



# Syntax and word-specific phonetics: the origins of the allomorphs of the Galician definite article

Francisco Dubert-García

Instituto da Lingua Galega-Universidade de Santiago de Compostela  
[francisco.dubert@usc.es](mailto:francisco.dubert@usc.es)

Submitted: 2 de diciembre de 2013. Accepted: 23 de mayo de 2014. Available on line: 30 de marzo de 2015

**Citation / Cómo citar este artículo:** Dubert-García, F. (2014). Syntax and word-specific phonetics: the origins of the allomorphs of the Galician definite article. *Loquens*, 1(2), e013. doi: <http://dx.doi.org/10.3989/loquens.2014.013>

**ABSTRACT:** The Galician definite article seems to be a case of a ditropic clitic which, phonologically, attaches to the preceding word while maintaining direct, constitutive syntactic relations with following words. The article has three allomorphs whose selection is conditioned by the ending of the preceding word. In this paper I study the historical origins of these allomorphs. They are the result of assimilatory processes due to constraints governing the structure of the syllable and intersyllabic contact; the domain of these constraints, however, was restricted to the definite article and the 3rd person accusative clitic pronoun, leaving unaffected other words having the same phonetic sequences in apparently similar contexts. In this paper I will demonstrate the convenience of explaining the origins of these alternations by means of usage based models designed to explore the role of frequency in the emergence of grammar and facilitate the natural combination of information from the language's semantic, syntactic and phonological components.

**Keywords:** Galician; syllable; clitics; word-specific phonetics; exemplar models.

**RESUMEN:** *Sintaxis y fonética particular de ciertas palabras: Orígenes de los alomorfos del artículo definido en gallego.*— El artículo definido gallego parece ser un clítico ditrópico: se liga fonológicamente a la palabra que le antecede, aunque mantiene relaciones constitutivas directas con las palabras que le siguen. Tiene tres alomorfos cuya distribución está condicionada por el segmento final de la palabra que lo precede. En este trabajo estudio los orígenes históricos de estos tres alomorfos, que son el resultado de procesos asimilatorios debidos a las restricciones que gobiernan la estructura de la sílaba y los contactos entre sílabas; en todo caso, la influencia de estas restricciones estaba claramente restringida al artículo definido y al pronombre átono acusativo de 3.<sup>a</sup> persona; estas restricciones no afectaron a otras palabras que tenían las mismas secuencias fonéticas en contextos aparentemente similares. En este artículo, intentaré mostrar la conveniencia de explicar los orígenes de las alternancias con las herramientas de los modelos de la lingüística basada en el uso, pues fueron concebidos para explorar el papel de la frecuencia en el surgimiento de la gramática y para permitir la combinación natural de información tomada de los componentes sintáctico, semántico y fonológico de la lengua.

**Palabras clave:** gallego; sílaba; clíticos; fonética particular de ciertas palabras; modelos de ejemplares

## 1. INTRODUCTION

The Galician definite article is an unstressed function word that determines nouns and nominalized units (Álvarez & Xove, 2002). Its morphophonemic intricacies have attracted the interest of numerous scholars: Álvarez Blanco (1983) and Dubert (2004) working in classic Structuralism, Dubert (2001) in Lexical Phonology, Fernández Rei (2002, 2005, 2012) in Lexical and Prosodic Phonology, and Kikuchi (2006) in Optimality Theory. As Kikuchi

has posited (2006, pp. 41–42), the present-day Galician definite article has three allomorphs /o, lo, no/ (see (1)) conditioned by the ending of the preceding word, although the article and the preceding word are not syntactically immediate constituents.

- (1) a) *[o]* pan “the bread”  
b) *como [o]* pan “I eat the bread”  
c) *cóme[lo]* pan “you eat the bread”  
d) *come[no]* pan “they eat the bread”

The allomorph /lo/ is selected when the preceding word ends in /r/ or /s/; after selecting /lo/, the trigger of the selection is deleted, so the presence of /lo/ becomes opaque; the allomorph /no/ is selected when the preceding word ends in a nasal and it is manifested by the alveolar feature of the nasal; the allomorph /o/ is selected elsewhere (see Section 2.2). Evidence shows that the selection of the allomorphs is made in the syntax, it is phonologically conditioned and triggers morphophonemic alternations. Dubert (2001) therefore introduces the current definite article in what Kaisse (1985) calls the *postlexical 1* component; Fernández Rei explains it as the outcome of precompiled phonological rules (Hayes, 1990).

This article will study the linguistic conditions that made possible the phonological changes that created the present-day alternations. I will assume that the linguistic conditions reflected the grammar at the time when the changes took place. The formalizations needed to describe the current alternations explain neither the changes, nor the conditions existing when the changes took place. This paper is historical since it describes changes and old states of language, but it is also synchronic since it tries to describe the original conditions that triggered the changes. Thus, my concern will be to examine the processes and conditions that produced the present-day alternations, not the present-day alternations *per se*.

The conditions of the present-day alternations differ somewhat between the standard written variety and vernacular dialects. The contexts of the alternations in the former are an arbitrary subset of those typical of the latter, which are lexically less restricted. The data for contemporary Galician are taken from the dialect of Santiago de Compostela (Dubert, 1998, 1999), from my own annotations as a dialectologist and from the *Atlas lingüístico galego. Vol. II. Morfología non verbal (ALGa2; García, Santamarina, Álvarez Blanco, Fernández Rei, & González González, 1995)*.

In Section 2 I present the properties of the present-day allomorphs of the article and their alternations; in Section 3 I present the origins of the alternations in Old Galician; in Section 4 I analyze problematic consonant sequences; in Section 5 the data are discussed in the light of usage based premises; finally, in Section 6 I present some conclusions.

I will follow the following graphic conventions: (.) syllable boundary, (#) morpheme boundary, (##) word boundary, (=) clitic boundary, (ω) phonological word, (&) contracted element, (V) any vowel, (C) any consonant, (X) any segment. Small capital letters will be used in glosses and double quotation marks (“”) with translations.

## 2. ALLOMORPHS OF THE DEFINITE ARTICLE IN CONTEMPORARY GALICIAN

In this section I will present the present-day Galician definite article, its allomorphs and their distribution; I will also classify the clitic based on its properties.

### 2.1. The syntactic position of the Galician definite article

The position of the article, illustrated in (2) (note that the spelling does not correspond to that of the standard variety), is strictly governed by syntactic conditions and is a by-product of historical developments: originally a demonstrative (the article is the outcome of Lat. *ille, illa, illud* “that”), it occupies the initial position inside its noun phrase (NP).

- (2) a) *tódalas casas*  
“All the houses”  
b) *da casa*  
OF&THE HOUSE  
“Of the house”  
c) *vexo os primos*  
SEE-1SG THE COUSINS  
“I see the cousins”  
d) *chegou o momento*  
TOOK-3SG THE MOMENT  
“The time has come”  
e) *ensiñábanno-los santos*  
SHOWED-3PL=DAT.1PL THE SAINTS  
“They showed us the saints”  
f) *van despoilos nenos*  
GO-3PL AFTERWARD THE BOYS  
“The boys go afterward”  
g) *collen despoilas patacas*  
TAKE-3PL AFTERWARD THE POTATOS  
“They collect the potatoes afterward”

The definite article may be preceded by the quantifiers *todo/toda/todos/todas* “all” (2a) and *ambos/ambas* “both”, prepositions (2b), verbs (2c, d), and clitic pronouns (2e). The definite article may also be preceded by words with which it has a more remote constitutive relation (2f, g). In no case is the article in an immediate constitutive relation with the preceding word; but stands in an adjacency relationship with it, i.e., an immediate sequential relation of direct contact.

Regarding stress, in (2) the article precedes stressed words, this probably being the more usual situation, except in cases where it is possible to attribute a pronominal function to the article, as in *os de Vigo* THE-M.PL FROM VIGO “those from Vigo”. The words preceding the article may be either unstressed (prepositions like *de* “of/from” in *da casa*, or conjunctions like *e* “and” in *e a nena* “and the girl”) or stressed (verbs as in 2c, d; adverbs as in 2f, g). Since the definite article is itself unstressed, it must lean on another stressed word of which it becomes a clitic. One might expect the definite article to be prosodically hosted as a *proclitic* within its noun phrase. Yet by virtue of adjacency, the article has established special phonological relations with units preceding it, as in (2). I shall present evidence that the definite article cliticizes on preceding words, becoming *enclitic*.

