Manuel Badal

High frequency as a morphomic coherence inhibitor: evidence from the evolution of some Catalan verbs[*]

Abstract: During the Middle Ages, a new verb class was formed within the Catalan 2nd conjugation, characterized by the presence of a velar augment (/g/ or /sk/) in the L-pattern (first-person present indicative and present subjunctive) and PYTA (preterite, old conditional and imperfect subjunctive). In this paper, we study three verbs that avoided, totally or partially, the analogical velarization undergone by most 2nd conjugation verbs: *haver* 'to have', *ser* 'to be', and *voler* 'to want'. Based on a corpus of Catalan works ranging from the 13th to the 19th century, we test the extent to which the L- and PYTA-patterns of these three verbs are psychologically real for speakers. The results confirm what some authors have already observed for other languages, pointing out that the high frequency of these specific verbs or the need for phonological congruence could inhibit the morphomic pattern or even block analogical levelling.

Keywords: augment, Catalan, morphology, morphome, velarization, verbs.

1 Introduction

During the Middle Ages, a new verb class was formed within the Catalan 2nd conjugation, characterized by the presence of a velar augment (/g/ or /sk/, depending on the verb) at certain cells of the paradigm (regarding this diachronic process, see, e.g., Pérez Saldanya 1998; Wheeler 2011; Badal 2021). The distribution of the velar can be

[*] This research, funded by the Ministerio de Universidades (MS21-121) and by the European Union (NextGenerationEU), is part of the project PID2020-113971GB-C21 (https://stel.ub.edu/gevad), funded by the Spanish Agencia Estatal de Investigación (MCIN/AEI/10.13039/501100011033). I thank Professor Lluís de Yzaguirre for his help with the verb form extraction using a LATEL (Laboratori de Tecnologies Lingüístiques) application from the Universitat Pompeu Fabra. I am also grateful for the comments and suggestions of Éva Buchi, Jesús Jiménez, Maria-Rosa Lloret, Manuel Pérez Saldanya and Clàudia Pons-Moll.

considered “morphomic” (as termed by Aronoff 1994), since the systematic appearance of this augment “cannot be aligned with any conceivable coherent semantic, syntactic, or phonological generalization” (Esher/O’Neill 2022, 351). The resulting layout from the analogical velarization process favored an implicative organization of the allomorphs, which facilitated the acquisition of the paradigm.

As illustrated in Table 1 with the verbs *beure* ‘to drink’ and *créixer* ‘to grow’, the velar augment usually appears in the L-pattern (formed by the first-person present indicative and the whole present subjunctive) and PYTA (formed by the preterite, the old conditional and the imperfect subjunctive),[1] and also in the past participle if this form adopts the regular ending *-ut*/*-uda* (except for the verb *vendre* ‘to sell’, which presents the non-velarized form *venut* ‘sold’).[2] Root-stressed past participles, as in the case of the verbs *dir* ‘to say’ or *prendre* ‘to take’, do not adopt the velar: *dit* ‘said’, *pres* ‘taken’.

		beure ‘to drink’ /g/	*créixer* ‘to grow’ /sk/
L-pattern	1sg. present indicative	*be**c*** [ˈbe**k**]	*cre**sc*** [ˈkɾe**sk**]
	Present subjunctive	3sg. *be**g**a* [ˈbe**ɣ**a]	3sg. *cre**sc**a* [ˈkɾe**sk**a]
PYTA	Preterite	3sg. *be**gu**é* [beˈ**ɣ**e]	3sg. *cre**squ**é* [kɾe**s**ˈ**k**e]

[1] The expression L-pattern refers to a distribution in which a distinctive form is shared uniquely by the whole of the present subjunctive together with the first-person singular present indicative (Maiden 2018, 84). PYTA, an acronym for the expression *perfecto y tiempos afines* ‘perfect and related tenses’ (Menéndez Pidal 1904), is formed by the Romance reflexes of Latin perfective forms (Maiden 2001). In Old Catalan, PYTA comprises the preterite, coming from Latin perfect indicative, the conditional ending in *-ra*, going back to Latin pluperfect indicative, and the imperfect subjunctive, coming from Latin pluperfect subjunctive.

[2] The imperative 2sg. and 2pl. usually adopt the velar in the verbs that present the *-s* mark in the 2sg. of the present subjunctive, as in *dir* ‘to say’ (2sg. *digues* [ˈdiɣes], 2pl. *digats* [diˈɣats] ~ *digau* [diˈɣaw]). In most verbs, such as *beure* ‘to drink’, these persons have the same form in the imperative as in the present indicative 3sg. and 2pl.; consequently, they do not present the velar augment: 2sg. *beu* [ˈbew], 2pl. *bevets* [beˈvets] ~ *beveu* [beˈvew].

	Conditional	3sg. *be**gu**era* [beˈ**ɣ**eɾa]	3sg. *cre**squ**era* [kɾe**s**ˈ**k**eɾa]
	Imperfect subjunctive	3sg. *be**gu**és* [beˈ**ɣ**es]	3sg. *cre**squ**és* [kɾe**s**ˈ**k**es]
Past participle		*be**g**ut* [beˈ**ɣ**ut]	*cre**sc**ut* [kɾe**s**ˈ**k**ut]

Table 1. L-pattern, PYTA-pattern, and past participle of *beure* 'to drink' (exponent of /g/ model) and *créixer* 'to grow' (exponent of /sk/ model) in Old Catalan (in the tenses in which all the forms are velarized, we only offer the 3sg.)

The distribution of the velar augment shown in Table 1 is the result of a diachronic change, which led to the "coalescence" (as termed by Maiden 2018, 292) of the L and PYTA morphomic patterns. Regarding the origin of the velar, Catalan has velarized forms that come either from the tenses of the Latin *infectum*, in particular the first-person present indicative and the present subjunctive, or from the Latin *perfectum*, specifically the preterite, the old conditional ending in *-ra* (coming from Latin pluperfect indicative) and the imperfect subjunctive (coming from Latin pluperfect subjunctive). In the forms coming from Latin *infectum* (i.e., the L-pattern), the velar was part of the root. This group of verbs split into three classes (Wheeler 2011, 189): *a*) verbs in which the /k/ was intervocalic, such as DĪCŌ > *dic* [ˈdik] 'I say'; *b*) verbs with a post-consonantal -G-, as in PLANGŌ > *planc* [ˈplaŋk] 'I feel sorry (for)', and *c*) root-inchoative verbs, which had the augment /sk/, like */ˈnasko/ > *nasc* [ˈnask] 'I am born'.[3]

In the tenses coming from Latin *perfectum* (i.e., PYTA), the velar augment was initially found in verbs with the perfect affix -U- (Pérez Saldanya 2013, 317), as in

[3] The outcoming forms of regular sound change are placed after the symbol >, while the symbol >> corresponds to system-morphological reorganizations, due to analogy. As for the IPA phonetic transcription, we adopt the Western Catalan forms (see IEC 2016).

DĒBUIT > *dec* [ˈdek] ‘s/he owed’ or in MOLUISSET > *molgués* [molˈɣes] ‘s/he ground’. The origin of the velar in these cases does not seem as clear as in the forms coming from Latin *infectum*. Most authors follow Meyer-Lübke’s proposal (1895, §278), according to which /g/ derives from the strengthening of the semivowel [w] that functioned as a perfect marker (-U-): *w* > g^w > *g*; however, Wheeler (2011, 195–198) disagrees with this account and follows Ronjat’s proposal (1937, §570). According to this view, the /g/ formative extended analogically from forms such as NOCUIT > *noc* [ˈnɔk] ‘s/he harmed’ or PLACUISSET > *plagues* [plaˈɣes] ‘s/he pleased’, in which the /g/ is undoubtedly etymological. On the other hand, the cases of the variant /sk/ in PYTA have two sources (Wheeler 2011, 198): either it derives from an earlier irregular metathesis process in certain -X- *perfectum* stems (as in VĪXIT > *[ˈvisket] > *visc* [ˈvisk] ‘s/he lived’), or it comes from the transfer of an L-pattern /sk/ stem to the PYTA set in root-inchoative verbs (originating from forms such as *[ˈnaskuit] > *nasc* [ˈnask] ‘s/he was born’).

Nevertheless, there are certain verbs that did not fully adopt this velar distribution pattern. In this paper, we analyze the causes that explain the apparently incoherent spread of velarization in three Catalan 2nd conjugation verbs with a remarkably high frequency: *haver* ‘to haver’, *ser* ‘to be’ and *voler* ‘to want’. Based on the results obtained from a corpus of texts ranging from the 13th to the 19th century, we test whether a “coherence” effect took place in the diachronic change of these verbs, in line with Maiden’s (2018, 3) prediction that “morphological innovations of any kind (e.g. analogical levelling of the alternation, analogical extension of the alternation, creation of novel alternants, introduction of suppletive forms into paradigms) that affect any one of the paradigmatic cells implicated in the alternation pattern equally and

always affect all the others." Our initial assumption is, thus, that if two morphomes are psychologically real for speakers (Maiden 2018, 13), all the forms associated to them must undergo the same analogical change. We will show that the corpus data contradict this hypothesis, though cast some light on the factors that can hinder the processes of morphological levelling and regularization.

