Journal of Portuguese Linguistics

Nkollo, M. (2022). Nasal epenthesis in preverbal accusative clitic pronouns. A variationist study of present-day dialectal European Portuguese. *Journal of Portuguese Linguistics*, 21:4, pp. 1–40. DOI: https://doi.org/10.5334/jpl.5890

Open Library of Humanities

Nasal epenthesis in preverbal accusative clitic pronouns. A variationist study of present-day dialectal European Portuguese

Mikołaj Nkollo, Adam Mickiewicz University, Poznań, PL, mikon74@amu.edu.pl

The paper investigates nasal epenthesis in vowel-initial preverbal 3rd person accusative pronouns in modern dialectal European Portuguese (EP). The study is underpinned by the data from the verbatim transcription section of CORDIAL-SIN, a dialectal corpus of contemporary EP. Speakers' behaviors are analyzed in the fifteen localities where variation is found. Different grammars (prosody-syntax mappings) are singled out, depending on whether the alveolar nasal is extended on preverbal clitic pronouns only or is found in other monosyllabic clitic words (definite articles and demonstrative pronouns) as well. Analogical extensions are demonstrated to be instrumental in inducing speakers to use the nasal-initial allomorph. The analysis points also to the varying realizations of proclisis triggers. Besides surfacing as a nasal diphthong, their last syllable frequently ends in a monophthong, its vocalic nucleus denasalizes or its quality gets altered. Finally, the historical profile of this sandhi process is approached. The change is argued to have spread from grammatically and communicatively unmarked contexts, close to orality. Rather than positing a continuous transmission of nasal epenthesis across generations, emphasis is placed on the consistency with which preverbal clitic pronouns were treated in various periods and in different communicative circumstances.

1. Introduction

Phonological processes operating across grammatical boundaries give rise to external sandhi phenomena (Sampson, 2016, p. 699). Their occurrence in European Portuguese (EP) is controlled by prosodic structure (Frota, 2013, p. 53). The aim of this paper is to offer a variationist study of such a phenomenon in present-day Portuguese dialects. The case in point involves the epenthesis of the alveolar nasal between two adjacent words, the second of which is a prosodically deficient monosyllable. In these sequences, the [n] occurs in between, and is syllabified as the onset of the second syllable. This sandhi phenomenon is assumed to reveal internalized grammars, i.e., phonology-syntax mappings, of dialect speakers. Clitic-verb sequences in these mappings are sometimes very different from how they are realized in standard EP.

Thus, the task is to reconstruct the underlying mechanisms instrumental in speakers' decision-making process. Phonological and morphosyntactic conditionings either prompt them to select the nasal-initial allomorph and extend it on preverbal pronominal clitics, or prevent them from doing so. Phonological conditioning involves the nature and the order of the disrupted segments. The alveolar nasal alters sequences involving a nasal diphthong and a subsequent syllable whose nucleus is an oral vowel. Once introduced, this segment becomes syllabified as the onset of the latter syllable. Morphosyntactic conditioning refers to the category of the two combined elements and the features they encode. The second in the sequence is the 3rd person accusative clitic pronoun. In standard EP, its forms do not vary in the preverbal position and are *o*, *a*, *os*, *as* ('him', 'her', 'them.MASC', 'them.FEM'). Expressions eligible to fill the first slot have more heterogeneous characteristics. Their common property is to coerce clitic pronouns into appearing preverbally (proclisis triggers), with the proviso that they themselves occur in the preverbal position (see Duarte & Matos, 2000). Sequences combining proclisis triggers and subsequent clitic pronouns are further referred to as trigger-clitic pronoun (TCP) sequences.

The paper is organized as follows: Section 2 is devoted to the phonological dimension of sequences with pronominal clitics and has a twofold structure. The first part is an overview of existing work on allomorphy involving clitic pronouns in EP and the proposals on domains for sandhi phenomena. As nasality is the ubiquitous phonological feature in the sequences dealt with in this paper, the status of nasals is addressed. In addition, this subsection contains a survey of word-final nasality in previous periods, including verbs and proclisis triggers. Section 3 is devoted to clitic placement at present and in the history of Portuguese. Section 4 presents the method and the procedural steps of data collection and processing. The presence of the alveolar nasal in various multi-word combinations occurring in CORDIAL-SIN interviews is introduced in Section 5. A closer look into these data reveals in what respects the phonological and morphosyntactic conditioning in the dialectal speech is different from standard EP. In Section 6, the variationist dimension of the nasal epenthesis across particular interview sites is approached. Concomitant phonological changes are subsequently addressed to further highlight the disjunction between

the dialects and standard EP phonology. In Section 7, the historical continuity hypothesis is discussed as a possible explanation of the presence of nasal-initial preverbal clitic pronouns in some of the localities listed in CORDIAL-SIN. A number of grammatical similarities between the scriptural production of previous eras and today's dialects are revealed. Obstacles intrinsic to this hypothesis are identified to show why it fails to capture speakers' behaviors successfully. A more moderate proposal, inspired by H. Andersen's actualization theory, is advanced to grasp the origin of divergent phonology-syntax mappings in present-day dialectal EP. The results are recapitulated in the closing section.

2. Phonological conditioning

One of the most debated points pertaining to EP pronominal clitics is their status as either words or affixes. Given the amount of their phonological idiosyncrasy (i.e., shifts not found anywhere else in the grammar), it is a contentious issue at which stage of the derivation the verb-clitic/clitic-verb sequences are generated.

2.1. Prosodic status of pronominal clitics and domains for sandhi phenomena

The debate was fuelled by Zwicky and Pullum's (1983, pp. 503–504) criteria advanced to circumscribe divergent properties of English *not* and *n't*. Some of their points prove also relevant for EP clitic pronouns: (i) unlike clitics, which are moderately selective about the category of their host, affixes are rigorous in selecting the category of the stem to which they attach; (ii) morphophonological properties that cannot be captured by general phonological rules in a given language are easily found in affixed words, but on a lesser scale in clitic-host combinations; (iii) unpredictable gaps are more characteristic of affix-stem than of clitic-host combinations (see also Miller & Monachesi, 2003, pp. 92–93). These observations seem to lend support to the affixhood of EP object pronouns. Zwicky and Pullum (1983) also point to the recursive addition of clitics, but not of affixes, to material already containing clitics, and to the fact that stem-affix combinations outrank host-clitic groups in that they are more often semantically irregular.

Yet, the affixhood hypothesis clashes with some of the distributional properties of clitic pronouns. Therefore, the idea of their word-internal insertion must be taken with reservations, if not jettisoned altogether (the point is extensively discussed in Vigário, 2003, pp. 131–155). Stress patterns are usually referred to in order to corroborate the non-affixal status of pronominal clitics. In EP, word stress can travel maximally to the third syllable from the right edge of the word (*three-syllable window*; Luís & Kaiser, 2016, p. 214). If clitic pronouns were parts of verbal forms, their attachment would coerce the stress into moving onwards, so that its placement would be compliant with the three-syllable window. Meanwhile, postverbal clitic pronouns do not alter the locus of the word stress in EP.

Aside from lexical stress, in numerous other cases word-internal allomorphy fails to come about if clitic pronouns, rather than inflectional or derivational affixes, are added to the stem. An intriguing case involves the insertion of the non-back glide between two heterosyllabic vowels, the first of which is the stressed [e]. It can be found in the first-person present indicative, the present subjunctive and most of the imperative forms of verbs whose infinitive ends in *-ear* (*despentear* 'to mess up one's hair' – *despenteio*, *chatear* 'to annoy' – *chateio*, *manusear* 'to handle' – *manuseio*, etc.) or, for some Latinate verbs, in *-iar* (*incendiar* 'to set on fire', *comerciar* 'to trade', *negociar* 'to have business affairs').[1] This process is blocked if the segment [e] is followed by a direct object pronoun. For example, *lê-o* 'read it' surfaces as [le.u], but never as [le.ju]. Thus, unlike the inflectional ending *-o*, accusative pronominal clitics do not act as word-internal syllables with respect to this sound change.

The affixhood vs. wordhood debate can be associated with the theoretical construct known as Strong Lexicalist Hypothesis (SLH): No syntactic rule can refer to elements of morphological structure (Dalrymple, 2001, pp. 83–84, 97–98). If speech production proceeds sequentially (the prerequisite to which is a set of hierarchically ordered modules; see Levelt, 1989; Labelle, 2001, pp. 156, 161–162), clitic attachment deviates from how 'normal' multi-word combinations are built. If clitic pronouns (and affixes) are parts of another expression, their processing must start very early. The syntactic module cannot operate until its input has been fully specified. This is tantamount to saying that all of the irregular properties of verb-clitic/clitic-verb combinations are word-internal and arise in the lexical module. Even more radically, SLH requires two separate entries to be stored in the mental lexicon of EP users: one for a bare verb and another one for sequences combining verb forms and clitic pronouns. The use of the latter entry becomes mandatory if arguments have not been planned as fully-fledged NPs in an earlier module. Therefore, clitic-verb/verb-clitic sequences must be memorized on a one-by-one basis, just as some stem-affix combinations are claimed to be.

In other proposals, remedial measures are introduced to obviate the need for positing an overgrown lexical module. One of them relies on the assumption that, instead of operating on raw sequences generated in the morphological and the syntactic component, the phonological interpretation has an indirect character. It is mediated by the prosodic structure (Frota, 2013, pp. 4–5), which involves a set of hierarchically organized domains: syllable (s), prosodic word (PW, or ω), phonological phrase (PP, or f), intonational phrase (IP, or I) and phonological utterance.

Syllable consists minimally of a nucleus, and maximally of an onset, nucleus, and coda. Taken together, the nucleus and the coda constitute the rhyme. The nucleus represents the peak

[1] In the paradigms of some of the *-iar* verbs, there are clause-mates, i.e., divergent forms filling a single cell (e.g., *comercio* alongside *comerceio*, *negocio* alongside *negoceio*, see Piel, 1989, pp. 222–223; and Thornton, 2011, pp. 362, 368, for the concept of clause-mates). In some of the corresponding nouns, the hiatus is likewise resolved by glide insertion (*manuseio* 'handling', *passeio* 'walk').

of sonority; less prominent segments that precede the nucleus belong to the onset. Those that follow the nucleus constitute the coda. The onset and the coda may remain empty (alternatively, the term onsetless is used for empty-onset syllables). If no segments are present in the coda, the syllable is open. For stress computation purposes, onset dominates the sound segments directly, whereas for the nucleus and the coda, the dominance is mediated via mora constituents. Therefore, the rhyme is the moraic component (Collischonn & Wetzels, 2016, p. 87).

Other prosodic elements – prosodic word group, and clitic group – are sometimes added to this list. These are the domains for sandhi phenomena in EP. For example, in Vigário's (2003, p. 125) typology of phonological processes, all of the postlexical phenomena make reference to the PW. By contrast, in word-internal phonological changes, some processes are active on this level as well (initial vowel realization, heterosyllabic [e]-centralization, hiatus-breaking glide insertion; see also Frota, 2013, pp. 83–88, for the experimental study of the last of these phenomena), whereas others (vowel reduction, theme vowel deletion, and final nasal diphthongization) are not.

Essentially, prosodic domains need not overlap with previously generated morphosyntactic units (their boundaries are not coincident).[2] Likewise, in the rules operative in a given domain, reference to other grammatical modules (or to semantic information) is allowed. This approach to clitic-verb/verb-clitic sequences does away with the need to combine expressions of both categories in the lexicon. Even if some phonological changes affect the form of the verb and the form of the clitic in an unpredictable way, compelling evidence is also found for the non-lexical derivation of verb-clitic sequences in EP.

One of the models stemming from Prosodic Phonology involves precompiled phrasal allomorphy. Precompilation lists allomorphs (in the lexicon) alongside contexts for their activation.[3] An alternative procedure involves the list of lexical rules for the instantiation of

[2] For example, syntactic segmentation notwithstanding, the length of intonational phrases is prosodically constrained: even if they may correspond to a single syntactic constituent, long IPs tend to split into smaller ones. Sequences featuring a balanced length are given preference over sequences juxtaposing long and short (short and long) IPs. The longest IP tends to be located at the rightmost edge of the prosodic utterance (Oliveira et al., 2017, p. 53). In brief, no one-to-one correspondence exists between the boundaries of syntactic and prosodic constituents. A summary of previous comparative work on Romance intonational phrasing focusing on the role of syntactic (constituency) and prosodic factors (branching, length) can be found in Frota, et al. (2007, pp. 133–134).

[3] Precompilation rules have been successfully applied in E. Fernández Rei's (2002) Ph.D. thesis to account for sandhi phenomena triggered by phonological clitics in present-day Galician (E. Fernández Rei's, 2002, pp. 216–217, 221–222, 241–242). Liquid-initial and nasal-initial clitic words in rural speech have been touched upon in her work as well. These variants are instances of allomorphy constrained by prosodic conditions. The /lo/ allomorph is expressly listed (precompiled) as both the clitic pronoun and the definite article. It is shown to occur after the preceding /s/ or /r/. The /no/ allomorph tends to appear after a nasal (including *algué-no viu* and *ningué-no viu*) or a glide. Empty-onset forms are found elsewhere. Importantly, a single set of cliticization rules is posited for both clitic pronouns and articles ("En definitiva, a clitización de artigos e pronomes persoais presenta semellanzas moi importantes, polo que non sería aventurado pensar que ocupan o mesmo lugar na gramática", E. Fernández Rei, 2002, p. 236, 'All in all, clitic attachment in articles and personal pronouns evinces many similarities. Thus, it would not be exaggerated to think that they occupy the same place in the grammar').

particular allomorphs. These rules are paired with suitable syntactic environments. Analyses conducted using this theoretical apparatus show that EP clitic pronouns are defectively prosodized morphosyntactic words. Despite occupying syntactic terminal nodes, they do not form autonomous PWs (Santos & Vigário, 2016, p. 130).