The phonological changes I will explore below show that, historically, the article was generally a proclitic only if it couldn’t be an enclitic. Thus, when not in initial position the definite article has tended to integrate phonologically with the preceding word. This has caused the defi-

nite article and the preceding words to interact phonologically, even though the article forms a syntactic unit with elements located to its right.

## 2.2. Present-day allomorphs

Morphophonemically the citation forms of the definite article are *o* THE-M.SG, *a* THE-F.SG, *os* THE-M.PL, and *as* THE-F.PL. However, it displays a set of allomorphs whose distribution is phonologically conditioned (Álvarez & Xove, 2002). Broadly speaking:

- a) the allomorphs *o*, *a*, *os*, *as* (i.e., the *first forms of the article* (1FA)) appear either in absolute initial position or preceded by a vowel;
- b) *lo*, *la*, *los*, *las* (i.e., the *second forms of the article* (2FA)) appear when the definite article is preceded by either /r/ or /s/;
- c) *no*, *na*, *nos*, *nas* (i.e., the *third forms of the article* (3FA)) appear after a nasal consonant, which will be manifested by alveolar realization of the nasal preceding the article (Álvarez Blanco, 1983; Dubert, 2001).<sup>1</sup>

The examples in (3) illustrate, in a similar phonological context, the different morphological behaviour of the article *a* THE-F.SG and the preposition *a* “to, human object marker” (where *a* might be any word starting with a vowel other than the article and the 3rd person of the accusative clitic, 3.ACC).

- (3)
- a) *vexo a rosa*  
SEE-1SG THE ROSE  
“I see the rose”
  - b) *vexo a Rosa*  
SEE-1SG TO ROSE  
“I see Rose”
  - c) *ve-la rosa < ves*  
SEE-2SG=THE ROSE  
“You see the rose”
  - d) *ves a Rosa < ves*  
SEE-2SG TO ROSE  
“You see Rose”
  - e) *ve-la rosa < ver*  
SEE-INF=THE ROSE<sup>2</sup>  
“To see the rose”
  - f) *ver a Rosa < ver*  
SEE-INF TO ROSE  
“To see Rose”
  - h) *ve[n] a rosa < ve[ŋ]*  
SEE-3PL=THE ROSE  
“They see the rose”

- i) *ve[ŋ] a Rosa < ve[ŋ]*  
SEE-3PL TO ROSE  
“They see Rose”

In (3a, b) *vexo a rosa* and *vexo a Rosa* are homophonous. In (3c), where the article is preceded by a word ending in /s/, /l/ replaces /s/: *ves*, but *ve-la rosa*; this alternation does not occur in (3d), however, where *a* is a preposition: *ves a Rosa*. The same applies to /r/ in (3e): with the article *a*, /l/ replaces /r/: *ve-la rosa*; but with the proposition, /r/ remains in (3f): *ver a Rosa*. In the most conservative dialects, when *a* is an article, the verb’s final nasal is alveolar, *ve[n=a]* *rosa* (3h), but with the preposition, the verb’s final nasal is realized as a velar, *ve[ŋ a]* *Rosa* (3i). The alternations of (3c, e) exemplify 2FA; the alternations of (3h), 3FA.<sup>2</sup>

These alternations are morphophonemic in character: a) the alternating segments /l/, /r/ and /s/ are phonologically distinctive: *mar* “sea”, *mal* “evil”; b) to describe the contexts of these alternations, syntactic information is required: the processes referred to only affect the expression of the definite article and 3.ACC clitics (see 3b, d, f, i). Even if, as Freixeiro Mato (2000) has claimed, the alternations are due to fast, sloppy speech (something that has not yet been demonstrated), one still needs to refer to this syntactic information in order to specify the context. Thus these [r]/[l] and [s]/[l] alternations do not constitute *phonetic processes* in the structuralist sense. Since the alternations depend on syntactic information and imply distinctive segments/features, in European classical structuralism the alternations should be classified as *morphophonemic* (Trubetskoy, 1928), or even *morphological* (Martinet, 1996). The allomorphy of the definite article displays what Spencer & Luís call *morphophonological idiosyncrasy*: “Alternations in phonological form that are not the result of completely regular phonological processes that affect all words or word combinations of a given phonological type” (2012, p. 109).<sup>3</sup>

The definite article is obligatorily contracted with certain preceding prepositions; note that these prepositions are unstressed:

- (4)
- a “to” > TO&THE-M.SG *ó*, F.SG *á*, M.PL *ós*, F.PL *ás*  
ca “than” > THAN&THE-M.SG *có*, F.SG *cá*, M.PL *cós*,  
F.PL *cás*  
de “of” > OF &THE-M.SG *do*, F.SG *da*, M.PL *dos*, F.PL *das*  
con “with” > WITH&THE-M.SG *co*, F.SG *ca*, M.PL *cos*,  
F.PL *coas*  
en “in” > IN&THE-M.SG *no*, F.SG *na*, M.PL *nos*, F.PL *nas*  
por “by” > BY&THE-M.SG *po*, F.SG *pa*, M.PL *pos*,  
F.PL *polas*

<sup>1</sup> A nasal at the right boundary of a word is realized as a velar when followed by a vowel belonging to the next word: [es’tãŋ a’i] ARE-3PL THERE “they are there”, [’forðŋ ’elis] WENT-3PL THEY “they went”. On the contrary, an alveolar nasal appears if the vowel of the second word belongs to 3rd person of the accusative clitic or to 3FA: [’fano] ‘do-3PL=3ACC.M.SG’ “they do it”.

<sup>2</sup> 3FA is rapidly disappearing in current Galician; the velar nasal [ŋ] is spreading to the pre-article position, so that *ven a rosa* “they see the rose” and *ven a Rosa* “They see Rose” are both pronounced [’beŋ ɐ ’rose] by most speakers. The alveolar nasal only remains in the case of the 3.ACC clitic: [’benə] “they see her”.

<sup>3</sup> Expressed in a generative framework, a point of interest in the morphology of the Galician definite article is the fact that the realization of the article is associated with clearly lexical phonological processes that must work after lexical insertion has taken place, because only after lexical insertion is the context of the lexical phonological rules satisfied (see Dubert, 2001; Fernández Rei, 2002; and Kikuchi, 2006).

These contractions are lexicalized: no present-day phonological rule produces *co* from *con* and *o*. For this reason, and given their frequency of use, these must all be stored in the lexicon.

The article may also be fused with a preceding unstressed /a/, as in (5):

- (5) a) *onda o can* > *ond[ɔ] can*  
 “Beside the dog”  
 b) *canda o neno* > *cand[ɔ] neno*  
 “With the boy”  
 c) *para o hospital* > *pr[ɔ] hospital*  
 “To the hospital”  
 d) *falta o diñeiro* > *falt[ɔ] diñeiro*  
 LACK-3SG THE MONEY  
 “There is not enough money”  
 e) *salta o neno* > *salt[ɔ] neno*  
 JUMP-3SG THE BOY  
 “The boy jumps”

In (5a, b, c) the article may fuse with the unstressed ending -/a/ of the prepositions; in (5d, e), it may fuse with the unstressed ending -/a/ of the verbs. All these coalescences are ordinary at normal speech rate (but not represented in the written language). Unlike the contractions in (4), the coalescences in (5) are not obligatory (Fernández Rei, 2002; Regueira, 1998).

Finally, the allomorphy of the definite article constitutes a (socio)linguistic variable, since most speakers may use either of the variants *ver o neno* (with 1FA) or *ve-lo neno* (with 2FA); however, they must say *polo neno* “by the boy”, rather than *por o neno* which sounds strange; likewise *tódolos* and *tódalas* ALL=THE. Many speakers for whom *ver o neno* alternates with *ve-lo neno* use *tódolos* and *tódalas* consistently rather than *todos os* and *todas as*, as shown in Table 1 (Dubert, 1999, pp. 107–111 describes the variable).

