

How to be a marginal phoneme

LÁSZLÓ FEJES* 

ELTE Research Centre for Linguistics, Hungary

Received: January 26, 2025 • Revised manuscript received: October 13, 2025 • Accepted: October 14, 2025

Published online: May 15, 2026

© 2025 The Author(s)



ABSTRACT

This paper provides an overview of marginal phonemes based on examples mainly from Uralic languages. It explores the reasons why it is not possible to give an exact definition of the term *marginal phoneme*. It also examines how marginal phonemes may be characterised from the point of view of the phoneme system, their relationship to allophonic phenomena, their connection to their phonological environments, and the possible processes through which they emerge. These conditions can serve as standpoints for the typologisation of marginal phonemes. Finally, some hypothetical generalisations on marginal phonemes are provided.

KEYWORDS

marginal phonemes, phoneme system, phoneme borrowing, allophony, phonological change

1. INTRODUCTION

Speech sounds contrasting lexemes or word forms are considered to be the realisations of distinct phonemes. The specification of the phoneme inventory is an essential part of the description of a language (cf. Hall 2013, 218). However, phonemes cannot be contrasted in every context (cf. Hall 2013, 231–237). The functional load of phonemes may be varied. Phonemes that are rarely contrasted in a language are labelled as *marginal phonemes* (cf. e.g. Katamba 1989, 29; Kiparsky 2018, 70; Suomi, Toivanen & Ylitalo 2008, 32) or *marginal segments* (cf. e.g. Gussenhoven & Jacobs 2017, 43). However, many comprehensive works, such as handbooks or introductions for linguistic students, do not even mention the term or the phenomenon.

* Corresponding author. E-mail: fejes.laszlo@nytud.elte.hu

This paper aims to establish a typology for marginal phonemes. The examples discussed are mostly, but not exclusively, taken from Uralic languages. This choice is justified not only by the author's expertise, but also by the lack of data from Uralic in the literature on marginal phonemes. For example, Hall (2013) presents a wide range of examples of marginal phonemes from various languages. Her catalogue suggests that the existence of marginal phonemes is not a marginal phenomenon cross-linguistically; instead, it seems that every language is expected to have some marginal phonemes. However, e.g. Uralic languages are sparsely represented among her examples: two cases from North Saami are mentioned based on Bals, Odden & Rice (2007) and Bye (2009), Enets is mentioned with reference to Anderson (2004), and Salminen (2007) is also included (Hall 2013, 217, 233, 239, 241, 248). However, the descriptions of Uralic languages usually mention some phonemes only occurring in loanwords, onomatopoeic words or just in a restricted set of morphemes (sometimes only in given phonological contexts); that is, phonemes with a low functional load.

The lack of data on Uralic languages can be explained by the fact that the issue of marginal phonemes has never been addressed in Uralic linguistics as a general problem. The extensive handbooks of Uralic linguistics rarely discuss the possible marginality of phonemes. In Sinor (1988), marginal vowels are only mentioned in the description of Hungarian (Szathmáry 1988, 199) and in the history of Proto-Finno-Permic (Sammallahti 1988, 524 – using the term *peripheral*). In Abondolo (1998c), Finnish (Abondolo 1998a, 152), Erzya (Zaicz 1998, 185), Hungarian (Abondolo 1998b, 430–431, 434) and Proto-Samoyedic (Janhunen 1998, 463) marginal phonemes are mentioned. In Abondolo & Valijärvi (2023b), in addition to Hungarian and Finnish (Abondolo & Valijärvi 2023a, 19, 23), only Veps (Grünthal 2023, 19, 23) marginal phonemes are referred to. Only a few chapters discuss marginal phonemes in Bakró-Nagy et al. (2022): in Lule Saami (Ylikoski 2022, 133, 134), in Northern Saami (Ante & Ylikoski 2022, 149, 150, 151) and in Eastern Khanty (Schön & Gugán 2022, 610). (In the description of Finnish, the restricted distribution of /b, d, g, ŋ/ and /ʔ/ is discussed, but the term *marginal phoneme* or something similar is not used for them. The marginal phoneme /ʃ/ is shown in parentheses in the consonant chart (cf. Laakso 2022, 256–257). Additionally, the marginality of vowel length in Veps (Grünthal 2022, 293) and palatalisation of alveolar consonants in Votic are also mentioned (the latter as *peripheral*). None of the subject indices of these handbooks list the term *marginal* (or *peripheral*) *phoneme*. The existence of marginal phonemes is well-known: these phonemes are often presented in parentheses in the phoneme charts. However, the phenomenon of marginality remains beyond the scope of interest.