The domain instrumental in accounting for the behavior of phonological clitics in EP is the PW. It usually includes a stem plus any of the adjacent affixes. If any of them forms an independent domain for lexical stress assignment, it constitutes a PW on its own. In other words, each PW bears exactly one lexical stress, which is allocated at its right edge (three-syllable window). PWs can be clustered together to form a higher level domain, namely Prosodic Word Group. That such cumulative units consist of successive PWs can be seen in certain compounds and abbreviations.

As phonological clitics are inherently unstressed, they must be part of one of the PWs onto which the adjacent expressions are mapped. In EP, clitic words are not all alike with respect to the PWs to which they belong. Proclitics prosodize with the host PW, thereby elongating it by one or two syllables at best. These syllables are simply adjoined at the PW's left edge. Phenomena revealing this type of attachment include the alternation in height of non-central non-high vowels, prosodic word initial stress, and emphatic stress assignment. With respect to all of them, prosodically deficient function words (including definite articles) behave like PW-initial segments (Vigário, 2003, p. 196). By contrast, postverbal clitic pronouns are incorporated into the preceding PW. The case is discussed by Vigário (2003, pp. 187–188) and Santos and Vigário (2016, p. 131). The postlexical sandhi phenomenon, deemed to best grasp the nature of incorporation, involves non-back final vowel deletion. It comes about at the PW's right edge. To be deleted, the non-back vowel must be PW-final, and occur in IP-medial position. For example, in *TOm**e**-o de maNHÃ* 'Take it in the morning', the enclitic pronoun prevents the host final vowel from occupying the PW-final position. Therefore, instead of being deleted, it must surface as a non-back glide. By contrast, in *CAbe-t**e** a TI* (*limpar a casa*) 'it's up to you (to clean the house)', the clitic final vowel becomes PW-final and is deleted ['kab.ta.'ti], whereas the IP continues.

Optional [ə] deletion in sentences like *JA t**e** aconteCEU* '(It) has already happened to you' (another realization involves the non-back glide [j]) is argued to be indicative of the fact that clitics are not enclitically attached to the preceding PW and that preverbal pronominal clitics are proclitic by nature (Santos & Vigário, 2016, p. 131). Besides advancing our understanding of the variationist dimension of present-day EP dialectal speech, nasal epenthesis in preverbal accusative clitic pronouns is significant in that it poses a challenge to this opinion, at least as far as EP dialects are concerned. If *no*, *na*, *nos*, *nas* variants are precompiled and listed as allomorphs alongside vowel-initial forms (e.g., Vigário, 1999, p. 589), then the conditions for

their instantiation must indicate that they prosodize with the preceding diphthong-final PW. The non-verbal category of the expression corresponding to the hosting PW must likewise be disclosed. In sum, despite occurring in the preverbal position, accusative clitic pronouns are assumed to be enclitically incorporated in the preceding expression. Rather than prosodically adjoining to the ensuing verb, they form a PW along with the left-hand-side proclisis trigger.

2.2. Nasal diphthongs

It is a long-established idea that nasal vowels represent an underlying tautosyllabic VN sequence. It dates back to at least the Prague School (Piggott, 1992, p. 38). This assumption was later fine-tuned by Mattoso Câmara Jr (1971, p. 33) to account for the status of nasal diphthongs in Portuguese. He posited a sequence that consists of an oral diphthong followed by a nasal archiphoneme. The nasality of the latter spreads into the preceding segment. Bisol (2008, p. 2) sustains the idea that both nasal vowels and nasal diphthongs rely upon VN in their underlying form, except that the initial sequence is VN for nasal vowels and VN-V for nasal diphthongs. According to her line of reasoning, the intervocalic nasal is incorporated into the vowels. This is why the V(nasal)V sequence fails to produce a hiatus. The intervocalic nasal is either lost or realized as the onset of the next syllable: *luna* > *lũa* > *lua* 'moon'; *um* one.MASC – *ũa* > *uma* one.FEM (see also Martins 2016a, pp. 7, 12). Moreover, Bisol (2008, p. 11) makes a distinction between two types of diphthongs in Portuguese: true diphthongs (*portão* 'gate', *põe* 'puts', *capitães* 'captains') correspond to two input vowels, one of which becomes a glide by syllabification. By contrast, false diphthongs (*viagem* 'trip', *homem* 'man', *som* 'sound') have only one vowel in the input. In present-day standard EP, verb final nasal diphthongs surface as nasal glides (e.g., *viram* ['vi.ʀɐ̃w̃]; Vigário, 2003, p. 76). The nasality of the vowels in nasal diphthongs is assumed to arise via the assimilation from the following tautosyllabic nasal consonant. Thus, the nasality of the glide carries over into the preceding vocalic segment. The two segments share the same value for the feature [back].

The ongoing debate over Portuguese nasal diphthongs revolves around the structure of the syllable they realize. For some scholars, these segments represent complex (or ramified) nuclei of open syllables (Mateus & d'Andrade, 2000; Martins, 2016a, p. 11, "Outras mudanças ocorreram que afetaram os ditongos decrescentes sem lhes alterar *a natureza de núcleos ramificados*" 'Other changes came about and affected falling diphthongs without altering *their nature of ramified nuclei*' [emphasis mine]). For others, falling nasal diphthongs are rhymes with both their nucleus and coda constituents realized. The glide is located in the syllable coda in this approach (Collishonn & Wetzels, 2016, pp. 89–90). The latter position is corroborated to some extent by word stress patterns of EP. Filled codas in nasal diphthongs are posited on account of the sameness of effect

on stress assignment exerted by both closed syllables and diphthongs. This line of reasoning hinges on how successfully the notion of syllable weight is mapped onto present-day Portuguese (many of the analyses conducted in this vein take into account Brazilian Portuguese).[4]

Finally, word-final diphthongization is argued to be a word-internal process (Vigário, 2003, pp. 77–78). Two apparent counterexamples – *irmã* 'sister' and *orfã* 'orphan.FEM', which end in a simple nasal vowel and where no nasal glide is present – are assumed to arise by the coalescence (Sampson, 2016, p. 669) of the stem-final vowel and the feminine class marker *-a*. No underlying final nasal segment is present in these words. Therefore, no nasal glide surfaces. In verbal morphology, a nasal diphthong is steadily upheld in the closing syllable, no matter whether a given verb form stands alone or is followed by other expressions, including clitic pronouns. This stability shows that the relevant domain for word-final nasal diphthongization is the morphological word rather than the prosodic word.

Historically, the [ɐ̃w̃] diphthong is known to have developed from word-final simple nasal vowels [õ] and [ã]. By contrast, [ẽj̃]/[ɐ̃j̃] diphthongs derive from the word-final [ẽ]. Generally, these evolutions do not stimulate much controversy in the literature (Sampson, 1983, pp. 34–37; Fagan, 1992, pp. 383–385). An exception to this is Mattos e Silva's (2006, p. 73) reconstruction, where three stages, rather than two, are singled out. The intermediate phase in the development from [õ] to [ɐ̃w̃] involves the [õw̃] nasal diphthong. Three late Latin endings – *-ane*, *-one*, *-anu* – are presumed to have produced the [õw̃] diphthong. The syncope of the intervocalic [n] is instrumental in triggering the nasal resonance of the preceding vowel. Mattos e Silva posits the following sequence of events to explain the advent of the [õw̃]: i) the [+ nasal] feature of [n] carries over onto the preceding vowel, i.e., [on] > [õn]; ii) rather than being simply deleted, [n] turns into a glide, i.e., [õn] > [õw]; iii) the [+ nasal] feature is inherited by the resulting semi-vowel, eventually producing [õw̃]. The emergence of [w̃] to replace [n] in the ending *-one* is traced back to the fact that both segments – the vowel and the resulting glide – share [+ back], [+ round] features.

Besides phonetic reconstructions, attempts were made to establish correspondences between the mutations of particular Latin endings and the spelling variants representing them

[4] In the grammars of Classical Latin, heavy (or bimoraic) syllables contrast with light (monomoraic) ones. Stress is allocated to heavy syllables, whereas light syllables are, by default, unstressed. Moraic contrasts are assumed to be inherent to languages featuring contrastive vowel length. In Classical Latin, open syllables containing long vowels, diphthongs, and closed syllables count as heavy. Open syllables with short vowels are light and cannot receive stress (with the notable exception of so called 'extrametrical' stress). A word is a paroxytone if its penultimate syllable is heavy. Proparoxytones occur if the penultimate syllables are light. These dependencies are mapped onto stress patterns of Portuguese. They account, for example, for the fact that the penultimate of proparoxytonic words is always an open syllable. Put differently, stress on the antepenultimate never coexists with closed syllables or diphthongs in the penultimate (this explains, for example, the likelihood of stress placement in *estômago* 'stomach' or *íngreme* 'steep', and the impossibility of *estômalgo*, *estômango*, *íngresme*, etc.).

in Galician-Portuguese documents. For example, Maia (1986, p. 602) identifies the following graphic developments for *-one*: *-õ*, *-on*, and *-om* (e.g., *condiçom*, *condiçõ*). Finally, the chronology of the emergence of word-final nasal diphthongs is variously reconstructed. Martins (2016a, pp. 11–12) localizes the advent of [ɐ̃w̃], [ẽj̃] and [ɐ̃j̃] in Old Portuguese. In so doing, she contradicts the opinion of Ramos (1985, p. 95), who concludes that at that stage the convergence of word-final simple nasal vowels [õ] and [ã] towards the nasal diphthong [ɐ̃w̃] was still under way. Essentially, these reconstructions are indicative of the fact that some of the CORDIAL-SIN realizations of the negative marker with a simple nasal vowel (i.e., {nũ = *não*}) in the rhyme should not be approached in terms of monophthongization. On the contrary, from the historical perspective, in present-day standard EP falling diphthongs represent the development from simple nasal vowels.

In the history of Portuguese, word-final nasality also had a direct bearing on clitic allomorphy in the preverbal position, which is the principal point of interest in the present paper. Pronominal clitics were subject to different forms of liaison with both the preceding and the ensuing sound material. In the latter variant, the encounter of pronouns' vocalic nuclei with vowel-initial verb forms resulted in vowel deletions. Elided forms in this configuration were optional. They helped circumvent the formation of heterosyllabic vocalic clusters. The point is discussed by Martins along with relevant examples (Martins, 1994, pp. 217–218), e.g., *sobligarõ* (< *se obligaram* '(they) took an obligation onto themselves'), *todo lhera prazer* (< *tudo lhe era prazer* 'everything was a pleasure for him'), etc. By contrast, fusions with the preceding sound material were exclusive to third-person direct object pronouns. In such configurations, *o*, *a*, *os*, *as* surfaced with an [n] onset following proclisis triggers ending in a nasal segment.

Finally, in medieval Portuguese, word-final nasality triggered liaisons of definite articles with the preceding expressions. The array of forms following nasal-final function words (in particular, the prepositions *com* 'with', *sem* 'without', and *em* 'in', and the negative *nem*) involved *no*, *na*, *nos*, and *nas* alongside *lo*, *la*, *los*, and *las*. Thus, in the manuscripts of Galicia and northern Portugal, aside from *cũ lo noſſo caſal* 'with our farmhouse', there are occurrences of *cõ no conuento* 'with the monastery' (Maia, 1986, p. 648). This is an intriguing case where allomorphy and free variation play along with each other. Whereas [n]-initial and [l]-initial forms to the exclusion of the vowel-initial one are an instance of complementary distribution, the selection of either of these two consonant-initial realizations was made on a free-choice basis.

3. Morphosyntactic conditioning

In present-day EP, the placement of clitic pronouns is syntactically conditioned, yielding a system which is distinct from what is found in other Romance languages: proclisis is not paired with the finiteness of the verb (see Duarte & Matos, 2000; among others), but depends on specific syntactic conditions.

3.1. Pronominal clitics in present-day EP

Enclisis is the default pattern in non-dependent finite clauses. Proclisis is conditioned by the elements that have the property of attracting the pronoun to the pre-verbal position. They constitute the first element of the sequences subject to analysis in this paper. They include: i) negation; ii) preverbal negative indefinite pronouns and adverbs (*ninguém*, *nunca*); iii) some other adverbs placed preverbally, such as *já* 'already', *ainda* 'still', among others; iv) preverbal quantified subjects. Proclisis occurs also in: v) clauses with a filled CP, e.g., *qu*-questions; and vi) subordinate clauses with overt complementizers. These distributional properties illustrate the syntactic behavior of clitics as heads, since they require a host of a specific grammatical category.

In standard EP grammar nasal epenthesis in accusative clitic pronouns is obligatory in enclisis only if the preceding verb form ends in a nasal diphthong. This constraint is nearly always respected in the speech of CORDIAL-SIN informants (see 1a). Some scholars impose an additional requirement on the verb saying that it must encode the [+ plural] feature (Luís & Kaiser, 2016, pp. 221–222). In the corpus, singular verb forms tend to combine with a vowel-initial accusative clitic pronoun even if their final syllable contains a suitable diphthongal segment (see 1b).

(1) (a) *Os dentes, às vezes, os rapazes*
ART.DEF.M.PL tooth.PL sometimes ART.DEF.M.PL guy.PL
tiravam-nos (*e aquilo eles caíam* {ɔ = *ao*} *mar*)
toss.IPFV.IND.3.PL = CL.ACC.3.M.PL
'The teeth, the guys tossed them sometimes (…)' (informant: Balduíno, male, Ribeiras, Horta district)

(b) *a gente* ***põe-o*** *todo em peças*
ART.DEF.F.SG people put.PRS.IND.3.SG = CL.ACC.3.M.SG all.M.SG in piece.PL
'People smash it all to pieces' (informant: Anteu, male, 65 yo, 3rd grade, Bade, Viana do Castelo district)

Moreover, nasal onsets are clitic-specific in today's standard EP. As discussed in 2.2, unlike in early stages of Portuguese, postverbal definite articles do not undergo allomorphy in present-day EP. They keep their form intact in spite of being homonymous with 3rd person accusative pronouns and no matter what the phonological profile of the syllable preceding them is (see 2a). Likewise, preverbal accusative pronouns are always vowel-initial in the standard grammar (see 2b).