**Table 1.** Indices of use of /lo/ with *por*, *todos-todas* and verb forms ending in /s/ or /r/. Adapted from Dubert (1999, pp. 107–111).

	index	tokens
Por + Art-Def	94.25%	377/400
Todos + Art-Def	71.43%	90/126 <sup>4</sup>
Verb <sub>NP</sub> [Art-Def	54.78%	342/575

### 2.3. The definite article as a clitic

From a syntactic point of view, then, the definite article is an unstressed *word form* corresponding to a *morphosyntactic word* (Matthews, 1991). Its location is governed by syntactic restrictions. However, it is an unstressed function word without stressed allomorphs, since the article lost its stress in its

historical development. As an unstressed unit, the article constitutes a phonological unit in conjunction with other stressed words (Fernández Rei, 2002) yet is still a word form. Having lost phonological wordhood, it has become a clitic.

Its syntactically determined location seems to transform the definite article into a *simple clitic*, an eroded form of an old phonological word that occupies the same position as the corresponding full form (cf. English ‘ve in *I’ve ~ I have been talking*); but unlike simple clitics, the definite article lacks stressed allomorphs (cf. English ‘ve with respect to *have*). Another property shared by the definite article and clitics is that neither presents clear restrictions when selecting a host: the article may be attached to verbs, stressed pronouns, unstressed pronouns, quantifiers, adverbs, prepositions, etc. (see (2)). However, the fact that its presence affects the basic form of its hosts (*comer* and *o pan*, but *come-lo pan*; *salta* and *o neno*, but *salt[ɔ] neno*), makes the definite article more like an affix (like English *n’t* in *don’t*; see Zwicky & Pullum, 1983) and detaches the article from the class of simple clitics.

Given its behaviour, the definite article may be considered a *ditropic clitic*, which means that “an element is syntactically associated with a phrase in one direction but phonologically bound to a word in the opposite direction” (Spencer & Luís, 2012, p. 66). Thus, the definite article doesn’t have a semantic link with the word to which it is morphologically attached, but with the syntactic unit to its right, which Cysouw denominates *attractor* (2005, p. 17). According to Cysouw, ditropic clitics are “a rare phenomenon, and an extensive search was needed to find examples among the world’s languages” (p. 31). Cysouw observes that “ditropic clitics summarise one of the more striking possibilities of human language, which should not be dismissed in the formulation of a theory of linguistic structure” (p. 34).

### 3. ON THE ORIGIN OF THE ALTERNATION AND SOME PROBLEMS

In its transformation into an article, beside stress Lat. *illum*, *illam* lost its initial vowel: *illum* > *lu* > *lo*. In the course of the ninth and tenth centuries Galician lost Latin word-internal intervocalic /l/, e.g., Lat. *uolare* > *voar* “fly” (Williams, 1962, pp. 68–74). When it appeared in intervocalic position (6a), the initial /l/ of the article *lo* underwent the same processes as word-internal intervocalic /l/s (6b) and was also lost. This development affected both the definite article and the 3.ACC clitic. Other words starting with /l/ (6c) were not affected, despite the similar context:

- (6) a) *como lo pan* > *como o pan*  
 EAT-1SG THE BREAD  
 “I eat the bread”  
 b) *moler* > *moer*

<sup>4</sup> In Santiago de Compostela, the *tódolos* forms are in competition with another set of lexicalized contractions: *tod[ɔ]s*, *tod[a:]s* (for *todos+os* and *todas+as* respectively), with a use index of 15.87% (20/126).



- “To grind”  
c) *este lugar*  
“This place”

Note that at the time when /l/ dropped the article (6a) must have had the aforementioned special phonological relationship with the preceding word, since its /l/ was treated as a word-internal intervocalic /l/ (6b): thus, prosodic and syntactic bracketing didn’t match. In (7) the square brackets [ ] represent phonological units, while the curly brackets { } represent syntactic units.

- (7) a) {[*como* {*lo*} *pan*]} > *como o pan*  
EAT-1SG THE BREAD  
“I eat the bread”  
b) {[*de* {*lo*} *pan*]} > *de o pan* > *do pan*  
OF&THE BREAD  
“Of the bread”

However, the /l/ of the definite article assimilated a preceding /r/ and /s/ (8a, b): *come/rlo/ pan* > *come/l:o/ pan* EAT-INF THE BREAD “to eat the bread”; *come/slo/ pan* > *come/l:o/ pan* EAT-2SG THE BREAD “you eat the bread”. The /l/ of the definite article assimilated to a preceding nasal (8c): *come/nlo/ pan* > *come/n:o/ pan* EAT-3SG THE BREAD “they eat the bread”. Long /l:/ and /n:/ were later simplified to /l/ and /n/, respectively (Ferreiro, 1995; Williams, 1962). This was not the case in (8d, e, f), where no assimilation has taken place. Here the word-internal clusters /sl/ and /nl/ remained unchanged, while the word internal /rl/ either remained unchanged or underwent metathesis:

- (8) a) *come[r lo] pan* > *come[l:o] pan* > *come-[lo] pan*  
b) *come[z lo] pan* > *come[l:o] pan* > *cóme-[lo] pan*  
c) *come[n lo] pan* > *come[n:o] pan* > *cóme-[no] pan*  
d) *me/rl/o* “blackbird” > *me/rl/o* ~ *me/lr/o*  
e) *ga/rl/opa* “jack plane” > *ga/rl/opa* ~ *ga/lr/opa*  
f) *e/nl/ear* “to tie up”, *ca/nl/e* “canal”, *ri/nl/e* “kidney”

Williams (1962) was aware of this situation and speaks about an intertonic context:

The fact that the primitive form of the Portuguese article, while apparently proclitic to the noun it modified, could be so closely attached to a preceding word ending in a vowel that the *l* of the article fell, or to a preceding word ending in a consonant that this consonant was assimilated to the *l* of the article, seems to indicate that the article developed most commonly in the intertonic position, that is, between two accented syllables, e.g., *amei o filho* (— — — —), but not necessarily adjacent to either of them, e.g., *todo o amor* (— — — —) (p. 140).

Speaking of medieval varieties of (Galician-) Portuguese, Mattos e Silva (1989) also says regarding 2FA:

Podem ocorrer depois das consoantes /s/ e /r/ que fecham sílaba final dos morfemas gramaticais independentes *após*, *depós*, *despós*, *ambas*, *todos*, *todas*, *per*, *por*, constituindo um único vocábulo fonético—e essa unidade fónica se revela na grafia em que não há espaço entre o artigo e o morfema a que se aglutina—pelo processo fonético assimilatório em que uma consoante adquire parcial ou totalmente (é esse o caso aqui) os traços da consoante que a segue (pp. 143–144).<sup>5</sup>

One may wonder why the /l/ in intervocalic position of both the definite article and the 3.ACC clitic behaved like a word-internal intervocalic /l/ (*como lo pan* > *como o pan* like *moler* > *moer*), yet when preceded by a consonant it behaved differently (*comer lo* > *comel lo* > *come-lo*, but *merlo* > *melro* ~ *merlo*). Why were there assimilations only in the cases of (8a, b, c), but not in the others? Note that in (7) /l/ behaved as if word-internal, but in (8) the word-internal /l/ of *merlo*, *garlopa*, *canle* behaved differently than the “word-internal” /l/ in *comen lo*, *comer lo*, *comes lo*. Secondly, in /n.l/, why was the assimilation progressive, /n.l/ > /n:/, while in /s.l/ and /r.l/ it was regressive, /s.l/ and /r.l/ > /l:/?

#### 4. SYLLABLE STRUCTURE, SONORITY AND ASSIMILATION PROCESSES

In this study I will only deal with the origins of the alternations referred to here as 2FA and 3FA in (2) and (8), but it was necessary to mention all these facts about intervocalic 1FA in (7) to show that the definite article had special, phonological relationships with the words that preceded it. These phonological relations are not only a feature of contemporary Galician but must already have been present in the present-day alternations originated. Let us now consider the nature of the phonological relations that ended up producing 2FA and 3FA in present-day Galician. First, I will present some strong tendencies in syllable structure; then, I will look at the Galician clusters /r.l/, /s.l/, /n.l/ in the light of those laws.

