas.
signs in highways Portuguese
the politicians gave now to JAE the role of supervising the signs in
'Politicians decided to give JAE the role of supervising Portuguese highway signs.'

(7) *V2 head of PWG*

A matrícula do meu novo carro é J(ota)A-18-18.
the licence-plate of-the my new car is JA-18-18
'The license plate of my new car is JA-18-18.'

(8) *V2 PhP head*⁸

Aquele professor também dava aulas aos mais jovens.
that teacher also gave classes to-the more young
'That teacher also taught younger students.'

(9) *WI=CL*

O jornalista perguntou a_{CL} Ana Moura pelo seu cachimbo novo.
the journalist asked to-the Ana Moura for his/her pipe new
'The journalist asked Ana Moura for his/her new pipe.'

(10) *WI=PW*

Este agricultor só plantava a_{PW} árvores baixas.
this farmer only planted trees small
'This farmer only planted small trees.'

A perceptual annotation with spectrographic inspection was performed using Praat 5.2.2 (Boersma & Weenink, 2007). Four tiers were created: orthographic, phonetic, Break Indices (BI), and segmental, according to P-ToBI (Frota et al. 2015), as illustrated in Figure 2.

⁸ Notice that, when PhP is restructured in IP, we considered annotating V2 IP head.

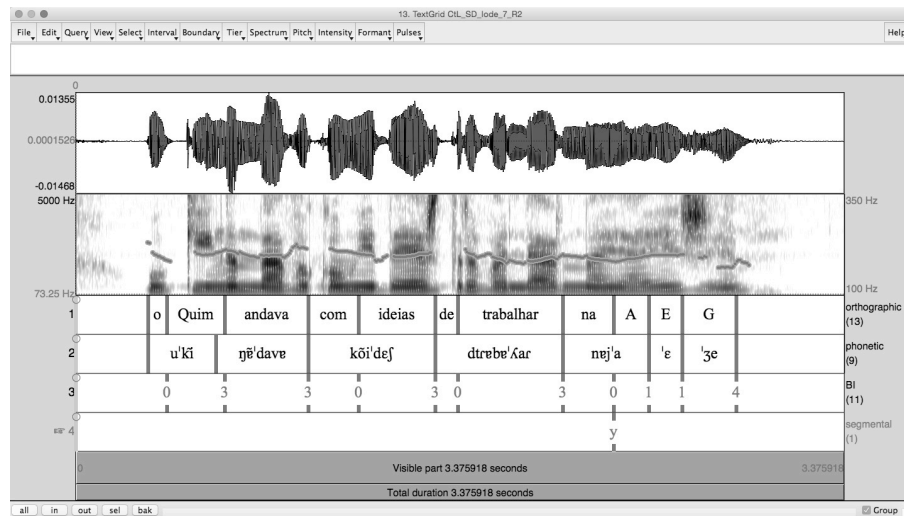


Figure 2: Example of annotation in *Praat* (target: *O Quim andava com ideias de trabalhar na AEG* ‘Quim wanted to work at AEG’; speaker: SD; region: Castro Laboreiro; speech style: reading task)

4. Results

In our data, all regions show instances of GI between central vowels, where the second one (V2) bears word stress. As expected, when V2 is unstressed, glides never appear. However, there is great variation with respect to the rate of insertion. In the following subsections we look at the results taking into consideration the following prosodic factors: (i) prosodic domain; (ii) V2 prominence level, and (iii) type of word (W1=CL or W1=PW) the first central vowel belongs to. In each case, geographic area and speech style are also considered. The effect of age is analyzed in a separate subsection.

4.1 Domain for glide insertion

For the inspection of the *prosodic domain* of GI we consider the data where W1 is a PW (data where W1 is a clitic are considered in subsection 4.3, below).

GI is by no means obligatory in any of the regions, irrespective of prosodic domain considerations (Fig. 3).⁹

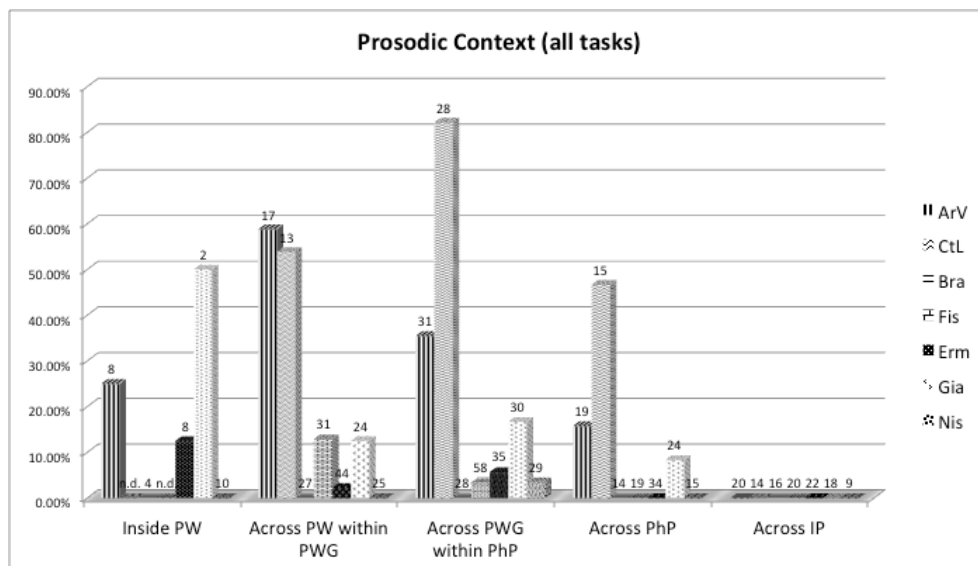


Fig. 3: Percent of glide insertion by prosodic context and region. V1 belongs to PW.

In addition, in all regions where the glide obtains when W1=PW, the IP domain always blocks insertion, unlike the lower domains. In some of the regions, furthermore, insertion is also attested inside PW. In other words,

⁹ In the graphs the number of potential contexts is given at the top of each bar.

the process is not limited to occur across words, as previous reports in the literature suggested, and when it happens across words, both items must be included within the same IP.

There are some regions that clearly insert more, namely Castro Laboreiro and Arcos de Valdevez, and areas where insertion is very rare or inexistent when W1=PW (all the other locations). Whereas in areas that insert more, insertion is found in all levels of prosodic hierarchy that are represented in the data, in the regions that insert less, GI tends to occur in lower domains, namely within, but not across PhP.

The same general observations hold if we consider read speech only (see Fig 4).

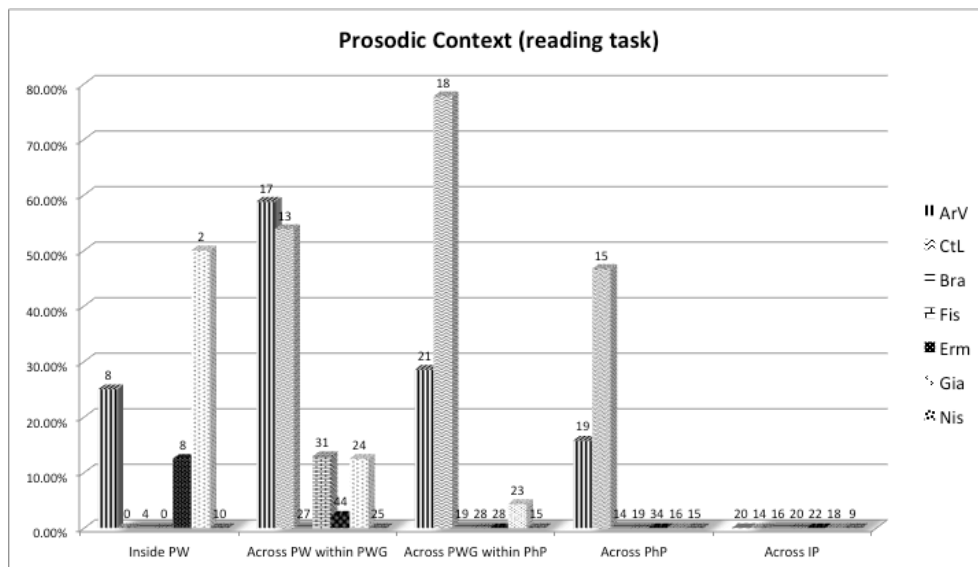


Fig. 4: Percent of glide insertion in the reading task by prosodic context and region. V1 belongs to PW.