The paper is organized as follows. In section 2, we set out the methodology followed to extract and analyze the cases. Next, in sections 3, 4, and 5, we present the traits of these verbs along with the counts extracted from our corpus and the data analysis. Finally, in section 6, we offer some conclusions.

2 Methodology

To extract the verb forms, we built a corpus comprising Catalan works ranging from the 13th to the 19th century (see Table 2). We chose this period to be able to study the velarized forms from the first Catalan texts until the promulgation of Fabra's orthography and grammar rules that form the basis of the Catalan norm;[4] the inclusion in the study of the 20th century would have introduced a clearly independent variable of the morphome influence: the impact of the rules, with possible effects especially on the analyzed forms that have not been admitted in the standard language, such as *vullc* [ˈvuʎk] 'I want'. In addition, all the corpus texts belong to the second half of each century, since, methodologically, the objective is to determine the evolution of verb forms at the end of each century. We recognize the risk of choosing only the second half of each century, since the velarization process may have started a little earlier than what has been documented in our corpus. However, our intention is to observe the

[4] Pompeu Fabra (1868–1948) is the author and promoter of modern-day Catalan. He wrote the *Normes ortogràfiques* (1913), the *Gramàtica catalana* (1918) and the *Diccionari general de la llengua catalana* (1932), among others.

overall evolution of verb forms; therefore, we consider our choice of time periods appropriate, since sudden changes would rarely occur half a century apart. For each century, our corpus contains 180,000 words: 90,000 of which are extracted from Western Catalan works, and the other half from Eastern Catalan works. This selection aims to get a balanced representation of the two major Catalan dialects. To obtain more homogeneous data, we decided to reduce the number of dialects in the corpus. Thus, most of the Western Catalan texts come from Valencian works, while most of the Eastern Catalan texts belong to Central Catalan. To optimize the verb form extraction, we used a LATEL (Laboratori de Tecnologies Lingüístiques) application from Pompeu Fabra University (De Yzaguirre 2018).[5] When the corpus data was insufficient, or when it seemed useful to corroborate the obtained results, we turned to the *Corpus Informatitzat del Català Antic* (hereinafter, *CICA*) for the Middle Ages, and to the *Corpus Informatitzat de la Gramàtica del Català Modern* (hereinafter, *CIGCMod*) for the early modern period.

Centuries	Western Catalan	Eastern Catalan
13th century	*Costums de Tortosa* (1272)	*Llibre de contemplació en Déu* (1271–1274)
14th century	*Epistolari de la València Medieval (I)* (1356–1399), *Els manuals de consells medievals de Xàtiva* (1376–1380)	*La fiyla del rey d'Ungria* (1350–1374), *Els quatre llibres de la reina Elionor de Sicília* (1367), *Corbatxo* (1397), *Lo somni* (1398)

[5] Access to the database is free, but a previous registration is mandatory.

15th century	*Tirant lo Blanch* (1490)	*Llibre de les solemnitats de Barcelona* (1450–1497), *Històries e conquestes dels reys de Aragó e comtes de Barcelona* (1495)
16th century	*Constitucions del convent de Sant Josep* (1550–1574), *Libre de Antiquitats* (1550–1598), *Textos de la Cancelleria Reial (Arxiu del Regne de València)* (1554–1585), *Les Memòries del cavaller valencià Gaspar Antist* (1555–1557), *Biografia de Català de Valeriola* (1568–1607)	*Llibre de les Solemnitats de Barcelona* (1564–1599), *Diari de Frederic Despalau* (1588–1600)
17th century	*Notícies de València i son regne* (1661–1679)	*Dietaris de la Generalitat de Catalunya* (1674–1689)
18th century	*Memòries d'un capellà del segle XVIII, Josep Esplugues* (1732–1782), *Col·loquis valencians atribuïts a Carles Leon* (1784–1797)	*Memòries d'un pagès del segle XVIII* (1753–1757), *Calaix de sastre* (1769–1786)
19th century	*Ensisam de totes herbes* (1891)	*Lo catalanisme* (1886)

Table 2. Works forming the corpus, divided by centuries (rows) and dialects (columns)

Once the counts were completed, to ascertain whether or not the distribution of the non-velarized and velarized forms ("velarized" variable) in relation to the centuries ("century" variable) was random, we conducted several chi-square tests with SPSS Statistics (IBM Corp. 2019). The chi-square test is based on the comparison of the bivariate frequencies obtained from the data (empirical frequencies) with the frequencies that would result if there were no association between the "velarized" and "century" variables (theoretical frequencies). The test produces two indicators: the χ^2 value for a two-variable distribution –non-velarized and velarized forms– in the periods in which the seven centuries are grouped (with three, four or six degrees of freedom, depending on the table model) and the asymptotic significance (*p*). The *p*-value is evaluated from the threshold of 0.05, as is usual in social sciences: when the *p*-value is less than 0.05, the probability that the elements have been randomly distributed according to the global frequency between the different groups is lower (less than 5%); in this case, the null hypothesis should be discarded, according to which the "century" variable does not influence the distribution of verb forms and it should be assumed that, on the contrary, the distribution of these forms in the different centuries varies significantly. On the other hand, if the *p*-value is greater than 0.05 or, in other words, if the probability of obtaining the real random distribution is greater than 5%, the null hypothesis will be accepted and it will be assumed that the verb forms have been distributed with the same criterion in the different centuries. After checking whether or not the forms are randomly organized, it is necessary to test the association between the variables. This data is determined using Cramér's V, with a parameter range of 0 to 1: 0.1 to 0.2 indicating a weak association; 0.2 to 0.4, a moderate one; 0.4 to 0.6, a

relatively strong one; 0.6 to 0.8, a strong one; and 0.8 to 1, a very strong one (Rea/Parker 2014, 219).

3 The verb *haver* 'to have'

The verb *haver* 'to have' is one of the most commonly used Catalan verbs, mainly as a perfect auxiliary (see IEC 2016), and presents certain irregularities in the paradigm, related both to vowel alternations in the root in some present indicative forms, and to the presence of a velar consonant in the PYTA cells. As illustrated in Table 3, regular sound change gave rise to three different roots for the present indicative forms, and a different one for the present subjunctive: in the first, *he*, *ha*- and *hav*-; in the second, only [ˈad͡ʒ-] (<haj-> ~ <hag->). Additionally, it should be noted that it did not present a uniform L-pattern, since the first person of the present indicative had a different root from the present subjunctive forms. In contrast, all PYTA cells presented a velar consonant (/g/) in Old Catalan, as illustrated in Table 4.

	1sg.	2sg.	3sg.	1pl.	2pl.	3pl.
PRS.IND	HABEŌ > *he*	HABĒS > *has*	HABET > *ha*	HABĒMUS > *havem*	HABĒTIS > *havets*	HABENT > *han*
PRS.SBJV	HABEAM > *haja*	HABEĀS > *hages*	HABEAT > *haja*	HABEĀMUS > *hajam*	HABEĀTIS > *hajats*	HABEANT > *hagen*

Table 3. Evolution from Latin to Catalan of the present indicative and the present subjunctive of the verb *habēre* > *haver* 'to have'

	1sg.	2sg.	3sg.	1pl.	2pl.	3pl.
PRT	*hac*	*haguist*	*hac*	*haguem*	*hagués*	*hagren*
COND	*hagra*	*hagres*	*hagra*	*hàgrem*	*hàgrets*	*hagren*
IPFV.SBJV	*hagués*	*haguesses*	*hagués*	*haguéssem*	*haguéssets*	*haguessen*

Table 4. PYTA morphome of the verb *haver* ‘to have’ in Old Catalan

In this paper, we study two different analogical changes affecting both patterns. Regarding the L-pattern, in our corpus, velarized forms are attested, but only in the present subjunctive. Due to this analogical change, the etymological root [ˈad͡ʒ-] (<haj-> ~ <hag->) is replaced by forms with the /g/ formative, like *haga* [ˈaɣa] ‘s/he have’ instead of *haja* [ˈad͡ʒa] ‘s/he have’. As mentioned above, the velar consonant was typical of PYTA forms, so there was an analogical extension from the forms coming from Latin *perfectum* to those of the present subjunctive, in a similar way to what happened in the case of verbs like *beure* ‘to drink’ or *deure* ‘to owe’, but without propagation towards the first person of the present indicative. As for PYTA forms, some cases of develarization are attested in our corpus, in which the root with the velar augment is replaced by a more regular one, *hav-* (e.g., *haguera* [aˈɣeɾa] >> *havera* [aˈveɾa] ‘s/he would have’), in a case of analogical levelling. In both cases, the aim is to date when this process begins and to see if it consolidates.