(2) (a) **Viram nas raparigas* (*Viram* ***as*** *raparigas*)
see.PRF.IND.3.PL ART.F.PL girls
'(they) saw the girls'

(b) *Quem* ***as*** ***viu****?* (**quem nas viu?*)
Q.who CL.ACC.3.F.PL see.PRF.IND.3.SG
'who saw them?'

3.2. Clitic placement in diachrony

In the history of EP, clitic placement had been different from what it is nowadays. In some of the structural environments, which were subsequently reanalyzed as obligatory enclisis contexts (from the first half of the 19th century on; some authors locate the change slightly earlier; see Galves et al., 2005; Namiuti, 2008, p. 53; Sandalo & Galves, 2013), pronominal clitics occurred either pre- or postverbally. Galves et al. (2005, pp. 42–43) single out five positions where the choice was made on an optional basis: root clauses with an overtly expressed preverbal subject, clitic left dislocation, sentences with a clause-initial non-VP adverb or non-argumental prepositional phrase. Additionally, the position of pronominal clitics varied in verb-first root clauses following a hierarchically subordinate clause and in verb-first root clauses following coordinating conjunctions *e* 'and', *pois* 'then', *mas* 'but'. Thus, if verb-clitic/clitic-verb variation contexts are considered to have been a separate mode of placement, pronominal clitics were able to occur in as many as five different positions with respect to the verb: enclisis, mesoclisis, non-obligatory proclisis, obligatory proclisis and interpolation (clitic-verb non-contiguity). The negative marker *não* had a special (albeit not exclusive) role in instantiating the last of these linear arrangements. Importantly, not only did the interpolation of *não* show up in obligatory proclisis contexts in negated clauses, but also in verb-clitic/clitic-verb ones (Namiuti-Temponi, 2006, p. 173). Obligatory proclisis contexts remained essentially unchanged throughout history. By contrast, the only structural position where enclisis has been intact involves the so called Tobler-Mussafia context (Hirschbühler & Labelle, 2000, p. 165): a definite ban on sentence-initial pronominal clitics. Indeed, violations of this constraint are really exceptional.

(3) ***Se*** ***exceptua*** *a* *Feſta de Aprezentaçaõ,* *quando*
REFL.3 make.exception.PRS.IND.3.SG DEF.F.SG feast of Presentation, when
ſe fas na Emfermaria …
REF.3.SG do.PRS.IND.3.SG in = DEF.F.SG hospital

'An exception is made for the Candlemas, if it is celebrated in the Hospital' (1749. *Compromisso da irmandade da Gloriosa Virgem e Martir Santa Cecilia*. Cap. IV. De como seraõ Admetidos …; file page 20 https://purl.pt/24981)

Finally, verb-clitic/clitic-verb variation contexts had varying outcomes in history. In alternate periods, preference was given either to the preverbal (its peak was reached in the 15th century) or to the postverbal placement. It seems that the revival of enclisis in free-choice environments had a popular inspiration. In her study of the speech of Gil Vicente's popular characters (16th century), Martins (2012, p. 98) shows that they tended to use an enclitic grammar.

4. Data and method

The analysis that follows is based on a single corpus of manageable size (600,000 words). CORDIAL-SIN (*Corpus Dialectal par Estudo da Sintaxe/Syntax-oriented Corpus of Portuguese Dialects*;

Martins (coord.) 2000–) is available on the website of the Linguistic Centre of the University of Lisbon in the open access format, including the verbatim transcription section. Recordings are not available. They had been taped between 1974 (Serpa) and 1997 (Vila Pouca do Campo, Unhais da Serra, Figueiró da Serra), depending on the locality. Their annotation was performed between 1999 and 2006 in the same research unit. CORDIAL-SIN is a collection of interview-based, geographically dispersed, dialectal speech material. Samples of spontaneous and semi-directed talk were collected at 42 sites, in both the mainland of Portugal and the archipelagoes. The informants are generally elderly (prevailingly over 60 yo.), have no or low literacy skills, and spent most of their lives close to the interview locations. In the transcripts, they have been anonymized using fictitious names in alphabetical order. Their sex is consistently revealed, but not their level of education or their age.

The sequences in the body of this paper are reproduced from the verbatim transcription section of the corpus. The section is based on speech recordings made to develop the *Atlas Linguístico-Etnográfico de Portugal e da Galiza* (ALEPG), the *Atlas Linguístico do Litoral Português* (ALLP) and the *Atlas Linguístico e Etnográfico dos Açores* (ALEAç). Each CORDIAL-SIN tapescript appears to have been compiled after a careful decision-making process – contentious fragments were double-checked to ensure that two annotators' opinions were concordant on what had been the exact sound sequence produced. Divergent perceptions are noted throughout, extending from more to less probable reconstructions of the intended sequence. Pauses, hesitations, slips of the tongue, reiterations, truncated words and all faulty speech occurrences are likewise painstakingly annotated. Interviews stored in the verbatim transcription section indicate the phonological sequencing of each non-standard case, be it word-internal or taking place across words.

The files are available in a machine-readable format as PDF documents. The analysis of the geographical variation in TCP sequences has been based on the tapescripts of all 42 interviews. The search was carried out in two stages with the aid of the advanced search function. First, all the occurrences of nine proclisis triggers were extracted. They were all required to end in a nasal diphthong. The list comprised *não* 'not', *quem* 'who', *ninguém* 'nobody', *sem* 'without', *também* 'also', *alguém* 'somebody', *bem* 'well, quite', *em* 'in' and *nem* 'not even'.

The retrieval of diphthong-final proclisis triggers gave 45 hits with nasal-initial preverbal accusative clitic pronouns. Consonant-initial onsets have been found in fifteen interviews. Nasal onsets never constitute the unique variant in any of these interviews. This is to say that in each locality, though not in the speech of every informant, consonant-initial and vowel-initial forms come together. In the same 15 interviews, 72 vowel-initial preverbal accusative clitics have been retrieved in phonologically and morphosyntactically equivalent conditions. In the remaining 27 localities, only vowel-initial realizations are attested. **Table 1** below cross-checks the two variants against seven proclisis triggers in the localities under discussion. Neither *alguém o/a/os/as* nor *alguém no/na/nos/nas* are present in the entire verbatim transcription section. Nor does *bem* appear as a proclisis trigger in these localities.

sequences with *n*-insertion	45	vowel-initial pronouns	72
quem no (*na, nos, nas*)	10	*quem o* (*a, os, as*)	2
não no (*na, nos, nas*)	27	*não o* (*a, os, as*)	56
ninguém no (*na, nos, nas*)	2	*ninguém o* (*a, os, as*)	3
nem no (*na, nos, nas*)	3	*nem o* (*a, os, as*)	1
também no (*na, nos, nas*)	–	*também o* (*a, os, as*)	10
sem no (*na, nos, nas*)	2	*sem o* (*a, os, as*)	–
em no (*na, nos, nas*)	1	*em* o (*a, os, as*)	–

Table 1: The distribution of TCP sequences in 15 CORDIAL-SIN interviews with divergent realizations of preverbal accusative 3rd person clitic pronoun.

Some local tendencies are conspicuous right away. The archipelago of the Azores and the North of the mainland of Portugal stand out as the areas where variation is most prominent (see **Figure 1**). It remains for future investigation to determine whether the same proclisis tiggers tend to exhibit similar ratios with each of the concurrent realizations of the accusative clitic.

In order to see in what other contexts clitic words were realized as nasal-initial, four subsequent queries were launched for *o, a, os* and *as* strings, defined as separate words. Even if the realization of linguistic units corresponding to these elements (demonstrative pronouns, pronominal clitics and definite articles) or to the environing expressions deviates from standard EP phonology, the verbatim transcription section of the corpus annotates each individual transgression, indicating what standard form had been intended by the speaker. For example*, e deixai* {PH|muˈʁelɐ = *morrer a*{fp}} *velha* (the abbreviation PH stands for phonetic and fp for filled pause; these symbols are omitted in the body of the paper). The separate word query prevented random, word-internal *os* or *as* strings from being displayed. The addition of the standard sequence targeted by the speaker made it possible to extract all the occurrences of the nasal-initial allomorph of any clitic word, e.g.*,* {PH|nũnu = *não o*} *sei*. Results reached after these two queries had to be further narrowed down and systematized in order to tease apart various grammatical categories (for example, to leave out of analysis the *no* contracted form corresponding to *em o*, etc.).

Essentially, further inquiry was restricted to the localities whose residents' speech contains at least one occurrence of nasal-initial preverbal 3rd person clitic pronoun. Such calculations are reckoned to circumscribe both the geography (distribution of different syntax-phonology mappings over different areas; Seiler, 2004, p. 394) of the nasal epenthesis and the minimal contrasts pervading informants' speech (random co-occurrence of mappings). The calculation of the number of informants owes its usefulness to the fact that the very presence of a non-canonical

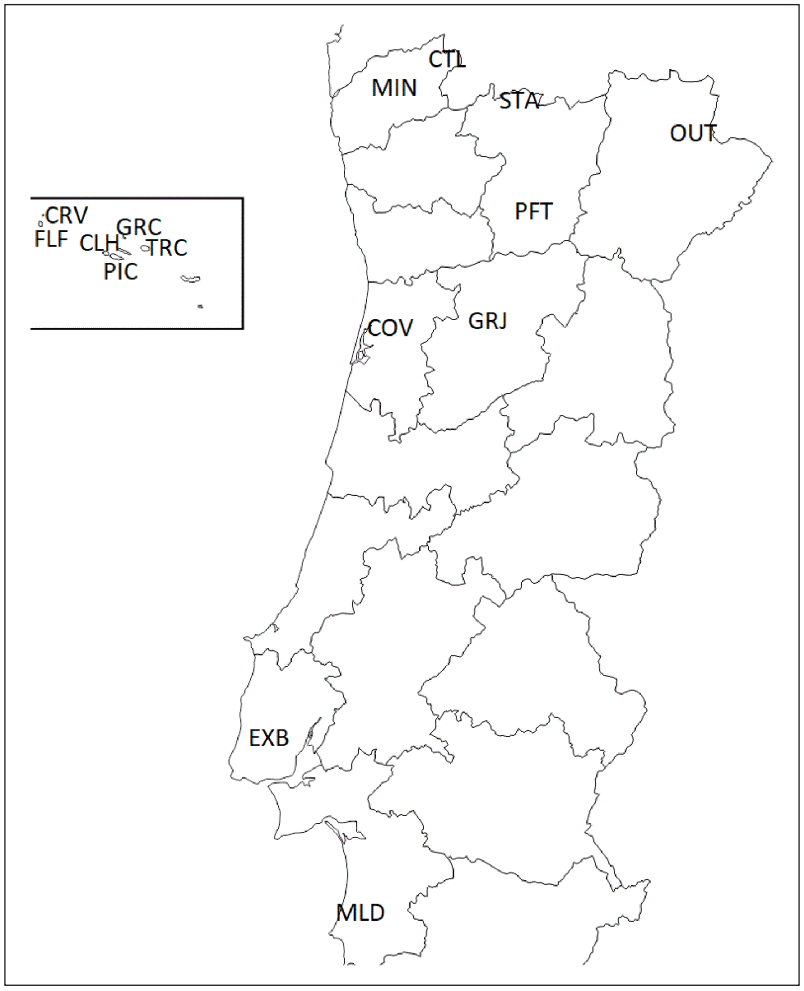

Figure 1: CORDIAL-SIN interview sites with [n]-initial preverbal accusative clitic pronouns: CLH Calheta (Angra-do-Heroísmo district); COV Covo (Aveiro); CRV Corvo (Horta); CTL Castro Laboreiro (Viana do Castelo); EXB Enxara do Bispo (Lisbon); FLF Fajãzinha (Horta); GRC Graciosa (Angra-do-Heroísmo); GRJ Granjal (Viseu); MIN Arcos de Valdevez, Bade, S. Lourenço da Montaria (Viana do Castelo); MLD Melides (Setúbal); OUT Outeiro (Bragança); PFT Perafita (Vila Real); PIC Bandeiras, Cais do Pico (Horta); STA Santo André (Vila Real); TRC Fontinhas (Angra-do-Heroísmo). Source: https://d-maps.com.

realization of the pronoun in a given interview site is not very enlightening and does not tell us much about the strength of the sound change.

5. Nasal epenthesis in clitic words in CORDIAL-SIN – an overview

Clitic allomorphy in standard EP notwithstanding, CORDIAL-SIN speakers behave differently with respect to nasal onsets in accusative pronouns. First of all, the corpus contains several instances of epenthesis where neither phonological nor morphosyntactic conditions are met.

In (4a) below, a nasal segment encroaches upon a context that, in standard EP phonology, requires a liaison between the final fricative of the first word and the initial vowel of the word that follows. Example (4b) resembles instances of external sandhi to which this paper is devoted (i.e., nasal onset in TCP sequences), except that the proclisis trigger *ainda* 'still' ends in an oral monophthong. Moreover, nasal-initial form in this sentence violates the constraint on the non-variability of preverbal clitic pronouns. Such cases eschew any attempt towards systematization.