##### 4.1. Preference laws for syllabic structure

According to Theo Vennemann (1988):

Every change in a language system is a local improvement, i.e., an improvement relative to a certain parameter. For instance, every syllable structure change is an improvement of syllable structure as defined by some preference law for syllabic structure. If a change worsens syllable structure, it is not a syllable structure

<sup>5</sup> [2FA] may occur after the consonants /s/ and /r/ at the end of the last syllable of the independent grammatical morphemes *após* “after”, *depós*, *despós* “afterward”, *ambas* “both-f.pl”, *todos* “all-m.pl”, *todas* “all-f.pl”, *per* “through”, *por* “through”, forming a single phonetic word—manifested in the spelling by the lack of a space between the article and the morpheme to which it is joined—through a phonetic assimilation process in which a consonant acquires either partially or (as in the present case) completely the features of the following consonant.

change, by which I mean a change motivated by syllable structure, but a change on some other parameter which merely happens also to affect syllable structure. (...) Since every change is an improvement, a change on a given parameter is motivated by inherent low degree of linguistic quality of the affected structures relative to that parameter (pp. 1–2).

Vennemann (1988) suggests a parameter called *Universal Consonantal Strength* (p. 9), defined as the “degree of deviation from unimpeded (voiced) air flow” (p. 8): see (9). Phonetic segments may be ordered in a scale with respect to this parameter: the higher the number, the stronger the segment.

- (9) 10 voiceless stops  
 9 voiced stops  
 8 voiceless fricatives  
 7 voiced fricatives  
 6 nasals  
 5 lateral liquids (of the *l* kind)  
 4 central liquids (of the *r* kind)  
 3 high vowels  
 2 mid vowels  
 1 low vowels

This scale mirrors that proposed by other scholars such as Ladefoged and Johnson (2010) or Laver (1994), who speak of sonority scales: roughly speaking, the more sonorous a segment is, the weaker it is; the less sonorous a segment is, the stronger it is. This is an incomplete scale since some segments such as affricates or approximants<sup>6</sup> are missing. According to Ladefoged and Johnson (2010), Laver (1994), and Vennemann (1988), some segments could occupy different positions in the scale depending on the language, the style, the speaker, etc. In (9) the rhotics are grouped together on the same level of strength: as I will show, there may be reasons for distributing rhotics among various points on the scale.

Next, Vennemann (1988) proposes a series of laws of preference for syllabic structure. Now these are preference laws (or constraints), not compulsory laws (or constraints). I will only discuss those of interest for our purposes. There is the head law, stated in (10):

- (10) The head law  
 “A syllable head is the more preferred: (a) the closer the number of speech sounds in the head is to one, (b) the greater the Consonantal Strength value of its onset, and (c) the more sharply the Consonantal Strength drops from the onset toward the Consonantal Strength of the following syllable nucleus” (Vennemann, 1988, pp. 13–14).

This law prefers syllables with a single consonant in the onset (e.g., /ta/) to syllables with either no onset (/a/), more than one segment in the onset (/tra/) or a weak onset (/ja/). This explains why we find strengthening processes like that of [ra.ja] *raia* “line” > [ra.ja], or the rhotacization of /l/ in Lat. *clauum* > *cravo* “nail”, in Galician. The coda law is stated in (11):

- (11) Coda law  
 “A syllable coda is the more preferred: (a) the smaller the number of speech sounds in the coda, (b) the less the Consonantal Strength of its offset, and (c) the more sharply the Consonantal Strength drops from the offset toward the preceding syllable nucleus” (Vennemann, 1988, p. 21).

This law prefers a syllable with no coda (e.g., /ta/) over a syllable with one (/tal/); if there is a coda, there is preference for a weak coda (/tas/) rather than a strong coda (/tak/). This law accounts for the vocalization of /l/ in Lat. *multum* > *moito* “a lot”; why in a syllable with a complex coda, the segment to the left is usually more sonorous than the segment to the right, as in *ins.tru.men.to* “instrument”; why in complex codas like *ns* in *ins.tru.men.to* the coda is simplified in popular Galician, and why it is the nasal and not the fricative segment that is lost: *is.tru.men.to*. Nasals are weaker than fricatives, so fricatives make better codas since the drop from the offset to the preceding syllable nucleus is sharper.

This distribution leads Hooper (1976, pp. 199–200) to consider that, in the syllable, there exist *strong positions*, those of onset, and *weak positions*, those of coda. The strength of onset positions is manifested by the fact that the strengthening processes of segments ([ra.ja] “line” > [ra.ja]) are more common in onset than coda positions; more phonological contrasts are possible in the onset than in the coda, since many phonological contrasts are neutralized in coda position, where we find less distinctive segments; likewise, the elision of segments is more usual in the coda than in the onset; and finally, assimilations are less common in the onset than in the coda: “As syllables are strung together, the end of one syllable (its weak position) immediately precedes the beginning of the next (its strong position)” (p. 200). If this is so, we should distinguish between segments that are strong by virtue of their features (as in the scale in 9) and segments that are strong by position. Inherently strong segments are preferred in strong position; inherently weak segments are preferred in weak position. When this situation doesn’t occur, non-optimal configurations are obtained that can be repaired following different strategies.

<sup>6</sup> According to Catford (1988), high vowels and approximants [β, ð, ɥ, j, w] may be grouped since they become fricatives if devoiced. I am assuming that approximants are weaker than laterals and rhotics, since approximants don’t imply any interruption of the airflow, noises, bursts, etc. (see Catford, 1988; Ladefoged & Johnson 2010; Ladefoged & Maddieson, 1996; and Laver, 1994). In this assumption, approximants are closer to vowels than laterals (which, albeit being approximants, are realized by interrupting the airflow by the centre the mouth). Laver (1994) locates affricates between fricatives and stops.

The next law, (12), contrasts with Hooper's (1976) proposal:

- (12) Strength assimilation law  
 "If Consonantal Strength is assimilated in a syllable contact, the Consonantal Strength of the stronger speech sound decreases" (Vennemann, 1988, p. 35).

Vennemann illustrates this law with numerous Romance and Germanic examples, e.g., *Carru* < *Carlo* "Charles" in Sicilian, *fòrru* < *forno* "oven" in Corsican. The law works in contexts of progressive and regressive assimilation: Lat. *met-sui* > *messui* "I have mowed", *sup-mos* > *summus* "highest". In these examples the weaker segment always imposes itself. The selection of the weaker segment is interesting precisely because it produces onsets with weaker segments (against the head law in (10)). Some sort of intervocalic lenition principle related to ease of articulation is probably also at work here.

Commenting on the exceptions to these laws, Vennemann (1988) refers to morphological boundaries, suggesting that "when the syllable boundary coincides with the morphological boundary, the direction of assimilation may be determined by a desire to protect the integrity of a stem or affix" (p. 39).

Finally, Vennemann (1988) proposes a law (13) related to the strength assimilation law, which agrees with the distribution between strong and weak positions, and which helps to trace the syllabic boundaries between them:

- (13) Contact law  
 "A syllable contact A<sup>s</sup>B is the more preferred, the less the Consonantal Strength of the offset A and the greater the Consonantal Strength of the onset B; more precisely—the greater the characteristic difference CS(B)–CS(A) between the Consonantal Strength of B and that of A" (p. 40).

Thus, a sequence of syllables like [ej.to] "heap" is preferred to a sequence like [is.to] "this", because the value of the difference between B[t] and A[j] in [ej.to] is higher than the value of the difference between B[t] and A[s] in [is.to]. If, according to (9), we assign a strength of 10 to [t], 8 to [s] and 1 to [j], and calculate the operation suggested in (13), the value of B–A in the case of B[t] and A[j] in [ej.to] is 9 (B[t] – A[j] = 10 – 1 = 9), while in B[t] and A[s] in [is.to] it is 2 (B[t] – A[s] = 10 – 8 = 2).

These rules or constraints internally characterize syllables in terms of a sonority profile across their constituent segments, and account for how "whole syllables would be able to be related to their neighbouring syllables

in an utterance in a hierarchy of relative syllabic prominence" (Laver, 1994, p. 505).

Thus, sequences where the inherent strength of the coda is higher than inherent strength of the next onset are non-optimal and may be corrected. Such corrections can be achieved by different means: metathesis, assimilation, elision of segments, etc. If assimilation is selected, following the assimilation law of (12), the stronger segment should assimilate to the weaker one unless the boundaries of the morpheme are protected.