This paper offers a comprehensive discussion in order to establish a typology of marginal phonemes by analysing data from various languages, primarily (but not exclusively) from Uralic languages. In Section 2, the notion of *marginal phoneme* is discussed, and it is pointed out that there is no clear definition of the term (nor is it feasible to define it clearly). Section 3 deals with the problem of the “surface” and “underlying” phonemes (phonemes in the structuralist and generative sense, respectively), and their relationship to marginality. Section 4 surveys the position of marginal phonemes in the phoneme system. Section 5 studies the relationship between marginal phonemes and allophony. Section 6 discusses the types of marginal phonemes according to the environments they occur in. Finally, Section 7 considers how marginal phonemes emerge.

2. THE NOTION OF MARGINAL PHONEME

The existence of marginal phonemes might be regarded as unexpected, as the low functional load of marginal phonemes could suggest that they tend to be eliminated from the language: in fact, they are frequently in free variation with some central (i.e. non-marginal) phonemes (cf. Section 5.1). [Matisoff \(1994, 121–122\)](#) argues that marginal phonetic features can be exploited for affective or symbolic purposes: “their very rarity makes them appropriate for grammatical or symbolic duty” (although his examples are not phonemes but tones). He also mentions that elements with low functional loads can help to avoid the overburdening of the system by the emergence of new homophones. Additionally, language contact and bilingualism can also support the survival of marginal phonemes.

In languages, the frequency of phonemes varies: there are always frequent and rare phonemes. The frequency of a phoneme may be strongly influenced by phonotactic restrictions. In many Uralic languages, certain vowel phonemes are only contrastive in the initial syllables of words, as in subsequent syllables, some of them cannot occur. This kind of restriction can concern subsets of natural classes ([Hall 2013, 242–243](#)): typically rounded vowels do not (or just exceptionally) occur in non-initial syllables (e.g. in Hill Mari, Komi or Ob-Ugric) or just some of them occur in this position (as /u/ in Standard Estonian or /o/ in Udmurt). Some phonemes may not occur word-initially (e.g. similarly to English, /ŋ/ in Mari or Ob-Ugric) or word-finally (as /o/ and /ø/ in Hungarian or non-coronal consonants¹ in Finnish). However, these phonemes are usually frequent enough, at least in the positions where they occur, not to be considered marginal.

Certain speech sounds may be contrastive in some contexts, while they are neutralised in others. In Moksha, /a/ and /æ/ are considered to be distinct phonemes, although they behave as allophones in several positions, and can only be contrasted in some others ([Bartens 1999, 28–30](#); [Kukhto 2018, 29](#)). E.g. only [æ] occurs between two palatalised alveolar consonants² or after a palatalised alveolar in a word-final position. Meanwhile, only [a] occurs after an unpalatalised alveolar (except for /s/, after which [æ] occurs as well), or even after a palatalised alveolar but followed by an unpalatalised one. (Stem-final /æ/ is even realised as [a] when the suffix begins with a non-palatalised alveolar, cf. [Bartens 1999, 30](#)). Nevertheless, in some other contexts, e.g. word-initially or next to non-alveolar consonants, /a/ and /æ/ are contrasted. Due to the relatively wide context in which they can be opposed to each other, these consonants are not considered marginal in the literature.

Studies on European ([Martindale et al. 1996](#)), Eurasian ([Tambovtsev & Martindale 2007](#)) and Australian ([Macklin-Cordes & Round 2020](#)) languages show that languages exhibit a similar distributional pattern: while the frequency of most phonemes is around the average, there are always some outliers – phonemes with much higher or much lower frequencies (cf. [Figure 1](#)).

Despite the widely known and often cited principle, “once a phoneme – always a phoneme”, descriptions of various languages often question the phonemic status of units which are contrastive only in a restricted set of environments. It is rather an exception when descriptions overtly discuss the marginality of certain phonemes. E.g. [Petr \(1986, 124–125\)](#) takes the idea that every

¹However, see Section 3.

²In this context, the palatal semivowel /j/ must also be regarded as a palatalised consonant.