In the (semi-)spontaneous tasks very few contexts obtained for GI, and in only two prosodic conditions, namely across PWG within PhP and

across PhP. In the prosodic condition with a relevant amount of data, i.e. across PWG within PhP, the rate of GI is clearly higher in all the regions where the context under observation obtained, in comparison with the same condition in the reading task (see Fig. 5).

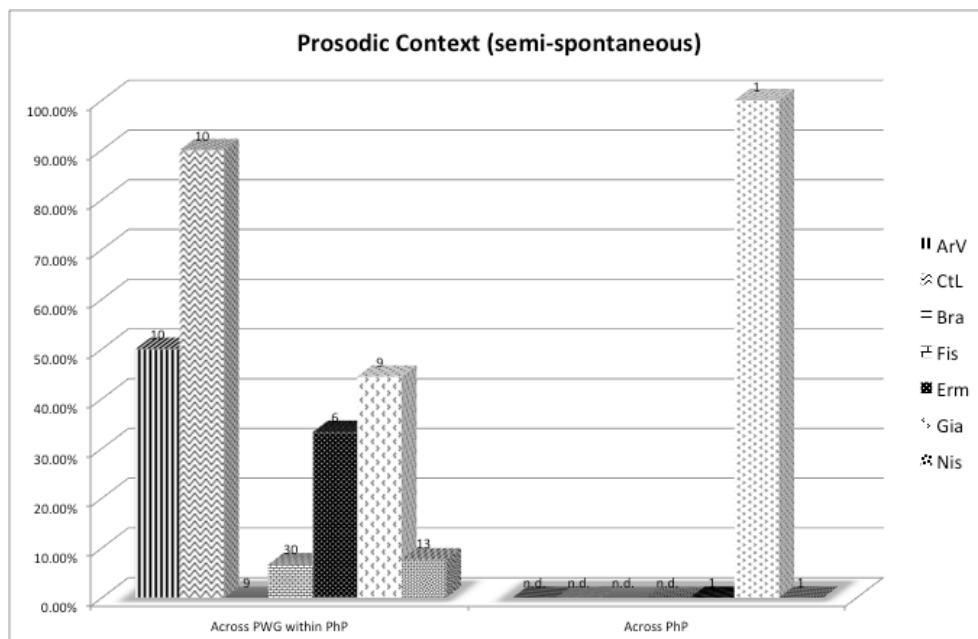


Fig. 5: Percent of glide insertion in the (semi-)spontaneous tasks by prosodic context and region. V1 belongs to PW.

Unlike in read speech, in (semi-)spontaneous tasks Nisa exhibits some amount of insertion, while Braga still lacks insertion, in the type of data considered here, i.e. when W1 is a PW.

4.2. Prominence configuration

In this subsection we will also consider the speech material where W1 corresponds to a PW, looking at each type of task in turn.

The *effect of higher levels of prominence on V2* in the reading task yields variable results across regions (see Fig. 6).

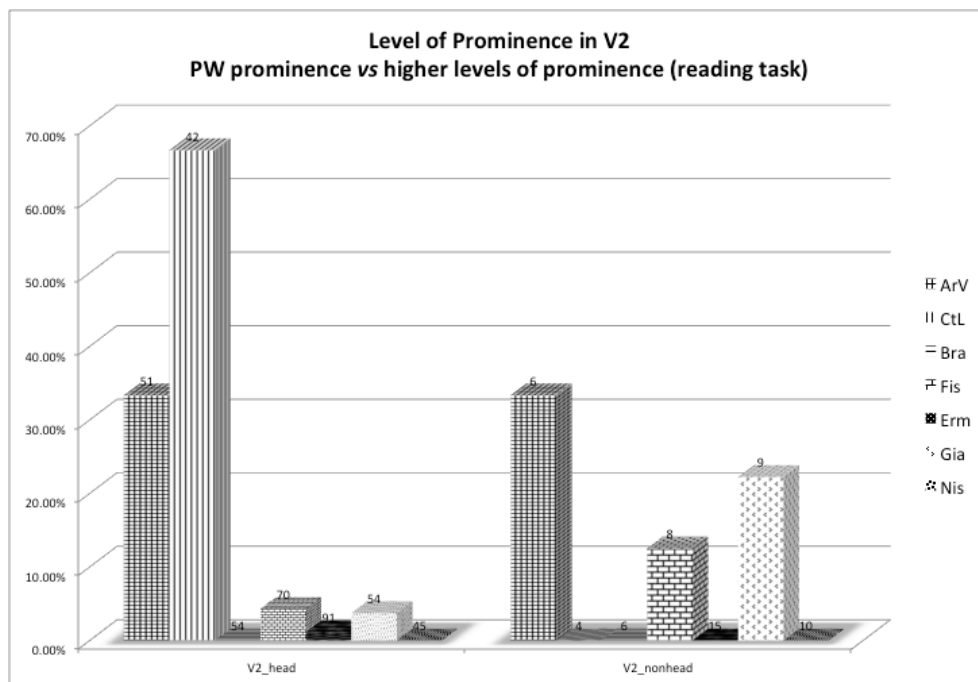


Fig. 6: Percent of glide insertion in the reading task by level of prominence in V2 (phrasal head/non-head) and region. V1 belongs to PW.

While in Castro Laboreiro there is a clear effect of the status of V2 in terms of prominence-level, with very high rates of insertion when V2 bears the prominence of a phrasal domain against no insertion when it is only the nucleus of the stressed syllable of the word, in other regions the reverse happens (Gião and Fiscal) or no effect is found (Arcos de Valdevez).

A different picture emerges if we look at specific levels of prominence. In all of the regions where this context was obtained, namely Arcos de Valdevez, Castro Laboreiro, Ermesinde and Gião, when V2 bears IP prominence GI is obligatory (see Fig. 7).

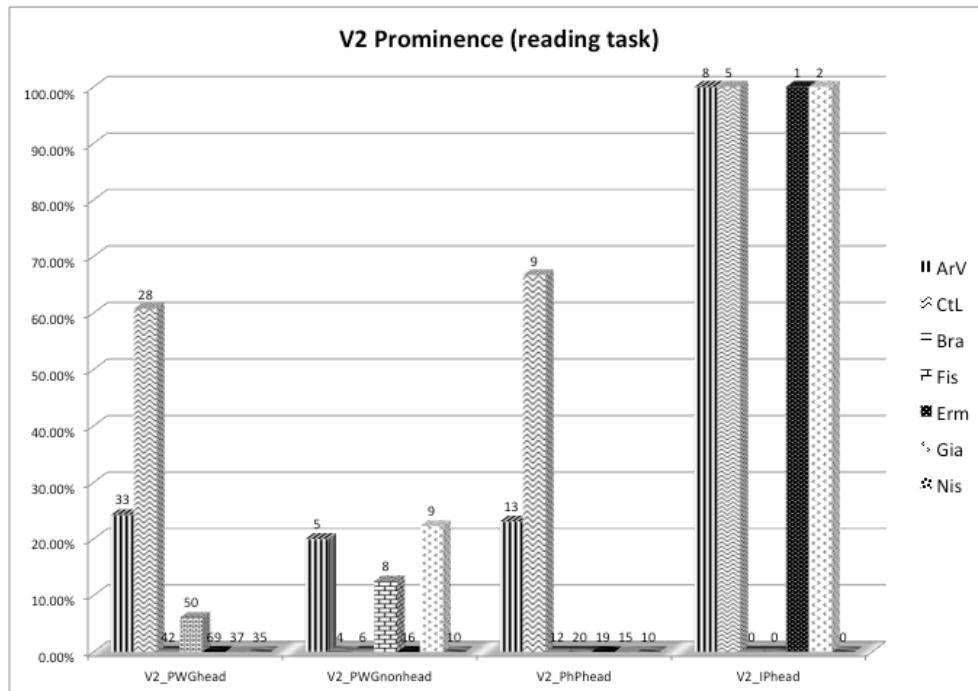


Fig. 7: Percent of glide insertion in the reading task by V2 level of prominence (head/non-head of different phrasal domains) and region. V1 belongs to PW.