## 4.2. Word internal sequences /V.IV/, /s.l/, /r.l/, /n.l/

Present-day Galician presents word-internal sequences of segments that are barely optimal, such as /s.l/: *eslora* "boat's length", *desleal* "disloyal"; /n.l/ *rinle* "kidney", *enleal* "to tangle"; /s.n/: *cisne* "swan", *trasno* "goblin"; /s.m/: *asmo* "unleavened", *chusma* "riffraff"; /n.r/: *tenro* "tender", *xenro* "son-in-law". In all these sequences, always considered heterosyllabic, the A segments of the coda are stronger than the B segments of the following onset. Galician tolerates such sequences fairly well,<sup>7</sup> however, as we already know, these sequences weren't treated the same way when the lateral belonged either to the definite article or to the 3.ACC clitic.

### 4.2.1. The sequence /s.l/

The *Vocabulario ortográfico da lingua galega* (VOL-Ga; Real Academia Galega/Instituto da Lingua Galega, 2004), a spelling dictionary containing more than 50,000 words, contains 105 lexemes with the sequence /s.l/ such as *desleal* "disloyal" and *esluir* "to dilute". These 105 lexemes may be reduced to 51 bases, for example grouping *islam*, *islamización* and *islamista* all under *islam*. Of these 51 bases:

- 23 contain the prefix *des-*, e.g., *desligar* "to untie".
- 3 contain the prefix *dis-*, e.g., *dislalia* "dyslalia".
- 4 contain the prefix *trans-*, e.g., *translación* "translation".
- 3 contain the prefix *tras-*, e.g., *traslado* "move".
- 8 contain the prefix *es-*, e.g., *eslumeacer* "to be eager".

Thus, out of the 51 bases containing /s.l/, 41 appear in complex words with a prefix ending in /s/. Only 8 bases contain /s.l/ as a root-internal sequence: *islam*, *eslavo* "Slavic", *slalom*, *slogan*, *lexislar* "to legislate". Among these words there are loanwords like *eslora* "boat's length", *slalom*, and foreign toponyms and ethnonyms like *islandés* "Icelandic". The only Latinate lexeme is the word *lexislar* "to legislate" and its derivatives, from

<sup>7</sup> /s/ in the coda followed by a voiced or a fricative segment requires correction in many dialects, tending to be replaced by /r/; this phenomenon is known as *rhotacism*.

the Latin compound *legis lator* “legislator”. *Lexislador* may be the only word directly inherited from Latin, since Latin /s.l/ is reduced to /l/ with compensatory lengthening of the preceding vowel (Monteil, 1970/2003, pp. 77, 128). *Eslavo* < *slavum* is a medieval Latinism.

Therefore, apart from derived words, word-internal /s.l/ hardly exists in the current Galician lexicon, at least within roots: it is only widespread when it includes a morphological boundary, /s.#l/. The fact that prefixes like /des/, /tras/ or /es/ may appear in other words (*desfacer* “undo”, *descarado* “shameless”, *trasacordo* “new deal”, etc.), backs up their integrity.

In the medieval lexicon, for which we have only limited data (see *Dicionario de dicionarios do galego medieval* [DDGM]; González Seoane, Álvarez de la Granja, & Boullón Agrelo, 2006–2012), we find words with the prefix /des/, such as *deslaçar* “to untie”; with the prefix /tras/: *trasladar* “to copy” (beside *tralladar*, *treladar*, *traladar*, *tralaçon*); with a supposed prefix /es/: *esleeçon* “election”, *esleedor* “elector”; and root-internal in *isla* “island”, *brosllador* “embroiderer”. Thus, /s.l/ appears in complex words, made up with *tras-*, *des-* and *es-*. The word *isla* is a Spanish loanword; *brosllar* “to embroider” comes from the Germanic *brūzdan*.

It is evident that, apart from *traladar* ~ *treladar* ~ *tralaçon*, the /s.l/ sequences that appeared in word internal position didn’t behave as /s.l/ did when /l/ was part of the definite article or the 3.ACC clitics. However, word-internal /s.l/ was a group that Galician phonology could tolerate; it is even found in present-day Galician.

#### 4.2.2. The sequence /r.l/

*VOLGa* (Real Academia Galega/Instituto da Lingua Galega, 2004) lists 101 lexemes containing word-internal /r.l/ which may be reduced to 53 bases:

- 5 containing the prefix *inter-*: *interlocutor*.
- 1 containing the prefix *super-*: *superlative* “superlative, extreme”.
- 25 loanwords from various periods and with differing degrees of integration, such as *abarloar* “to berth” (< Fr. *par lof*), *charla* “chat” (< It. *ciarlare*), etc. It also appears in foreign ethnonyms and toponyms such as *irlandés* “Irish”.
- 22 are inherited from Latin or onomatopoeic, with root internal /r.l/: *burla* “mockery” (< *burrūla*), *birlo* “skittle” (< *pirūlu*), *chirlar* “to cheep” (onomatopoeic).

Latin /r.l/ was an unstable group that evolved according to the Vennemann laws: *\*sterla* > *stella* “star”, *\*agerlos* > *agellus* “little field” (Bassols de Climent, 1983, p. 210; Monteil, 1970/2003, p. 94).

Many of these words maintain specific patterns of change in Western Galician. For example, the metathesis /r.l/ > /l.r/ is common: *galropa* < *garlopa* “jack plane”, *melro* < *merlo* “blackbird”, *bilro* < *birlo* “skittle”. These

metatheses appear not to be conditioned by stress (compare *bilro* and *calróu* “winkle”) and are accompanied by a change in the manner of articulation of the rhotic segment, which shifts from a tap in coda position /r.l/, *garlopa*, to a trill in onset, /l.r/, *galropa*. In dialectal Galician, /r.l/ also alternates with /l.r/ and /r/: *porlón* ~ *polrón* ~ *porrón* “red gunnard”, *espirrar* ~ *esbilrar* “to sneeze”.

The sequence /r.l/ appears well-formed according to Vennemann’s laws (with the weaker segment in the A position and the stronger in the B position) and should not need to be improved from the point of view of syllabic structure; the resulting sequence, /l.r/, however, seems to be ill-formed if /r/ has the same value on the strength scale as /r/ (a stronger consonant in the A position followed by a weaker one in the B position). In fact, the traditional grouping of the rhotics under one and the same label is misleading because it suggests that, because they constitute a single phonological class, they should occupy the same place on a phonetic scale of strength/sonority. Ladefoged and Maddieson (1996) say that rhotics are a class defined by family resemblances determined more by their phonological relationships than by internal similarities. Thus, the alveolar trills may be [+sonorant], while uvular trills may end as dorsal fricatives [-sonorant], or even glottal segments (this is the case of Brazilian Portuguese). Hooper (1976, p. 208) proposes that the Spanish trill /r/ is stronger than the tap /r/ on the strength scale; she locates /r/ between voiced stops (strength 7 on her scale) and fricatives (strength 5 on her scale), but places the tap /r/ between the semivowels (strength 1) and the laterals /l/ (strength 3). If this classification is right and applicable to Galician, the change /r.l/ to /l.r/ improves the quality of the sequence: /r.l/ B[l] – A[r] would be (with the values of Hooper’s scale) 3–2 = 1, while /l.r/ B[r]–A[l] would be 6–3 = 3. Note that /l.r/ seems to derive from words that in a previous stage had /r.l/ and possibly /r/; it seems that the change /r.l/ > /l.r/ (*garlopa* > *galropa*) opens the door to the aforementioned variation /l.r/ ~ /r/ (*espirrar* “sneeze” > *esbilrar*).

Although Galician /r.l/ is not very common and seems to be unstable, loanwords with /r.l/ were easily adopted; witness Galician forms such as *barlovento* “windward”, *abarloar* “to berth”, *charla* “chat”, *escarlata* “scarlet”, etc. The older loanwords may have undergone the change /r.l/ > /l.r/: *galropa*, *abalroar*, *pelrón*. In the medieval language (DDGM) /r.l/ appears in *arlota* “prostitute”, *girlanda* “garland”, *canterlada* “torn”; /l.r/ appears in *bulrrador* “impostor”, *palreiro* “talkative”. Thus, the alternation /r.l/ ~ /l.r/ was already found in the Middle Ages.

None of these root-internal instances of /r.l/ underwent the same development as /r.l/ when /l/ belongs to the definite article.