Additionally, some marginal phonemes occur only in onomatopoeic words and in loanwords. It can be argued that neither these words nor these phonemes are part of the language. E.g. some dialects of Hungarian spoken in Transylvania and Western (Romanian) Moldavia have a bilabial trill phoneme occurring in just a handful of words ([by̥t̪] øk] ‘cricket’, [bys:øg] ‘to sneeze’ (Gálffy 2005, 85; Juhász 2012, 102). Standard Hungarian has an initial /t/ in these words, and some dialects have an initial cluster /pr/. The latter forms are occasionally used even in the standard language. One can argue that [b] is just an emphatic realisation of the /pr/ cluster in these onomatopoeic words (or at least words of onomatopoeic origin), and not the realisation of an independent phoneme. A similar restricted use of [b] is also reported from Erzya (Rueter 2010, 61) and Komi-Permyak (Wichmann & Uotila 1942, XII, 362). Its phonemic status is even more questionable in Finnish, where it occurs only in interjections like [tb:], [tpb:] or [bb:] (Karlsson 1983, 33). The phonotactic structure of these words does not resemble Finnish words, which always contain at least one vowel. Additionally, it is worth considering that this phoneme is cross-linguistically rare.

Likewise, in the case of marginal phonemes found in onomatopoeic words, it can be argued that putative phonemes occurring exclusively in loanwords are merely illusions caused by the interference of distinct languages (code-switching), especially when bilingualism is widespread. The case is extremely difficult when a loanword is borrowed from a dialect of the same language, and not from a distinct language. However, when it is accepted that a given phoneme occurring only in onomatopoeic and/or loanwords belongs to the language, its marginality can be questioned if it occurs relatively frequently, especially in loanwords from languages that speakers of the target language typically do not speak. E.g. the phonemic status of /f/ can hardly be questioned in any Slavic language. However, since /f/ almost never occurs in words of Slavic origin,⁵ its marginality might be assumed. A counterargument against this supposition is that /f/ today is relatively frequent in these languages.

In the following, those phonemes that are either extremely rare or less rare but limited to specific phonological environments or confined to certain layers of the vocabulary (most typically loanwords and onomatopoeic words) will be considered marginal.

3. “SURFACE PHONEMES” VS. “UNDERLYING PHONEMES”

Before identifying marginal phonemes, a specific issue must be considered. In linguistics, two slightly different notions of *phoneme* are used. The first, which can be labelled as the *surface phoneme* or the *phoneme in the structuralist sense of the term*, distinguishes phonemes based on whether they are contrastive on the surface. The second, which can be called *underlying phoneme* or the *phoneme in the generative sense of the word*, also considers the alternations in which the segment participates (these are called *archiphonemes* in the structuralist frameworks).

For example, in Standard Komi coda /v/ generally alternates with onset /t̪/: /t̪ov/ ‘soul’ : /t̪ot̪-is/ ‘soul-POSS.3SG’. However, Russian loanwords ending in /v/ are left out from this

⁵The rare exceptions are due to rare changes like **upřevati* ‘to trust, to hope’, cf. BCMS *ufati*, Old Czech *úfati*, Polish *ufać*.

alternation: /ov/ ‘surname’⁶ : /ov-is/ ‘surname-POSS.3SG’. This case is easy to analyse, assuming that words ending in an alternating /v/ in fact end in an underlying /t/, and are realised as [v] when they enter a coda. However, there are also Russian loanwords ending in a non-alternating /t/, e.g. /pədvət/ ‘basement’. Suppose the aim is to analyse this difference phonologically. In that case, it has to be supposed that there are two underlying /t/ phonemes (archiphonemes) in Standard Komi: an immutable one and an alternating one. This analysis is problematic for three main reasons.

First, in this case, it is impossible to determine how (in which features) the two /t/s differ from each other.

Second, it is impossible to decide which /t/ is present in the onset. Still, without deciding it, it is impossible to settle the frequencies of the two /t/s; consequently, it is impossible to decide whether one of them is marginal or not.

