Papers in Historical Phonology
http://journals.ed.ac.uk/pihph
ISSN 2399-6714
Volume 8, 1-15 (2023)
DOI: 10.2218/pihph.8.2023.8932

Consonant clusters and verb stems: making sense of distributional gaps

ANDRÁS CSER
Hungarian Research Centre for Linguistics; Pázmány Péter Catholic University

Abstract

This paper investigates an apparent gap in the distribution of nasal + stop clusters, as well as certain aspects of the diachronic emergence of this gap, in Latin and Hungarian. The phenomenon investigated is the absence of a frequent consonant cluster ([nt] in Latin, [ŋk] in Hungarian) from a position at the end of verb stems. An important property of the missing consonant cluster in both languages is that it also functions as a person marker in the verbal inflection. It is argued that in Latin this gap is functionally motivated: it represents a case of syntagmatic pressure to avoid repeating the same sequence at too close an interval. In Hungarian, by contrast, the absence of [ŋk] from verb stem-final position is arguably unrelated to the identical phonological form of the 1Plural affix and is simply the result of accidents of diachronic development.

1 Introduction

This paper examines a hitherto unnoticed gap in the distribution of the most frequent consonant cluster, [nt], in Classical Latin; to wit, this cluster is never found at the end of verb stems. While the incidence of consonant clusters in any particular environment does not, and is not expected to, necessarily mirror their incidence in the language in general, the absence of [nt] in this position is remarkable for a variety of reasons to be discussed in detail below. I shall argue that this gap is not of the accidental kind, and it cannot be derived from developments in the historical phonology of early Latin: it has a functional motivation in the fact that [nt] is the only consonant cluster in Latin that has a clear morphological function as the exponent of 3Plural ('plural' is henceforth 'Plur', and 'singular' is 'Sing').

Interestingly, a similar state of affairs appears to obtain in an unrelated — and morphologically very different — language, Hungarian, where no verb stem ends in [ŋk], even though this is one of the most frequent clusters in the language. This cluster also happens to be the only one that is a person marker (viz. of 1Plur) in the language. Nevertheless, in Hungarian, as opposed to Latin, the absence of [ŋk] from verb stem final position seems to result from etymological accidents rather than any systemic motivation.

While overlaps and ambiguities arising among the phonological forms of morphological markers, or of stems, have been extensively studied (e.g., Bertram, Schreuder & Baayen 2000, Bertram, Laine, Baayen, Schreuder & Hyönä 2000, Baayen, Schreuder, de Jong & Krott 2002, Baerman 2011, Sauerland & Bobaljik 2013, Kaplan & Muratain 2015), similar overlaps and ambiguities arising between the phonological form of stems and affixes have not (with a few notable exceptions such as Plag, Zimmermann & Kunter 2015). This paper thus represents the results of a preliminary investigation into these phenomena.

For a dead language any serious discussion of its structure is corpus-based in a trivial sense. For Latin several web-based corpora exist but none of these is specifically dedicated to a clear and readily usable presentation of phonological patterns. The ones that I primarily used for the quantitative aspects of the research reported in this paper are the corpus of the Packard Humanities Institute (latin.packhum.org) and the online dictionary of the Perseus Digital Library (perseus.tufts.edu). The former is a corpus of virtually all Classical Latin texts up to 200 AD (as well as many later texts); the latter is an online dictionary based on the material of Lewis & Short (1879). We also made extensive use of the LiLa Knowledge Base (lila-erc.eu), whose own lexical material is partly based on Lewis & Short (1879). Since Latin spelling in its standardised form is fairly consistent regarding the marking of consonantal segments (see Cser 2020, 13–43 for details), patterns involving most consonant clusters lend themselves to corpus-based analysis much more readily than patterns involving glides or vowels; the latter would be intractable without massive manual labour.

For Hungarian I used two resources. One is the Hungarian National Corpus (mnsz.nytud.hu; Oravecz, Váradi & Sass 2014), which is a morphosyntactically annotated corpus of texts representing Modern (standard and literary) Hungarian, now exceeding 1 billion words. The other is Szószablya (szotar.mokk.bme.hu/szoszablya), a morphologically analysed lexical and textual database. All quantitative data adduced in this paper for Hungarian derive from these corpora.