#### 4.2.3. The sequence /n.l/

In *VOLGa* (Real Academia Galega/Instituto da Lingua Galega, 2004), 52 lexemes contain the group /n.l/. We can reduce them to 31 bases:



- 1 with the prefix *circun-*: *circunloquio* “circumlocution”.
- 16 with the prefix *en-*: *enlear* “to tie up”.
- 3 are foreigner toponyms or ethnonyms: *groenland-és* “Greenlander”.
- 11 contain root-internal /n.l/: *arrenlar* “to tire”, *canle* “canal”, *penla* “ladle”.

The root-internal sequence is the outcome of historical phonological processes. First, Latin intervocalic *-n* dropped after nasalizing the preceding vowel; then the two vowels in hiatus were merged, yielding a nasal vowel (Ferreiro, 1995, pp. 127–128). These nasal vowels produced nasal codas (14) in some dialects; in other dialects, the vowels simply lost the nasal feature.

- (14) Lat. *canale* “canal” > *cāale* > *cāle* > *canle* ~ *cal*  
 Lat. *fontanella* “little spring” > *fontāela* > *fontēla*  
 > *fontenla* ~ *fontela*

Excluding complex words, Galician inherited no Latin root-internal sequences like /n.l/ on account of Latin’s tendency to change /n.l/ to /l.l/ (which also accords with the drift described by Vennemann, 1988): Lat. *corolla* “garland” < \**coron(o)la*, Lat. *uillum* “a cup of wine” < \**uin(o)lom* (Bassols de Climent, 1983, p. 210).

Changes like those in (14) created non-optimal sequences, with a nasal coda (A) stronger than a lateral onset (B): *fontenla* “little spring”. Maybe the loss of nasal vowels (vowels where nasality is a distinctive feature) may explain this behaviour. Nevertheless, while the loss of the nasal feature is a natural phenomenon (*fontēla* > *fontela*) because oral vowels are more natural than nasal ones, the creation of a nasal consonant (*fontēla* > *fontenla*) seems to be a marked phonological phenomenon that generates a marked syllable sequence, first because a checked syllable is created from an open syllable; secondly, because the sequence of segments created is problematical, since a coda A is created that is stronger than the following onset B. Although a less natural phenomenon, some words without the nasal vowel contributed to the /n.l/ schema, e.g., *ponla* “branch” < Lat. *pūlla*, *trenla* “rope” < \**tragēlla*.

None of these developments coincides with what is found when the /l/ belongs to the definite article.

## 5. DISCUSSION

As the outcome of a *regular phonetic law*, Latin /l/ and /n/ might have disappeared anywhere in intervocalic position, in which case the loss of intervocalic /l/ and /n/ would have occurred not only within words but also between words. Instead, word-initial or word-final /l/ and /n/ were kept: *este lado* “this side”, *esta nube* “this cloud”, *non é* “it isn’t”. In terms of current prosodic phonology, the elision of /l/ and /n/ was applied within a certain prosodic domain, probably the *phonological word*. If so, elision may have been the outcome of a *domain span rule* (Nespor & Vogel, 2007, p. 15) such as (15):

- (15)  $l, n > \emptyset / [XV\_VX]_{\omega}, \omega = \text{phonological word}$

Since the /l/ of the definite article was lost when preceded by a vowel, we may assume that (15) also refers to environments containing the article; since the phonological word is a domain established on phonological grounds, and (15) operated within it, it would seem that the elision rule must be postlexical. All we need is to formulate the algorithm that defines the prosodic word in such a way that it includes the article (for present-day Galician, see Fernández Rei, 2002).

However, if we assume that every definite article constitutes a phonological word in conjunction with its preceding host, we have seen that, in the case of /s.l/, /n.l/ and /r.l/, different changes took place depending on whether /l/ was part of the definite article (where assimilation occurred) or not (where there was no assimilation). It seems odd that, where the article is concerned, /V##l/ should have constituted part of one prosodic domain but /C##l/, a different one.

Intervocalic deletion or postconsonantal assimilation occurred even when the article was preceded by unstressed function words, giving rise to the contractions in (4):

- a) /V##l/ contexts

*a lo pan* > *a o pan* > *ó pan*  
*de lo pan* > *de o pan* > *do pan*  
*como lo pan* > *como o pan*

- b) /C##l/ contexts

*por lo pan* > *pol lo pan* > *polo pan*  
*en lo pan* > *en no pan* > *no pan*  
*comer lo pan* > *comel lo pan* > *come-lo pan*

Thus, [*a ## lo*], [*de ## lo*], [*en ## lo*] must form part of the same prosodic domain as [*como ## lo*], [*comen ## lo*], [*comer ## lo*], where the article was preceded by a stressed word.

Nespor and Vogel (2007, pp. 27–60) make a distinction between two kinds of rules: those which operate only with phonological information (prosodic domains are phonological information) and those which depend on morphological or syntactic information. Both kinds of rules operate both inside and between words: “A distinction must be drawn between those phonological processes that are to be handled by purely phonological rules and those that are to be handled by rules that also take other than phonological information into account” (p. 33).

They only consider processes of the first kind to belong to Prosodic Phonology. The early assimilations /s.l/, /r.l/ > /l:/ and /n.l/ > /n:/ seem to have been morphosyntactically conditioned from the start. This is not surprising since the definite article and the 3.ACC clitic are the only unstressed monosyllabic word forms starting with /l/. Thus, according to Nespor & Vogel (2007), the rules responsible for the postconsonantal assimilations must be included in a component other than Prosodic Phonology and hence in a different component from the intervocalic elision of *come lo pan* or *de lo pan*. However, we saw that

the changes that occurred in the context of the postconsonantal definite article had a clear phonetic motivation based on strength/sonority scales and optimal syllabic configurations: this appears to be what Pierrehumbert (2002) calls *word-specific phonetics*, but across syntactic boundaries. Locating the assimilation rules in a different component does not seem to explain why the assimilations took place only with definite articles (and the 3.ACC clitics). The special component in which these special assimilation rules need to be located is an outcome of the assimilation, not an explanation for the change; indeed, the reason why we need this special component is due to the morphological conditioning of the assimilations.

For the present-day morphology of the Galician article Dubert (2001), Fernández Rei (2002), and Kikuchi (2006) propose an allomorph /lo, la, los, las/, i.e., 2FA, which is selected by the ending of the preceding word; once 2FA is selected, the consonants /s/ and /r/ drop out: *comer lo pan* > *come-lo pan*. Fernández Rei and Dubert think that the elision is morphophonemically conditioned. Kikuchi (pp. 45–46) attempts to make the rule more phonological by suggesting that words such as *desleal* “disloyal” or *merlo* “blackbird” keep the groups /s.l/ and /r.l/ in present-day Galician because DOMAIN-CONTIGUITY (the faithfulness constraint against morpheme-internal deletion) is ranked higher than OCP[+cont] (an OCP constraint that impedes a sequence of two continuous segments). In the case of 2FA, however, the faithfulness constraint MAXC is ranked lower than OCP[+cont], producing the elision of /s/ and /r/.<sup>8</sup> His phonology, however, needs to recognise the morphological boundaries of words. In much the same way, if we were to apply the same reasoning to the time at which the old assimilations took place, we might say that the faithfulness constraints blocked the change in *desleal*, but not in *comer lo pan*; in the case of the article, the constraints based on strength produced the assimilation because they were ranked higher than the faithfulness constraints. But again, why was this the case? Why were the faithfulness constraints stronger inside the morphological word but weaker when the /l/ of the article was involved? In *islam*, *merlo* and *fontenla*, /s.l/, /r.l/, and /n.l/ are permanently stored in the lexicon, continuously violating certain constraints and containing non-optimal lexical syllabic configurations, while in the case of the article, the sequences /s.l/, /r.l/, and /n.l/ must be constructed on the fly each time they are needed, disappearing just after its use (at least in a generative framework where the constructions with the article should not be memorized, but only exist momentarily, when used).