Third, such an approach can lead to the radical expansion of the phoneme system in languages with a complex morphophonology. In Standard Hungarian, some mid vowels exhibit syncope when they would otherwise occur in an open syllable: /bokor/ ‘bush’ : /bokr-ok/ ‘bush-PL’. However, it is not always the case: /motor/ ‘engine’ : /motor-ok/ ‘engine-PL’. In the analysis described above, this difference in behaviour requires duplication of the /o/ phoneme. Additionally, /o/ usually undergoes vowel harmony: /hɑːz-hoz/ ‘house-ALL’ : /tyːz-höz/ ‘fire-ALL’ : /keːz-hez/ ‘hand-ALL’. But there is at least one exception: /hɒt-kor/ ‘six-TEMP’ : /øt-kor/ ‘five-TEMP’ : /neːʒ-kor/ ‘four-TEMP’. So, the non-alternating /o/ has to be duplicated again (although the further suffixation of such forms is very restricted, sometimes it is possible: /neːʒ-kor-i/ ‘four-TEMP-ADJZ’). In vowel harmony, not just (potential) targets can behave in different ways, but also triggers: /ʃiːn/ ‘(train) track’ : /ʃiːn-ek/ ‘track-PL’, but /ʃiːp/ ‘whistle’ : /ʃiːp-ok/ ‘whistle-PL’, which can be an argument for differentiating two /i/s (/i/s, /e/s, even an extra /ɛ/ for forms like /derek-ɒ/ ‘waist-POSS.3SG’).

Such an approach inevitably results in huge (underlying) phoneme systems, several marginal phonemes included. In fact, these “marginal phonemes” will not be phonemes which are rarely contrastive but phonemes which occur in morphemes with a less regular (or completely irregular) morphology. Therefore, underlying phonemes (archiphonemes) should not be classified as marginal, and the irregular behaviour should be attributed to the morphemes instead of the supposed underlying phonemes.

However, at least one case must be considered. In Finnish, there is a phenomenon called *jäännöslopuke* (~ *residual ending*, earlier *loppuhenkonen* ~ *final aspiration*, cf. Hakulinen 1941, 46–48). It refers to a segment which occurs only morpheme-finally and which is realised as a glottal stop utterance-finally or before words beginning with a vowel (even as a geminate in the latter case). However, when it is followed by a (word- or suffix-initial) consonant, it completely assimilates to that. This assimilation process is called *rajageminaatio* (~ *boundary gemination*), *rajakahdennus* (~ *boundary duplication*) or *loppukahdennus* (~ *final duplication* etc.).

At first sight, it seems to be evident to suppose a phonemic glottal stop (/ʔ/), which assimilates to the consonants after it (and is dropped word-internally before vowels). Despite that, some descriptions of Finnish ignore the existence of such a phoneme (cf. Karlsson 1983, 34; Häkkinen 1983, 41) or even explicitly deny its phonemic status (Hakulinen et al. 2004, § 3). Nonetheless, several morphophonological phenomena are simpler to explain if we assume that

⁶From the widespread ending /-ov/ in Russian surnames.

there is an “invisible” consonant in these morphemes, e.g. it seems to close the syllable and induce the weak grade of gradable stems: /tarve?/ ‘need:SG.NOM’ : /tarpe:n/ ‘need:SG.GEN’ (cf. /turpa/ ‘muzzle:SG.NOM’ : /turvan/ ‘muzzle:SG.GEN’). Additionally, the phenomena attested at the word boundary are most simply explained by assuming a consonantal segment. E.g. the difference between the female name /anna/ ‘Ann(e)’ and the second person singular imperative (or connegative) form of the verb ‘give’ (/anta:/ ‘give:INF’ /anna?/ ‘give:IMP.2SG’ (or ‘give:CONNEG’; realised with a glottal stop or an assimilated consonant at the end) can hardly be distinguished in any way other than assuming that these forms end in a consonant. This may lead to the assumption of a phonemic /?/, which should undoubtedly be considered marginal. In fact, many descriptions of Finnish implicitly admit the existence of some segment at the end of words by writing forms with *jäänöslopuke* with a superscript *x* (°) at the end of such words, e.g. *tarve^x* ‘need:SG.NOM’. This way of transcription suggests the presence of a virtual phoneme. Interestingly, while (Hakulinen et al. 2004, § 3) categorically denies the phonemic status of [?], (Hakulinen et al. 2004, § 34) states that there is an “abstract morphophoneme” at the end of the discussed forms, which surfaces in various ways. The mainstream handbooks of Uralistics usually consider /?/ to be a marginal phoneme in Finnish: cf. Abondolo (1998a, 152); Abondolo & Valijärvi (2023a, 23–25); Laakso (2022, 256–257).