(4) (a) {mɐnu = *mas o*} *cão* *nunca* *apareceu*
but = ART.DEF.M.SG dog never appear.PRF.IND.3.SG
'but the dog has never appeared'; informant Alarico; male; no further data; Castro Laboreiro, Viana do Castelo district)

(b) {'ĩdɐ = *Ainda*} {nɐz = *as*} *há*.
'still = CL.ACC.F.SG be-there.PRS.IND.3.SG
'They are still there' (informant: Hortense, female, no further data; Santo André, Vila Real district)

Putting aside confusing examples of this kind, in CORDIAL-SIN there frequently appear nasal onsets in defiance of the morphosyntactic constraints defined in the introductory section. As a result, the nasal segment is projected upon any syntactic combination that exhibits a suitable arrangement of syllables across words (nasal diphthong – oral vowel). This is by far the largest group in the typology of non-canonical (i.e., departing from standard EP grammar) nasal epentheses documented in the entire corpus. The sequences included in this group are, among others: adverb – definite article (see 5a), verb – demonstrative pronoun (see 5b), preposition – definite article (see 5c), indefinite pronoun – verb (5d)

(5) (a) *Então,* *olhe* *que* *ela* [*nem*] {ˈnɐ̃jnɐ = *nem a*} *pensão*
Well note.IMP.3.SG COMP she.SBJ.F not even = ART.DEF.F.SG maintenance
paga!
pay.PRS.IND.3.SG
'Well, she does not pay even for the maintenance' (informant Arquibaldo, male, 75 yo, can read, Covo, Aveiro district)

(b) *que* *faziam* {nuʃ = *os*} *que* *estavam* *nos*
REL do.IPFV.IND.3.PL = CL.DEM.M.PL REL be.IPFV.IND.3.PL in-ART.DEF.M.PL
hospitais
hospital.PL
'that were doing those who were in the hospitals' (informant Agar, female, 80 yo, can read and write, Perafita, Vila Real district)

(c) *A* *gente* *ficámos* {sɐ̃ʃnɐ = *sem as*} *culturas*
DEF.ART.F.SG people stay.PRF.IND.1.PL without = ART.DEF.F.PL crop.PL
'we were left without our harvest crops' (informant Dulcina; female; 70 yo; illiterate; Vila Pouca do Campo; Coimbra district)

(d) {nĩ'gɐ̃jnɐsɐ̃'derɐ̃w = *ninguém acenderam*} *as* *luzes*
nobody = turn-on.PRF.IND.3.PL ART.DEF.F.PL light.PL
'Nobody turned on the lights' (informant Herodiano, male; 64 yo; 4th grade completed; Aljustrel, Beja district; NB. note the devious S–V concord in number)

Amid the contexts exhibiting this type of epenthesis, verb – definite article combinations are most recurrent (see 6a–b).[5] In what follows, variable realizations in this configuration will help pin down with more accuracy CORDIAL-SIN informants' behaviors towards nasal onset in preverbal clitic pronouns.

(6) (a) *Depois* *quando* {'veɐ̃jnɐ = *vêem a*} *pessoa …*
Afterwards, when see.PRS.IND.3.PL = ART.DEF.F.SG person
'then, when they see someone' (informant: Emanuel, male, 73 yo, Granjal, Viseu district)

(b) {'terɐ̃nɐ = *terem a*} *eira* *feita*
have.INF.3.PL = ART.DEF.F.SG threshing-floor done.PST.PTCP.F.SG
'(so that) they have the threshing floor ready' (informant: Cirilo; male, 66 yo, illiterate; Luzianes, Beja district)

The above examples point to the perceptual salience of the nasal segment in this particular phonological environment. The concept of salience has often been used in different areas of linguistics to refer to features or patterns that are more noticeable to speakers (see Auer et al., 1998; Kerswill & Williams, 2002; a.o.). Compared to the degree of intrinsic prominence of the left- and right-hand side segments, the alveolar nasal occupies a low position on the sonority scale (Anderson, 2002). Therefore, it introduces a clear-cut acoustic boundary (a robust formant transition) between two adjacent peaks. These correspond to the nuclei of two successive syllables (the sonority value of the glide in the rhyme of the first syllable is likewise high ranked). As soon as the alveolar nasal is inserted at a syllable edge, a sharp phonetic contrast disrupts the protracted sonority of the two prominent heterosyllabic segments clustered together (see Selkirk, 1984, pp. 114–116; Luís, 2018, pp. 255–256, discusses a comparable case in the Korlai creole conjugation). That is why the alveolar nasal segment is easily perceived and stored in memory in such environments.

[5] Consonant-initial definite articles are found in 28 out of 42 CORDIAL-SIN interviews. The entire corpus contains 1652 occurrences of this sequence, with a verb form ending in a nasal diphthong. They are distributed over 184 informants. Fifty-one of them (27.71%) produce at least one nasal-initial definite article. In sum, there are 219 attestations of nasal onsets, i.e., 13.25% of the global production of the sequence (Nkollo, 2020, pp. 290–291).

The second type of non-canonical nasal epenthesis appears in contexts where morphosyntactic criteria for the addition of the nasal consonant are fulfilled, but which fail to exhibit suitable phonological structure. In (7) below, the enclitic accusative has a non-empty onset in spite of the fact that the preceding verb form ends in a simple oral vowel. The alveolar nasal appears to have been automated as part of a subsequent clitic word regardless of sound characteristics (oral and monosegmental) of the left-hand side material.

(7) *Que* *eu* {'vinɐ = *vi-a*} *vir* *por* *lá*
COMP I.SBJ see.PRF.IND.1.SG = CL.ACC.3.F.SG come.INF over there
'that I saw her coming that way' (informant Guilherme, male, 87yo; Montalvo, Santarém district)

Finally, the particular status of nasal onsets in the dialectal EP stands out sometimes when they are subtracted. This applies to contexts where subtraction is totally unexpected and not found elsewhere (i.e., in the speech of other interviewees or in standard EP).

(8) (a) *apanham-nos* *e* *tiram-os*
catch.PRS.IND.3.PL = CL.ACC.3.M.PL and remove.PRS.IND.3.PL CL.ACC.3.M.PL
dali *para* *fora*
from-there to outside
'They catch them and throw them outside' (informant: Celisa, female, no further data; Fontinhas, Angra do Heroísmo district)

(b) *Estavam todos na cama que ele*
be.IPFV.IND.3.PL all.M.PL in = DEF.ART.F.SG bed COMP he.SUBJ
enganaram-o
cheat.PRF.IND.3.PL CL.ACC.3.M.SG
'All of them were in the bed and they tripped him up' (Antipas, male, no further data; Arcos de Valdevez, Viana do Castelo district)

Unlike liquid-initial allomorphic variants of the 3rd person accusative pronoun, which are consistently realized by CORDIAL-SIN informants in suitable morphosyntactic and phonological conditions (for example, there are no attestations of [r]-final infinitives followed by vowel-initial 3rd person pronouns), the alveolar nasal is occasionally deleted in enclisis. Some of these deletions are puzzling, given that a canonical form is used inside the same sentence (see *apanham-nos e tiram-os* in 8a).

6. Variationist analysis

This part has a twofold structure. It brings together local dimension of the variation and speakers' individual choices. Proportional frequency is thought to be a manifold method to make the point on the preliminary data. It corresponds to the percentage of TCP sequences with a nasal

onset or to the percentage of speakers employing this variant, where the total number of the occurrences of the sequence, both standard and non-standard (provided they are compliant with the morphosyntactic conditioning specified above), or the total number of speakers is 100% (Hoekstra & Versloot, 2019, pp. 31–32). Before moving on to the variables, a short methodological disclaimer is needed. In spite of its advantages in capturing phonology-syntax mappings in TCP sequences, proportional frequency count applied to CORDIAL-SIN data is a risk-laden endeavor. The size of the corpus on a par with the relative scarcity of relevant configurations (117 TCP sequences) may prove a stumbling block to probative conclusions. Moreover, in some of the surveyed localities there are as few as three or four sequences with suitable phonological and morphosyntactic characteristics. To circumvent this hindrance, bare numbers are disclosed throughout.

6.1. Local distributions and individual behaviors

The first variable is designed to measure how the interview sites are different in the number of informants employing nasal epenthesis. The result is indicative of whether nasal allomorph is a socially widespread mode of realization in a given interview site or is rather isolated (see Seiler, 2006, p. 167). It also measures the degree of uniformity among the local speakers in the processing of the configuration at work. To this end, the percentage of informants producing nasal-initial pronouns needs to be calculated. The count in **Table 2** takes into consideration the overall number of informants producing the TCP sequence.

On average, no snowball effect can be found. This is to say that seldom does the presence of a consonant-initial preverbal accusative pronoun in the speech of an informant carry over to the rest of the speakers in the same locality. The higher the number of the interviewees producing the TCP sequence, the lower the percentage of onset-inserting ones is. The three localities boasting the 100% score involve maximally two (CLH, FLF) informants producing the relevant configuration. In those cases where the nasal allomorph is present in the speech of the majority of interviewees or where the two groups are equal (CTL, COV, GRC), the overall number of speakers is likewise low, reaching maximally four persons (CTL). Besides, individual behaviors tend to be balanced, with there being nearly as many individuals who only rely on canonical sequences as there are those who produce at least one consonant-initial preverbal accusative pronoun.

The second variable is designed to measure the strength of the tendency to employ nasal onsets across interview sites, once a TCP sequence is present there (Seiler, 2006, p. 168). In other words, the ratio of consonant-initial preverbal accusative pronouns to vowel-initial ones needs to be calculated for every locality. As said above, in none of the places where the interviews were conducted are nasal-initial preverbal accusative pronouns employed as the unique variant. Yet, in some places (**Table 3**), they outnumber the corresponding canonical forms.

Locality	Number of speakers producing the TCP sequence	Group 1: N (nasal onset)	Group 2: N (empty onset only)	Percentage of Group 1
CLH, FLF	2	2	0	100
EXB	1	1	0	
CRV, GRJ, PFT, TRC	3	2	1	66.66
CTL	4	2	2	50
COV, GRV	2	1	1	
OUT	7	3	4	42.85
MLD	5	2	3	40
PIC	6	2	4	33.33
STA	3	1	2	
MIN	5	1	4	20
Total. N = 51		26	25	50.98

Table 2: Percentage of informants employing nasal-initial preverbal clitic pronouns in CORDIAL-SIN interviews (the list of abbreviated names of localities, see **Figure 1**).

Locality	N (nasal-initial)	N (vowel-initial)	Ratio of nasal-initial to vowel-initial forms
FLF	5	1	1 : 0.2
GRC	3	1	1 : 0.33
CRV	2	1	1 : 0.5
CLH	4	3	1 : 0.75
CTL, STA, TRC	2	2	1 : 1
MLD	5	7	1 : 1.4
GRJ	4	7	1 :1.75

(Contd.)

Locality	N (nasal-initial)	N (vowel-initial)	Ratio of nasal-initial to vowel-initial forms
PFT, OUT	3	6	1 : 2
COV	6	14	1 : 2.33
PIC	2	9	1 : 4.5
EXB	1	5	1 : 5
MIN	1	6	1 : 6
	45	72	1 : 1.6

Table 3: Ratio of nasal-initial to vowel-initial preverbal accusative pronouns in TCP sequences across CORDIAL-SIN interview sites (the list of abbreviated names of localities, see **Figure 1**).

All four localities (CLH, CRV, FLF, GRC) with prevailing non-canonical realizations are clustered in the archipelago of the Azores. At the same time, the number of TCP sequences documented in the interviews conducted there is low, thus calling for some caution in drawing conclusions. In two interviews (EXB, MIN), a single occurrence of consonant-initial variant invites even more caution. Tellingly, in the localities where the highest numbers of TCP sequences are found (COV, GRJ, MLD, PIC), the vowel-initial variant is clearly preferred.

Another variable is speaker-oriented. It characterizes individual grammars of CORDIAL-SIN speakers by measuring the position of TCP sequences amid the contexts where phonological conditions are met for nasal epenthesis in prosodically deficient words. Thus, particular informants' speech will be checked for the same non-canonical forms in syntactically different conditions (mainly verb – definite article plus rare occurrences of verb – demonstrative pronoun and adverb – definite article). This procedure is believed to be more justifiable than describing localities in terms of the presence of the nasal epenthesis in other contexts. The uniqueness of TCP sequences in being the locus for onset insertion can be taken as a proof for a special status of clitic pronouns in the grammar of CORDIAL-SIN informants. Special status is meant here to refer to speakers' awareness that expressions belonging to other categories, even if homonymous with accusative clitics, are precluded from undergoing the epenthesis. Conversely, the presence of nasal onsets in other phonologically comparable settings lends support to the hypothesis of a phonology-driven individual grammar, with an apparently indiscriminate treatment of clitic words of any category in various syntactic contexts.

In four of the 15 localities where the relevant nasal epentheses are found, the only context where they appear is the TPC sequence, i.e., proclisis trigger ending in a nasal segment followed by 3rd person direct object pronouns (CLH, EXB, FLF, GRC; in sum, six informants with nasal-initial

preverbal accusative pronouns, one – with canonical forms only). These epentheses fail to appear in other configurations with clitic words. In the remaining eleven localities, 44 informants produce the TPC sequence. Twenty of them produce at least one nasal-initial preverbal clitic pronoun and 24 do not. These two groups show divergent behaviors towards nasal onsets in other configurations where favorable phonological conditions for the nasal epenthesis are met. The respective figures are as follows: 13 speakers in the former group employ the nasal onset also in other contexts (ratio 1: 0.538; see 5a–b above repeated as 9a–b below), and only six in the latter (ratio 1: 3; five interview sites: GRJ, OUT, PFT, STA and PIC with two such interviewees).