3.1 Present subjunctive results

To show the counts extracted from the corpus, we use a cross table consisting of two variables: with the centuries in the rows (some of them grouped), and the “velarized” variable in the columns, under which the cases are divided into two columns, depending on whether or not they present the velar consonant. According to the corpus data, the percentage distribution between non-velarized (e.g., *haja* ‘s/he have’) and velarized forms (e.g., *haga* ‘s/he have’) hardly varies over the period studied in the present subjunctive of the verb *haver* ‘to have’ (see Table 5). The test reveals that there are no significant differences in the distribution of the verb forms between the centuries ($\chi^2_{(3)} = 4.990$, $p = 0.172$, Cramér’s V = 0.107).

			Non-velarized	Velarized
Period	13th–14th centuries	Count	195	3
		% within century	98.5%	1.5%
	15th century	Count	75	1
		% within century	98.7%	1.3%
	16th century	Count	32	—
		% within century	**100%**	0%
	17th–19th centuries	Count	120	6
		% within century	95.2%	4.8%

Table 5. Cross table of the present subjunctive of the verb *haver* 'to have'

Except for the 16th century, we document some velarized verb forms, such as those of (1). However, these cases are always in the minority with respect to the etymological forms (2), since the velarized forms do not exceed 5% of total cases in any period.

(1) a. que cascun d'aquells *haga* béns seus propris apartats ['that each of those have goods in their own sections'] (*Xàtiva* 53 [14th century])

b. d'aquí que la evolució de las ideas *hagui* de ser constant é indefinida ['hence the evolution of ideas must be constant and indefinite'] (*Lo catalanisme* 149 [19th century])

(2) a. Al que dius no és rahó que *aja* alguna valor ['What you say is not reason that has any value'] (*Corbatxo* 11 [14th century])

b. no's cregui pas que s'*hagi* ja fet lo darrer pronunciament ['don't think that the last pronouncement has already been made'] (*Lo catalanisme* 17 [19th century])

All in all, we see how, over the centuries, the number of velarized forms is meager compared to the number of etymological ones. Therefore, neither the presence of the velar consonant in the PYTA forms nor the velarization of the L-pattern of verbs such as *caure* ‘to fall’ or *deure* ‘to owe’, which also had an etymological root ending in /dʒ/ (e.g., *caja* [ˈkad͡ʒa] ‘s/he fall’ or *deja* [ˈded͡ʒa] ‘s/he owe’) and a velarized PYTA (e.g., *caigués* [kajˈɣes] ‘s/he fell’ or *dec* ‘s/he owed’), influenced the verb *haver* ‘to have’. The maintenance of the etymological forms in the present subjunctive could be explained by this verb’s high frequency, which is an important factor in preserving allomorphs and blocking linguistic change (Bybee/Brewer 1980, 218; Anshen/Aronoff 1988; Rainer 1988; Booij 1997, 43). Additionally, since the first person of the present indicative did not adopt the velar consonant, it could not exert the same force as it did in the case of verbs such as *caure* ‘to fall’ or *deure* ‘to owe’, in which /g/ was first introduced into the first-person present indicative and then extended to the present subjunctive because of the class-stability principle (Wurzel 1987).[6]

3.2 PYTA results

The preterite and the imperfect subjunctive counts are grouped (see Table 6) since they present a unitary behavior: all the documented forms of these two tenses are velarized, as illustrated in (3).

		Non-velarized	Velarized
Period	13th–14th centuries	—	203

[6] According to this principle, the introduction of the velar in the first person of the present indicative should entail the velarization of the whole present subjunctive, due to the implicative relationship established between these forms (Bybee 1985, 68–74). Following this association, when the first person of the present indicative presents a specific root, this allomorph normally reappears in the whole of the present subjunctive (Pérez Saldanya et al. 2004, 97).

	15th century	—	391
	16th century	—	79
	17th–19th centuries	—	287

Table 6. Cross table of the preterite and the imperfect subjunctive of the verb *haver* 'to have'

(3) a. ab un vidrier que li fou malcriat que li *hagué* de paguar ab la gineta ['with a glazier who was spoiled and had to pay him with the rider'] (*Despalau* 103 [16th century])

b. quant se'n manaren las eugas *agesan* tanbé pascut agafar las bacas ['when the mares were sent away they would also have grazed and taken the cows'] (*Memòries d'un pagès* 49 [18th century])

The only PYTA tense in which cases of analogical levelling are attested is the old conditional. In this tense, the percentage distribution between non-velarized (e.g., *havera* 's/he would have') and velarized forms (e.g., *haguera* 's/he would have') varies throughout the studied period (see Table 7), with significant differences in the distribution of verb forms between centuries, due to the change in trend in the 19th century ($\chi^2_{(6)} = 112.586$, $p < 0.001$, Cramér's V = 0.784).

			Non-velarized	Velarized
Period	13th century	Count	—	8
		% within century	0%	**100%**
	14th century	Count	—	58
		% within century	0%	**100%**
		Count	—	20

	15th century	% within century	0%	**100%**
	16th century	Count	—	8
		% within century	0%	**100%**
	17th century	Count	—	16
		% within century	0%	**100%**
	18th century	Count	—	44
		% within century	0%	**100%**
	19th century	Count	19	10
		% within century	65.5%	34.5%

Table 7. Cross table of the conditional of the verb *haver* 'to have'

Until the 18th century, all documented forms present the velar augment (4).

(4) a. car la dita ciutat no m'*haguera* consellat que jo us faés mal ['for the said city would not have advised me to harm you'] (*Lo somni* 126 [14th century])

b. ab la facultat de demanar als oficials el vagatges que *haguera* cada ú de menester ['with the power to ask the officers for the luggage that each one needs'] (*Esplugues* 60 [18th century])

In the 19th century, there is a statistically relevant change, since forms with the regular stem *hav-* (5a) are documented for the first time and they become at least as common (65.5% of cases) as velarized forms (5b), which, according to their frequency, seem to have taken the first step towards recession (34.5% of cases).

(5) a. Eixe fadrinet que pasa per lo carrer tan templát, si no fora per Tomasa, ningú l'*havera* miràt ['That handsome boy walking down the street, if it weren't for Tomasa, no one would have looked at him'] (*Ensisam* 35 [19th century])

b. y dasta *haguera* donat la sang de les sehues venes si *haguera* segut menester ['and even would have given the blood from his veins if it had been necessary'] (*Ensisam* 237 [19th century])

From the 17th century onwards, all the conditional cases (both types of forms) belong to Valencian works. It should be noted that the forms coming from the old conditional disappeared in general Catalan (apart from specific cases) and were only preserved in Valencian but used as an imperfect subjunctive (Ridruejo 1985). So, according to our corpus data, in most Catalan varieties, the velar consonant is maintained in the PYTA tenses (e.g., *hagué* [aˈɣe] 's/he had' or *hagués* [aˈɣes] '(that) s/he had'), especially in the preterite and in the original imperfect subjunctive. As mentioned in the previous section, the high frequency of this verb probably favors the maintenance of original velarized forms, even though they are discordant with the rest of the forms found in the paradigm, which do not have a velar consonant (e.g., *ha* [ˈa] 's/he has', *havia* [aˈvia] 's/he had', *haurà* [awˈra] 's/he will have'). However, there are varieties, such as southern Valencian, in which there was a root levelling process in the /-ra/ forms coming from the old conditional, like *haguera* [aˈɣera] 's/he would have' or *hagueres* [aˈɣeres] 'you would have'. In this variety, the original forms of the imperfect subjunctive, such as *hagués* [aˈɣes] 's/he had' or *haguesses* [aˈɣeses] 'you had', were no longer used, and the preterite had lost a great deal of vitality. Consequently, the PYTA morphome was blurred, since the only tense used at this time was the old conditional, which was then used as an imperfect subjunctive. Therefore, the fact that, of the three PYTA tenses, only the old conditional was spontaneously used, made it more susceptible to undergoing an analogical levelling process, according to the uniformity principle, which favors paradigms without allomorphs (Mayerthaler 1987,

49). Even so, the forms with the velar consonant, such as *haguera* 's/he would have', are profusely documented at the beginning of the 20th century (see Alcover/Moll 1929–1932).

4 The verb *ser* 'to be'

Ser 'to be' is one of most commonly used and, at the same time, most irregular Catalan verbs, to the extent that it has up to six different suppletive roots, each of which can present several allomorphs with vowel and consonant alternations (see IEC 2016, 270). These singularities are reflected in the present indicative and in the present subjunctive. As happens with the verb *haver* 'to have', in the case of *ser* 'to be', it is not possible to articulate a homogeneous L-pattern either, because the first person of the present indicative and the present subjunctive forms present different roots: *so-* and *si-*, respectively. As illustrated in Table 8, some of the forms of these two tenses come from Classical Latin, and others, marked with an asterisk, from Vulgar Latin.