The paper is structured in the following way. In 2 the most relevant phonological aspects of nasal + stop clusters in general are briefly surveyed. In 3 the Latin data are presented and analysed and an explanation for the gap in stem-final patterns is offered. In 4 the Hungarian data are presented and analysed and a different explanation is offered for a seemingly similar pattern. Finally, 5 concludes the paper.

2 Some general remarks on nasal + stop clusters

It is well known that combinations of a nasal and a stop (in this order), where the nasal tends cross-linguistically to be homorganic with the following stop, are among the most frequent non-initial consonant clusters in the languages of the world (for typological background see Côté 2000, Vallé, Rossato & Rousset 2009, Gordon 2016, 97 ff., with some generalisations going back to Greenberg 1965). They are found even in languages which allow a highly restricted set of syntagmatic combinations of consonants, such as Italian (Krämer 2009, 137 ff.) or Japanese (Labrune 2012, 140 ff.). In both Latin and Hungarian nasal + stop clusters are found between vowels as well as finally, though in the latter position their set is somewhat restricted; furthermore, word-internally they can be followed by a consonant, again coming from a restricted set. In this paper we will not be concerned with the syllabification of such sequences; for the languages discussed here there is ample literature on syllable structure and the analysis of consonant clusters (for recent summaries see Cser 2020 on Latin, Siptár & Törkenczy 2009 on Hungarian).

In harmony with cross-linguistic generalisations regarding place of articulation, within the set of consonant clusters of the kind discussed here, the coronals [nt nd] are more frequent than those found at other places of articulation (for exact figures see below). The second most frequent place of articulation in both languages is velar; this is followed by the labial place of articulation. Hungarian also has a contrastive palatal place for stops and nasals; the palatal nasal + stop clusters have the smallest lexical, as well as textual, frequency.

In terms of voicing, the general tendency is for voiceless stops to be more frequent lexically than voiced stops. The reverse is true in post-nasal position in languages that have an active or a recently defunct post-nasal voicing process (such as Modern Greek; Hayes 1996, Kümmel 2007, 53 ff.). Neither Classical Latin nor Modern Hungarian has such a process; and no such process can be reconstructed for Latin diachronically either. In the prehistory of Hungarian, however, the post-nasal voicing of stops and subsequent loss of the nasal itself was a regular change which contributed significantly to the emergence of the voiced stop series in the language (Proto-Finno-Ugric **kunta* 'clan' > Hu *had* 'war band', Proto-Finno-Ugric **lońća* > Hu *lágy* [laːɟ] 'soft', Proto-Ugric **ämpV* > Hu *eb* 'dog', Proto-Finno-Ugric **muŋkV* 'body' > Hu *mag* 'seed, core, self' etc.; see Lakó 1968, Kálmán 1970 and Benkő 1993 svv.). While this change has long been defunct, and its ultimate outcome is a set of stops, not clusters, in Modern Hungarian the proportion of voiced stops in post-nasal position is significantly higher *vis-à-vis* voiceless stops than the proportion of voiced vs. voiceless stops in the lexicon in general (for exact figures again see below).

3 The Latin data

The patterning of the nasal + stop clusters in Latin is very similar in terms of lexical and textual frequency. The numbers for both types of frequency are shown in table 1 as well as the charts below (figures 1–4). Data for lexical frequency are derived from the Perseus Database; data for textual frequency are derived from the Packard Humanities Institute Database. Note that the textual frequency of [nt] is inflated by the 161,463 occurrences of final *-nt*, which is the person marker of 3Plur in the active voice, and 37,084 occurrences of final *-ntur*, which is the person marker of 3Plur in the passive voice. With all and only such verb forms ending in *-nt(ur)*, the textual frequency of these sequences is high in a trivial way but, since the citation form of Latin verbs does not include the 3Plur form, their lexical frequency is zero, i.e. they do not contribute to the lexical frequency numbers of [nt] in any way. However, even if we subtract the suffixal occurrences of [nt], the textual frequency of this cluster would still be 304,770 — by far the highest number for any cluster in Latin.

The only other nasal + stop cluster that is found word-finally is [ŋk]. It occurs in various forms of a handful of deictics whose textual frequency is high (27,166, included in the data for this cluster below). The lexical frequency of final [ŋk] is 10, also included in the number below. None of the nasal + stop clusters occur word-initially for phonotactic reasons.