(9) (a) *Então, olhe que ela* [*nem*] {ˈnɐ̃jnɐ = *nem a*} *pensão*
Well note.IMP.3.SG COMP she.SBJ.F not even = ART.DEF.F.SG maintenance
paga!
pay.PRS.IND.3.SG
'Well, she does not pay even for the maintenance' (informant Arquibaldo, male, 75 yo, can read, COV)

(b) *que faziam* {nuʃ = *os*} *que estavam nos*
REL do.IPFV.IND.3.PL = CL.DEM.M.PL REL be.IPFV.IND.3.PL in-ART.DEF.M.PL
hospitais
hospital.PL
'that were doing those who were in the hospitals' (informant Agar, female, 80 yo, can read and write, PFT)

These figures show an overall avoidance of nasal onsets in other contexts in the speech of the informants who only use vowel-initial clitic pronouns in TCP sequences. By contrast, consonant-initial preverbal accusative pronouns are moderate attractors for the selection of the nasal-initial allomorph in other combinations involving clitic words and exhibiting suitable phonological structure. The 26 informants of this group are evenly divided: one group extends the alveolar nasal to other prosodically deficient words, whereas the second does not. Be that as it may, no implicational hierarchy emerges in either direction (i.e., nasal epenthesis in TCP sequences > nasal epenthesis elsewhere, or the other way round). **Table 4** represents the variation in phonology-syntax mappings in CORDIAL-SIN speakers established on the basis of the above calculation.

With respect to vowel-initial phonological clitics, in the 15 localities subject to the count, an entirely canonical individual grammar is the best represented option (group 3), even if it hardly exceeds one third of the interviewees. This group does not lend itself to further description. The speech of its members exhibits no variation whatsoever. As far as vowel-initial clitic words are concerned, everything is compliant with standard EP grammar. Conversely, the behavior of the speakers belonging to the three other groups evinces an interesting interplay of methods of processing prosodically deficient monosyllables. These groups of informants instantiate what

Group	Nasal-initial preverbal accusative pronouns (TCP sequences)	Nasal epenthesis in other prosodically deficient words (vowel-initial in standard EP)	Number of informants (%). N = 51
1	+	+	13 (25.49)
2	+	–	13 (25.49)
3	–	–	19 (37.25)
4	–	+	6 (11.76)

Table 4: Nasal epenthesis in clitic words of different categories in individual speech production of CORDIAL-SIN informants (only localities with nasal-initial preverbal accusative pronouns).

Harris and Campbell (1995, pp. 97–119) and Andersen (2001, p. 230, 2006, pp. 69–70) refer to as extension, i.e., a more or less intentional change for which no particular pragmatic motivation can be defined. One of the contexts in which it may take place is where a given morphophonemic rule is projected onto a new morphosyntactic environment and the innovation is immediately actualized in usage (see also Bybee & Torres Cacoullos, 2008, p. 400). The description that follows relies on the assumption that each time when speakers treat clitic pronouns and other clitic words differently with respect to nasal epenthesis, they must be aware that they represent distinct morphosyntactic categories (Nkollo, 2020, p. 294). Likewise, it should be borne in mind that in the individual speech production of CORDIAL-SIN informants canonical and non-canonical variants are systematically interwoven. Thus, the presence of an irregular nasal onset in a given clitic word must not be taken as a proof of an outright relinquishment of the standard grammar.

In group 2 nasal onsets are extended onto preverbal accusative pronouns. At the same time, other similar-sounding clitic words are consistently vowel-initial. Neither favorable phonological conditions nor the homonymy of accusative pronouns and definite articles prove sufficient to compel the informants into extending onsets beyond the TCP sequence. Therefore, a tacit assumption to the effect that only clitic pronouns are allowed to take a nasal onset must be part of this group's internalized grammar. Nasal epenthesis in preverbal accusative clitics is thought to have come about by analogy with the alternations in the postverbal domain in phonologically equivalent conditions. Thus, the only impaired segment of standard EP clitic grammar in this group of informants pertains to the preverbal vs. postverbal divide. Class-specific treatment of clitic pronouns is maintained throughout.

At first glance, the utterances in group 1 were produced as though its members had been oblivious of morphosyntactic distinctions between various classes of clitic words, all of which are treated alike with respect to nasal onsets. In accusative clitic pronouns, initial nasal segments are

introduced haphazardly, irrespective of whether they are postverbal or not. In other monosyllables, the nasal diphthong – oral vowel sequence seemingly acts as a sufficient cue to proceed with nasal epenthesis as well. Still, this method of processing clitic words must not be necessarily equated with a purely phonology-based grammar or the lack of any grammar whatsoever. As said above, besides the TCP sequence, it is the verb – definite article combination that is the most frequent locus for nasal onsets across words. The postverbal position is assumed here to be the overriding criterion. Thus, the linear ordering of clitic words induces speakers to process definite articles the same way as postverbal accusative pronouns. Obviously, the homonymy is another contributing factor to the category-blind nasal epenthesis.

Group 4 is peculiar in that its members behave as if they were aware of the ban on nasal epenthesis in pronominal clitics in the preverbal domain. Thus, nasal-initial allomorphs in 3rd person accusative pronouns are rigorously restricted to enclisis. Curiously, in their speech this parameter does not clash with instances of nasal epenthesis outside TCP sequences. Nasal onsets are unwarily produced only if the preceding word is a verb and contains suitable sound material. Unlike in group 2, where the informants were rigorous in applying a class-specific treatment of clitic pronouns, group 4 members compute the conditions for the use of the nasal-initial allomorph on the basis of the postverbal placement of clitic words. If a verb form happens to be diphthong-final, the nasal segment appears and all other parameters are played down.

The last variable characterizes individual phonology-syntax mappings with respect to TCP sequences only (see **Table 5**). It relies on the simple count of occurrences of nasal-initial preverbal accusative clitics in the global production of TCP sequences by a given informant. As two informants employ exclusively nasal-initial pronouns, the ratio cannot be calculated. The comparison of scores helps measure the strength (or the degree of entrenchment) of the non-canonical realization in individual grammars. As indicated above, there are 26 speakers producing nasal-initial preverbal accusative clitics. Ten of them produce merely a single TCP sequence, in which the pronoun happens to exhibit a nasal onset. They will be left out of the count, as a single occurrence tells us nothing about variation or its absence.

Again, the preference given to non-canonical forms is paired with a poor attestation of TCP sequences in the speech of particular CORDIAL-SIN informants. Conversely, the more frequent this combination is, the lower is the share of nasal-initial allomorphs. If the numbers of vowel-initial and consonant-initial variants are equal, speakers also produce only a small aggregate number of TCP sequences. Thus, once used, a nasal onset in the preverbal accusative pronoun is not mass-reproduced in CORDIAL-SIN interviews. Only as an exception do nasal-initial clitic pronouns play the role of attractors for nasal epenthesis in other TCP sequences produced by a given speaker. On average, this sandhi phenomenon appears to be no more than a fanciful realization, as it were, in an otherwise standard individual grammar.

Informant (locality)	Number of nasal-initial accusative pronouns in TCP sequences	Total number of TCP sequences	%
Flórido (GRC)	3	3	100
Gotardo (STA)	2	2	
Amélia (FLF), Galiano (MLD)	4	5	80
Heraclides (CLH)	2	3	66.66
Idalinda (CLH)	2	4	50
Emanuel (GRJ), Anselmina (MIN), Celestina (PIC), Atos (OUT), Célisa (TRC)	1	2	
Agar (PFT)	2	5	40
Ercília (GRJ)	3	8	37.5
Ácia (PFT)	1	3	33.33
Arquibaldo (COV)	6	19	31.57
Borromeu (EXB)	1	6	16.66

Table 5: Frequency of nasal-initial accusative pronouns in TCP sequences in the individual speech production of CORDIAL-SIN informants.

6.2. Concomitant phonological processes

Aside from displaying varying shapes of the accusative clitic pronoun, TCP sequences in CORDIAL-SIN are even more markedly subject to a series of concomitant phonological processes. Whereas some of them leave the skeleton of the sequence intact, the others alter it to the point where it becomes necessary to redefine the phonological characteristics outlined in the introduction. One of these processes involves monophthongization. Technically, the glide in the first of the juxtaposed syllables is deleted, leaving the preceding vowel as the only element of the rhyme. Monophthongization is found in 53 out of 117 TCP sequences in the 15 localities with two concurrent realizations of the preverbal pronoun. It is also regular in other configurations in phonologically equivalent conditions. According to the annotations in the verbatim transcription section, monophthongization occurs more frequently in syllables preceding nasal-initial accusative pronouns (23 out of 45 TCP sequences). The corresponding result for sequences with vowel-initial pronouns works out at 30 out of 72 occurrences. The examples below illustrate the point in each of the TCP variants.

(10) (a) *Mas, por exemplo, para quem* {nɐ̃nɐʃ = *não as*} *tenha* *aqui*
but, for instance, for one-who.REL NEG = CL.ACC.3.F.SG have.PRS.SBJV.3.SG here
'For example, for someone who does not have them here' (informant Idalinda, female, 62 yo, 4th grade, CLH)

(b) *E eu se tiver* *consciência,* [{nɐ̃ = *não*}] *não a*
and I.SBJ if have.FUT.SBJV.1.SG awareness NEG = CL.ACC.3.F.SG
engano
cheat.PRS.IND.1.SG
'And if I am aware, I will not trip her up' (informant Borromeu, male, illiterate, EXB)

Two interpretations are to be envisaged. According to one, if the first syllable has a monosegmental rhyme, more propitious conditions are provided for an onset to be added to the subsequent clitic word. Conversely, if the shape of the pronoun is planned prior to the production of the TCP sequence, then the nasal onset retroactively triggers the deletion of the glide in the preceding syllable. Thus, there is a backward effect of one of the segments onto the realization of the preceding sound material. A non-left-to-right production of TCP sequences is the prerequisite for this scenario.

Further observation of CORDIAL-SIN data demonstrates that in the vast majority of cases monophthongization is paired with another phonological process. Once the glide in the first syllable is deleted, vowel quality tends to get altered, too. The most frequent alteration affects the nucleus of *não* [nɐ̃w̃], which surfaces eventually as [nũ], as can be seen in (11a). This change is found in both of the variants. In TCP sequences with consonant-initial pronouns, the alteration of vowel quality takes place in 17 out of 45 cases. Before vowel-initial pronouns, it affects 26 out of 72 sequences. The correlation between monophthongization and the shift in the vocalic timbre is unilateral. Monophthongs are produced irrespective of whether the vowel gets altered or not (see 10a-b and 11a-b), but change in vowel quality is nearly always superimposed onto monophthongs. There is, however, a single instance of vowel alteration paired with a diphthongal mode of the realization of the rhyme (see 11c).

(11) (a) *que* *ele* {nũ = ***não***} ***a*** *matasse*
COMP he.SBJ NEG CL.ACC.3.F.SG kill.IPFV.SBJ.3.SG
'that he didn't kill her' (Belchior, male, 53 yo, PIC)

(b) *Há* {kɐ̃nɐʃ = ***quem as***} *tinja* *em preto*
be-there.PRS.IND.3.SG one-who.REL = CL.ACC.3.F.PL dye.PRS.SBJV.3.SG in black
'Some dye them black' (informant Emanuel, male, 73 yo, GRJ)

(c) *houve* {kɐ̃jnɐz = ***quem as***} {ɐplɐ̃'tow = *plantou*}
be-there.PRF.IND.3.SG one- WHO.REL plant.PRF.IND.3.SG
'There was someone who had planted it' (informant Amélia, female, 63yo, 3rd grade completed, FLF)

The processes affecting TCP sequences discussed thus far have had no erosive effect on the phonological conditioning. Two concurrent realizations of the sequence remain essentially independent of monophthongization and alteration of vowel quality. By contrast, the third process alters the sequence to the point where an adjustment of preliminary conditions may become necessary. Four TCP sequences are subject to the loss of the nasal resonance in the nucleus of the first syllable. This shift is invariably paired with the nasal epenthesis in subsequent clitic pronouns. In addition, the loss of nasality in the rhyme of the first syllable is always concomitant to monophthongization and alteration of vowel quality. This mode of realization is found only in the mainland of Portugal. The districts are Aveiro (2 occurrences), Vila Real and Bragança (see 12a–d). In each of the cases, it is the negative marker *não* that is distorted. On occasion, the loss of the nasal resonance is found also in other syntactic configurations (see the cleft sentence in 13).

(12) (a) {ˈnunɐʃ = *Não as*} *furávamos*.
NEG = CL.ACC.3.F.PL drill.IPFV.IND.1.PL
'We were not drilling them' (informant Gotardo, male, no further data, STA)

(b) *Se* {ˈnunɐ = *não a*} *roubarem*!
If NEG = CL.ACC.3.F.SG rob.FUT.SBJV.3.PL
'If they do not rob her' (informant Camila, female, no further data; OUT)

(c) *Sabe*? {ˈnunɐ = *Não a*} *queria* *porque ele…*
know.PRS.IND.3.SG NEG = CL.ACC.3.F.SG want.IPFV.IND.1.SG because he.SBJ
'You know, I didn't want it because he …' (informant Arquibaldo, male, 75 yo, can read, COV)

(d) *E* {ˈnunɐ = *não a*} *queria*!
and NEG = CL.ACC.3.F.SG want.IPFV.IND.1.SG
'And I didn't want it / her' (Arquibaldo)

(13) *e eles é que* {ɛɾɐnuʃ = *eram os*} *patrões*.
and they.SBJ be.PRS.IND.3.SG that be.IPFV.IND.3.PL = DEF.ART.M.PL boss.PL
'THEY were the bosses' (Arquibaldo)

This cumulative phonological shift poses a challenge to the previously defined phonological side of the use of the alveolar nasal. Neither of the prerequisites (the diphthongal rhyme or its nasal resonance) is met in the first of the two juxtaposed syllables. Morphosyntactic constraints notwithstanding, a simple oral vowel ending the preceding word does not warrant the alveolar nasal any more than any other consonant. Proof of this is two-word sequences with a clitic element in which CORDIAL-SIN speakers insert intervocalic glides or liquids across words (see 14a–b).