This seems to indicate that the higher level of prominence matters. Levels of prominence below the IP, by contrast, only seem to be relevant in Castro Laboreiro.

As shown in Fig. 8, in (semi-)spontaneous speech the higher level of prominence on V2 (PhP) clearly favors insertion (to the exception of Castro Laboreiro, where insertion always occurred, but there was only one context for insertion when V2 is the head of PWG).

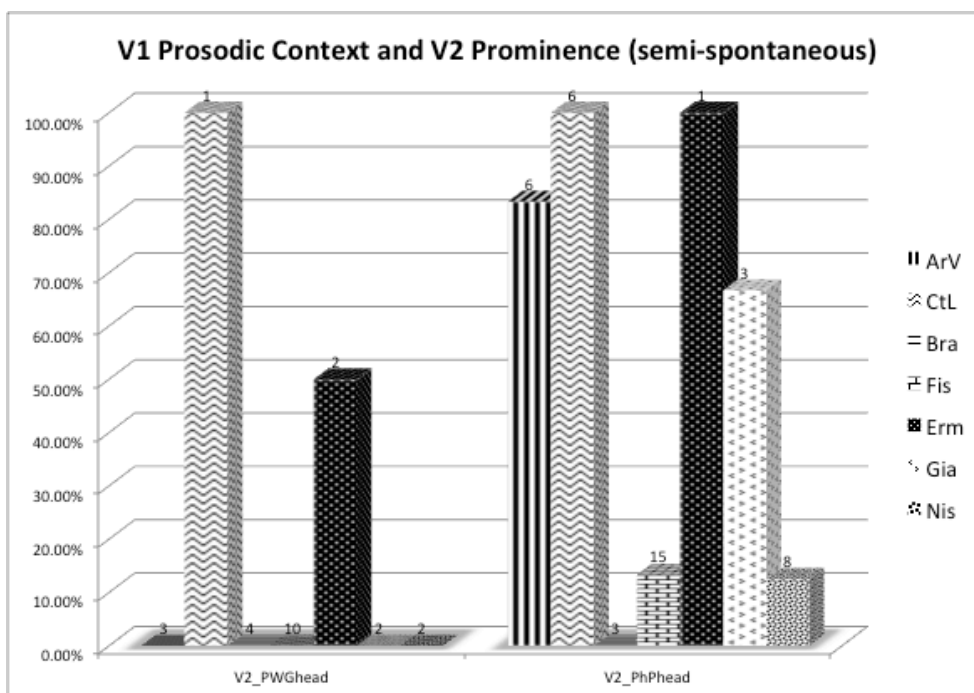


Fig. 8 – Percent of glide insertion in the (semi-)spontaneous tasks by V2 level of prominence (head/non-head of different phrasal domains) and region. V1 belongs to PW.

4.3 Phonological status of Word 1

The *phonological status of W1 (CL or PW)* also seems to play a role in GI. As shown in Fig. 9, for the read speech materials, in some regions GI only occurs when W1 is a CL, namely in Braga, Ermesinde and Nisa, and the rate of GI in the condition W1=CL is always higher than in the condition W1=PW.

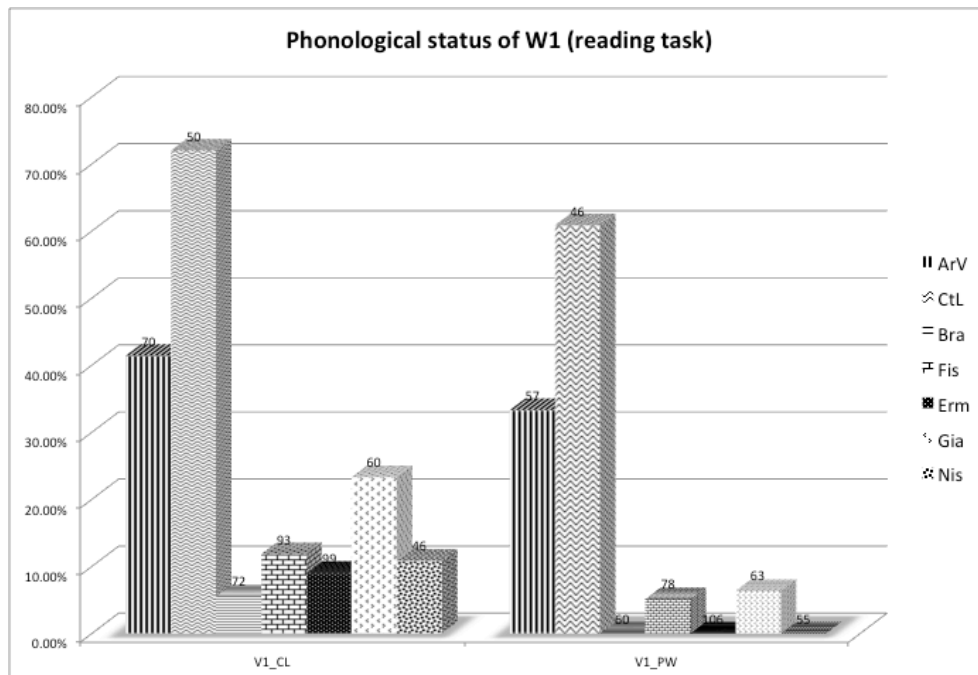


Fig. 9: Percent of glide insertion in the reading task by type of word (PW or CL) and region.

Furthermore, in those locations where insertion occurs with both types of words, the proportion of GI when W1 is a CL is significantly higher than when W1 is a PW ($Z=-2.366, p=.008$).

(Semi-)spontaneous data show the exact same tendency in the regions where the two word categories appear and where there is no ceiling effect (notice that in Ermesinde and Nisa there are no targets with W1=CL and in Castro Laboreiro there is always insertion) (see Fig. 10).

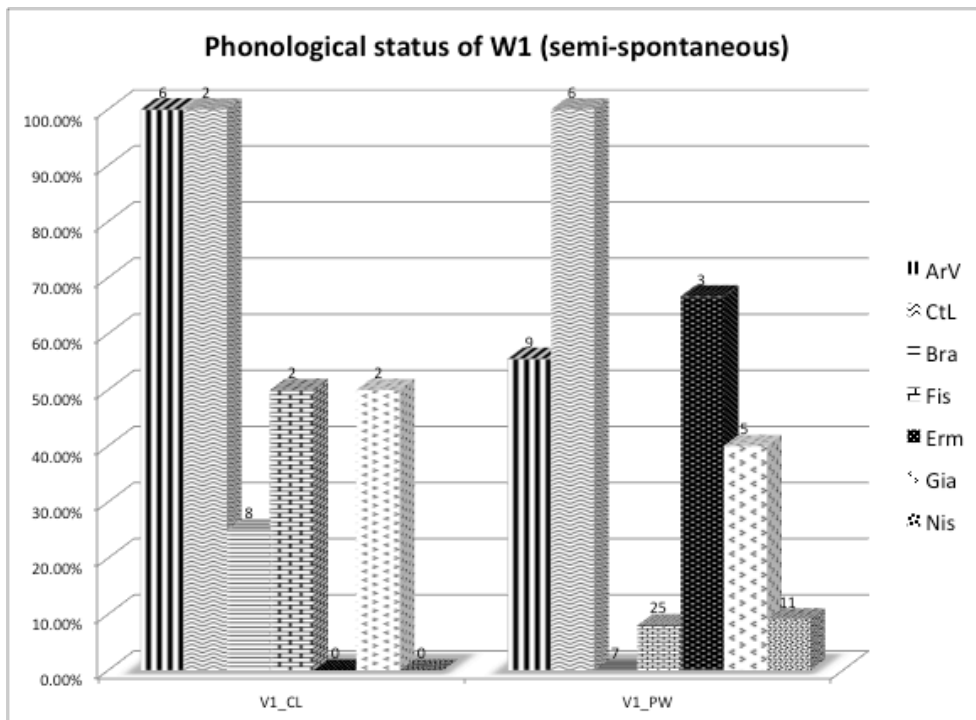


Fig. 10: Percent of glide insertion in the (semi-)spontaneous tasks by type of word (PW or CL) and region.