Besides calculating lexical as well as textual frequency, I also narrowed the search in a third step to show the lexical frequency of such clusters within verb stems.[1] This is just to demonstrate that verbs show, by and large, very similar ratios between the individual clusters to those found in the lexicon at large — as would be expected by default. The relevant numbers can be seen in column 4 of Table 1.

[1] Throughout the paper by 'verb stem' I mean what is called the present (or infectum) stem of verbs. We will not be concerned with any of the forms based on the perfect stem, the only other stem on which finite verb forms are based. However, the claims made in this paper are not contradicted by any verb forms based on the perfect stem.

	Lexical	Textual	Lexical in verbs	Verb stem-final
nt	3623	503317	387	0
nd	1679	181214	306	119
mp	1219	97902	162	13
mb	412	13527	69	18
ŋk	1524	105648	241	8
ŋg	737	45000	205	127

Table 1: The frequency of nasal + stop clusters in Latin.

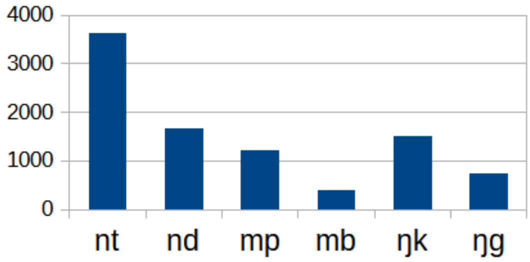

Figure 1: The lexical frequency of nasal + stop clusters in Latin.

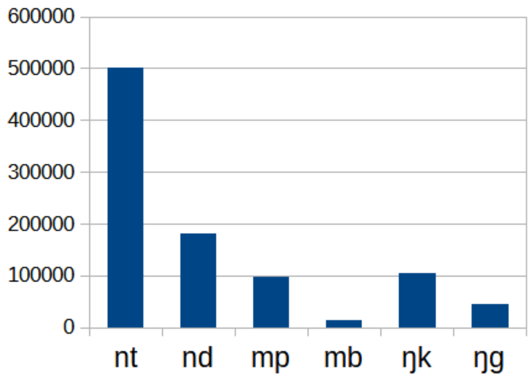

Figure 2: The textual frequency of nasal + stop clusters in Latin.

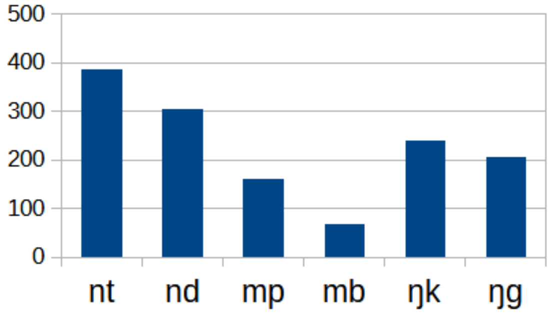

Figure 3: The lexical frequency of nasal + stop clusters within verbs in Latin.

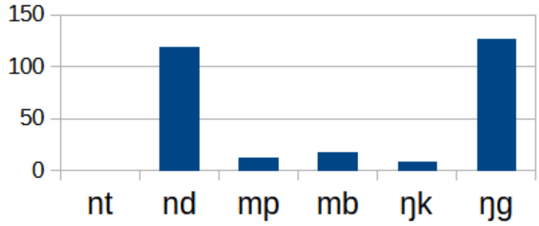

Figure 4: The lexical frequency of verb stems ending in nasal + stop clusters in Latin.

As can be seen from table 1 and figures 1–3, at all three places of articulation nasal + voiceless stop clusters outnumber nasal + voiced stop clusters. Coronals are the most frequent, and labials are collectively the least frequent. By contrast, the occurrence of nasal + stop clusters at the end of verb stems shows a markedly different pattern (see column 5 of table 1 and figure 4).

On the one hand, there is a very clear preponderance of stems ending in nasal + voiced stop as compared to nasal + voiceless stop clusters. What is even more interesting is the complete absence of the most frequent cluster [nt] from stem-final position. Here we shall look at the latter observation in some detail and leave the marked asymmetry between [ŋg] vs. [ŋk], and the less marked but still noticeable asymmetry between [mb] vs. [mp] aside. Why is it a fact requiring at least a putative explanation that no Latin verb stem ends in [nt]? What reason do we have to assume that this is not simply an accidental gap? I believe the following observations can be adduced here.

(1) The cluster [nt] is the single most frequent consonant cluster in the language in terms of lexical as well as textual frequency.