Probably, one of the explanations of the assimilations comes from the fact that the frequency of use of the sequences /s.l/, /r.l/ and /n.l/ when /l/ forms part of the definite article is much greater than the frequency of use of these sequences in word-form-internal position: the use of the construction [Verb-/s##l/-Def-Art] is overwhelmingly more common than the individual use of isolated words like *esluir*, *desligar*, *eslora*, which contain the same sequences. Most word forms corresponding to the 2SG, 1PL and 2PL of verbs end in /s/; most noun phrases begin with a definite article. To the frequency of [Verb-/s##l/-Def-Art] we may add other common constructions such as those in (16):

- (16) a) *Despoi-los dous falaron*  
AFTERWARDS=THE TWO SPOKE-3PL  
“Afterwards, they both spoke”  
b) *Ante-las mulleres vestían así*  
BEFORE=THE WOMEN DRESSED-3PL.IMP SO  
“Before, women used to dress that way”  
c) *As mulleres e mailos homes*  
THE WOMEN AND MORE=THE MEN  
“The women and the men”  
d) *Nó-los dous imos a Muros*  
WE-THE TWO GO-2PL TO MUROS  
“We are both going to Muros”  
e) *Dámosvo-las grazas*  
GIVE-1PL=DAT.2PL=THE THANKS  
“We thank you”

All these constructions may be reduced to a more abstract one, [/s##l/-Def-Art], which is even more usual (in fact, [Verb-/s##l/-Def-Art] is a more specific instantiation of both [/s##l/-Def-Art] and [Verb /l/-Def-Art]). It seems that it is frequency of use that helps to maintain the contraction *polo(s)*, *pola(s)* BY=THE-M, BY-THE-F “by the” and the sequence *tódolos*, *tódalas* ALL=THE-M.PL, ALL=THE-F.PL “all the” which are today probably lexicalized (see Section 2.2), while variation between *comer o pan* and *come-lo pan* is possible.

We do not have lemmatized corpora for either medieval or contemporary spoken Galician; thus, it is very difficult to obtain reliable statistics. Bugarín López (2007)<sup>9</sup> is a frequencies dictionary that can be used as an approximation. Some figures are presented in Table 2.

The frequency of /s.l/, /r.l/, and /n.l/ where /l/ belonged to the definite article caused these sequences to be adapted to the syllabic restrictions described by Venne-mann. This could explain why even though we find /r.l/ in the same prosodic domain in *comer lo (pan)* and *merlo*, we have different outcomes. The high ranking of DOMAIN-

<sup>8</sup> Kikuchi (2005) argues that /l/ is a continuous segment since /b, d, g/ have continuous allophones after /l/. Thus, the sequences /s.l/ and /r.l/ would violate OCP, since /s/ and /r/ are continuous. However, Kikuchi overlooks the fact that while /lb/ and /lg/ are realized as [lβ] and [lɣ], respectively, /ld/ is not realized as [lð] but as [ld], which may mean that /l/ is [-cont] (see Dubert, 2005, for the distribution and history of voiced stops and approximants in Galician).

<sup>9</sup> Bugarín López’s (2007) corpus contains 2,064,874 tokens obtained from written and oral sources. Unfortunately, because of the way it was conceived it does not represent language used by speakers in natural interactions. Due to the way entries are organised, it is also not always possible to identify the morphosyntactic representation of the forms presented. Figures given from this source should therefore be treated with caution, as a rough approximation.

CONTIGUITY in the word internal sequences might have derived from low frequency of use; the high ranking of constraints based on syllabic structure and strength in the case of the definite article might have been the outcome of high frequency of use.

**Table 2.** Frequencies of use of some items containing /s.l/, /r.l/ vs. *tódolos-tódalas*, 2FA, *polo-pola*, and Art-Def. (from Bugarín López, 2007).

item	token no.	frequency index
/s.l/ <i>esluír</i>	5	0.000242149%
/s.l/ <i>desleal</i>	8	0.000387438%
/r.l/ <i>garlopa</i>	9	0.000425868%
/r.l/ <i>charla</i>	19	0.000920165%
/s.l/ <i>deslizar</i>	26	0.001259173%
/r.l/ <i>burla</i>	36	0.001743471%
/s.l/ <i>trasladar</i>	59	0.002857355%
/r.l/ <i>merlo</i>	72	0.003486941%
<i>tódolos-tódalas</i>	1,668	0.080780804%
2FA (excluding <i>polo</i> )	9,757	0.472528957%
<i>polo(s)-pola(s)</i>	9,807	0.474950444%
Art-Def	273,289	13.23531477%

It is known that frequency of use is a determining factor in the organization and reorganization of grammar, as Bybee (2001, 2010) shows. Grammar and grammatical structures emerge from use. Frequency and adjacency are powerful explanatory tools to account for the historical development of the morphology of the Galician definite article. Units that tend to appear together may end up forming new, more complex units (this is the origin of the contractions in (4)). Strictly linguistic notions such as domains, contexts and strength may not suffice to explain the structure of the language, or at least to account for the phonological changes we have witnessed.

In the usage based model, it is assumed that every time a linguistic item is used it leaves a trace in the speaker's memory (Bybee, 2001; Langacker, 2000; Pierrehumbert, 2002): the more frequent the use, the stronger the trace. If the constructions [s###l/-Def-Art], [r###l/-Def-Art], and [n###l/-Def-Art] were frequent enough, they might have been memorized and automatized as chunks, in such a way that they may have been treated as units and adapted to the pressures of the strength scales.

On the other hand, according to Venneman (1988) the assimilations found in /s.l/ and /n.l/ should have favoured the weaker segment, so that both sequences ought to have merged into /l/; this is true of /s.l/ (*comes lo* > *comelo*), but not of /n.l/, where the stronger segment assimilated the weaker, giving /n/ (*comen lo* > *comen no* > *comeno*). However, there exists a clear association between the syntactic label Verb, the morphosyntactic property 3PL and a nasal segment at the end of the word form of 3PL

verbs; this could have constituted a schema that strongly links all these features (Langacker, 2000; Taylor, 2002). Besides, it should be noted that indices of frequency of use of 3PL forms are high; the common use of 3PL, expressed with [+nasal], reinforces the link between the properties and the nasal feature, through the lexical connections that *comen lo* establishes with *comen*. This could be another expression of faithfulness constraints, which may be the combined results of both frequency and the analogical pressures of strong schemata like this one. That may be the reason why the less optimal option was selected regarding the direction of assimilation in /n.l/: it seemed preferable to maintain the expression with a nasal segment of the 3PL. Furthermore, if /s.l/ and /n.l/ had both developed to /l/, the word forms of the 2SG and 3PL of verbs would have become homophones.

In this section I have been using concepts such as *constructions* and *schemata*, taken from a cognitive, functionalist, non-derivative model of grammar. This model is not syntactocentric in the sense of Jackendoff (2002), and does not allow two processes of lexical insertion in a syntactic structure previously designed by the syntax, or an opposition between surface and underlying forms; hence neither rules nor constraints are needed to link the two levels. In this model, schemata and constructions freely combine information from different components (phonology, syntax and semantics), so different levels of phonology are not necessary: some schemata contain only phonological information, while others contain information from other components of the grammar. The model naturally incorporates pure phonological processes and morphosyntactically conditioned phonological processes, without needing to distinguish between *lexical*, *postlexical 1* and *postlexical 2* levels. In fact, the model was conceived to incorporate information about syntax, semantics and phonology in a single schema. The contexts that block 2FA in present-day Galician (described meticulously by Fernández Rei, 2002, pp. 223–232) show how difficult it is to deal with these exceptions without referencing syntactic information.

## 6. CONCLUSIONS

In this article I have discussed the linguistic conditions which allowed the phonological changes that created morphological alternations in the Galician definite article. I have assumed that these linguistic conditions reflected the grammar at the time when the changes took place. The definite article has three allomorphs whose selection depends on the final segment of the preceding word. The current Galician definite article resembles a simple clitic in that it may attach to different word classes and its position is syntactically determined; however, like an affix, it has no stressed allomorphs and its phonological form may be affected by the host. It seems to be a di-tropic clitic, phonologically attached to the preceding word but maintaining direct syntactic relations with the following words. I have set out to describe the linguistic



state in which those allomorphs developed in Old Galician, between the ninth and tenth centuries.

The three allomorphs are the historical outcomes of phonetic assimilations; these assimilations were word-specific, since they only affected the definite article and the 3.ACC clitic. In order to explain them, we are forced to contemplate interaction between purely linguistic notions such as syllabic structure, segmental changes, scales of segmental strength, word classes, schemata, constructions, etc., and the frequency of use of the linguistic items involved in the assimilations. After presenting some of Vennemann's laws regarding syllabic structure, I propose that an explanation using an exemplar model should be possible, since this model is designed to combine in a single scheme or construction precise phonetic, morphosyntactic and semantic information and to permit interaction between different units stored in memory, which in generative models have to be created "on the fly".