Nonetheless, these contradictions can be reconciled by assuming the presence of a central phoneme, not a marginal one. In Finnish, only dental consonants can appear as stem-final consonants. In the rare cases when a non-dental consonant occurs word-finally, it must change in a way: the word-final labial nasal surfaces as [n]: [avain] ‘key:SG.NOM’, cf. [avaimen] ‘key:SG.GEN’. This means that a word-final central but not dental phoneme can be assumed in words with *jäänöslopuke* instead of /?/. The most likely candidates are the velar stop /k/ and the glottal /h/. In fact, these are the historical sources of the *jäänöslopuke* in most cases. The most appropriate analysis for *jäänöslopuke* and *rajageminaatio* seems to be the assumption that these morphemes still end in /k/ or /h/ underlyingly. Consequently, [?] is not the default realisation of a marginal phoneme, but a specific word-final (but not assimilated) realisation of a word-final /k/ or /h/ (which is assimilated to a following consonant, if there is one). Evidently, this analysis contradicts the generalisation that only dental consonants can surface word-finally. It is possible to analyse [?] as belonging to the following word, but this is hardly feasible with geminates, as native Finnish words usually do not begin with a cluster, and never with a geminate. So, an exception must be acknowledged to the general ban of non-dental consonants word-finally: when they are identical with the initial consonant of the next word, and when they are realised as a glottal stop (which can be comprehended as a weakened form of [k] or [h]).⁷ Nonetheless, it still lacks an explanation why morpheme-final /k/ or /h/ must be deleted before vowels: otherwise they occur intervocally.

4. FROM THE POINT OF VIEW OF THE SYSTEM

In general, the phoneme system is conceived as an organised set with a network of oppositions, expressed with the presence or the absence of given features (and some characteristics usually

⁷In fact, some unadapted loanwords can also end in non-dental consonants, e.g. /dōup(:)in/ ‘doping’ (Hakulinen et al. 2004, § 3).

not analysed with features, e.g. length). A marginal phoneme can fit the phoneme system well: such cases are discussed in Section 4.1. However, there are also cases when some distinctive features of the marginal phonemes are not applied in the phoneme system otherwise: such examples are discussed in Section 4.2.

4.1. Filling the gaps

In some cases, phoneme systems have gaps: that is, a certain combination of otherwise distinctive features is not used. Languages tend to fill these gaps during their history. However, if some contact languages have a sound corresponding to the missing phoneme, loanwords can be adopted (and adapted) in a way that the loanword contains a phoneme filling the gap.

For example, Standard Czech has a five-vowel system with the distinction of front-back and three heights. Additionally, vowels can be short and long. However, long /o:/ does not occur in native words; it is attested only in words of foreign origin (see Table 1, cf. Petr 1986, 33, 124, 128, 130). It is noteworthy that Standard Czech has a diphthong /ou/ occurring in native stems, but also /au/ and /eu/, occurring in loanwords and some emphatic words.

The situation is similar in Standard Latvian, which has a front/back opposition between high and low vowels, and three distinct heights among front vowels (Table 2). Additionally, length is also distinctive. However, back mid vowels do not occur in native words, although the diphthong /uo/ (in addition to several other diphthongs) is widely used. Still, short /o/ and long /o:/ only occur in loanwords (cf. Bergmane et al. 1959, 24–26; Fennel & Gelsen 1980, xxvi; Mathiasen [1997], 28–29, 33; Prauliņš 2012, 10–12).⁸

A similar case is the adoption of /f/ in several Slavic or Baltic (and, supposedly, many more) languages.⁹ These have /v/ in native words, and the voiced–voiceless distinction is common among obstruents. However, /f/ does not occur in native words, only in loanwords and onomatopoeic words. The issue is complicated by the fact that [f] also occurs as an allophone of /v/. In a position where /v/ and /f/ are neutralised into [f], the phone can be analysed as the realisation of both the voiced and the voiceless phoneme even in native words, e.g. Polish *wczoraj* [ft͡ɕɔraj] ‘yesterday’ can be both analysed as /vt͡ɕɔraj/ and /ft͡ɕɔraj/. Moreover, there

Table 1. The vowel system of Standard Czech

i, i:		u, u:
ɛ, ɛ:		o, (o:)
	a, a:	

⁸In the spelling of Latvian, *o* is generally read as /uo/ (e.g. *valoda* [valuoda] ‘language’), but it also marks /o/ and /o:/ in words of foreign origin (e.g. *motors* [moto:rs] ‘engine’).