However, here the difference between insertion with W1=CL and with W1=PW is not significant ($Z=-1.352, p=.219$).

Again, (semi-)spontaneous tasks exhibit more GI than read speech, also in the data where W1 is a CL. Furthermore, in both types of speech, in Braga GI is observed only when W1 is a CL. In general, the data also show a tendency for more insertion in the rural location when compared to the urban location of the same region.

4.4. Age as a factor

For the inspection of the *effect of age* we shall conflate the locations under investigation into two regions: Northwestern (including all the data points but Nisa) and Central (Nisa). This allows us to have a general view of the effect of age, with the major tendencies emerging. The interaction of this variable with the other variables is evaluated in the next subsection. We first analyze the general data, and then consider the differences across speech style.

The data in Fig. 11 show a clear difference in the rate of insertion by the older and younger subjects, in both regions. Globally, considering the data from both regions, a significant difference is observed between the younger and the older age groups in terms of rate of glide insertion ($U=52.00, p=.033$).

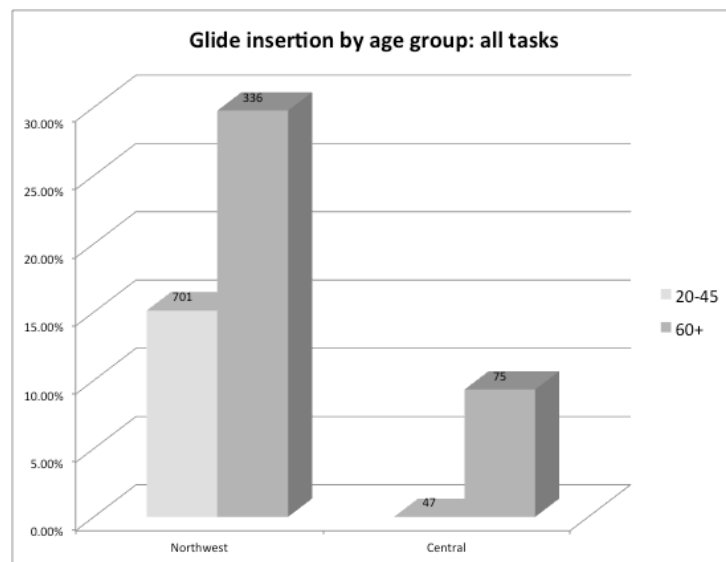


Fig. 11 – Percent of glide insertion by age group in the Northwestern and the Central regions.

Figures 12 and 13 below show that both speech styles contribute equally to the general tendency.

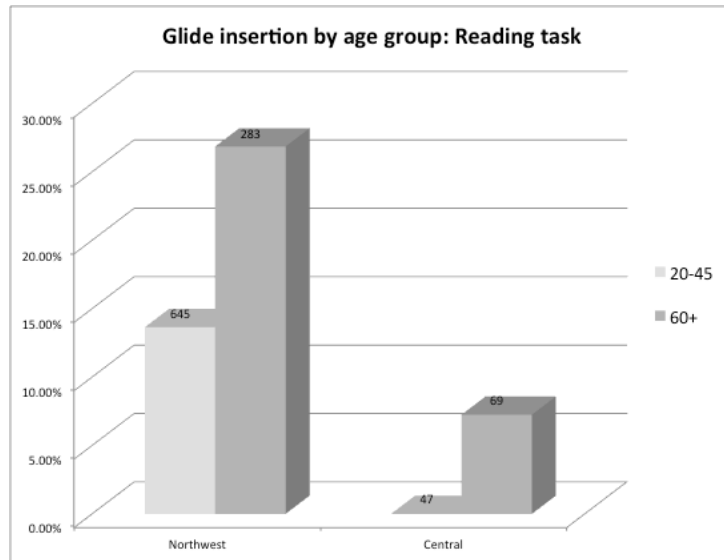


Fig. 12: Percent of glide insertion in the reading task by age group in the Northwestern and the Central regions.

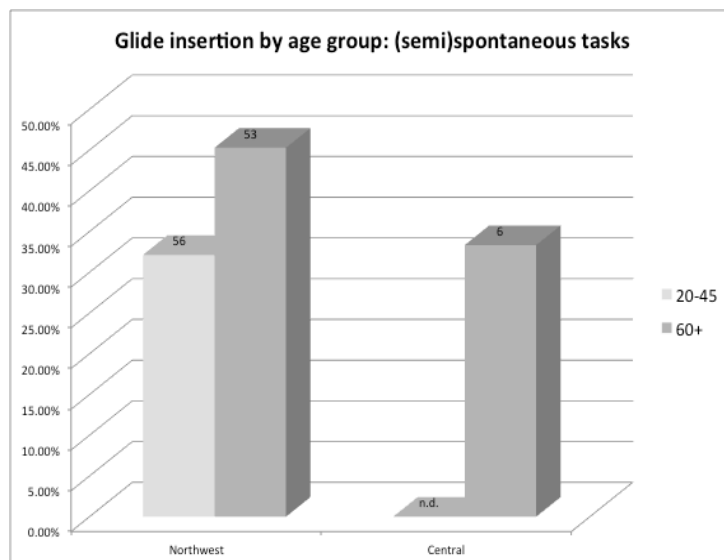


Fig. 13: Percent of glide insertion in the (semi-)spontaneous tasks by age group in the Northwestern and the Central regions.

Because there were no contexts for GI in the data from the younger group of the Central region, the effect of age cannot be observed in (semi-)spontaneous speech for this region. Like in the previous subsections, the data show a systematic increase in the proportion of GI in the (semi-)spontaneous tasks in comparison with the reading materials.

4.5. Predictive power of linguistic and extralinguistic factors

In order to observe which (if any) of the variables influence GI, we have run a binary logistic regression model (e.g. Bayley, 2013), with GI as a dependent variable (0=absence of GI, 1=presence of GI) and the following 6 factors: region (urban and rural areas were conflated in each district, and hence, 4 regions were considered, namely, Arcos de Valdevez and Castro Laboreiro, Braga and Fiscal, Ermesinde and Gião, and Nisa); task (reading and (semi-)spontaneous tasks); age (20-45 and 60+); prosodic status of W1 (W1=PW and W1=CL), prosodic context (below PhP, above PWG and below IP, and across IP), and prominence (V2 bears PW stress, V2 is prominent at phrasal levels below IP, V2 is prominent at the IP level)¹⁰. Additionally, we also wanted to observe if a given factor, even if not relevant when considered in isolation, might predict GI when in interaction with other factors. To this end, we combined each extralinguistic factor with

¹⁰ For prosodic context and prominence, another category had to be labeled for the statistic analysis purpose – ‘not relevant’. This label was used for glide insertion where W1 is a CL.

the linguistic ones, forming 9 possible interactions (region*prosodic status of W1, region*prosodic context, region*prominence, task*prosodic status of W1, task*prosodic context, task*prominence, age*prosodic status of W1, age*prosodic context, age*prominence). Using the stepwise method, the model inserted each of our predictors (interactions included) at a time and 5 steps were performed (out of 15 possible predictors). This means that only 5 predictors were considered to be good predictors of the occurrence of the phenomenon under analysis, accounting for 87% of the data (see Table 2).

Classification Table^a

Observed			Predicted		
			GI		Percentage Correct
			no	yes	
Step 1	GI	no	1002	73	93.2
		yes	117	96	45.1
		Overall Percentage			85.2
Step 2	GI	no	1004	71	93.4
		yes	106	107	50.2
		Overall Percentage			86.3
Step 3	GI	no	1050	25	97.7
		yes	145	68	31.9
		Overall Percentage			86.8
Step 4	GI	no	1050	25	97.7
		yes	144	69	32.4
		Overall Percentage			86.9
Step 5	GI	no	1050	25	97.7
		yes	143	70	32.9
		Overall Percentage			87.0

Table 2. Classification table (the cut value is .500).