(2) Consonant-stems (in other words third conjugation verbs, including heteroclitic verbs[2]) are the second most numerous class in the Latin lexicon, preceded only by *a*-stems (first conjugation verbs). In the LiLa Knowledge Base 24% of all verbs are marked as belonging to the class of consonant-stems; my own count of verbs in the (non-searchable) Lewis & Short dictionary gave a higher approximate number closer to 40%. The latter is probably closer to the correct number for Classical Latin, since LiLa includes many post-classical and medieval verbs, and neologisms were mostly created in the first conjugation (e.g., *acquieto* 'to acquiesce').

(3) In Proto-Indo-European, an infix/suffix -*n*- was used to form present stems from roots (Clackson 2007, 153–154). This affix has survived in many verbs in the Old Indo-European languages, and in Latin it proved particularly stable before stops (*jungo* 'join', *cumbo* 'lie', *vinco* 'win' etc., see Leumann 1977, 533–535, Weiss 2020, 431).

(4) The number of *t*-final roots, from which -*nt*- final present stems could potentially be formed, was considerable in Proto-Indo-European. In particular, in Rix, Kümmel, Zehnder, Lipp & Schirmer (2001) there are 43 *t*-final and 16 *tH*-final roots, to which the 4 *nt*-final roots can be added. These together make up 6.02% of the roots in the entire *Lexicon*.

(5) There are two verbs in Latin that are reconstructed as deriving historically from *t*-final roots accompanied by the nasal present affix or including a lexical [n]: *pando* 'open' from PIE **peth*$_2$- and *mando* 'eat' from PIE **menth*$_2$- (see de Vaan 2008 s.vv.). As can be seen, they do not appear in Latin as -*nt*-final stems but show a voiced reflex for the stop, which is clearly not the result of a *[nt] > [nd] sound change since such a change cannot be reconstructed for the prehistory of Latin. The etymologies of these two verbs present special problems which will not be pursued further; a detailed discussion is found in Schrijver (1991, 222, 498–504).[3]

In the light of these facts I believe it is reasonable to look for an explanation for the apparent gap in the distribution of nasal + stop clusters. There is one explanation that offers itself quite clearly for why there are no verb stems that end in [nt]. This cluster is the only one that functions as an affix in the entire inflectional morphology of Latin; as was said above, it is the 3Plur person marker on verbs in the form -*nt* in the

---

[2] The term heteroclitic verb means those verbs whose forms are systematically identical partly to those of 'pure' consonant stem verbs, partly to those of *i*-stem (fourth conjugation) verbs. In traditional grammatical terminology these verbs are called *ĭ*-stems; Cser (2015 and 2020) as well as Kaye (2015) call these verbs heteroclitic.

[3] See also Weiss (2020, 183). The hypothesis advocated there is that a sound change *[tn] > [nd] is responsible for the forms in question. However, these two verbs are the sole examples adduced for the change.

active, and in the form -*ntur* in the passive voice.[4] After consonant-final stems and stems ending in [i] it appears as -*unt(ur)* (e.g. *agunt* 'they do', *veniunt* 'they come'). It may thus be surmised that the absence of *nt*-final stems is a manifestation of the tendency to avoid morphologically functional phonological sequences such as ⁺-*ntunt(ur)*.

A handful of verb stems do include — but do not end in — the cluster [nt]. These fall into two types. One type comprises verbs that are prefixed with a nasal-final element *in*- or *con*-, or with *inter*-, e.g. *intendo* 'stretch out', *contego* 'cover', *intercedo* 'intervene'.[5] In these the cluster appears in the verb as a result of morphological composition and is relatively far from the ending. The other type comprises verbs whose present stem itself includes [nt]. A few of these are *i*-stems (*sentio* 'feel', *mentior* 'lie' and their prefixed forms); the majority are *a*-stems, where the [t] is often (part of) a derivational suffix (*canto* 'sing' from *cano* 'sing', *cruento* 'make bloody' from *cruentus* 'bloody'). In the former subtype there are always two vowels between the two occurrences of the sequence [nt] (*sentiunt* 'they feel'); in the latter the most sonorous vowel separates them (*cantant* 'they sing'), which is historically long (Old Latin *cantānt*). They thus represent the closest approximation between two occurrences of the same cluster that the morphophonological system appears to permit.