	1sg.	2sg.	3sg.	1pl.	2pl.	3pl.
PRS.IND	SUM > *so*	ES > *es*	EST > *és*	SUMUS > *som*	*SUTIS > *sots*	SUNT > *són*
PRS.SBJV	*SIAM > *sia*	*SIAS > *sies*	*SIAT > *sia*	*SIAMUS > *siam*	*SIATIS > *siats*	*SIANT > *sien*

Table 8. Evolution from Latin to Catalan of the present indicative and the present subjunctive of the verb *ser* 'to be'

As for the first-person present indicative, apart from the etymological form, *so* 'I am', in Old Catalan, there are two other non-velarized forms, *som* (syncretic with the 1pl.) and *son* (syncretic with the 3pl.). In the counts, all the non-velarized forms are grouped in the same cell, since our goal is to date the appearance of the velarized form *soc* [ˈsok]

‘I am’, and to see when it becomes the most widely used form. Regarding the present subjunctive forms, the analogical velarization must be related to the need to resolve the morphophonological problems of Old Catalan forms (Pérez Saldanya 1998, 84): the subjunctive ending *-a* was preceded by the root-final vowel (*-i*), so that a hiatus was formed. In this way, the addition of the velar augment made it possible to undo the hiatus and to give the syllable a simpler structure (Hualde 1992, 383; Pérez Saldanya 1998, 84), with a consonantal onset: e.g., *si.a* [ˈsia] >> *si.ga* [ˈsiɣa] ‘s/he be’.

With respect to PYTA forms, we only analyze the third person of the preterite, since it is the only form coming from Latin *perfectum* in which we document velarized forms. It is, therefore, an atypical case, since cases of analogical velarization usually affect not only the entire tense, but also the entire morphomic pattern. Etymologically, the PYTA forms of *ser* had no velar augment (see Table 9). Apart from the etymological form *fo* ‘s/he was’, this person presents other forms as *fon* and *fou*. The origin of *fon* is not very clear, since there is no other form of the third person of the preterite of any verb with the *-n* ending that could have served as a base for its formation. The variant *fou*, on the other hand, was formed analogically on other *u*-ending preterites, such as *feu* ‘s/he did’ or *viu* ‘s/he saw’. The velarized variant *fonc* [ˈfoŋk] ‘s/he was’, extensively documented over the centuries, took as its base the form *fon*, to which it added the velar consonant by analogy to forms such as *tenc* [ˈteŋk] ‘s/he had’ or *venc* [ˈveŋk] ‘s/he came’.

	1sg.	2sg.	3sg.	1pl.	2pl.	3pl.
PRT	*fui*	*fuïst*/*fost*	*fo*	*fom*	*fos*	*foren*
COND	*fora*	*fores*	*fora*	*fórem*	*fórets*	*foren*

IPFV.SBJV	*fos*	*fosses*	*fos*	*fóssem*	*fóssets*	*fossen*

Table 9. PYTA morphome of the verb *ser* in Old Catalan

4.1 First-person present indicative results

In the first-person present indicative of *ser* 'to be', the percentage distribution between non-velarized (e.g., *so*, *son* 'I am') and velarized forms (*soc* 'I am') varies notably throughout the studied period (see Table 10). The results obtained from the corpus reveal that there are significant differences in the distribution of verb forms between the centuries ($\chi^2_{(6)} = 244.435$, $p < 0.001$, Cramér's V = 0.910), with an almost total association between the "century" variable and the "velarized" variable.

			Non-velarized	Velarized
Period	13th century	Count	34	—
		% within century	**100%**	0%
	14th century	Count	59	1
		% within century	98.3%	1.7%
	15th century	Count	102	—
		% within century	**100%**	0%
	16th century	Count	1	—
		% within century	**100%**	0%
	17th century	Count	6	—
		% within century	**100%**	0%
	18th century	Count	10	3
		% within century	76.9%	23.1%

	19th century	Count	7	72
		% within century	8.9%	91.1%

Table 10. Cross table of the first-person present indicative of the verb *ser* 'to be'

Until the 17th century, the attested forms do not present the /g/ consonant (6a), except for one case from the 14th century (6b).

(6) a. E despuys que·n tal esser me *só* vist, tostemps per estrem me *só* entristit e alterat. ['And after I am seen in such a being, I am always extremely saddened and disturbed.'] (*Corbatxo* 18 [14th century])

b. Déu qui sol lo cor dels hòmens veu, sap si yo *sóc* dolent e panident del mal comès ['God who alone sees the heart of men, knows if I am bad and penitent for the wrong done'] (*Corbatxo* 88 [14th century])

In the 18th century, however, the percentage distribution of verb forms begins to vary with respect to previous centuries, since, even though the etymological forms are still the most frequent (7a), 23.1% of the verb forms already present the velar consonant (7b).

(7) a. que jo *só* donador en la vila de St. Pera Pescador, qual és lo de St. Andreu ['that I am a donor in the town of St. Pere Pescador, which is St. Andreu'] (*Memòries d'un pagès* 67 [18th century])

b. No parle en xansa, caferro!, que *sóc* home molt formal. ['Don't talk in slang, stupid! I'm a very formal man'] (*Col·loquis de Carles Leon* [18th century])

Finally, in the 19th century, there is a great change, since the velarized forms comes to represent 91.1% of the cases (8a), so that the etymological forms becomes residual (8b).

(8) a. A vore un oncle molt ric men aní'l dichòus pasát; pero com yo *soc* un pòbre, es quedá tot afrontát ['I went to see a very rich uncle last Thursday; but as I am a poor man, he was all ashamed'] (*Ensisam* 18 [19th century])

b. y crega lo que li dic, perque no *só* amic de trampa, qu'es la pura veritat ['and believe what I tell you, because I am not a friend of deceit, which is the pure truth'] (*Ensisam* 291 [19th century])

According to our corpus data, the introduction of /g/ in the first-person present indicative is quite late: we must wait until the 18th century to find the first velarized forms (except for the case from the 14th century) and until the 19th for them to generalize within the paradigm. This delay with respect to the rest of the velarized verbs of the 2nd conjugation, like *beure* 'to drink' or *moldre* 'to grind', may be because, in *ser* 'to be', there was no homophony problem between the first and third person of the present indicative.[7] In addition, as with *haver* 'to have', it is a widely used verb, which may also favor the maintenance of etymological forms. However, the velarized forms ended up being consolidated in the 19th century, probably due to convergence[8] with most verbs of the 2nd conjugation, which had already adopted the velar augment in the first person of the present indicative centuries earlier.

4.2 Present subjunctive results

Regarding the present subjunctive of *ser* 'to be', the percentage distribution between non-velarized (e.g., *sia* 's/he be') and velarized forms (e.g., *siga* 's/he be') varies

[7] As a result of regular sound change, a set of 2nd conjugation Catalan verbs with a root-final *-u-* [w], like *beure* 'to drink', or *-l-*, like *moldre* 'to grind', presented syncretism between the first and third person of the present indicative: e.g., 1sg. *beu* [ˈbew] 'I drink' vs. 3sg. *beu* [ˈbew] 's/he drinks'. In these verbs, the velarization of the first-person present indicative was already consolidated in the 14th century, because it served to resolve the homophony problem between these two forms: e.g., 1sg. *beu* >> *bec* [ˈbek] 'I drink'.

[8] According to the convergence principle, the set of paradigmatic cells affected by a change tends over time to acquire certain common phonological characteristics across all verbs in which they occur (Maiden 2005, 139–140), in this case, /g/ as a first-person marker in 2nd conjugation verbs.

significantly over the analyzed centuries (see Table 11), with significant differences in the distribution ($\chi^2_{(6)} = 3323.054$, $p < 0.001$, Cramér's V = 0.959). Cramér's V, again, with a value of almost 1, indicates that the association between the "century" variable and the "velarized" variable is almost perfect.

			Non-velarized	Velarized
Period	13th century	Count	1757	1
		% within century	99.9%	0.1%
	14th century	Count	831	—
		% within century	**100%**	0%
	15th century	Count	247	—
		% within century	**100%**	0%
	16th century	Count	398	—
		% within century	**100%**	0%
	17th century	Count	74	—
		% within century	**100%**	0%
	18th century	Count	36	28
		% within century	56.3%	43.8%
	19th century	Count	3	237
		% within century	1.3%	98.8%

Table 11. Cross table of the present subjunctive of the verb *ser*

From the 13th to the 17th century, only one case with the velar augment is documented (9a); the rest of the extracted forms from these centuries maintain the etymological form without /g/ (9b).

(9) a. si alcun vendrà a la cort al primer manament que a él serà fet e que tenga aquí cort e *siga* en la cort [‘if anyone will come to the court at the first commandment that will be made to him and let him have court here and be in the court’] (*Costums de Tortosa* 49 [13th century])

b. e com los embaxadors seran passats, facen axí com hé dit fins que *sien* davant la altesa vostra [‘and as the ambassadors will be passed, they do as I have said until they are before your highness’] (*Tirant* 109 [15th century])

As in the first person of the present indicative, in the 18th century, the situation changed remarkably since the etymological forms (10a) and the velarized ones (10b) were distributed in a balanced manner. However, the non-velarized cases continued to be more frequent, representing 56.3% of the total forms for the century.