(14) (a) {‘ajuʒ = *Há-os*} *de ferro*, {‘ajuʒ = *há-os*}
be-there.PRS.IND.3.SG = CL.DEM.M.PL of iron be-there.PRS.IND.3.SG = CL.DEM.M.PL
de pau, {nu = *não*} *é*?
of stick, isn’t
‘Some are of iron, some are of wood, aren’t they’? (informant Astreia, female, no further data, OUT)

(b) *A gente* {PH|’kɔβriɬɐʒ = *cobre-as*} *bem*
ART.DEF.F.SG people cover.PRS.IND.3.SG = CL.ACC.3.F.PL well
‘We cover them well / People cover them well’ (informant Brás, male, 53yo, TRC)

Ergo the selection of the alveolar nasal must have been planned beforehand, i.e., before preparing words in speech production. This is not in keeping with the view that sound segments are activated sequentially, from left to right. One might wonder, then, if the production of TCP sequences in (12a–d) is really a feedforward process, with neatly staged phases: lexical selection, morphological and phonological encoding, phonetic encoding, and neuromuscular control of articulators (Levelt et al., 1999, pp. 3–4).

Two explanations for the nasal epenthesis after a simple oral vowel in TCP sequences can be advanced. These explanations are not mutually exclusive, although one of them would call for a thorough experimentally based confirmation. Generally, CORDIAL-SIN offers data produced by elderly speakers. Research over accuracy and speech rate point to substantial differences depending on the speaker’s age. In the research task designed and carried out by Bilodeau-Mercure and Tremblay (2016), speakers’ orofacial physiology – lip and tongue muscular flexibility and endurance, as well as tactile sensitivity – has been investigated. Older informants (66–85 yo) were demonstrated to have a reduced speech rate compared to younger speakers. In articulation, they were shown to run up against particular difficulties producing nasal vowels. They were trying to remedy this by either slowing down or producing an oral vowel – nasal consonant sequence instead of nasal vowels. This effect has been blamed on an age-related motor disorder, namely the decline of lip endurance.

As for CORDIAL-SIN data, the age of the remaining denasalizing informants is not revealed, except for the informant Arquibaldo. Yet many of the interviewees are around 60 years old or older. If the result of the experiment is correct, then the use of the nasal-initial allomorph can be seen as an attempt to make up for the loss of nasal resonance in the rhyme of the preceding syllable. Investigating this effect in CORDIAL-SIN speakers at present is probably an unfeasible task. Yet, examining other elderly informants is a possible direction for further research.

Another explanation for the selection of the alveolar nasal rather than of any other consonant is cue-based. Nasal onsets are obligatory in enclisis after plural verb forms ending in a nasal diphthong. Their perceptual salience in combinations involving clitic pronouns reaches such a

degree that they can be extended over the preverbal domain even after a monosegmental and denasalized rhyme.

7. Discussion

Explanations formulated in terms of varying syntax-phonology mappings presuppose speakers' individually selected and varying methods of language processing. This approach is synchronic, and so independent of the previous developments of the TCP sequence. Meanwhile, the distribution of nasal-initial preverbal pronouns across particular interview sites is uneven, thus inviting history-based explanations of the origin of the variation. Non-canonical forms are clustered in two areas. One of them comprises northern municipalities and the northernmost ridges of the mainland of Portugal. The gist of the historical continuity hypothesis is the presence of consonant-initial preverbal accusative pronouns in the scriptural tradition of the Entre-Minho-e-Douro region and southern Galicia, with the attestations dating back to as early as the 13th to 16th centuries. Obviously, the reliance on historical data in accounting for present-day variation hinges essentially on how transmission of linguistic habits is conceived of.

7.1. Historical continuity gone astray

In accordance with the competing grammars approach (Kroch, 2001; Galves & Kroch, 2016, pp. 487–488), individuals displaying variation in the realization of the TCP sequence, must have instantiated at some point different phonology-syntax mappings in a kind of internalized diglossia. Competition arises when speakers have been exposed to both a vernacular (with the nasal-initial allomorph) and a learned grammar (vowel-initial variant). A sufficient amount of evidence is a prerequisite to developing two simultaneous and alternative rankings of parameters. Conversely, if a single realization of either of the forms is steadily maintained in individual speech production, no diglossia can be applied. In this approach, language acquisition (parameter setting; Battye & Roberts, 1995, p. 8) is the decisive phase in bringing about grammatical change. Individual grammars are computed on the basis of both human biological endowment and a multitude of cues available to new generations of language users via their parents' speech (Roberts, 2007, p. 124). Innovations can be equated with parameter resettings, i.e., micro-scale transitions between stable successive individual grammars. The principal mechanism prompting the change is the faulty imitation (by abductive reasoning) of the older generation's speech habits. As the internalized grammar of their parents is not directly available to children, its reconstruction is unavoidably imperfect. Thus, infant grammar moves away from the original one. Observable change is explicitly declared to be piecemeal, gradual, and chaotic (Lightfoot, 1999, p. 105), which portrays well the actual usage of preverbal accusative clitics by CORDIAL-SIN informants.

At present, the viability of this approach will be matched against the interplay of vowel- and consonant-initial allomorphs documented in history.

In her work devoted to Galician-Portuguese written texts, Maia (1986, p. 671) describes nasal-initial accusative pronouns as forms that always occur after a nasal-final word (poor attestation of vowel-initial forms after the preceding negative marker, e.g., *nõ as aguardando* 'without waiting for them' is noted accordingly). An assimilatory process, actuated by the nasal segment in the final syllable of the preceding word, is declared to have affected *lo*, *la*, *los*, *las* pronominal forms rather than empty-onset ones (see Martins, 1994, p. 214, "*lo*, *los*, *la*, *las* são as formas do pronome acusativo átono que, do ponto de vista evolutivo, precedem as formas *o*, *os*, *a*, *as*" 'from the evolutionary point of view, *lo*, *los*, *la*, *las* are atonic accusative pronominal forms that are earlier than *o*, *os*, *a*, *as*'). Thus, two consonant-initial variants used to exist in parallel before the advent and the subsequent spread of vowel-initial forms.

Tellingly, Maia's analysis contains no reference to the syntactic category of the preceding expression. This is tantamount to saying that the choice of either consonant-initial or vowel-initial realization must have been made in the phonological module. Morphosyntactic conditioning played virtually no role in the selection of the onset. The only incidence of syntax involved the linear ordering of expressions containing suitable sound material. Thus, inflected verb forms were but one of the possible types of hosts for nasal-initial accusative clitics. Examples quoted by Maia (1986, p. 672) leave no doubt as to the category-blind attachment of clitic pronouns with nasal onsets: *e nõnaſ vẽderdes nẽ doardes* 'and you will neither sell them nor give them as a dowry', *nẽ nas dar a noble caualeyro* 'nor to give them to the gentle knight'; *dalos o abbade a quen nos laure* 'the abbot gives them to whoever can cultivate them', *poſerõno todo* 'they put all of it' (see also Martins 1994, pp. 217–218). This analysis appears, then, to provide a compelling evidence of historical ties between old speech habits and their present-day outcomes in the North of Portugal and in Galicia.

Still, the scriptural production of the subsequent centuries reveals the decline in the number of diphthong-final or nasal-final words that allowed nasal epenthesis in ensuing preverbal pronouns. In the 16th-century *Crónica dos Reis de Bisnaga*, concurrent realizations are interspersed only after the relative/interrogative *quem* and the negative marker *não*. The query for the right-hand side context of *sem* and *nem* has brought hits involving vowel-initial accusative pronouns only. These results are indicative of an increasing role of morphosyntactic conditioning in the processing of preverbal clitic pronouns. Conversely, phonological criteria ceased to be felt as crucial. The examples below, retrieved from the *Corpus Informatizado do Português Medieval* (Xavier, 1998, coord.), illustrate the point.

(15) (a) … *tomou* *as* *terras* *a* ***quem*** ***as*** *tinha* *contra*
take.PRF.IND.3.SG ART land.PL to who.REL CL.ACC.3.F.PL AUX.IND.3.SG against
razão *tomadas* *a* *elrey*
reason take.PST.PTCP.F.PL to king
'…took the parcels of land from whom had abusively got hold of them from the King' (Título 4)

(b) …, *não* *sey* ***quem no*** *possa* *contar*
NEG know.PRS.IND.1.SG who = CL.ACC.3.M.SG can.PRS.SBJV.3.SG tell.INF
pera *que* *seja* *crido*
so that be.PRS.SBJV.3.SG trust.PST.PTCP.M.SG
'I don't know who could tell it so that it/he is believed' (Título 8)

(c) … *o* *que* *aveis de* *fazer* *ojee,* ***não*** ***o***
DEM REL have.to.PRS.IND.2.PL do.INF today NEG CL.ACC.3.M.SG
deixeis *pera* *amanhaa*
leave.IMP.2.PL for tomorrow
'what you are supposed to do today don't leave it until tomorrow'(Título 4)

(d) …, *e* ***não no*** *poderão* *tirar* *d* *este*
and NEG = CL.ACC.3.M.SG can.PFV.IND.3.PL draw.INF from DEM.PROX.M.SG
proposyto *todollos* *seus* *gramdes*
intention all.M.PL = ART.DEF.M.PL POSS.3.PL nobleman.PL
'And none of his noblemen was able to drive him away from this idea' (Título 20)

The primacy of the morphosyntactic conditioning at the expense of the phonology-based one reached its peak in the 17th century.[6] The query of the resources of the Biblioteca Nacional de Portugal – Biblioteca Nacional Digital (BND), encompassing 11 integral texts published between 1614 and 1697,[7] brought a single occurrence of nasal-initial 3rd person pronoun in the preverbal

[6] The loss of nasal onsets in preverbal clitic pronouns in that period has to be put in correspondence with interpolation (clitic-verb non-adjacency). In Classical Portuguese, this linear pattern involved a single expression – the negative marker *não*. It used to occur adjacent to expressions checked for Tense to produce sentential negation (Matos, 1999, Matos, 2003, pp. 774–776). Interpolation has never become obligatory. Instead, clitic-verb adjacency and non-adjacency could be applied interchangeably, according to authors' preferences (Martins, 1994, p. 307). The reversal of the order of syllables (*não o* > *o não*) did away with phonological conditions enabling speakers to fill the onset of preverbal pronouns (Nkollo, 2021, pp. 484–485, demonstrates that 3rd person accusative clitics tended to be paired with interpolation at a larger scale than the other pronouns, whose distribution over *não*-cl-Verb and cl-*não*-Verb sequences was more balanced). The effect of the *o não* order on the form taken by preverbal clitic pronouns was immediate, as shown by example (16).

[7] The following texts have been used: 1614 Nicolao Agostinho. *Relaçam summaria da vida* … https://purl.pt/11490; 1625 Bertolameu Guerreiro *Jornada dos vassalos da coroa de Portugal* https://purl.pt/17352; 1639 *Regimento dos Juizes das Aldeas* … https://purl.pt/14985; 1640 Diogo de Santana. *Relaçam verdadeira do milagroso* … https://purl.pt/16740; 1645 Jorge da Costa. *Sermaõ do jubileo geral* … https://purl.pt/20760; 1650 Bento Teixeira Feio. *Relaçam do naufragio que fizeram* … https://purl.pt/28100; 1660 Padre Antonio Vieira. *Copia de huma carta para ElRey* …

domain. It comes from a compilation of judicial regulations, the *Regimento dos Juizes das Aldeas …* (1639), which displays an antiquated grammar (multiple interpolations, i.e., more than one constituent separating the clitic from the verb; see Martins 2016b, pp. 421–423, the inflected infinitive used with an injunctive value; Carvalho 2015, pp. 669, 673, and variable past participles in compound tense forms).

(16) … ***naõ*** ***no*** ***fazendo*** *deſde o* *dia* *que* *lhe*
NEG = CL.ACC.3.M.SG do.PRS.PTCP since ART.DEF.M.SG day REL CL.DAT.3.SG
puzerem *a* *dita* *pena …*
put.FUT.SBJV.3.PL ART.DEF.F.SG say.PST.PTCP.F.SG punisment
'… not doing this from the day when he receives this punishment' (p. 6)

Thus, dwindling preverbal contexts in which nasal epenthesis used to be documented pose several problems to the continuity hypothesis. Obviously, the disjunction between the speech production of the inhabitants of remote areas of the Entre Minho e Douro region and the grammar of the then urban centers might not be excluded. Likewise, a given grammatical habit is not unlikely to resume after a period of dormancy (see D'Alessandro, 2014 for the unexpected advent of *annə*, borrowed from the Italian compound tense auxiliary, to replace an almost extinct impersonal pronoun *nomə* < Lat. HOMO in Abruzzese, a southern Italian variety spoken in the central part of the region). Be that as it may, the scarcity of attestations of the nasal-initial allomorph in the scriptural production of the 17th-century EP must not be overlooked. For the time being, an uninterrupted transmission of the nasal epenthesis in preverbal pronouns over successive generations of speakers is no more than a guesswork.