All this could still be called speculative. However, if one looks at the other two person markers that are always polysegmental, their phonological forms are, for all intents and purposes, completely absent from verb stems. In particular, this means the following. The 1Plur active ending is -*mus*; and the sequence [mus] is only found in a single verb *mussito* 'murmur, mutter, be silent', clearly an onomatopoeic formation. For completeness I also checked the earlier form of the ending *-*mos*, but the corresponding sequence is not found in verb stems at all. The 2Plur ending is -*tis*; and the sequence [tis] is again not found in verb stems at all. The earlier form of the ending was *-*tes*; and even [tes] is found only in two verbs: *testor* 'testify' is a denominative verb from *testis* 'witness', and *petesso* 'strive', which occurs only 5 times in Classical Latin, 4 out of these in hexametre-final position where metrical exigency precludes the usual form *peto* of the same verb. It thus seems to be a valid generalisation that phonological sequences that are identical in composition to polysegmental verb endings are avoided in verb stems: for the two

---

[4] Indeed, no affix in Latin inflectional morphology even creates a consonant cluster by combining with a stem, except for the NomSing suffix -*s* (*dap-s* 'meal'); if the formation of present and perfect stems is analysed as belonging to inflectional morphology, one may add perfect -*s*- (*scrip-s-* 'write') and the present nasal (*fra-n-g-* 'break' with a nasal infix, *sper-n-* 'despise' with a nasal suffix).

[5] On general aspects of prefixation in Latin see Cser (2020, 153–183).

trisegmental sequences this is virtually exceptionless, whereas the bisegmental sequence [nt] is strictly avoided in stem-final position only.[6]

4 The Hungarian data

The lexical and textual frequency of nasal + stop clusters in Hungarian is shown in table 2. Textual frequency data were culled from the Hungarian National Corpus (mnsz.nytud.hu; Oravecz et al. 2014); lexical frequency data were culled from the Szószablya database (szotar.mokk.bme.hu/szoszablya). As with Latin, here too I have added a column showing the lexical frequency of such clusters within verb stems. And here too this demonstrates that verbs show, by and large, very similar ratios between the individual clusters to those found in the lexicon at large.

	Lexical	Textual	Lexical in verbs	Verb stem-final
nt	3575	32263670	663	320
nd	2652	22384995	491	22
mp	656	2285900	127	0
mb	1079	6518284	173	0
ɲc	59	77685	6	0
ɲɟ	225	614990	24	0
ŋk	1402	14227718	220	0
ŋg	2044	6561471	502	182

Table 2: The frequency of nasal + stop clusters in Hungarian

[6] The monosegmental endings are those of the active singular: 1Sing *-o* or *-m*, 2Sing *-s* and 3Sing *-t*. In certain environments the latter two also have vowel-initial alternants, similar to the polysegmental endings (for details see Cser 2015 and 2020, 124–152). I did not check for the distributions of these single segments in verb stems. For a general discussion of repetition avoidance in language see Walter (2007).

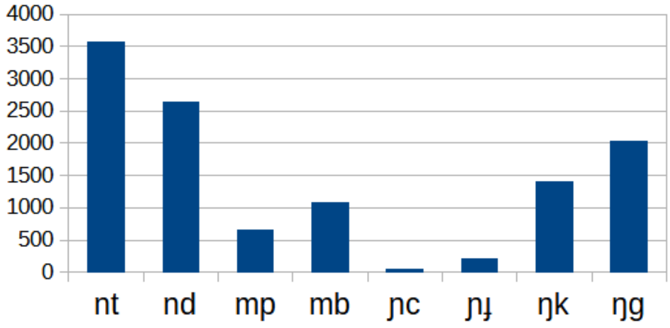

Figure 5: The lexical frequency of nasal + stop clusters in Hungarian

As can be seen from figure 5, with the exception of [nt] vs. [nd], voiced stops are lexically more frequent after nasals than voiceless stops. The different places of articulation show the same order of decreasing frequency as in Latin, with palatal added at the end of the list (coronal >> velar >> labial >> palatal).