(10) a. acabada la missa, transporta lo Santíssim Sacrament (que *sia* alabat per a sempre) des de lo Altar major [‘after the Mass, transport the Blessed Sacrament (may it be praised forever) from the High Altar’] (*Esplugues* 53 [18th century])

b. establit ab capítol que el producte de ell *siga* per mitat de la església per a gastos [‘established with chapter that the product of it be for half of the church for expenses’] (*Esplugues* 40 [18th century])

Finally, in the 19th century, there was an abrupt change, and the velarized forms rose to 98.8% (11a). With respect to the etymological forms, only a few cases are documented (11b).

(11) a. *siguin* las que *siguin* la forma de son gobern y las institucions que lo regeixin [‘whatever the form of their government and the institutions that govern it’] (*Lo catalanisme* 52 [19th century])

b. Y així el primer *sía* Enoc quil escriure prinsipiá, ó *sia* Set [‘And so Enoch was the first to write, or Seth’] (*Ensisam* 310 [19th century])

In the end, the velarization process of the present subjunctive of *ser* ‘to be’ has a similar chronology to the first-person present indicative, despite the fact that they do not constitute a homogeneous L-pattern because of having different roots (*so-* in front of *si-*). In the present subjunctive, as in the case of the first-person present indicative, we have to wait until the 18th century to attest velarized forms, and until the 19th century for them to generalize. In the case of the present subjunctive, the change may be due to two factors: the first related to the presence of the velar in the first person of the present indicative, since, as mentioned in Section 3.1, the L-pattern forms tend to share specific traits; secondly, the analogical velarization of the present subjunctive also aims to resolve morphophonological problems, as mentioned above: the velar consonant allows the hiatus of the etymological forms to be undone: e.g., *si.a* >> *si.ga* ‘s/he be’.

4.3 Third-person preterite results

When doing the count of the third-person preterite of *ser* ‘to be’, the forms *fo*, *fon* and *fou* ‘s/he was’ are grouped together in the ‘non-velarized’ cell, given that the interest in this tense is to check the velarization of this person. According to the corpus data, the percentage distribution between non-velarized (*fo, fon, fou* ‘s/he was’) and velarized forms (*fonc* ‘s/he was’) varies significantly throughout the studied period (see Table 12). The corpus results show that there are significant differences in the arrangement of the verb forms between the centuries ($\chi^2_{(6)} = 1251.345$, $p < 0.001$, Cramér’s V = 0.655). However, the distribution of the forms presents some complex twists that need to be explained.

			Non-velarized	Velarized
Period	13th century	Count	157	—
		% within century	**100%**	0%
	14th century	Count	344	11
		% within century	96.9%	3.1%
	15th century	Count	1117	57
		% within century	95.1%	4.9%
	16th century	Count	234	261
		% within century	47.3%	52.7%
	17th century	Count	50	285
		% within century	14.9%	85.1%
	18th century	Count	210	125
		% within century	62.7%	37.3%
	19th century	Count	57	11
		% within century	83.8%	16.2%

Table 12. Cross table of the third person of the preterite of *ser* 'to be'

In the 13th century, only the etymological form *fo* 's/he was' is documented, as illustrated in (12).

(12) a. Presa *fo* la ciutat de Tortosa de mans de sarrayns per lo molt alt et noble et [h]onrrat en Ramon Berenguer ['The city of Tortosa was taken from the Saracens by the very high and noble and honorable Ramon Berenguer'] (*Costums de Tortosa* 1 [13th century])

b. per aysó, Seyner, *fo* necessaria causa que la potencia ymaginativa fos en home ['therefore, Lord, it was necessary for the imaginative power to be in man'] (*Llibre de contemplació en Déu* XLII – [6] [13th century])

In the 14th and 15th centuries, the situation is quite equivalent: the non-velarized forms (13a) are the majority since they account for 96.9% and 95.1% of the cases, respectively. Forms with the velar consonant, like the one in (13b), are still rare in the texts.

(13) a. Sa fila és molt pus bela que no *fo* se mare ['Her daughter is much more beautiful than her mother was'] (*Filla* 31 [14th century])

b. si los cavallers observaven aquell segons la fi per què *fonch* instituÿt e ordenat ['if the knights observed that according to the end for which it was instituted and ordered'] (*Tirant* 75 [15th century])

In the 16th century, there is a rather abrupt change, as the frequency of velarized forms (14a) equals that of non-velarized ones (14b): the former represent 52.7%, while the latter, 47.3%.

(14) a. *Fonch* convidat Sa Magestat per al dia de Sant Vicent Màrtir ['His Majesty was invited for the day of Saint Vincent the Martyr'] (*Antiquitats* 238 [16th century])

b. Ý *fon* forsat que los alabarders de la guarda de don Bernardino de Càrdenes […] ['And it was forced that the halberds of the guard of Don Bernardino de Càrdenes […]'] (*Antist* 250 [16th century])

In the 17th century, the velarized forms (15a) reach their peak, representing 85.1% of the cases, compared to the non-velarized forms (15b), which barely reach

15% of the cases. From this point onwards, one might expect the non-velarized forms to disappear.

(15) a. Y després de estar assentats, *fonch* donada dita benvinguda a dit excel·lentíssim senyor [‘And after they were seated, the said welcome was given to the said most excellent sir’] (*Dietaris Gen. Cat.* 19 [17th century])

b. Y li *fon* forçós lo exir dos o tres vegades als balcons per a que el poble el poguera veure [‘And he had to go out on the balconies two or three times so that the people could see him’] (*Notícies de València i son regne* [802] [17th century])

However, in the 18th century, the trend is reversed, as the forms without the velar become dominant again, with 62.7% of cases (16a). The velarized forms, however, remain quite frequent since they remain in almost 40% of the total forms of the century (16b).

(16) a. elet ý sobrestant de la obra ó *fon* en lo efecte solament el Dr. Phèlix Martí [‘elected and overseer of the work was actually only Dr. Phèlix Martí’] (*Esplugues* 65 [18th century])

b. ab què *fonch* molt amolestat, com se dirà més abant [‘with which he was greatly annoyed, as will be said later’] (*Memòries d’un pagès* 88 [18th century])

In the 19th century, the regression of velarized forms (17a) is even more accentuated since the non-velarized forms exceed 80% of the cases (17b). It is worth noting the appearance, in the Valencian text, of two cases of *sigué* [siˈɣe] ‘s/he was’, like the one in (17c), formed analogically on the root of the present subjunctive (*si-*) with the velar consonant.

(17) a. *fonc* Enoc, fill de Cai, net del primer pare Adam [‘it was Enoch, son of Cain, grandson of the first father Adam’] (*Ensisam* 309 [19th century])

b. En Santángel *fou* evidentment un dels més ferms puntals de Colombo [‘Santàngel was obviously one of Columbus’ staunchest props’] (*Lo catalanisme* 57 [19th century])

c. vives dilichènsies fiu per pendre veu de vosté: difisil no me *sigué* [‘I did many things to take your voice: it was not difficult for me’] (*Ensisam* 186 [19th century])

To sum up, based on the data provided by the corpus, the percentages of the third person of the preterite of *ser* ‘to be’ undergo significant changes over the course of the studied period. In this respect, until the 15th century, the non-velarized forms are predominant. However, *fonc* ‘s/he was’ equals the frequency of the non-velarized forms in the 16th century and reaches its peak in the 17th century. However, in the 18th century, the trend is reversed again and the forms without the /g/ once more become the most common. This trend is definitively consolidated in the 19th century, when the velarized forms become the minority. In our opinion, it is complicated for a form like *fonc* to succeed and consolidate: the fact that only one form of the preterite presented the velar consonant, while the rest remained unaffected by the analogical change, went against the principle of coherence, according to which an alteration that affects one PYTA root in one point of the paradigm ends up affecting the rest of the persons and tenses of the pattern; consequently, since no other person of the tense followed the same evolution, *fonc* went definitively into recession in the 19th century.

An additional remarkable fact is revealed when comparing the behavior of this form with what happens in the case of the other verbs of the 2nd conjugation that also

underwent a process of analogical velarization, such as *dir* 'to say' or *prendre* 'to take'. In these verbs, it is precisely the third form of the preterite that takes the longest to adopt the velar, not only within the morphome, but within the same tense (see Wheeler 2012). This is because, within the PYTA, it is the least marked form. In the case of *ser* 'to be', the opposite happens: it is precisely the less marked form that is the only one subjected to the process of analogical velarization, contrary to what is expected.

5 The verb *voler* 'to want'

Voler 'to want' is another highly frequent Catalan verb, partly because it is usually used as a modal verb and when followed by an infinitive it presents behaviors typical of auxiliary verbs (see IEC 2016, 947). It is an intricate verb because, in modern-day Catalan, it presents significant dialectal divergences in the L-pattern due to analogical velarization, as we will detail below. The study of these divergences challenges the theory of linguistic change, because the results show that, in each variety, a different principle of naturalness prevails (on the hierarchy of naturalness principles, see Mayerthaler 1987, 52). In Old Catalan, this verb had a homogeneous L-pattern, since these forms presented a root-final /ʎ/: *vull-* [ˈvuʎ] (see Table 13). Furthermore, the PYTA forms initially had the velar consonant, which was added to a root ending in an alveolar lateral: *vol-* (see Table 14).