By looking at the factors/interactions inserted in each step and their significance values (see Table 3, Appendix I), we conclude that region*prosodic context, task*prominence level, age*prominence level,

prosodic status of W1, and task are good predictors of GI ($p < .05$). This means that all extralinguistic factors influence GI, but only task affects the occurrence of this phenomenon when considered in isolation. In other words, the other factors are relevant only when in interaction with specific linguistic factors. A similar observation is made for linguistic factors: all of them are relevant, but only the prosodic status of W1 affects the occurrence of GI when considered in isolation.

These statistical results confirm several of our observations in sections 4.1-4.3. In section 4.1, we have seen that some regions insert more (Castro Laboreiro, Arcos de Valdevez, Gião), and in all prosodic contexts, whereas other regions insert less (Ermesinde, Fiscal), and in this case glides are inserted only in lower prosodic domains. The observation that in regions that insert more the process is found in several prosodic domains, including higher levels, whereas in regions that insert less, glides appear only in lower domains is confirmed by the binary logistic regression model in the significant effect of the interaction region*prosodic context on the occurrence of GI.

In section 4.2, we have seen that higher levels of prominence also seem to matter. However, a few differences were observed between speech styles, at least some of which possibly due to differences between the two modalities in the potential contexts obtained for GI. In read speech, to the exception of Castro Laboreiro, where phrasal prominence in V2 always matters, in all regions only IP-prominence clearly induces the presence of

glide (recall that in the regions where this context obtains, glide was always inserted). In (semi-)spontaneous tasks, GI also clearly seems promoted by the highest level of prominence that occurred. However, the highest level of prominence that appeared in this speech style is PhP prominence. This may explain the significant effect of the interaction task*prominence level in our binary regressive model.

In section 4.3 we showed that GI is more frequent when V1 is a clitic than when it is a PW. Our model also confirmed the statistical relevance of the prosodic status of W1.

We have observed in sections 4.1-4.3 that in the various conditions, (semi-)spontaneous speech showed more GI than read speech. The statistical significance of this difference is confirmed by the fact that task was selected as a good predictor by our statistical model.

The last predictor selected by the model shows an interaction between age and prominence. At this point, we do not have enough elements to interpret this result, and thus we leave it for future investigation.

5. General discussion

In our data, GI occurs between central vowels *across words*, as previously described (Lopo, 1895; Pereira, 1908; Santos, 1897; Segura, 2013, a.o.). However, our results show that not only the process is bound by

the IP domain, but it may also occur *within PW* (e.g. Faáto>Fa[j]áto).¹¹ Although there is great variation in the frequency of occurrence of the glide, in all of the regions under observation the process seems to be active, but is not obligatory. Despite the fact that not all regions display GI in every prosodic condition, the data are compatible with the view that GI is the result of an optional span rule bound by the IP (Frota, 2000; Nespors & Vogel, 1986/2007), as represented in (10). These are new pieces of information, adding to the general characterization of GI in the dialects of EP.

- (10) $0 \rightarrow [j] / (\dots V_1 _ V_2 \dots)_{IP} (opc.)$
 (where V1 and V2 are central vowels and V2 bears word stress)

In this sense, the process is similar to most segmental phenomena described for SEP (Frota, 1995, 2000), namely, semivocalization of round vowels followed by a vowel and other processes of resyllabification – e.g., (*O músic[w] africano cantou várias canções*)_{IP} vs (*O bailarín[u]/*[w]*)_{IP} (*antes de partir*)_{IP}...). Portuguese thus remains a language where no segmental processes are bound by the phonological phrase, unlike other languages, like Standard Italian or Bengali (Nespors & Vogel, 1986/2007 and Hayes & Lahiri, 1991, respectively).

¹¹ The words containing adjacent sequences of central vowels are rare in EP. However, this sequence is possible word internally and participants seem to have treated the relevant items in our corpus in a native-like way, since GI operated. Nevertheless, we cannot discard the possibility that these invented nouns may have triggered specific behavior in hiatus resolution, possibly, disfavoring GI.

The contexts for GI vary greatly in our data. Nevertheless, the statistical analysis has shown that both linguistic and extralinguistic factors seem to play a role in the rate of GI, namely, prosodic status of W1 and task, taken in isolation, and prosodic domain and prominence level, in interaction with extralinguistic factors (specifically, prosodic domain in interaction with region, prominence level with task, and prominence level with age).

The amount of variation found across prosodic conditions, the small overall rate of GI in most regions, and the sensitivity to language external factors suggest that the process is either regressing or stigmatized (Harris, 1991; Labov, 1966, 1972; Mithun, 1992, a.o.). In particular, in most of the regions studied GI is rather infrequent, older speakers produce more insertion than younger subjects and the frequency of insertion is much higher in informal speech style than in the more formal one. In addition in some areas such as Braga, GI almost only emerged in less controlled speech. This, in our view, suggests that in more formal speech styles, speakers may use the features that are known to characterize SEP, and inhibit those that are specific to local varieties (a pattern that is commonly found in sociolinguistic research since Labov, 1966, 1972). In other words, we may hypothesize that in the dialectal areas under observation, communities are bi-dialectal, and linguistic and extralinguistic factors may press in different directions when it comes to using features of each dialect: while phonological constraints of local, dominant dialect would favor

epenthesis, external constraints (e.g. Standard prestige) might press towards inhibition of GI insertion (Weinreich, Labov, & Herzog, 1968, p. 159). This may call for a constraint-based analysis where linguistic-internal constraints and linguistic-external constraints may interact and be assigned variable weights (Coetzee, 2009, et seq.). We leave this as a direction for future research.

The results indicate that higher levels of prominence impact on the frequency of GI. In particular, when V2 bears IP prominence glide is always inserted. Statistical analysis corroborated the relevance of V2 level of prominence in interaction with task and age. Importantly, an effect of sentence nuclear stress on V2 on other hiatus resolution processes was also found in many other Romance languages (see the review in Cabré & Prieto 2005).

The prosodic status of W1 also seems to matter, since insertion is significantly more frequent when W1 is a clitic, and in one of the regions insertion is only found in that word condition (Braga). A distinction in the rate of GI insertion as a function of W1 status, in particular with no insertion when W1 is a PW and insertion when it is a CL, may be interpreted as a *frequency effect* explainable by models such as Bybee's (2001) and Coetzee's (2013). Special phonology, often reflecting older stages of the language are frequent in Portuguese host-clitic combinations (e.g. *comem-no*, '(they) eat-it', instead of 'comem-o') or clitic clusters (e.g. *na* 'in-the_{FemSg}', instead of *em a*), and in many other languages as well (e.g.

Anderson, 2005; Hayes, 1990; Mascaró, 2007; Napoli & Nevis, 1987; Peperkamp, 1997; Vigário, 2003). The reports found for GI insertion to break a hiatus in Galician seem to suggest that the process may be restricted to word sequences involving clitics (as in *comprabaio* instead of *comprabao*, ‘(he) buy_{Past-it_{MascSg}}’ from Álvarez & Xove, 2002, p. 38). Our results seem to indicate, nevertheless, that no such lexicalization has occurred in most regions, since the presence of the glide is not obligatory in the regions where there is little insertion, and in the regions where there is more insertion, namely in Castro Laboreiro and Arcos de Valdevez, the glide occurs very frequently with W1=PW as well. In the region where insertion only occurs when W1 is a CL, i.e. Braga, more investigation is required in order to determine if this is indeed the only context for the emergence of glide presently in this area. We may notice, nevertheless, that in the rural area close to Braga (Fiscal), although insertion is not frequent in the condition W1=PW it does occur, which may suggest that GI insertion has not yet lexicalized in the region.

A difference in the prosodic contexts for GI insertion was observed across regions, which appears to be related to the rate of insertion. In the regions that insert more, namely Castro Laboreiro and Arcos de Valdevez, the glide is found in domains of prosodic hierarchy of various levels, whereas in the areas with less insertion, GI insertion takes place in lower domains and does not apply across PhP. That prosodic rules tend to be more obligatory in lower domains like has been observed in other languages, for

instance in Dutch (Booij, 1995). These results may suggest, in addition, that when a prosodic rule is regressing, the loss starts at higher levels. Yet another possibility, nevertheless, is that it is harder for speakers to avoid stigmatized dialectal variants at lower prosodic domains. In our view, this issue certainly calls for further investigation.