⁹Although the adoption of /f/ is common, the circumstances are different, e.g. in some languages /v/ has a /u/ allophone instead of [f] (Sussex & Cubberley 2006, 144–145). Especially in these cases, it is also questionable whether, instead of the fricative /v/, an approximant /ʋ/ should be supposed. Nonetheless, the presented analysis holds for Standard Czech, Polish, Russian, Bulgarian and Macedonian.

Table 2. The vowel system of Standard Latvian

i, i:	u, u:
e, e:	(o, o:)
æ, æ:	ɑ, ɑ:

are hardly any synchronic arguments supporting the former analysis. Consequently, it seems that the adoption of /f/ in loanwords seemingly leads to its emergence in several native words, while these words did not change.

4.2. Growing “knobs”

Speaking about gaps, we should also speak about outlying phonemes, distinguished by features which are otherwise not distinctive in the language: in this paper, these will be called *knobs*. An eye-catching example of the phenomenon is Standard Finnish. Standard Finnish hardly exploits the feature of voicing: in native words, obstruents are usually voiceless, and sonorants are voiced. Most of the marginal phonemes of Standard Finnish are the result of the exploitation of the distinction in voicing.

A very special exception to the generalisation on non-contrastivity of voicing is the phoneme /d/, which is a voiced obstruent, contrasted to /t/. Originally, it was a /ð/ (although also marginally, /p/ was present in Finnish as well); its quality has changed due to Swedish influence. Its distribution is very restricted: it occurs only word-internally, in an intervocalic position or following /h/. In morphological terms, it occurs only in the weak grade of gradable stems (originally in the onset of closed or even numbered syllables, e.g. /katu/ ‘street.NOM.SG’ : /kadu-n/ ‘street-GEN.SG’) and some alternating, historically gradable suffixes. It was lost in most of the dialects and has been replaced by a hiatus, a flap or some other sound (/l, r, j, v, h/ etc.); meanwhile, it developed into /d/, and it was preserved in Standard Finnish due to standardisation and kept alive artificially.

Additionally, /d/ also occurs in other positions in loanwords, where it may even appear as a geminate that undergoes gradation: /vedɑ:/ ‘Vedda.NOM.SG (member of an indigenous people of Sri Lanka)’ : /vedɑ-n/ ‘Vedda-GEN.SG’. Two other voiced plosives also occur in loanwords: /b/ and /g/. These also participate in gradation: /mobɑtɑ?/ ‘mob.INF’ : /mobɑ:n/ ‘mob-PRS.1SG’, /blogɑtɑ?/ ‘blog.INF’ : /blogɑ:n/ ‘blog-PRS.1SG’. Additionally, /z/ (/penzɑ/ ‘Penza’) and /ʒ/ (/ʒiguli/ ‘Zhiguli’) may occur in foreign words (first of all, in proper names). Among these, /ʒ/ is an extreme outlier as the corresponding voiceless sibilant, /s/, is also marginal in Finnish: postalveolar consonants do not occur in native stems and suffixes (Setälä 1902², 1–3; Hakulinen et al. 2004, §3, §6). On these consonants, see also Section 5.1.

The voiceless labial fricative /f/ is also restricted to loanwords. Seemingly, it is opposed to a voiceless labial fricative /v/ (cf. Setälä 1902², 4). However, according to phonetic measurements, the latter is rather an approximant [ʋ] (cf. Bolla 1985, 138–141, 236), and it is frequently analysed as a semivowel (cf. Karlsson 1983, 34–36; Hakulinen et al. 2004, §§3–4). The main arguments for such an analysis are that, similarly to /j/, the labiodental “semivowel” is never geminated and never occurs as a first element of clusters (Karlsson 1983, 35–36) – at least not in

native words, cf. /klovni/ ‘clown’ (Hakulinen et al. 2004, §31). (The voiceless labial fricative /f/ can be geminated: /lef:a/ ‘movie’, /prof:a/ ‘professor’, even in a complex cluster: /golfi-a/ ‘golf-PART’, cf. Hakulinen et al. 2004, §28.) Still, sometimes /v/ (or /ʋ/ instead, possibly even /w/) and /f/ are discussed as a voiced–voiceless pair, cf. Nuutinen (1994, 77).