	1sg.	2sg.	3sg.	1pl.	2pl.	3pl.
PRS.IND	*VOLEO > *vull*	*VOLES > *vols*	*VOLET > *vol*	*VOLEMUS > *volem*	*VOLETIS > *volets*	*VOLENT > *volen*
PRS.SBJV	*VOLEAM > *vulla*	*VOLEAS > *vulles*	*VOLEAT > *vulla*	*VOLEAMUS > *vullam*	*VOLEATIS > *vullats*	*VOLEANT > *vullen*

Table 13. Evolution from Vulgar Latin to Old Catalan of the present indicative and the present subjunctive of the verb *voler* 'to want' (the forms that constitute the L-pattern are shaded).

	1sg.	2sg.	3sg.	1pl.	2pl.	3pl.
PRT	*volc*	*volguist*	*volc*	*volguem*	*volgués*	*volgren*
COND	*volgra*	*volgres*	*volgra*	*vólgrem*	*vólgrets*	*volgren*
IPFV.SBJV	*volgués*	*volguesses*	*volgués*	*volguéssem*	*volguéssets*	*volguessen*

Table 14. PYTA morphome of the verb *voler* 'to want' in Old Catalan

Therefore, the verb *voler* 'to want' has identical traits to some of the 2nd conjugation verbs in which the velar augment analogically extended from the PYTA- to the L-pattern, such as *valdre* 'to cost': an L-pattern with a root-final /ʎ/ and a velarized PYTA with /g/. However, the velarization of the first person of the present indicative and of the present subjunctive presents divergences between Catalan varieties, which distinguishes this verb from those mentioned above. Currently, this verb has two L-pattern paradigms, one for general Catalan (see Table 15), and another specifically for Valencian (see Table 16).

	1sg.	2sg.	3sg.	1pl.	2pl.	3pl.
PRS.IND	*vull*	*vols*	*vol*	*volem*	*voleu*	*volen*
PRS.SBJV	*vulgui*	*vulguis*	*vulgui*	*vulguem*	*vulgueu*	*vulguin*

Table 15. Present indicative and present subjunctive of the verb *voler* 'to want' in general Catalan

	1sg.	2sg.	3sg.	1pl.	2pl.	3pl.
PRS.IND	*vullc*	*vols*	*vol*	*volem*	*voleu*	*volen*

PRS.SBJV	*vullga*	*vullgues*	*vullga*	*vullgam*	*vullgau*	*vullguen*

Table 16. Present indicative and present subjunctive of the verb *voler* 'to want' in Valencian (the velarized forms, which constitute the L-pattern, are shaded).

In the first model, typical of general Catalan, the most important feature is that the L-pattern is blurred, since the first person of the present indicative still maintains the etymological form without /g/, *vull* 'I want', while the whole of the present subjunctive presents the velar consonant with a new root within the paradigm, *vul-*. Therefore, there are two differences between the first person of the present indicative and the present subjunctive; firstly, the root-final consonant, and, secondly, the presence of the velar, which appears only in the present subjunctive. In Valencian, on the other hand, the L-pattern uniformity is maintained: both the first person of the present indicative and the present subjunctive preserve the root *vull-*, with the palatal lateral; moreover, all the pattern forms have the velar augment. It should be noted that the combination of palatal consonant and velar consonant (e.g., *vullc* [ˈvuʎk] 'I want', *vullga* [ˈvuʎɣa] 's/he want') is anomalous in general Catalan and constitutes a loosening of the language phonological restrictions, because sequences of affricate or prepalatal or palatal plus consonant are not found morpheme-internally, except for geminate /ʎʎ/ (Wheeler 1984, §3.4; Hualde 1992, 381; IEC 2016, 25).

Therefore, in the L-pattern evolution of *voler* 'to want', there is a naturalness conflict resolved differently according to the dialectal variety. The conflict occurs between the class-stability principle, regarding the morphome coherence, and that of phonological congruity,[9] given that, in Catalan, there are in principle no contacts

[9] The system-congruity principle states that "any inflectional system of an individual language has specific structural properties characterizing its structural typology and thus distinguishing it from the inflectional systems of other languages" (Wurzel 1987, 62).

between palatal and velar consonants. In general Catalan, the principle of phonological congruity prevails, while, in Valencian, the class-stability principle takes precedence, favoring the coherence of the pattern. All this variability must be considered, especially when analyzing the velarization of the first-person present indicative, given that this phenomenon is limited to Valencian.

5.1 First-person present indicative results

As mentioned, in the first-person present indicative counts of *voler* 'I want', we only consider data from Valencian texts, since it is the only variety in which the velar consonant is introduced.[10] As illustrated in Table 17, the percentage distribution between non-velarized (*vull* 'I want') and velarized forms (*vullc* 'I want') varies noticeably throughout the studied period, with significant differences in the distribution of the forms between the centuries ($\chi^2_{(4)} = 83.782$, $p < 0.001$, Cramér's V = 0.794).

			Non-velarized	Velarized
Period	13th century	Count	—	—
	14th century	Count	1	—
		% within century	**100%**	0%
	15th century	Count	59	—
		% within century	**100%**	0%
	16th century	Count	1	—
		% within century	**100%**	0%

[10] In other Catalan varieties, like Balearic or North-Western, velarized forms such as [ˈvujk] or [ˈbujk] 'I want' are sporadically attested (see Alcover/Moll 1929–1932). However, these dialects are not included in the corpus since, as mentioned in the methodology, the works we consider come from the two major Catalan varieties, i.e., Central and Valencian.

	17th century	Count	—	—
	18th century	Count	5	4
		% within century	55.6%	44.4%
	19th century	Count	12	51
		% within century	19%	81%

Table 17. Cross table of the first-person present indicative of the verb *voler* 'to want'

As for the cases distribution, until the 16th century, only etymological forms are documented, as illustrated in (18).

(18) a. Et per açò *vull* que puxa ésser penyorada et venudes les penyores ['And that's why I want it to be able to be pledged and the pledges sold'] (*Xàtiva* 58 [14th century])

b. La una estava de l'altar del Sant Ladre a la trona de la Epístola, *vull* dir, al pilar junct a la dita trona ['One was from the altar of the Holy Thief on the throne of the Epistle, I mean, on the pillar next to said throne'] (*Antiquitats* 208 [16th century])

In the 18th century, there is a significant change because, for the first time, forms with the velar consonant are attested, such as the examples in (19a), which are distributed in a balanced way with the etymological forms (19b): the former represent 44.4%, while the latter constitute 55.6%.

(19) a. En lo portal de Serrans, o m'esperes o t'espere, perquè jo em *vullc* acostar i estar prop del tauladet ['In the Serrans portal, either you wait for me or I wait for you, because I want to get closer and be near the roof'] (*Col·loquis Carles Leon* [18th century])

b. quant tinga complet el any de la sua celebración *vull* dir [‘when the year of its celebration is complete I mean’] (*Esplugues* 185 [18th century])

In the 19th century, velarized forms become the most frequent (81%), like the ones in (20a), while etymological forms only represent 19% of cases (20b).

(20) a. A la mar men *vullc* anar á vore les cosidores [‘I want to go to the sea to see the seamstresses’] (*Ensisam* 11 [19th century])

b. yo també huapa la *vull* y qu’em porte algun diner [‘I also want her to be pretty and bring me some money’] (*Ensisam* 57 [19th century])

All in all, the results indicate that the velarization of the first person of the present indicative of *voler* ‘to want’ is quite late. From the second half of the 17th century, we have no data in the corpus, so we have resorted to the *CIGCMod*, in which 115 cases of *vull* ‘I want’ (21a) and only one of *vullc* ‘I want’ (21b) are documented. As for the 18th century, 47 cases of *vull* ‘I want’ (22a) and 53 occurrences of *vullc* ‘I want’ (22b) are attested, so that it seems that, in this century, the analogical forms are already as frequent as the etymological ones.