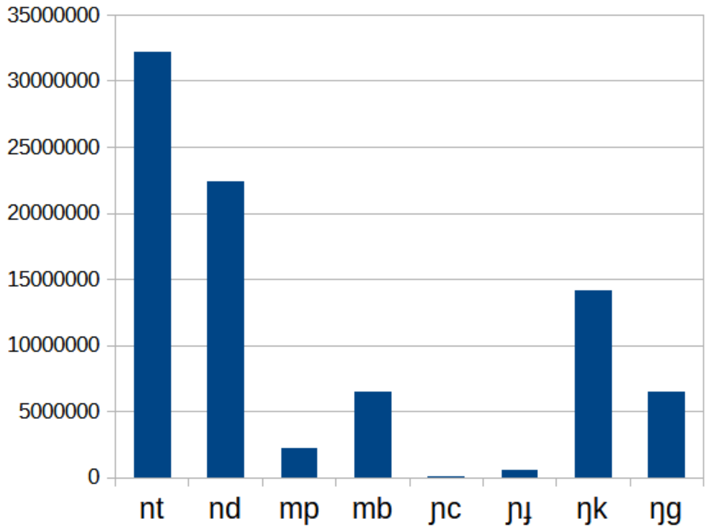

Figure 6: The textual frequency of nasal + stop clusters in Hungarian

The textual frequency of the same clusters shows by and large the same relations, as shown in figure6, with the exception of [ŋk], which is more frequent than [ŋg] — its numbers are increased by the 1Plur affix [ŋk] (*lát-unk* 'we see', *hát-unk* 'our back', *előtt-ünk* 'before us' etc.).[7]

---

[7] The hyphen in the Hungarian data indicates a morpheme boundary, it is not part of the orthographic representation.

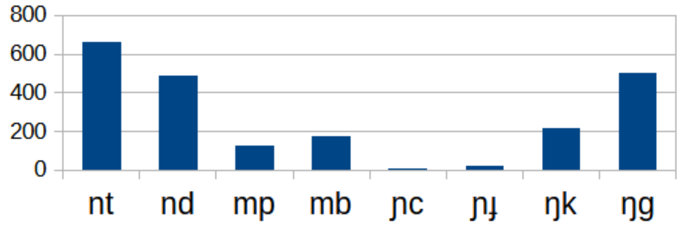

Figure 7: The lexical frequency of nasal + stop clusters within verbs in Hungarian

The overall lexical frequency relations within verbs, shown in figure 7, map those of the lexicon in general fairly clearly. Here the only exception is that [ŋg] is not only more than twice as frequent as [ŋk], it is also more frequent than [nd].

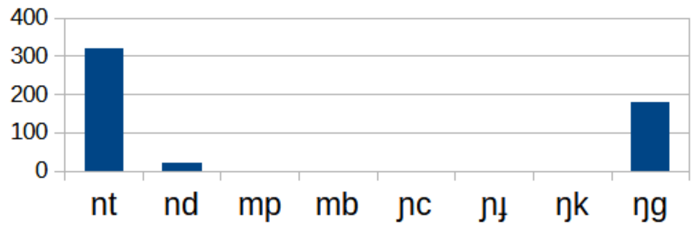

Figure 8: The lexical frequency of verb stems ending in nasal + stop clusters in Hungarian

Finally, the last column of Table 2, and figure 8, show the nasal + stop clusters that are found at the end of verb stems. It is notable that out of the eight clusters only three are found in this position; there are no labials, no palatals, and no [ŋk]. The two dominant stem-final clusters are [nt] and [ŋg], with [nd] represented only by a small number of lexemes. On the one hand, it is clear that only those three clusters that are the most frequent in the language overall are found in stem-final position. On the other hand, the absence of the majority of the relevant clusters is explicable with reference to morphological factors different from those that I invoked in the case of Latin, where the single most frequent cluster was never found in the same position.

In Hungarian, many of the [nt] and [ŋg]-final stems are morphologically complex, or at least represent typical morphological patterns. One derivational affix that is relevant in this context is -V*nt* (*csett-int* 'snap fingers (once)', *legy-int* 'wave with hand (once)', cf. *csatt-og* 'give snapping sound (continuously)', *legy-ez* 'fan (continuously)'). Another one is factitive or transitivising -*t*, which is added to many absolute stems, among them nasal-final ones (*men-t* 'save', *ron-t* 'damage, destroy', *ön-t* 'pour (transitive)', cf. *men-ekül* 'flee', *rom-lik* 'be damaged', *öm-lik* 'pour (intransitive)'). These are responsible for a large part of the set of [nt]-final verb stems.