## ACKNOWLEDGEMENTS

This article forms part of the project 'Cambio lingüístico no galego actual', Ministerio de Economía y Competitividad del Gobierno de España, FFI2012-33845; it was also supported by the Xunta de Galicia, GRC2013/40, and the ERDF. I wish to express my gratitude for the comments by Alan King, Elisa Fernández Rei and two anonymous reviewers of this article.

## REFERENCES

- Álvarez Blanco, R. (1983). O artigo en galego. Morfoloxía. *Verba: Anuario Galego de Filoloxía*, 10, 169–182.
- Álvarez, R., & Xove, X. (2002). *Gramática da lingua galega*. Vigo: Galaxia.
- Bassols de Climent, M. (1983). *Fonética latina*. Madrid: CSIC.
- Bugarín López, M. X. (2007). *O galego fundamental. Diccionario de frecuencias*. A Coruña: Fundación Pedro Barrié de la Maza.
- Bybee, J. (2001). *Phonology and language use*. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511612886>
- Bybee, J. (2010). *Language, usage and cognition*. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511750526>
- Catford, J. C. (1988). *A practical introduction to phonetics*. Oxford: Clarendon Press.
- Cysouw, M. (2005). Morphology in the wrong place. A survey of preposed enclitics. In W. U. Dressler, D. Kastovsky, O. E. Pfeiffer, & F. Rainer (Eds.), *Morphology and its demarcation* (pp. 17–37). Amsterdam: John Benjamins. <http://dx.doi.org/10.1075/cilt.264.02cys>
- Dubert García, F. (1998). *A fala de Santiago de Compostela. Estudio xeolingüístico. Apéndice. Textos*. Santiago de Compostela: Universidade de Santiago de Compostela.
- Dubert García, F. (1999). *Aspectos do galego de Santiago de Compostela*. Santiago de Compostela: Universidade de Santiago de Compostela.
- Dubert García, F. (2001). A alomorfía do artigo definido galego á luz da fonoloxía prosódica. In X. L. Regueira Fernández & A. Veiga (Eds.), *Da gramática ó dicionario. Estudos de lingüística galega* (pp. 91–105). Santiago de Compostela: Universidade de Santiago de Compostela.
- Dubert García, F. (2004). Achegas sobre a morfoloxía do artigo determinado e do pronome átono acusativo de terceira persoa. In R. Álvarez Blanco, Fernández Rei, F., & A. Santamarina (Eds.), *A lingua galega: Historia e Actualidade. Volume IV* (pp. 59–69). Santiago de Compostela: Consello da Cultura Galega & Instituto da Lingua Galega.
- Dubert García, F. (2005). Problemas na orixe da alternancia dos alófonos dos fonemas /b d g/. In A. I. Boullón Agrelo, X. L. Couceiro Fernández, & F. Fernández Rei (Eds.), *As tebras alumeadas. Estudos filolóxicos ofrecidos en homenaxe a Ramón Lorenzo* (pp. 323–334). Santiago de Compostela: Universidade de Santiago de Compostela.
- Fernández Rei, E. (2002). *Regras fonolóxicas posléxicas e regras precompiladas de alomorfía sintagmática: Dominios prosódicos en galego*. Santiago de Compostela: Universidade de Santiago de Compostela.
- Fernández Rei, E. (2005). A dirección de clitización fonolóxica no galego actual. In A. I. Boullón Agrelo, X. L. Couceiro Fernández, & F. Fernández Rei (Eds.), *As tebras alumeadas. Estudos filolóxicos ofrecidos en homenaxe a Ramón Lorenzo* (pp. 335–346). Santiago de Compostela: Universidade de Santiago de Compostela.
- Fernández Rei, E. (2012). A clitización fonolóxica en galego e en portugués brasileiro. In X. Lagares & H. Monteagudo (Eds.), *Galego e Português Brasileiro: História, variação e mudança* (pp. 105–127). Niterói-Rio de Janeiro: Universidade General Fluminense.
- Ferreiro, M. (1995). *Gramática histórica galega. Volume I. Fonética y Morfosintaxe*. Santiago de Compostela: Laiovento.
- Freixeiro Mato, X. R. (2000). *Gramática da lingua galega, II: Morfosintaxe*. Vigo: A Nosa Terra.
- García, C., Santamarina, A., Álvarez Blanco, R., Fernández Rei, F., & González González, M. R. (1995). *Atlas lingüístico galego. Vol. II. Morfología non verbal (ALGa2)*. A Coruña: Fundación Pedro Barrié de la Maza.
- González Seoane, E. (Coord.), Álvarez de la Granja, M. A., & Boullón Agrelo, A. I. (2006–2012). *Dicionario de dicionarios do galego medieval (DDGM)*. Instituto da Lingua Galega. Available at <http://sli.uvigo.es/DDGM/>
- Hayes, B. (1990). Precompiled phrasal phonology. In S. Inkelas & D. Zec (Eds.), *The phonology-syntax connection* (pp. 85–108). Chicago, IL: The University of Chicago Press.
- Hooper, J. B. (1976). *An introduction to natural generative phonology*. New York, NY: Academic Press.
- Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press. <http://dx.doi.org/10.1093/acprof:oso/9780198270126.001.0001>
- Kaisse, E. M. (1985). *Connected speech: The interaction of syntax and phonology*. New York, NY: Academic Press.
- Kikuchi, S. (2005). The emergence of the unmarked in Galician plural formation. *On'in Kenkyu [Phonological studies]*, 8, 17–24.
- Kikuchi, S. (2006). On Galician definite article allomorphy. *On'in Kenkyu [Phonological studies]*, 9, 41–48.
- Ladefoged, P., & Johnson, K. (2010). *A Course in phonetics* (6th ed.). Boston, MA: Wadsworth/Cengage Learning.
- Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Oxford: Blackwell.
- Langacker, R. W. (2000). A dynamic usage-based model. In M. Barlow & S. Kemmer (Eds.), *Usage based models of language* (pp. 1–63). Stanford, CA: Center for the Study of Language and Information.
- Laver, J. (1994). *Principles of phonetics*. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781139166621>
- Martinet, A. (1996). *Éléments de linguistique générale* (4th. ed.). Paris: Armand Colin.
- Matthews, P. H. (1991). *Morphology* (2nd. ed.). Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781139166485>
- Mattos e Silva, R. V. (1989). *Estruturas trecentistas: Elementos para uma gramática do português arcaico*. Lisboa: Imprensa Nacional-Casa da Moeda.
- Monteil, P. (2003). *Elementos de fonética y morfología del latín* (C. Fernández Martínez, Trans.). Universidad de Sevilla (Original work published 1970).



- Nespor, M., & Vogel, I. (2007). *Prosodic phonology. With a new foreword*. Berlin: Mouton de Gruyter.
- Pierrehumbert, J. B. (2002). Word-specific phonetics. In C. Gussenhoven & N. Warner (Eds.), *Laboratory Phonology 7* (pp. 101–140). Mouton de Gruyter, Berlin. <http://dx.doi.org/10.1515/9783110197105.101>
- Real Academia Galega/Instituto da Lingua Galega (2004). *Vocabulario ortográfico da lingua galega (VOLGa)* [corpus]. Retrieved from <http://www.realacademiagalega.org/recursos-volg>
- Regueira, X. L. (Ed.) (1998). *Os sons da lingua*. Vigo: Xerais.
- Spencer, A., & Luís, A. R. (2012). *Clitics. An Introduction*. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781139033763>
- Taylor, J. R. (2002). *Cognitive grammar*. Oxford: Oxford University Press.
- Trubetskoy, N. S. (1929). Sur la ‘morphologie’. *Travaux du Cercle Linguistique de Prague*, 1, 85–88.
- Vennemann, T. (1988). *Preference laws for syllable structure and the explanation of sound change: With special reference to German, Germanic, Italian, and Latin*. Berlin: Mouton de Gruyter.
- Williams, E. B. (1962). *From Latin to Portuguese. Historical phonology and morphology of the Portuguese language* (2nd. ed.). Philadelphia: University of Pennsylvania Press.
- Zwicky, A. M., & Pullum, G. K. (1983). Cliticization vs. inflection: English *n't*. *Language*, 59(3), 502–513. <http://dx.doi.org/10.2307/413900>