(21) a. ý no·t temes de res que no et *vull* fer mal [‘and don’t be afraid of anything I don’t want to hurt you’] (*Vint processos criminals d’Albalat de la Ribera (1611–1666) 1* 83 [17th century])

b. e ha de fer y tot ho arregle, y així no *vullch* ser cansat [‘and I have to do and fix everything, and so I don’t want to be tired’] (*Sermó de la invenció de la creu (I)* 422 [17th century])

(22) a. La fona *vull* esclafir, ya que la tinc en les mans. [‘I want to crush the sling, since I have it in my hands.’] (*Canya i cordell* 260 [18th century])

b. al llaurador que la contà, que yo no *vullch* moure bronquines [‘to the farmer who tells it, that I don’t want to make a fuss’] (*Rondalla de rondalles* 188 [18th century])

If the velarization chronology of *voler* ‘to want’ is compared with other verbs with a root-final /ʎ/ in the L-pattern, such as *valdre* ‘to cost’, it is observed that, in these verbs, the velarized forms are already in the majority in the 15th century; however, in the case of *voler* ‘to want’, we have to wait three more centuries for it to become the most frequent form (and only in Valencian texts). It should be noted that *voler* ‘to want’ is quite a frequent verb and, as mentioned several times throughout this paper, high frequency favors the maintenance of irregularities or atypicalities. In addition, in verbs such as *valdre* ‘to cost’, the analogical velarization leads to root alteration, because /ʎ/ is replaced by /l/ (e.g., *valla* [ˈvaʎa] >> *valga* [ˈvalɣa] ‘s/he cost’), since the combination of palatal plus velar consonant is generally avoided in Catalan, as mentioned above. The desire to maintain the palatality in *voler* ‘to want’ may have been another cause of the delay of the first-person present indicative velarization. Besides, the principle of transparency does not come into play, as it did in verbs like *beure* ‘to drink’ or *moldre* ‘to grind’, since there is no need to resolve any syncretism between the first and third person of the present indicative (1sg. *vull* ‘I want’ vs. 3sg. *vol* ‘s/he wants’).

5.2 Present subjunctive results

In the analysis of the present subjunctive of *voler* ‘to want’, the data from the entire corpus is included, given that the velarization of this tense occurs in all Catalan varieties. The distribution of non-velarized (e.g., *vulla* [ˈvuʎa] ‘s/he want’) and velarized forms (e.g., *vulga* [ˈvulɣa] ~ *vullga* [ˈvuʎɣa] ‘s/he want’) varies significantly

during the studied period (see Table 18), with clear differences in the distribution of the verb forms between the centuries ($\chi^2_{(6)} = 356.474$, $p < 0.001$, Cramér's = 0.983).

			Non-velarized	Velarized
Period	13th century	Count	110	—
		% within century	**100%**	0%
	14th century	Count	73	—
		% within century	**100%**	0%
	15th century	Count	56	—
		% within century	**100%**	0%
	16th century	Count	21	—
		% within century	**100%**	0%
	17th century	Count	1	—
		% within century	**100%**	0%
	18th century	Count	3	17
		% within century	15%	85%
	19th century	Count	—	88
		% within century	0%	**100%**

Table 18. Cross table of the present subjunctive of the verb *voler* 'to want'

As in the first person of the present indicative, until the 17th century, we only find etymological forms without the velar consonant, like the ones illustrated in (23).

(23) a. que per sí metexs o per éls *vulen* conprar e els serà assignada peynora ['that for themselves or for them they want to buy and a pledge will be assigned to them'] (*Costums de Tortosa* 31 [13th century])

b. E, si per ventura serà alguna que per supèrbia no *vulla* dir la sua culpa [‘And, if it happens to be someone who, out of pride, does not want to say their fault’] (*Sant Josep* 68 [16th century])

However, in the 18th century, the distribution changes drastically and velarized forms increase to 85% (24a), although some etymological forms are still documented (24b).

(24) a. ý este suplix lo que falta, la sera ý demés gasto que *vullga* fer [‘and this makes up for what is missing, the wax and the rest of the expenses he wants to make’] (*Esplugues* 179 [18th century])

b. Déu *vulla* dar molts anys de vida a bonas personas [‘God should want to give many years of life to good people’] (*Memòries d’un pagès* 62 [18th century])

In the 19th century, the analogical velarization of the present subjunctive is definitively consolidated, since the only forms attested in the corpus present the velar consonant, like the instances in (25).

(25) a. li rese pera qu’em *vullgues* ó pera que yo t’olvide [‘I pray to him so that you love me or I forget you’] (*Ensisam* 17 [19th century])

b. las forsas se dispersarán y anirán á engroixir los partits extrems que *vulguin* aprofitarsen [‘the forces will disperse and the extreme parties that want to take advantage of it will grow stronger’] (*Lo catalanisme* 91 [19th century])

As with the first-person present indicative, the appearance of velarized forms in the present subjunctive in our corpus is late: we must wait until the 18th century. However, in the *CICA*, a case of *vulga* ‘I/he/she want’ is already attested in the first half of the 14th century (26a) and another of *vullga* ‘I/he/she want’ in the first half of the 15th century (26b); nevertheless, at that time, the etymological forms were clearly

in the majority: 57 occurrences of *vulla* 'I/he/she want' in the first half of the 14th century and 234 in the first half of the 15th century (26c).

(26) a. la qual cosa jo no ho *vulga* oir ni aver presa ferma d'ell tro que de vós n'agués manament ['which I don't want to hear or have a firm decision from him until I have a command from you'] (*Pergamins, processos i cartes reials 5* doc. 169 [14th century])

b. que per sostenir justícia algú hi *vullgue* treballar per amor de Déu ['that in order to maintain justice someone wants to work there for God's sake'] (*Sermons* [St. Vicent Ferrer] VI – 10 [15th century])

c. e jo t'hauré tots temps per cara, e, encara que no u *vulla*, seràs dona mia ['and I will always have you in consideration, and, even if I don't want it, you will be my wife'] (*Decameró (1st part)* 190 [15th century])

Regarding the 17th century, as only one case has been extracted from our corpus, we have also consulted the *CIGCMod*. In this period, as in the case of the first person of the present indicative in Valencian, the etymological forms continue to be hegemonic; thus, 10 cases of *vullga* 'I/he/she want' (27a) and 6 of *vulga* 'I/he/she want' (27b) are documented, compared to 209 occurrences of *vulla* 'I/he/she want' (27c).

(27) a. e demà per tot dia diga y alegue lo que *vullga* ['and tomorrow for the whole day say and allege what you want'] (*Tres processos de Moncada* 291 [17th century])

b. per sa infinita bondat y clamentia nos *vulga* preseruar de tals mals ['for his infinite goodness and mercy he wants to preserve us from such evils'] (*Manual de novells ardits* 52 [17th century])

c. y en cas no tingue o no *vulla* tenir capellà [‘and in case he does not have or does not want to have a priest’] (*Diplomatari Andorra* 51 [17th century])

In short, the velarization of the present subjunctive already seems consolidated in the second half of the 18th century, both in Valencian and in general Catalan, whereas the analogical velarization of the first-person present indicative is limited to Valencian. In this variety, both the uniformity principle and the coherence and convergence principles come into play. The uniformity principle operates because the L-pattern remains homogeneous, with the same root (*vull-*) and the presence of the velar augment in all the morphome cells. Secondly, the principles of coherence and convergence operate, since the analogical velarization also ends up being consolidated in all the pattern cells and adopts the same mark (/g/) as most 2nd conjugation verbs in these forms.

The evolution undergone by general Catalan presents two drawbacks, since neither of the principles operates. Firstly, the analogical velarization of the present subjunctive entails the creation of a new allomorph (*vul-*), while the first-person present indicative preserves the etymological *vull* ‘I want’; therefore, this change is contrary to the uniformity principle, since it generates a new allomorph in this verb paradigm. The second issue is related to the L-pattern blurring, to the detriment of the coherence of the diachronic change: analogical velarization affects only one part of the pattern, the present subjunctive, while the first person of the present indicative remains unaltered. According to this principle, the change should have affected this last person as well; however, it has remained unaltered in general Catalan to this day. This change leads us to modify slightly what Maiden (2018, 122–123) proposes regarding the coherence of the diachronic change and the psychological reality of the morphome: according to

Maiden, the verbs in which the coherence of the morphomic pattern is blurred are those that never had it initially, such as Romance reflexes of DARE 'to give', ESSE 'to be' or STĀRE 'to stand'. However, the verb *voler* 'to want' did present a homogeneous L-pattern in Old Catalan, because of regular sound change (see Table 13). The phonological markedness of combining a palatal consonant with a velar may have hindered this change. But verbs such as *valdre* 'to cost' resolved this problem by modifying the etymological root and replacing it with one ending in *-l*, coming from PYTA (*valc* [ˈvalk] 'I/he/she cost', *valgués* [valˈɣes] 'I/he/she cost'): e.g., *valla* [ˈvaʎa] >> *valga* [ˈvalɣa] 'I/he/she cost'. This solution, however, seems not to have worked for the first-person present indicative of *voler* 'to want' in general Catalan, probably due, again, to its high frequency.

6 Conclusions

Despite the analogical velarization process that entailed the creation of a new verb class within the Catalan 2nd conjugation, there are some verbs, such as *haver* 'to have', *ser* 'to be' or *voler* 'to want', which did not strictly undergo this regularization.

In the case of *haver* 'to have', some velarized forms in the present subjunctive are attested, although they never became statistically significant. The high frequency of *haver*, used mainly as a perfect auxiliary verb, probably favored the preservation of irregular forms over the centuries (Bybee/Brewer 1980, 218; Anshen/Aronoff 1988; Rainer 1988; Booij 1997, 43). With respect to PYTA, the only tense where changes are observed is the old conditional, with some non-velarized forms in the 19th century documented in Valencian. The obsolescence of the preterite and imperfect subjunctive might have pushed the old conditional forms to undergo a levelling root process, because PYTA had been masked in that variety. However, this analogical change is not

consolidated at the end of the studied period, and presents dialectal variation, since the forms with the velar consonant are widely attested at the beginning of the 20th century.