In geographic terms, the regions that exhibit more insertion are the most Northern in the Northwest area (Castro Laboreiro and Arcos de Valdevez), whereas the most southern location, Nisa, is amongst the regions that insert less. This is not unexpected considering that these regions are, respectively, the farther and the closest to the standard EP areas, where insertion does not occur.

Besides that, areas that are important urban centers, namely Braga and Ermesinde, very close to Porto, exhibit very low rates of insertion. We have seen, moreover, that in general rural locations insert more than the urban locations in the same region. This is compatible with the view that the process is regressing, the rural and more peripheral areas being more conservative and thus possibly losing it more slowly. More data is needed to determine how consistent these observations are within the geographical space where GI insertion is found (Oliveira, 2016).

6. Final remarks

In this paper we show that GI insertion to break a hiatus is an optional process that spans the IP domain. Insertion occurs *across words*, as previously described, but it is bound by the IP domain, and may also apply *within PW* (e.g. Faáto>Fa[j]áto).

In all of the regions analyzed (Arcos de Valdevez, Castro Laboreiro, Braga, Fiscal, Ermesinde, Gião, Nisa) the process is active. However, there is great variation in the frequency of occurrence of insertion. A number of linguistic and extralinguistic factors seem to favor/inhibit GI insertion, and speakers from different geographical areas appear to vary in the weight assigned to each of these factors. Using a binary logistic regression model with GI insertion as a dependent variable we tested the predicting power of 6 factors: region (4 regions were considered, namely, Arcos de Valdevez and Castro Laboreiro, Braga and Fiscal, Ermesinde and Gião, and Nisa); task (reading and (semi-)spontaneous tasks); age (20-45 and 60+); prosodic status of W1 (W1=PW and W1=CL), prosodic context (below PhP, above PWG and below IP, and across IP), and prominence (V2 bears PW stress or higher levels of prominence). Results showed that all factors considered seem to contribute to explaining the presence of the glide, some considered in isolation, like prosodic status of W1 and task, and some in interaction with others, such as prosodic context in interaction with region, and prominence level in interaction with task and in interaction with age.

The fact that GI insertion is significantly less frequent in younger subjects, in the most southern region of our sample, in the more formal

speech style and in more central urban centers may indicate that we are witnessing a change in progress and/or GI insertion is a stigmatized feature that is characteristic of non-dominant dialects. This topic requires further investigation, both enlarging the regions studied and using other means for collecting relevant information, such as speakers' intuitions or more elaborate psycholinguistic methods.

The process was still found in Nisa, the most southern area where GI insertion has been identified (Segura, 2013). The rate of GI insertion is nevertheless much lower than in the Northwestern locations, and it is virtually confined to word sequences where W1=CL. This may suggest that the process is disappearing from this area. Further study in this region could be very informative as to the role that factors like the prosodic status of the words involved and/or their frequency of occurrence may play in the pathway of loss of a phrasal prosodic rule.

Despite the fact that GI insertion to break a hiatus in the Portuguese dialects still requires further investigation and several questions remain open, this research increases our knowledge of the mapping of prosodic variation in EP, and we believe it may give new insights to the understanding of the phonology of hiatus cross-linguistically, and more in general the mechanisms involved in phonological variation and change.

Acknowledgements

This work is part of an ongoing investigation within the project InAPoP – Interactive Atlas of the Prosody of Portuguese (PTDC/CLE-LIN/119787/2010) and was also partially funded by UID/LIN/00214/2013. The data analyzed here form a subset of the materials considered in Oliveira (2016). Oliveira et al. (2014) is a preliminary, shorter version of this work. We would like to thank all subjects that were willing to participate in the recordings, as well as the various local authorities that have helped us in several ways. We are also grateful to Bruno Neto and Nádia Barros who were involved in the collection of data analyzed here. Previous versions of this work have been presented at the *1st Workshop of the Project Interactive Atlas of the Prosody of Portuguese* (University of Lisbon, Lisbon, 27th June 2013), at *I International Symposium on Variation in Portuguese* (University of Minho, Braga, 28th-30th April 2014) and *22nd Manchester Phonology Meeting* (University of Manchester, Manchester, 29th-31th May 2014). We would like to thank the audiences of these meetings for valuable comments, specially Celeste Rodrigues, Heglyn Pimenta, Nancy Hall, José Ignacio Hualde and Carlos Gussenhoven. We are also grateful to Xosé Luís Rgueira for relevant information on glide insertion in Galician and Sónia Frota for her contribution in several stages of this research.

References

- Aguiar, J. (2008). *Unidades e Processos Fonológicos no falar da região da Terra Quente: contributos para a Linguística Forense* (Unpublished master's thesis), Universidade do Minho, Braga, Portugal.
- Alba, M. (2006). Accounting for variability in the production of Spanish vowel sequences. In N. Sagarra & A. J. Toribio (Eds.), *Selected proceedings of the 9th Hispanic linguistics symposium* (pp. 273-285). Somerville, MA: Cascadilla Proceedings Project.
- Álvarez, R., & Xove, X. (2002). *Gramática da lingua Galega*. Vigo: Editorial Galaxia.
- Anderson, S. (2005). *Aspects of the theory of clitics*. Oxford: Oxford University Press.
- Anderson, A., Bader, M., Bard, E., Boyle, E., Doherty, G. M., Garrod, S., Isard, S., Kowtko, J., McAllister, J., Miller, J., Sotillo, C., Thompson, H. S., & Weinert, R. (1991). The HCRC map task corpus. *Language and Speech*, 34, 351-366.
- Barros, N. (2014). *Fraseamento prosódico em Português: Uma análise entoacional de construções parentéticas e tópicos em duas variedades do Português Europeu* (Unpublished master's thesis), Universidade de Lisboa, Lisboa, Portugal.

- Bayley, R. (2013). The quantitative paradigm. In J. K. Chambers & N. Schilling (Eds.), *The handbook of language variation and change* (2nd ed., pp. 85-107). New Jersey: Wiley-Blackwell.
- Bergman, R. (1971). Vowel sandhi and word division in Igede. *Journal of West African Languages*, 8, 13-26.
- Boersma, P., & Weenink, D. (2007). *Praat – doing phonetics by computer*. Version 5.3.56 [www.praat.org].
- Boléo, M. de P., & Silva, M. H. S. (1961). O mapa dos dialectos e falares de Portugal continental. *Boletim de Filologia*, III, 85-112.
- Booij, G. (1995). *The phonology of Dutch*. Oxford: Clarendon Press.
- Brissos, F. (2015). Dialectos Portugueses do centro-Sul: corpus de fenómenos e revisão do problema da (des)unidade. In *Zeitschrift für romanisch Philologie*, 131(4), 999-1041. doi: 10.1515/zrp-2015-0071.
- Bybee, J. (2001). Frequency effects on French liaison. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 337-359). Amsterdam: Benjamins.
- Cabré, T., & Prieto, P. (2005). Positional and metrical prominence effects on vowel sandhi in Catalan. In S. Frota, M. Vigário & M. J. Freitas (Eds.), *Prosodies* (pp. 123-157). Mouton de Gruyter: The Hague.
- Casali, R. (1997). Vowel elision in hiatus contexts: Which vowel goes? *Language*, 73, 493–533.

- Casali, R. (2011) Hiatus resolution. In M. van Oostendorp, C. Ewen, E. Hume & K. Rice. (Eds.), *The blackwell companion to phonology* (Vol. 3, pp. 1434-1460). Malden, MA & Oxford: Wiley-Blackwell.
- Chitoran, I., & Hualde, J.I. (2007) From hiatus to diphthong: The evolution of vowel sequences in Romance. *Phonology*, 24, 37–75.
- Cintra, L. F. (1971). Nova proposta de classificação dos dialectos galego-portugueses. *Boletim de Filologia*, 22, 81-116.
- Coetsee, A. (2009). An integrated grammatical/non-grammatical model of phonological variation. In Y.-S. Kang, J.-Y. Yoon, H. Yoo, S.-W. Tang, Y.-S. Kang, Y. Jang, C. Kim, K.-A. Kim & H.-K. Kang (Eds.), *Current issues in linguistic interfaces* (Vol. 2, pp. 267-294.). Seoul: Hankookmunhwasa.
- Colina, S. (1997). Epenthesis and deletion in Galician: An optimality-theoretic approach. In F. Martínez-Gil & A. Morales-Front (Eds.), *Issues in the phonology and morphology of the major Iberian languages* (pp. 235-267). Washington, DC: Georgetown UP.
- Cruz, M. (2013). *Prosodic variation in European Portuguese: Phrasing, intonation and rhythm in Central-Southern varieties* (Unpublished doctoral dissertation). Universidade de Lisboa, Lisboa, Portugal.
- Elimelech, B. (1976). A tonal grammar of Etsako. (UCLA *Working Papers in Phonetics*, 35.) Los Angeles: University of California.