5. ALLOPHONY

In what follows, two issues concerning the connections between marginal phonemes and allophony are discussed. On the one hand, marginal phonemes are often realised in the same way as a central phoneme, resulting in vacillating realisations.¹⁰ This variability can lead to the question of their phonemic status. Such issues are discussed in Section 5.1. On the other hand, the default realisation of marginal phonemes can coincide with a frequent but not default realisation of some central phonemes. Examples for this issue are discussed in Section 5.2.

5.1. The instability of marginal phonemes

Nuutinen (1994) argues that [d], [b], [g] do not have a phonemic status in Finnish. According to him, [d] in native words is generally omitted in the spoken language, while in other cases, these consonants are replaced by the corresponding voiceless sounds. He also argues against accepting [ʃ] as a phoneme, as it is generally replaced by [s] (he does not even mention [z] and [ʒ]). The existence of the Finnish contrast in voicing is questioned by Nuutinen. He argues that Finns tend to write voiced consonants in foreign words in which the consonant is voiceless (e.g. *tobless-baari* ‘topless bar’); additionally, some Finnish loanwords contain voiced obstruents while they contain voiceless obstruents in the source language (e.g. [bi:si] ‘song’ < English *piece*, [fudis] ~ [futis] ‘football’ < Swedish *fotboll* [fu:tɔl]).¹¹ He recalls an anecdote according to which peasants tried to imitate the speech of educated people by voicing the plosives. According to Nuutinen (1994), originally Finnish adopted foreign words by substituting voiceless obstruents for voiced ones; however, at some point in time, this strategy became stigmatised and considered to be a sign of ignorance and lack of education. This happened although most speakers did not use voiced obstruents, and even educated speakers struggled with the consistent use of them.¹² Contrary to the argumentation of Nuutinen (1994), all these facts demonstrate that at least a group of (educated) speakers of Finnish use the voiced obstruents. Although the

¹⁰Here, vacillation means that two or more forms are acceptable in the same meaning; however, they are not completely free variants. The probability of the use of the various forms might depend on factors like the lexeme, the grammatical context, the register, or the idiolect of the speaker. However, these factors do not determine the form, and different forms might be used in the same discourse by the same speaker.

¹¹Nuutinen (1994, 70) also mentions the examples *sgabat* and *sgidi*, which are difficult to document from contemporary Finnish. However, *skabat* [skabat] ‘race; exam; things’ (< Swedish *kapp*- ‘to race’) *skidi* [skidi] ‘child, kid’ (< English *kid*) are documented (e.g. Karttunen 1979, 226–228); both are clearly hyperurbanisms, but the latter has not much to do with hyperurbanistic voiced obstruents. Karttunen 1979 does not include any words beginning with *sg*. (Although Karttunen 1979 is outdated today, and due to the rapid changes in slang, it must have been somewhat outdated in Nuutinen’s times, it must be considered relatively state-of-the-art for the era it was written in).

¹²The phenomenon is not restricted to voiced plosives: the substitution of [ʋ] for [f] (and [hu] for [fi]) and the simplification of consonant clusters were similarly stigmatised.

definition of Standard Finnish is debatable, the speech of this educated and prestigious social class must play a key role in the specification of the standard. Additionally, even though there is a variation between the voiced and voiceless consonants, this variation is only attested in certain words: while *bussi* ‘bus’ is pronounced both as [bus:i] and [pus:i], *pussi* ‘(plastic/paper) bag’ is always pronounced as [pus:i], and never as *[bus:i]. That is, these obstruents must be analysed as distinct phonemes: the ones which can be pronounced both as voiced and voiceless are phonemically voiced; those which can be realised only as voiceless are phonemically voiceless.¹³