In the case of *ser* 'to be', the first person of the present indicative and the present subjunctive present a very similar chronology. The velarization of the form of the first person of the present indicative is justified by convergence with most 2nd conjugation verbs, which adopted /g/ in forms with this tense-mood. In the present subjunctive, the change could be due to two factors: the influence exerted by the first person of the present indicative; and the tendency to repair the hiatus of the etymological forms (e.g., *si.a* >> *si.ga* 's/he be'). With regard to PYTA, the third person of the preterite form is the only one coming from Latin *perfectum* in which we document velarized cases. It is difficult to come up with a convincing explanation for the evolution undergone by this form. The final recession of velarized forms might be attributed to the need for coherence: since, over the centuries, no other PYTA form adopts the /g/, the velarization process is blocked, as a single PYTA cell with a velar augment is unusual.

With respect to *voler* 'to want', the L-pattern resulting from regular sound change undergoes an alteration driven by analogical velarization. While Valencian still preserves a homogeneous pattern, the morphome is obscured in general Catalan. The first person of the present indicative adopts the velar in Valencian only, whereas the analogical velarization of the present subjunctive is widespread in Catalan. In the evolution of *voler* in Valencian, the principles of uniformity and class-stability take precedence over phonological congruence, given that the uniformity of the L-pattern is preserved: all forms share the same root-final palatal and present the velar consonant, despite this unusual combination. However, in Catalan in general, the phonological congruence hinders the morphome homogenization, leading to the L-pattern obscuring,

and the generation of a new irregularity: while the first person of the present indicative retains the etymological form, the present subjunctive forms adopt a new allomorph to which the velar is added.

To sum up, although in most cases the evolution of morphomic patterns is usually coherent because all morphome cells undergo the same change(s), this research brings to light cases in which this tendency is obscured. More specifically, the results of the study corroborate what some authors have already observed for other phenomena: the more a verb is used, the less susceptible it is to morphological regularization patterns. The data under analysis in this paper, thus, cast light on the fact that frequency can sometimes exert greater pressure than morphomic coherence in verbal diachronic evolution.

7 Bibliography

Alcover, Antoni M./Moll, Francesc de B., *La flexió verbal en els dialectes catalans*, Anuari de l'oficina romànica de lingüística i literatura 2 (1929), 79–184; 3 (1930), 73–168; 4 (1931), 9–104; 5 (1932), 9–72 (= http://espm.iec.cat/fvdc). [last access: April 25th, 2023]

Anshen, Frank/Aronoff, Mark, *Producing morphologically complex words*, Linguistics 26 (1988), 641–655.

Aronoff, Mark, *Morphology by itself: stems and inflectional classes*, Cambridge/MA, MIT Press, 1994.

Badal, Manuel, *Morfologia verbal catalana: anàlisi diacrònica dels verbs velaritzats*, Doctoral dissertation, València, Universitat de València, 2021.

Booij, Geert, *Autonomous morphology and paradigmatic relations*, Yearbook of morphology 1996 (1997), 35–53.

Bybee, Joan, *Morphology: a study of the relation between meaning and form*, Amsterdam, Benjamins, 1985.

Bybee, Joan/Brewer, Mary Alexandra, *Explanation in morphophonemics: changes in Provençal and Spanish preterite forms*, Lingua 52:3–4 (1980), 201–242.

CICA = Torruella, Joan/Pérez Saldanya, Manuel/Martines, Josep (edd.), *Corpus Informatitzat del Català Antic*, 2009–, http://www.cica.cat. [last access: March 15th, 2023]

CIGCMod = Martines, Josep/Martines, Vicent (edd.), *Corpus informatitzat de la gramàtica del català modern*, Alacant, ISIC-IVITRA.

De Yzaguirre, Lluís, *Banc analític d'antigalles lèxiques*, Barcelona, Universitat Pompeu Fabra, 2018, http://latel.upf.edu/cgi-bin/badal.cgi. [last access: February 10th, 2023]

Esher, Louise/O'Neill, Paul, *The autonomy of morphology*, in: Ledgeway, Adam/Maiden, Martin (edd.), *The Cambridge handbook of Romance linguistics*, Cambridge, Cambridge University Press, 2022, 346–370.

Fabra, Pompeu, *Normes ortogràfiques*, Barcelona, Institut d'Estudis Catalans, 1913.

Fabra, Pompeu, *Gramàtica catalana*, Barcelona, Institut d'Estudis Catalans, 1918.

Fabra, Pompeu, *Diccionari general de la llengua catalana*, Barcelona, Llibreria Catalònia, 1932.

Hualde, José Ignacio, *Catalan*, New York, Routledge, 1992.

IBM Corp., *IBM SPSS Statistics for Windows (26.0)*, Armonk, NY, IBM Corp, 2019.

IEC 2016 = Institut d'Estudis Catalans, *Gramàtica de la llengua catalana*, Barcelona, Institut d'Estudis Catalans, 2016.

Maiden, Martin, *A strange affinity: "perfecto y tiempos afines"*, Bulletin of Hispanic studies 78 (2001), 441–464.

Maiden, Martin, *Morphological autonomy and diachrony*, Yearbook of morphology 2004 (2005), 137–175.

Maiden, Martin, *The Romance verb: morphomic structure and diachrony*, Oxford, Oxford University Press, 2018.

Mayerthaler, Willi, *System-independent morphological naturalness*, in: Dressler, Wolfgang U./Mayerthaler, Willi/Panagl, Oswald/Wurzel, Wolfgang U. (edd.), *Leitmotifs in natural morphology*, Amsterdam, Benjamins, 1987, 25–58.

Menéndez Pidal, Ramón, *Manual elemental de gramática histórica española*, Madrid, Librería General de Victoriano Suárez, 1904.

Meyer-Lübke, Wilhelm, *Grammaire des langues romanes*, vol. 2: *Morphologie*, Paris, Welter, 1895.

Pérez Saldanya, Manuel, *Del llatí al català: morfosintaxi verbal històrica*, València, Universitat de València, 1998.

Pérez Saldanya, Manuel, *"Que sa il·lustre senyoria cullga o faça cullir dits fruits": una aproximació històrica als verbs velaritzats*, in: Clua, Esteve/Lloret, Maria-Rosa (edd.), *Qüestions de morfologia flexiva i lèxica del català: volum d'homenatge a Joaquim Viaplana*, Alacant, Institut Interuniversitari de Filologia Valenciana, 2013, 313–333.

Pérez Saldanya, Manuel/Sifre Gómez, Manuel/Todolí i Cervera, Júlia, *Morfologia catalana*, Barcelona, Editorial UOC, 2004.

Rainer, Franz, *Towards a theory of blocking: the case of Italian and German quality nouns*, Yearbook of morphology 1 (1988), 155–186.

Rea, Louis M./Parker, Richard A., *Designing and conducting survey research: a comprehensive guide*, San Francisco, Jossey-Bass, 2014.

Ridruejo, Emilio, *La forma verbal en "-ra" en valenciano*, in: *Linguistique comparée et typologie des langues romanes. Actes du XVII^{ème} Congrès international de linguistique et philologie romanes (Aix-en-Provence 1983)*, vol. 2, Aix-en-Provence, Université de Provence, 1985, 437–448.

Ronjat, Jules, *Grammaire istorique des parlers provençaux modernes*, vol. 3: *Morphologie et formation des mots, notes de syntaxe*, Montpellier, Société des langues romanes, 1937.

Wheeler, Max W., *La conjugació valenciana: geografia, diacronia i psicologia*, in: Casanova, Emili (ed.), *Estudis en memòria del professor Manuel Sanchis Guarner: estudis de llengua i literatura catalanes*, València, Universitat de València, 1984, 409–419.

Wheeler, Max W., *The evolution of a morphome in Catalan verb inflection*, in: Maiden, Martin/Smith, John Charles/Goldbach, Maria/Hinzelin, Marc-Olivier (edd.), *Morphological autonomy: perspectives from Romance inflectional morphology*, Oxford, Oxford University Press, 2011, 182–209.

Wheeler, Max W., *La morfologia verbal del "Llibre dels feits" en el context de l'evolució de la llengua*, in: Hauf, Albert (ed.), *El Llibre dels feits. Aproximació crítica*, València, Publicacions de l'Acadèmia Valenciana de la Llengua, 2012, 155–176.

Wurzel, Wolfgang U., *System-dependent morphological naturalness in inflection*, in: Dressler, Wolfgang U./Mayerthaler, Willi/Panagl, Oswald/Wurzel, Wolfgang U. (edd.), *Leitmotifs in natural morphology*, Amsterdam, Benjamins, 1987, 59–96.