Another hindrance has to do with the archipelago of the Azores, the second area with a noteworthy presence of nasal-initial preverbal clitics. Written records of the history of the vernacular clitic grammar on particular islands are badly missing. Starting from the 15th century, the Azores are known to have been populated by settlers coming prevailingly from mainland Portugal. "A maioria dos povoadores seria proveniente de Portugal continental, merecendo particular referência os contributos do Algarve, Alentejo e Minho; no entanto, tais contributos não terão ocorrido de igual modo nas diferentes ilhas" (Lima, 2009, p. 229, 'the majority of settlers appear to have come from the mainland of Portugal, with Algarve, Alentejo and Minho having provided most significant numbers of immigrants; nevertheless, newcomers from these

https://purl.pt/16556; 1686 P. Antonio Carvalho da Costa. *Compendio Geografico …* https://purl.pt/13823; 1692 Vicente Barbosa C.R. *Compendio da relaçam …* https://purl.pt/12025; 1697 Padre M. Ayres de Almeyda. *Sermam do acto da fee que se celebrou …* https://purl.pt/13488 and 1735 *Cartas* do P. Antonio Vieyra https://digital.bbm.usp.br/handle/bbm/4527. The last text is a posthumous edition of the original epistolary work of the Jesuit father dating entirely from the 17th century. This text, the only one in the list, has been retrieved from a Brazil-based source.

regions are thought to have not been evenly divided among different islands'; translation – MN). The historical continuity hypothesis would require the speech habits of the majority of migratory waves to have been superseded by the grammar of the settlers from the Minho region. As for now, no evidence for this direction of grammatical change is available.

Putting aside the lack of sufficiently probative documentation across various places and periods, the representativeness of the historical sources proves an even more serious obstacle. What is more, the analyses highlighting language acquisition as the primary driver of language change and conducted in terms of internalized diglossia and parameter reranking require samples of infant-directed speech in successive generations of speakers to be systematically compared with one another. In an attempt to remedy the lack of necessary data, Kodner (2019, pp. 4–12) tries to infer the characteristics of infant-directed speech on the basis of modern and historical literary corpora. His results rely on the sameness of frequency values in both types of sources. In particular, high token frequency elements are expected to be similar irrespective of the addressees and communicative circumstances. Moreover, the extension of learning algorithm sampled from the two kinds of data consistently points to the same learning outcomes. Therefore, the core cues acquired in childhood are thought to be plausibly reconstructible on the basis of non-infant corpora. Meanwhile, Kodner's pioneering study seems to be concerned primarily with the acquisition of the lexicon. It remains for future investigation to see whether these results can successfully grasp the evolution of TCP sequences in previous generations of EP learners.

7.2. Nasal epenthesis as an externally-motivated analogical change

Instead of capturing variation documented in CORDIAL-SIN in terms of historical continuity, a more moderate solution is advanced here, making use of the concept of markedness. Rather than positing the transmission of nasal epenthesis in preverbal accusative clitics from one generation to another, it emphasizes the sameness of the processing of TCP sequences at various moments in the history and in various communicative circumstances. This theoretical apparatus has been inspired by H. Andersen's actualization theory.

For Andersen, grammatical innovations are either internally (from below) or externally motivated (from above). "Typically in internally motivated change, an innovation is extended to unmarked contexts earlier than to corresponding marked contexts" (Andersen, 2008, p. 36). If speech act circumstances are considered, unmarked contexts correspond to spontaneous communication and informal text genres close to orality. In morphosyntax, unmarked contexts are routinely basic and most frequent ones, e.g., main clauses as opposed to subordinate clauses, simple vs. compound tenses. Solutions initiated in these contexts can subsequently gain ascendancy in marked situations and marked morphosyntactic settings, thus ceasing to be felt as informal or colloquial. Innovations from above, on the other hand, depart from marked contexts,

both pragmatically and grammatically. They subsequently exert pressure (i.e., gain currency and become largely adopted in a given speech community) on spontaneous communication. This scenario materializes in a situation of language contact (e.g., transfers from the standard language into the dialects, syntactic calques from a foreign language of prestige; see Andersen, 2006, pp. 68–70). In morphosyntax, more elaborate solutions tend to spill over simpler contexts. With the passing of time, the elevated character of these innovations is no longer perceived. At this point, they become unmarked (Andersen, 2001, p. 238).

In this paper, communicatively unmarked contexts correspond to the speech production of CORDIAL-SIN informants. Onset insertion in preverbal accusative pronouns is likewise a feature proper to an oral, unmarked register. By contrast, the absence of nasal-initial forms is thought to be a learned variant (externally-motivated innovation), induced by contact with standard EP. Present-day Portuguese dialects must not be seen as repositories for phonological layers characteristic of Galician-Portuguese or Middle Portuguese texts. Instead, the reverse is thought to be the case: nasal epenthesis in preverbal accusative pronouns documented in medieval and Renaissance texts is thought to have reflected speech habits characteristic of the oral register. Obviously, the degree of communicative markedness of Galician-Portuguese and Portuguese texts prior to the 17th century outranks the markedness of spontaneous dialectal speech. Still, since these texts contain instances of nasal epenthesis, their authors must have been exposed to two divergent grammars, with empty onsets alongside nasal onsets. Thus, pragmatically, the use of nasal-initial allomorphs in Galician-Portuguese or Middle Portuguese is considered an innovation advancing from below.

To conclude, it remains to ascertain whether consonant-initial preverbal accusative pronouns are unmarked also in morphosyntax. The answer depends on how variation in clitic placement in today's EP is characterized, and that is not a consensual issue. Some authors claim that proclisis is triggered by the properties of the (higher) left periphery hosting complementizers and topicalized or focused elements (Madeira, 1992; Ledgeway, 2016, p. 1013). Other scholars point to proclisis triggers as the elements that have the property of attracting the clitic to the pre-Infl domain (Duarte & Matos, 2000). As a result, for some scholars proclisis is a less marked pattern and enclisis corresponds to a more complex derivation, with the V movement to a higher functional category. In other proposals, enclisis is primary and proclisis is computationally costlier (a last resort Movement; see Duarte & Matos, 2000, pp. 135–137), because it is activated by specific syntactic elements.

The puzzle can be solved if data from EP L1 acquisition are looked into. Enclisis is the dominant placement in children until approximately forty-two months of age (Duarte et al., 1995). In that period, enclisis is nearly indiscriminately applied, regardless of the presence of proclisis triggers in a sentence (Costa et al., 2015, pp. 14–17; Costa et al., 2016, pp. 440–446).

Therefore, enclisis, alongside phonological shifts accompanying it, is thought to be the default solution. By contrast, proclisis is the marked pattern. Nasal epenthesis in preverbal accusative pronouns is an overgeneralization of a sandhi phenomenon from enclisis (unmarked context) in syntactically more complex environments (proclisis; marked context). Thus, in morphosyntax the innovation progresses from below as well. Syntactic conditions under which vowel-initial preverbal accusative pronouns must be used after diphthong-final words require an explicit learning. This variant appears to be problematic for some CORDIAL-SIN informants.

Concluding remarks

In Portuguese linguistics, dialectal speech production has attracted scholars' interest and inspired a number of analyses conducted within the generative framework. Obviously, syntax represents the dominant thread in these studies. Thus, there is still room for phenomena that do not fall directly within the scope of syntax, and whose analysis is carried out with the aid of a different theoretical apparatus. Nasal epenthesis across words is thought to be one of such issues.

Moreover, in this study the corpus has been used to analyze the non-canonical realization of prosodically deficient pronouns, to determine the geographical distribution of phonological distortions and the concomitant intra-speaker variation. Speech samples stored in the verbatim transcription section track segment after segment how speakers capitalize on a set of available structures to go beyond the bounds of the codified grammar of their language. The decision on whether a nasal epenthesis materializes in a given TCP sequence is taken on the basis of the following available structures: (i) sound series made of a nasal diphthong and a subsequent simple oral vowel; (ii) homonymy of accusative clitic pronouns and definite articles; (iii) prosodic deficiency of both classes of expressions; (iv) obligatory nasal epenthesis in postverbal accusative pronouns after diphthong-final syllables. These cues prove instrumental in inducing speakers to add the alveolar nasal in preverbal clitic pronouns. All in all, the abusive nasal epenthesis is an analogically induced extension.

As for the historical clues to present-day nasal-initial preverbal accusative pronouns, rather than positing continuous transmission of speech habits, the paper emphasizes the sameness in the treatment of these expressions at various moments and in different communicative circumstances. It remains to be seen how the analogical mechanisms singled out above interact with other variables – the type of the proclisis trigger (preference given to nasal-initial variant after *quem*), the speaker's age and education level. Concomitant phonological processes, especially the loss of nasal resonance, may bring new insights into the nature of the relationship, thus offering a more comprehensive outlook on the variation documented in CORDIAL-SIN clitic pronouns.

Competing interests

The author has no competing interests to declare.

References

Andersen, H. (2001). Actualization and the (Uni)directionality of change. In H. Andersen (Ed.), *Actualization: Linguistic Change in Progress. Papers from a workshop held at the 14th International Conference on Historical Linguistics, Vancouver, B.C. 14 August 1999*. Amsterdam-Philadelphia: John Benjamins, 225–248. DOI: https://doi.org/10.1075/cilt.219.11and

Andersen, H. (2006). Synchrony, diachrony, and evolution. In O. Nedergaard Thomsen (Ed.), *Competing Models of Linguistic Change: Evolution and beyond* (pp. 59–90). Amsterdam-Philadelphia: John Benjamins. DOI: https://doi.org/10.1075/cilt.279.07and

Andersen, H. (2008). Grammaticalization in a speaker-oriented theory of change. In Th. Eythórsson (Ed.), *Grammatical Change and Linguistic Theory: The Rosendal Papers* (pp. 11–40). Amsterdam-Philadelphia: John Benjamins. DOI: https://doi.org/10.1075/la.113.02and

Anderson, R. T. (2002). Onset clusters and the sonority sequencing principle in Spanish: A treatment efficacy study. In F. Windsor, M. L. Kelley & N. Hewlett (Eds.), *Investigations in clinical phonetics and linguistics* (pp. 213–224). Mahwah, NJ: Lawrence Erlbaum Associates.

Auer, P., Barden, B., & **Grosskopf, B.** (1998). Subjective and objective parameters determining 'salience' in long-term dialect accommodation. *Journal of Sociolinguisitcs, 2*(2), 163–187. DOI: https://doi.org/10.1111/1467-9481.00039

Battye, A., & **Roberts, I.** (1995). Introduction. In A. Battye & I. Roberts (Eds.), *Clause Structure and Language Change* (pp. 3–28). Oxford: Oxford University Press.

Bilodeau-Mercure, M., & **Tremblay, P.** (2016). Age Differences in Sequential Speech Production: Articulatory and Physiological Factors. *Journal of the American Geriatrics Society, 64*(11), e177–e182. DOI: https://doi.org/10.1111/jgs.14491

Bisol, L. (2008). The contrastive nasality in Portuguese, *Rutgers Optimality Archive*. Available at http://roa.rutgers.edu (23 September 2021).

Bybee, J., & **Torres Cacoullos, R.** (2008). Phonological and Grammatical Variation in Exemplar Models. *Studies in Hispanic and Lusophone Linguistics, 1*(2), 399–413. DOI: https://doi.org/10.1515/shll-2008-1026

Carvalho, M. J. (2015). Para a história do infinitivo flexionado português: uma abordagem semântico-pragmática [Towards the History of Portuguese Inflected Infinitive: a semantic-pragmatic approach]. *Zeitschrift für romanische Philologie, 131*(3), 664–689. DOI: https://doi.org/10.1515/zrp-2015-0047

Collischonn, G., & **Wetzels W. L.** (2016). Syllable Structure. In W. L. Wetzels, S. Menuzzi & J. Costa (Eds.), *The Handbook of Portuguese Linguistics* (pp. 86–106). West Sussex: John Wiley & Sons. DOI: https://doi.org/10.1002/9781118791844.ch6

Costa, J., Fiéis, A., & **Lobo, M.** (2015). Input variability and late acquisition: Clitic misplacement in European Portuguese. *Lingua, 161*, 10–26. DOI: https://doi.org/10.1016/j.lingua.2014.05.009

Costa, J., Fiéis, A., & **Lobo, M.** (2016). A aquisição dos pronomes clíticos no português L1 [The Acquisition of Clitic Pronouns in Portuguese L1]. In A. M. Martins & E. Carrilho (Eds.), *Manual de Linguística Portuguesa* (pp. 431–452). Berlin-Boston: De Gruyter. DOI: https://doi.org/10.1515/9783110368840-018

D'Alessandro, R. (2014). Death and contact-induced rebirth of impersonal pronouns. A case study, *Probus, 26*(2), 249–274. DOI: https://doi.org/10.1515/probus-2014-0009

Dalrymple, M. (2001). *Lexical Functional Grammar*. New York: Academic Press. DOI: https://doi.org/10.1163/9781849500104

Duarte, I., & **Matos, G.** (2000). Romance Clitics and the Minimalist Program. In J. Costa (Ed.), *Portuguese Syntax. New Comparative Studies* (pp. 116–142). Oxford: Oxford University Press.

Duarte, I., Matos, G., & **Faria, I. H.** (1995). Specificity of European Portuguese Clitics in Romance. In I. H. Faria & M. J. Freitas (Eds.), *Studies on the Acquisition of Portuguese* (pp. 129–154). Lisboa: APL/Colibri.

Fagan, D. (1992). On the history of Portuguese /u̯, ēi/: The case for spontaneous diphthongization. *WORD, 43*(3), 383–397. DOI: https://doi.org/10.1080/00437956.1992.12098315

Fernández Rei, E. (2002). *Regras fonóloxicas posléxicas e regras precompiladas de alomorfia sintagmática: domínios prosódicos en galego* [Postlexical Phonological Rules and Precompiled Rules of Sintagmatic Allomorphy: prosodic domains in Galician], PhD Thesis, The University of Santiago de Compostela, Facultade de Filoloxía.