Even though there are monomorphemic [ŋg]-final stems (*pang* 'stagnate', *fing-* 'fart'), their number is inflated by a verb-forming affix [Vŋg] (*düh-öng* 'rage', cf. *düh* 'anger') as well as by a (continuous) reflexive or detransitivising affix [g] (*ren-g* 'shake (continuously) (intransitive)', *kon-g* 'give hollow sound', cf. *ren-dül* 'shake (once) (intransitive)', *ren-dít* 'shake (once) (transitive)', *kon-dul* 'give hollow sound (once)', *kon-dít* 'ring a bell (once)'). There is even an absolute stem that can give rise to both a [ŋg]-final and a [nt]-final stem (e.g. *rán-g* 'twitch', *rán-t* 'tug').

The small set of [nd]-final stems consists of synchronically monomorphemic *mond* 'say' and a handful of verbs with the non-productive affix -*and/-end* (e.g. *örv-end* 'be happy', cf. *ör-ül* 'id.'). While other stops are also found in verb-forming patterns, they are incapable of forming clusters with other consonants, or are always accompanied by other affixes and are thus not in stem-final position (cf. the [d] element in *ren-dül, ren-dít* above, or the [k] in the verb *bán-kód-* 'be sad' vs. *bán-t* 'hurt, sadden').

In sum, then, it appears that we have no reason to assume a specific morphological or morphophonological reason for the absence of [ŋk]-final verb stems in Hungarian. This gap simply results from the etymological and morphological accidents which shrink the set of available consonant clusters to the three most frequent in the language overall.

5 Conclusion

This paper has looked at an apparent gap in the distribution of nasal + stop clusters in Latin and Hungarian. The phenomenon that was investigated in particular is the absence of a relatively frequent consonant cluster ([nt] in Latin, [ŋk] in Hungarian — the former the single most frequent consonant cluster in Latin) from a position at the end of verb stems. The interesting property of the missing consonant cluster in both languages is that it also happens to be the only one to function as a person marker in the verbal inflection of the respective language. I have argued that in Latin this represents a case of syntagmatic pressure to avoid repeating the same sequence at too close an interval. The absence of sequences identical to the other two polysegmental verbal endings may be seen as corroborating evidence for this functionally motivated gap.

As opposed to Latin, in Hungarian it seems that the absence of [ŋk] from verb stem-final position is unrelated to the identical phonological form of the 1Plur affix. It appears to be more of a historical accident in that there are simply no diachronic or morphological pathways to such a stem-final cluster. As a consequence, in Hungarian only clusters that are *more* frequent lexically than [ŋk] are found in this position; whereas in Latin only clusters that are *less* frequent lexically than [nt] are found in this position.

Comments invited

PiHPh relies on post-publication review of the papers that it publishes. If you have any comments on this piece, please add them to its comments site. You are encouraged to consult this site after reading the paper, as there may be comments from other readers there, and replies from the author. This paper's site is here:

https://doi.org/10.2218/pihph.8.2023.8932

Acknowledgements

The research reported in this paper has been supported by the Secretariat of the Eötvös Loránd Research Network.

Author contact details

András Cser
Hungarian Research Centre for Linguistics & Pázmány Péter Catholic University
Benczúr u. 33.
Budapest 1068
Hungary

cser.andras@nytud.hu

References

Baayen, Harald, Robert Schreuder, Nivja de Jong & Andrea Krott. 2002. Dutch inflection: the rules that prove the exception. In: Nooteboom, Sieg et al. (eds) *Storage and computation in the language faculty. studies in theoretical psycholinguistics* 61–92. https://doi.org/10.1007/978-94-010-0355-1_3

Baerman, Matthew. 2011. Defectiveness and homophony avoidance. *Journal of Linguistics* 47. 1–29.

Benkő, Loránd (ed.). 1993 *Etymologisches Wörterbuch des Ungarischen*. Budapest: Akadémiai.

Bertram, Raymond, Robert Schreuder & Harald Baayen. 2000. The balance of storage and computation in morphological processing: the role of word formation type, affixal homonymy, and productivity. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26(2). 489–511.

Bertram, Raymond, Matti Laine, Harald Baayen, Robert Schreuder & Jukka Hyönä. 2000. Affixal Homonymy triggers full-form storage, even with inflected words, even in a morphologically rich language. *Cognition* 74. B13-25. https://doi.org/10.1016/S0010-0277(99)00068-2.

Clackson, James. 2007. *Indo-European linguistics: an introduction.* Cambridge: Cambridge University Press.