- Elordieta, G., Frota, S., & Vigário, M. (2005). Subjects, objects and intonational phrasing in Spanish and Portuguese. *Studia Linguistica*, 59(2-3), 110-143.
- Florêncio, M. (2001). *Dialecto Alentejano - contributos para o seu estudo*. Lisboa: Colibri.
- Frota, S. (1995). Os domínios prosódicos e o Português Europeu: Fenómenos de sandhi. In *Actas do X encontro nacional da Associação Portuguesa de Linguística* (pp. 221-237). Lisboa: APL/Colibri.
- Frota, S. (2000). *Prosody and focus in European Portuguese. Phonological phrasing and intonation*. New York: Garland Publishing.
- Frota, S. (Coord) (2012-2015). *Interactive atlas of the prosody of portuguese*. [<http://labfon.letras.ulisboa.pt/InAPoP/>]
- Frota, S. (2014). The intonational phonology of European Portuguese. In S. Jun (Ed.), *Prosodic typology II* (pp. 6-42). Oxford: Oxford University Press.
- Frota, S., Oliveira, P., Cruz, M., & Vigário, M. (2015). *P-ToBI: tools for the transcription of Portuguese prosody*. Lisboa: Laboratório de Fonética, CLUL/FLUL. [<http://labfon.letras.ulisboa.pt/InAPoP/P-ToBI/>]
- Frota, S., Cruz, M., Fernandes-Svartman, F., Collischonn, G., Fonseca, A., Serra, C., Oliveira, P., & Vigário, M. (2015). Intonational variation in Portuguese: European and Brazilian varieties. In S. Frota & P. Prieto (Eds.), *Intonation in romance* (pp. 235-283). Oxford: Oxford University Press.

- Frota, S., Cruz, M., Castelo, J., Barros, N., Crespo-Sendra, V., & Vigário, M. (2016). Tune or text? Tune-text accommodation strategies in Portuguese. In *Proceedings of the speech prosody 2016*. Boston: Boston University.
- Grice, M. (1995). *The intonation of interrogation in Palermo Italian; Implications for intonation theory*. Tübingen: Niemeyer.
- Hall, N. (2011). Vowel epenthesis. In M. van Oostendorp, C. Ewen, E. Hume & K. Rice. (Eds.), *The blackwell companion to phonology* (Vol. 3, pp. 1576-1596). Malden, MA & Oxford: Wiley-Blackwell.
- Hall, N. (2013). Acoustic differences between lexical and epenthetic vowels in Lebanese Arabic. *Journal of Phonetics*, 41(2), 133-143.
- Harris, J. (1991). Conservatism versus substratal transfer in Irish English. In Peter Trudgill & J.K. Chambers (Eds.), *Dialects of English: Studies in grammatical variation* (pp. 192-213). Harlow, Essex: Longman.
- Hayes, B. (1990). Precompiled phrasal phonology. In S. Inkelas & D. Zec (Eds.), *The Phonology-syntax connection* (pp. 85-108). Chicago: University of Chicago Press.
- Hayes, B., & Lahiri, A. (1991). Bengali intonational phonology. *Natural Language Linguistic Theory*, 9(1), 47–96.
- Labov, W. (1966). *The social stratification of English in New York city*. Washington: Center for Applied Linguistics.

- Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press. Technology, Cambridge. [New York: Garland Press, 1980].
- Kutsch Lojenga, C. (1994). Ngiti: A central-sudanic language of Zaire. *Nilo-Saharan: Linguistic analyses and documentation 9*. Köln: Rüdiger Köppe.
- Lopo, J. (1895). Linguagem popular de Valpaços. *Revista lusitana*, III, 326.
- Maia, C. (1975). Os falares do Algarve (inovação e conservação). *Revista Portuguesa de Filologia*, XVII(I,II), 37-205.
- Martins, A. M., & Vitorino, G. (1989). Palatalisation et vélarisation conditionnées de la voyelle tonique dans certains dialectes portugais: Evolutions identiques dans l'espace roman. In *Espaces romans: Etudes de dialectologie et de géolinguistique offertes à Gaston Tuaille* (Vol. 2, pp. 330-356). Grenoble: Ellug, Université Stendhal.
- Mascaró, J. (2007). External allomorphy and lexical representation. *Linguistic Inquiry*, 38(4), 715–735.
- Mithun, M. (1992). The substratum in grammar and discourse. In E. Jahr (Ed.), *Language contact: Theoretical and empirical studies* (pp. 103-115). Berlin/New York: Mouton de Gruyter.
- Napoli, D., & Nevis, J. (1987). Inflected prepositions in Italian. *Phonology Yearbook*, 4, 195-209.
- Nespor, M., & Vogel, I. (1986/2007). *Prosodic phonology* (2nd ed.). Berlin/New York: Mouton de Gruyter.

- Oliveira, P. (2016). *A prosódia da iode epentética para quebrar hiatos em variedades do Português* (Unpublished master's thesis). University of Lisbon.
- Oliveira, P., Paulino, N., Cruz, M., & Vigário, M. (2014). Onde ainda([j])há o fenómeno. Contributo para o estudo da inserção de glide entre vogais centrais. In A. Moreno, F. Silva, I. Falé, I. Pereira & J. Veloso (Eds.), *Textos seleccionados do XXIX encontro da Associação Portuguesa de Linguística* (pp. 419-436). APL: Colibri.
- Paulino, N., & Frota, S. (2015). Variação prosódica no Português Europeu: Análise comparada de fenómenos de sândi vocálico. In *Textos seleccionados do XXX encontro nacional da Associação Portuguesa de Linguística* (pp. 435-448). Porto: APL.
- Payne, D. L. (1981). *The phonology and morphology of Axininca Campa*. Summer institute of linguistics publications in linguistics. University of Texas: Arlington.
- Peperkamp, S. (1997). *Prosodic words*. HIL Dissertations 34. The Hague: Holland Academic Graphics.
- Pereira, A. (1908). Tradições populares e linguagem de Vila Real. *Revista Lusitana*, IX, 229-258.
- Rodrigues, C. (2003). *Lisboa e Braga: Fonologia e variação*. Lisboa: FCG/FCT.

- Roseano, P., Vanrell, M. M., & Prieto, P. (2015). Intonational phonology of Friulian and its dialects. In S. Frota & P. Prieto (Eds.), *Intonation in romance* (pp. 101-139). Oxford: Oxford University Press.
- Santos, F. (1897). Linguagem popular de Trancoso. *Revista Lusitana*, V, 161-173.
- Segura, L. (2013) Variedades dialectais do Português Europeu. In E. Raposo, M. B. Nascimento, M. A. Mota, L. Seguro & A. Mendes (Eds.), *Gramática do Português* (Vol. 1, pp. 85-142). Lisboa: Fundação Calouste Gulbenkian/Centro de Linguística da Universidade de Lisboa.
- Selkirk, E. (1984). *Phonology and syntax: The relation between sound and structure*. Cambridge: MIT Press.
- Selkirk, E. (1986). On derived domains in sentence phonology. *Phonology Yearbook*, 3(37), 1-405.
- Tenani, L. (2002). *Domínios prosódicos no Português* (Unpublished Doctoral Thesis), Universidade de Campinas, São Paulo, Brazil.
- Vasconcellos, J. L. (1890-1892). Dialectos Alentejanos. *Revista Lusitana*, II, 15-45.
- Vasconcellos, J. L. (1896). Dialectos Alentejanos. *Revista Lusitana*, IV, 13-77.
- Vasconcellos, J. L. ([1901]1987). *Esquisse d'une dialectologie portugaise* (3rd ed. 1987). Lisboa: Instituto Nacional de Investigação Científica/Centro de Linguística da Universidade de Lisboa.