Frota, S. (2013). *Prosody and Focus in European Portuguese Phonological Phrasing and Intonation*. London-New York: Routledge Taylor & Francis Group. DOI: https://doi.org/10.4324/9781315054384

Frota S., D'Imperio, M., Elordieta, G., Prieto, P., & **Vigário, M.** (2007). The phonetics and phonology of intonational phrasing in Romance. In P. Prieto, J. Mascaró & M.-J. Solé (Eds.), *Segmental and Prosodic Issues in Romance Phonology* (pp. 131–153). Amsterdam-Philadelphia: John Benjmains. DOI: https://doi.org/10.1075/cilt.282.10fro

Galves, C., Britto, H., & **Sousa, M.** (2005). The Change in Clitic Placement from Classical to Modern European Portuguese: Results from the Tycho Brahe Corpus. *Journal of Portuguese Linguistics, 4*(1), 39–67. DOI: https://doi.org/10.5334/jpl.166

Galves, C., & **Kroch, A.** (2016). Main Syntactic Changes from a Principle-and-Parameters View. In W. L. Wetzels, S. Menuzzi & J. Costa (Eds.), *The Handbook of Portuguese Linguistics* (pp. 487–503). West Sussex: John Wiley & Sons. DOI: https://doi.org/10.1002/9781118791844.ch27

Harris, A., & **Campbell, L.** (1995). *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511620553

Hirschbühler, P., & **Labelle, M.** (2000). Evolving Tober-Mussafia effects in the placement of French clitics. In S. Dworkin & D. Wanner (Eds.), *New Approaches to Old Problems. Issues in*

Romance Historical Linguistics (pp. 165–182). Amsterdam-Philadelphia: John Benjamins. DOI: https://doi.org/10.1075/cilt.210.15hir

Hoekstra, E., & **Versloot, A. P.** (2019). Factors promoting the retention of irregularity. On the interplay of salience, absolute frequency and proportional frequency in West Frisian plural morphology. *Morphology*, *29*, 31–50. DOI: https://doi.org/10.1007/s11525-018-9334-2

Kerswill, P., & **Williams, A.** (2002). 'Salience' as an explanatory factor in language change: Evidence from dialect leveling in urban England. In M. C. Jones & E. Esch (Eds.), *Language Change. The Interplay of Internal, External and Extra-Linguistic Factors* (*Contributions to the Sociology of Language 86*) (pp. 81–110). Berlin–New York: De Gruyter Mouton. DOI: https://doi.org/10.1515/9783110892598.81

Kodner, J. (2019). Estimating child linguistic experience from historical corpora. *Glossa: a journal of general linguistics, 4*(1), 122. DOI: https://doi.org/10.5334/gjgl.926

Kroch, A. (2001). Syntactic Change. In M. Baltin & Ch. Collins (Eds.), *The Handbook of Contemporary Syntactic Theory* (pp. 699–729). Malden, Mass./Oxford: Blackwell. DOI: https://doi.org/10.1002/9780470756416.ch22

Labelle, M. (2001). Trente ans de psycholinguistique [Thirty Years of Psycholinguistics]. *Revue québécoise de linguistique, 30*(1), 155–176. DOI: https://doi.org/10.7202/000516ar

Ledgeway, A. (2016). Clausal Complementation. In A. Ledgeway & M. Maiden (Eds.), *The Oxford Guide to the Romance Languages* (pp. 1013–1028). Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199677108.003.0040

Levelt, W. (1989). *Speaking: from intention to articulation*. Cambridge, Mass.: MIT Press.

Levelt, W., Roelofs, A., & **Meyer, A.** (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*(1), 1–38. DOI: https://doi.org/10.1017/S0140525X99001776

Lightfoot, D. (1999). *The development of language: Acquisition, change, and evolution*. Oxford: Wiley-Blackwell Publishers.

Lima, M. (2008). Povoamento e história demográfica dos Açores: o contributo da genética [The Settlement and the Demographic History of the Azores]. *Boletim do Núcleo Cultural da Horta, 17*, 227–241.

Luís, A. R. (2018). Casos de reanálise morfológica envolvendo estrutura morfotática e saliência fonética [Cases of Morphological Reanalysis Involving Morphotactic Structure and Phonetic Salience]. In J. Veloso, J. Guimarães, P. Silvano & R. Sousa-Silva (Eds.), *A linguística em diálogo. Volume comemorativo dos 40 anos do Centro de Linguística da Universidade do Porto* (pp. 247–264). Porto: Centro de Linguística da Universidade do Porto.

Luís, A. R., & **Kaiser, G.** (2016). Clitic Pronouns: Phonology, Morphology, and Syntax. In W. L. Wetzels, S. Menuzzi & J. Costa (Eds.), *The Handbook of Portuguese Linguistics* (pp. 210–233). West Sussex: John Wiley & Sons. DOI: https://doi.org/10.1002/9781118791844.ch12

Madeira, A. (1992). On Clitic Placement in European Portuguese. In H. van de Koot (Ed.), *University College London Working Papers in Linguistics 4*, 91–122. London: University College.

Maia, C. de A. (1986). *História do Galego-Português. Estado linguistico da Galiza e do Noroeste de Portugal desde o século XIII ao século XVI* (*Com referência à situação do galego moderno*) [The History of Galician-Portuguese. Linguistic situation of Galicia and North-Western Portugal from the 12th to the 16th century with a reference to the status of Modern Galician]. Coimbra: Instituto Nacional de Investigação Científica.

Martins, A. M. (1994). *Clíticos na História do Português* [Clitics in the History of Portuguese], PhD Thesis, University of Lisbon.

Martins, A. M. (2012). Clíticos na história do português à luz do teatro vicentino [Clitics in the History of Portuguese in the light of Gil Vicente's theatre]. *Estudos de Lingüística Galega, 3*, 83–109.

Martins, A. M. (2016a). Introdução: O português numa perspetiva diacrónica e comparativa [Introduction: a Diachronic and Comparative Outlook on Portuguese]. In A. M. Martins & E. Carrilho (Eds.), *Manual de Linguística Portuguesa* (pp. 1–40). Berlin-Boston: De Gruyter. DOI: https://doi.org/10.1515/9783110368840-003

Martins, A. M. (2016b). A colocação dos pronomes clíticos em sincronia e diacronia [Clitic Placement in Synchrony and Diachrony]. In A. M. Martins & E. Carrilho (Eds.), *Manual de Linguística Portuguesa* (pp. 410–430). Berlin-Boston: De Gruyter. DOI: https://doi.org/10.1515/9783110368840-017

Mateus, M., H., M., & **d'Andrade, E.** (2000). *The Phonology of Portuguese*. Oxford: Oxford University Press.

Matos, G. (1999). Negative Concord and the Scope of Negation. *Catalan Working Papers in Linguistics, 7*, 175–190.

Matos, G. (2003). Aspectos sintáticos da negação [Syntactic Aspects of Negation]. In M. H. M. Mateus, A. M. Brito, I. Duarte & I. H. Faria (Eds.), *Gramática da Língua Portuguesa* (6th ed., pp. 767–793). Lisboa: Caminho.

Mattos e Silva, R.V. (2006). *O Português Arcaico*: *fonologia, morfologia e sintaxe* [Old Portuguese: Phonology, Morphology and Syntax]. São Paulo: Editora Contexto.

Mattoso Câmara, Jr, J. (1971). *Problemas de linguística descritiva* [The Issues in Descriptive Linguistics] (4th ed.). Petropolis: Editora Vozes.

Miller, Ph., & **Monachesi, P.** (2003). Les pronoms clitiques dans les langues romanes [Clitic Pronouns in Romance Languages]. In D. Godard & A. Abeillé (Eds.), *Les langues romanes, problèmes de la phrase simple* (pp. 53–106). Paris: CNRS.

Namiuti, C. (2008). *Aspectos da história gramatical do português. Interpolação, negação e mudança [Aspects of Portuguese Grammar's History: Interpolation, Negation and Change]*, PhD thesis, Universidade Estadual de Campinas – Instituto de Estudos da Linguagem.

Namiuti-Temponi, C. (2006). Um estudo sobre o fenômeno da interpolação de constituintes na história do português [The Study on Constituents' Interpolation in the History of Portuguese]. *Cadernos de Estudos Lingüísticos, 48*(2), 171–194. DOI: https://doi.org/10.20396/cel.v48i2.8637177

Nkollo, M. (2020). Onset insertion across words in present-day dialectal European Portuguese. A case for competing grammars? *Revista Internacional de Lingüística Iberoamericana XVIII, 2*(36), 281–307.

Nkollo, M. (2021). Padrões de não-adjacência clítico-verbo no português clássico. A sintaxe nas interfaces com a fonologia [Patterns of clitic-verb non-adjacency in Classical Portuguese. Syntax-Phonology Interfaces]. In L. Schøsler, J. Härmä & J. Lindschouw (Eds.), *Actes du XXIX^e^ Congrès international de linguistique et de philologie romanes (Copenhague, 1–6 Juillet 2019)* (Vol. 1, pp. 479–490). Strasbourg: Editions de Linguistique et de Philologie.

Oliveira, P., Cruz, M., Paulino, N., & **Vigário, M.** (2017). Glide insertion to break a hiatus across words in European Portuguese. The role of prosodic, geographic and sociolinguistic factors. In P. Barbosa, M. da Conceição de Paiva & C. Rodrigues (Eds.), *Studies on Variation in Portuguese* (pp. 50–79). Amsterdam-Philadelphia: John Benjamins. DOI: https://doi.org/10.1075/ihll.14.02oli

Piel, J-M. (1989). A flexão verbal do português (Estudo de morfologia histórica) [Portuguese Verb Inflection: a study in the historical morphology]. In J.-M. Piel (Ed.), *Estudos de linguística Galego-Portuguesa* (pp. 213–244). Lisboa: Imprensa Nacional-Casa de Moeda.

Piggott, G. L. (1992). Variability in feature dependency: The case of nasality. *Natural Language and Linguistic Theory, 10*, 33–77. DOI: https://doi.org/10.1007/BF00135358

Ramos, M. A. (1985). Nota Lingüística: critérios de edição: normas de transcrição [A Linguistic Note: editorship's criteria, transcription norms]. In E. Gonçalves & M. A. Ramos (Eds.), *A lírica galego-portuguesa*: *textos escolhidos* (pp. 81–127). Lisboa: Editorial Comunicação.

Roberts, I. (2007). *Diachronic Syntax*. Oxford: Oxford University Press.

Sampson, R. (1983). The Origin of Portuguese *-ão*, *Zeitschrift für Romanische Philologie, 99*(1–2), 33–68. DOI: https://doi.org/10.1515/zrph.1983.99.1-2.33

Sampson, R. (2016). Sandhi phenomena. In A. Ledgeway & M. Maiden (Eds.), *The Oxford Guide to the Romance Languages* (pp. 669–680). Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199677108.003.0040

Sandalo, F., & **Galves, C.** (2013). Clitic placement and grammaticalization in Portuguese. In Ch. Meklenborg Salvesen & H. P. Helland (Eds.), *Challenging Clitics* (pp. 119–134). Amsterdam-Philadelphia: John Benjmains. DOI: https://doi.org/10.1075/la.206.05san

Santos, R. S., & **Vigário, M.** (2016). The Phonology–Syntax Interface. In W. L. Wetzels, S. Menuzzi & J. Costa (Eds.), *The Handbook of Portuguese Linguistics* (pp. 125–140). West Sussex: John Wiley & Sons. DOI: https://doi.org/10.1002/9781118791844.ch8

Seiler, G. (2004). On three types of dialect variation and their implications for linguistic theory. Evidence from verb clusters in Swiss German dialects. In B. Kortmann (Ed.), *Dialectology meets Typology. Dialect Grammar from a Cross-Linguistic Perspective* (pp. 367–399). Berlin-New York: Mouton de Gruyter.

Seiler, G. (2006). The role of functional factors in language change. An evolutionary approach. In O. Nedergaard Thomsen (Ed.), *Competing Models of Linguistic Change: Evolution and beyond* (pp. 163–182). Amsterdam-Philadelphia: John Benjamins. DOI: https://doi.org/10.1075/cilt.279.11sei

Selkirk, E. (1984). On the major class features and syllable theory. In M. Aronoff & R. T. Oehrle (Eds.), *Language Sound Structure: Studies in Phonology* (pp. 107–136). Cambridge, Mass: MIT Press.

Thornton, A. (2011). Overabundance (Multiple Forms Realizing the Same Cell): A Non-Canonical Phenomenon in Italian Verb Morphology. In M. Maiden, J.-Ch. Smith, M. Goldbach & M.-O. Hinzelin (Eds.), *Morphological Autonomy. Perspectives From Romance Inflectional*

Morphology (pp. 358–381). Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199589982.003.0017

Vigário, M. (1999). Clitização no português europeu: uma operação pós-lexical [Clitic Attachment in European Portuguese – a postlexical operation]. In *Actas do XVI Encontro Nacional da Associação Portuguesa de Linguística* (pp. 577–598). Braga.

Vigário, M. (2003). *The Prosodic Word in European Portuguese*. Berlin-New York: De Gruyter. DOI: https://doi.org/10.1515/9783110900927

Zwicky, A. M., & **Pullum, G.** (1983). Cliticization vs. Inflection. English *n't*. *Language, 59*(3), 502–513. DOI: https://doi.org/10.2307/413900

Corpora

Martins, A. M. (coord.). (2010). *CORDIAL-SIN: Syntax-oriented Corpus of Portuguese Dialects*, Lisbon: Centre of Linguistics of the University of Lisbon. Available at URL http://www.clul.ul.pt/en/resources/411-cordial-corpus [Last access September 17 2021].

Xavier, M. F. (coord.). (1998). *CIPM*: *Corpus Informatizado do Português Medieval* [Electronic Corpus of Medieval Portuguese]. Lisbon: Linguistics Centre, Universidade Nova de Lisboa. Available at URL http://cipm.fcsh.unl.pt/ [Last access February 19 2021].

Digital sources

Biblioteca Nacional Digital (BND): https://bndigital.bnportugal.gov.pt/

https://d-maps.com