Côté, Marie-Hélène. 2000. *Consonant cluster phonotactics: a perceptual approach.* Ph.D. dissertation, MIT.

Cser, András. 2015. The nature of phonological conditioning in Latin inflectional allomorphy. *Acta Linguistica Hungarica* 62. 1–35.

Cser, András. 2020. *The phonology of Classical Latin*. Oxford: Wiley / Blackwell.

Gordon, Matthew K. 2016. *Phonological typology*. Oxford: Oxford University Press.

Greenberg, Joseph. 1965. Some generalizations concerning initial and final consonant sequences. *Linguistics* 18. 5–34.

Hayes, Bruce. 1996. Phonetic mechanisms of post-nasal voicing. *Journal of The Acoustical Society of America* 100. 2693. https://doi.org/10.1121/1.417048

Kálmán, Béla. 1972. Hungarian historical phonology. In Benkő, Lóránd & Imre Samu (eds.) *The Hungarian language*. The Hague: Mouton. 49–84. https://doi.org/10.1515/9783110880236-003

Kaplan, Abby & Yuka Muratani. 2015. Categorical and gradient homophony avoidance: evidence from Japanese. *Laboratory Phonology* 6. 167–195. https://doi.org/10.1515/lp-2015-0006

Kaye, Steven. 2015. *Conjugation class from Latin to Romance: heteroclisis in diachrony and synchrony*. DPhil thesis, University of Oxford.

Krämer, Martin. 2009. The phonology of Italian. Oxford: Oxford University Press.

Kümmel, Martin Joachim. 2007. *Konsonantenwandel: Bausteine zu einer Typologie des Lautwandels und ihre Konsequenzen für die vergleichende Rekonstruktion*. Wiesbaden: Reichert.

Labrune, Laurence. 2012. *The phonology of Japanese*. Oxford: Oxford University Press.

Lakó, György. 1968. *Proto-Finno-Ugric sources of the Hungarian phonetic stock*. Bloomington: Indiana University Press.

Leumann, Manu. 1977. *Lateinische Laut- und Formenlehre*. München: Beck.

Lewis, Charlton T. & Charles Short. 1879. *Latin Dictionary*. Oxford: Clarendon Press.

Oravecz, Csaba, Váradi Tamás, Sass Bálint. 2014. The Hungarian Gigaword Corpus. In: Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Hrafn Loftsson, Bente Maegaard, Joseph Mariani, Asunción Moreno, Jan Odijk & Stelios Piperidis (eds.), *LREC 2014 — Ninth International Conference On Language Resources And Evaluation*. Lisbon: European Language Resources Association. 1719–1723.

Plag, Ingo, Julia Zimmermann & Gero Kunter. 2015. Homophony and morphology: the acoustics of word-final S in English. *Journal of Linguistics* 53: 1–36. DOI: 10.1017/S0022226715000183.

Rix, Helmut, Martin Kümmel, Thomas Zehnder, Reiner Lipp & Brigitte Schirmer. 2001. *Lexikon der indogermanischenVerben: die Wurzeln und ihre Primärstammbildungen*. Wiesbaden: Reichert.

Sauerland, Uli & Jonathan David Bobaljik. 2013. Syncretism distribution modeling: accidental homophony as a random event. In Goto Nobu, Otaki Koichi, Sato Atsushi & Takita Kensuke (eds.), *Proceedings of GLOW in Asia 9*. Tsu, Japan University of Mie 31–53.

Schrijver, Peter. 1991. *The Reflexes of the PIE Laryngeals in Latin.* Amsterdam / Atlanta, Ga.: Rodopi.

Siptár, Péter & Törkenczy Miklós. 2009. *The phonology of Hungarian.* Oxford: Oxford University Press.

Vaan, Michiel de. 2008. *Etymological dictionary of Latin and the other Italic languages.* Leiden: Brill.

Vallée, Nathalie, Solange Rossato & Isabelle Rousset. 2009. Favoured syllabic patterns in the world's languages and sensorimotor constraints. In François Pellegrino, Egidio Marsico, Ioana Chitoran, and Christophe Coupé (eds.), *Approaches to phonological complexity*. Berlin: Walter de Gruyter. 111–40.

Walter, Mary Ann. 2007. *Repetition avoidance in human language*. PhD dissertation, MIT.

Weiss, Michael. 2020. *Outline of the Historical and Comparative Grammar of Latin.* Ann Arbor: Beech Stave Press.