- Vigário, M. (2003). *The prosodic word in European Portuguese*. Berlin/New York: Mouton de Gruyter.
- Vigário, M. (2010). Prosodic structure between the prosodic word and phonological phrase: Recursive nodes or an independent domain? *The Linguistic Review*, 27(4), 485-530.
- Weinreich, U., Labov, W., & Herzog, M. (1968). Empirical foundations for a theory of language change. In W. Lehman & Y. Malkiel (Eds.), *Directions for historical linguistics* (pp. 95-188). Austin: University of Texas Press.

Appendix I

INSERT HERE TABLE 3, ATTACHED

Table 3: Variables in the Equation.

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a ProsCont * Region			235,264	9	,000	
ProsCont(1) by Region(1)	3,901	,422	85,497	1	,000	49,434
ProsCont(1) by Region(2)	1,447	,451	10,313	1	,001	4,251
ProsCont(1) by Region(3)	1,759	,442	15,830	1	,000	5,810
ProsCont(2) by Region(1)	3,551	,508	48,869	1	,000	34,857
ProsCont(2) by Region(2)	,827	,643	1,653	1	,199	2,286
ProsCont(2) by Region(3)	1,012	,572	3,136	1	,077	2,752
ProsCont(3) by Region(1)	3,509	,435	64,956	1	,000	33,405
ProsCont(3) by Region(2)	-,683	,809	,713	1	,399	,505
ProsCont(3) by Region(3)	,787	,529	2,211	1	,137	2,196
Constant	-3,551	,383	85,818	1	,000	,029
Step 2 ^b Prominence_level * Task			66,201	3	,000	
Prominence_level(1) by Task(1)	-1,991	,345	33,291	1	,000	,137
Prominence_level(2) by Task(1)	-3,007	,677	19,738	1	,000	,049
Prominence_level(3) by Task(1)	-2,732	,386	50,207	1	,000	,065
ProsCont * Region			210,928	9	,000	
ProsCont(1) by Region(1)	3,962	,446	78,867	1	,000	52,577
ProsCont(1) by Region(2)	1,368	,466	8,609	1	,003	3,928

	ProsCont(1) by Region(3)	1,841	,463	15,790	1	,000	6,302
	ProsCont(2) by Region(1)	3,929	,588	44,673	1	,000	50,871
	ProsCont(2) by Region(2)	1,715	,700	6,000	1	,014	5,554
	ProsCont(2) by Region(3)	1,714	,629	7,417	1	,006	5,549
	ProsCont(3) by Region(1)	3,848	,550	48,982	1	,000	46,920
	ProsCont(3) by Region(2)	-1,327	,844	2,472	1	,116	,265
	ProsCont(3) by Region(3)	,827	,592	1,950	1	,163	2,287
	Constant	-1,760	,449	15,387	1	,000	,172
Step 3 ^c	Prominence_level * Task			53,243	3	,000	
	Prominence_level(1) by Task(1)	-1,829	,384	22,706	1	,000	,161
	Prominence_level(2) by Task(1)	-2,004	,827	5,868	1	,015	,135
	Prominence_level(3) by Task(1)	-2,751	,420	42,856	1	,000	,064
	Age * Prominence_level			52,976	3	,000	
	Age(1) by Prominence_level(1)	-1,832	,275	44,288	1	,000	,160
	Age(1) by Prominence_level(2)	-2,970	1,290	5,304	1	,021	,051
	Age(1) by Prominence_level(3)	-,778	,345	5,085	1	,024	,459
	ProsCont * Region			206,704	9	,000	
	ProsCont(1) by Region(1)	4,772	,499	91,295	1	,000	118,181
	ProsCont(1) by Region(2)	1,644	,482	11,619	1	,001	5,174
	ProsCont(1) by Region(3)	2,308	,483	22,788	1	,000	10,053

	ProsCont(2) by Region(1)	4,263	,623	46,797	1	,000	71,046
	ProsCont(2) by Region(2)	1,786	,736	5,882	1	,015	5,964
	ProsCont(2) by Region(3)	1,711	,647	6,992	1	,008	5,534
	ProsCont(3) by Region(1)	3,985	,567	49,453	1	,000	53,768
	ProsCont(3) by Region(2)	-1,459	,851	2,941	1	,086	,232
	ProsCont(3) by Region(3)	,762	,602	1,602	1	,206	2,143
	Constant	-1,313	,469	7,834	1	,005	,269
Step 4 ^d	W1(1)	3,461	1,127	9,437	1	,002	31,834
	Prominence_level * Task			53,966	3	,000	
	Prominence_level(1) by Task(1)	-2,127	,412	26,678	1	,000	,119
	Prominence_level(2) by Task(1)	-1,901	,862	4,860	1	,027	,149
	Prominence_level(3) by Task(1)	-2,776	,443	39,196	1	,000	,062
	Age * Prominence_level			50,904	3	,000	
	Age(1) by Prominence_level(1)	-1,833	,278	43,456	1	,000	,160
	Age(1) by Prominence_level(2)	-3,067	1,302	5,545	1	,019	,047
	Age(1) by Prominence_level(3)	-,621	,361	2,967	1	,085	,537
	ProsCont * Region			188,101	9	,000	
	ProsCont(1) by Region(1)	3,320	,540	37,747	1	,000	27,658
	ProsCont(1) by Region(2)	,146	,533	,075	1	,783	1,158
	ProsCont(1) by Region(3)	,872	,525	2,765	1	,096	2,392

	ProsCont(2) by Region(1)	5,986	1,119	28,607	1	,000	397,833
	ProsCont(2) by Region(2)	3,455	1,190	8,429	1	,004	31,672
	ProsCont(2) by Region(3)	3,438	1,137	9,140	1	,003	31,117
	ProsCont(3) by Region(1)	5,610	1,083	26,842	1	,000	273,275
	ProsCont(3) by Region(2)	,210	1,252	,028	1	,867	1,233
	ProsCont(3) by Region(3)	2,422	1,099	4,859	1	,027	11,272
	Constant	-3,031	1,034	8,586	1	,003	,048
Step 5 ^e	Task(1)	20,023	9126,916	,000	1	,998	496667789,585
	W1(1)	3,479	1,126	9,548	1	,002	32,434
	Prominence_level * Task			2,029	3	,566	
	Prominence_level(1) by Task(1)	-22,004	9126,916	,000	1	,998	,000
	Prominence_level(2) by Task(1)	-21,471	9126,916	,000	1	,998	,000
	Prominence_level(3) by Task(1)	-22,494	9126,916	,000	1	,998	,000
	Age * Prominence_level			48,463	3	,000	
	Age(1) by Prominence_level(1)	-1,800	,278	41,964	1	,000	,165
	Age(1) by Prominence_level(2)	-3,079	1,311	5,520	1	,019	,046
	Age(1) by Prominence_level(3)	-,512	,366	1,954	1	,162	,600
	ProsCont * Region			181,628	9	,000	
	ProsCont(1) by Region(1)	3,285	,537	37,402	1	,000	26,704
	ProsCont(1) by Region(2)	,151	,529	,082	1	,775	1,163
	ProsCont(1) by Region(3)	,856	,522	2,687	1	,101	2,353

ProsCont(2) by Region(1)	5,723	1,129	25,699	1	,000	305,704
ProsCont(2) by Region(2)	3,226	1,191	7,340	1	,007	25,186
ProsCont(2) by Region(3)	3,039	1,174	6,697	1	,010	20,886
ProsCont(3) by Region(1)	5,395	1,081	24,891	1	,000	220,295
ProsCont(3) by Region(2)	,276	1,250	,049	1	,825	1,318
ProsCont(3) by Region(3)	2,283	1,104	4,276	1	,039	9,803
Constant	-3,185	1,037	9,426	1	,002	,041

a. Variable(s) entered on step 1: ProsCont * Region .

b. Variable(s) entered on step 2: Prominence_level * Task .

c. Variable(s) entered on step 3: Age * Prominence_level .

d. Variable(s) entered on step 4: W1.

e. Variable(s) entered on step 5: Task.