It seems that this kind of instability is typical for those phonemes which appear in loanwords. A similar case is Hungarian /akade:mia/, which is usually pronounced with common Hungarian [ɑ]s ([akade:mia]) but pronounced also as [akade:mia] by some conservative, educated speakers (Nádasdy & Siptár 1989, 11–12, Keresztes 1993, 49). Nádasdy & Siptár (1989, 13) argue that there is only one phoneme /a/, which usually surfaces as rounded, but some words belong to a lexical category in which it surfaces unaltered. According to Keresztes (1993, 49), pronouncing [a] is just a way of attitudinisation. Nonetheless, it is clear that this kind of pronunciation only occurs in a stratum of foreign words (in other morphemes, it indicates that the speaker is the speaker of the Palóc dialect, in which [a] corresponds to Standard [ɑ] regularly, and [ɑ] never surfaces). It is very difficult, if possible at all, to prove that the lack of the labialisation of the short back low vowel is due to these words’ lexical markedness and not to the fact that they contain marginal phonemes. Sometimes even a restricted contrast can be demonstrated: [hal:] can be both ‘hear.PRS.3SG’ and ‘lobby.NOM.SG’, but [hal:] can only be the latter in Standard Hungarian (although this kind of pronunciation is stylistically marked).

A similar example comes from Votic. According to Ariste (1968, 2), the vowel /i/ is a phoneme which appears only in loanwords from Russian. In contrast to this, Lauerma (1993, 53) argues that Russian [i] is replaced by various other vowels (/ə, i, u/) even in Russian loanwords; and therefore, [i] must be considered an allophone even in words where it appears. However, it is unpredictable where [i] may appear at least as an alternative to some other vowel; therefore, it cannot be an allophone. It is more appropriate to assume a phonemic /i/ which can also be realised in a way typical of some other phoneme.

To sum up, the mere fact that the appearance of a marginal sound is inconsistent does not mean that it does not represent a marginal phoneme. On the contrary, when a frequent sound in a given position can freely alternate with a rare one in some morphemes but not in others, it must be concluded that the alterable sound represents some other phoneme than the non-alterable. The marginality of a phoneme can manifest itself not only in its rarity but also in its instability, that is, the omission of the realisation of its specific features. The fact that different speakers use the rare sound in different words does not automatically question the analysis of these sounds as the realisation of marginal phonemes. It is normal for speakers with the same phoneme system to utilise different phonemes in some words; the fact that a sound may be contrastive in different words for different speakers does not question its contrastivity.

¹³It may seem that hyperurbanistic voicing suggests that even voiceless obstruents can be realised with voicing. However, in such cases, it is not a voiceless phoneme realised as a voiced one, but the speaker is unfamiliar with the word and assumes a voiced obstruent in the word. A parallel case is *skidi* [skidi] ‘child, kid’ (< English *kid*), in which it was not /k/ realised as [sk] but speakers assumed it to begin with /sk/. The reason might be that they were used to foreign words beginning with clusters relatively often (especially compared to native words).

5.2. Restricted contrastivity of frequent allophones

The examples in Section 5.1 showed cases when rare sounds represent marginal phonemes which are not always realised as that rare sound, but sometimes identically to a default realisation of a central phoneme. In this section, the typical realisations of the marginal phonemes are relatively frequent, but in most cases, the sounds are the realisations of central phonemes and only rarely of marginal phonemes.

A clear example of that is Standard Erzya, in which /i/ and /e/ are retracted after non-palatalised coronals and are realised as [i] and [ɘ], respectively. The sound [i] also occurs after non-coronal consonants or word-initially in loanwords or onomatopoeic words (Rueter 2010, 60): /vivod/ ‘conclusion’, /kalmik/ ‘Kalmyk’, /irnems/ ‘howl menacingly’ (vs. /irnems/ ‘howl, yowl’), /kirgams/ ‘curry, comb, brush’ (vs. /kirgams/ ‘clean of (twigs, knots)’). The example of /kirgams/ is interesting because it is a Tatar loanword. While practically all Erzya speakers speak Russian, the knowledge of Tatar is rather exceptional among them nowadays: consequently, this kind of pronunciation cannot be attributed to interference.

Rueter (2010, 61) includes a phoneme /i/ in the vowel system, and the transcription applied by him suggests that he analyses all the occurrences of [i] as the realisation of this phoneme. However, as the appearance of [i] is completely neutralised and predictable after coronals (that is, it never occurs after palatalised ones and it is in a complementary distribution with [i], which never occurs after non-palatalised ones), it seems to be more appropriate to analyse [i] as the allophone of /i/ there (with a special regard to [i] ~ [i] initial suffixes).¹⁴ Consequently, while [i] is a relatively frequent sound in Erzya, it is the realisation of /i/ in most of the cases, and of /i/ just in a restricted set of cases.

Another example is Finnish /ŋ/. As in many other languages, [ŋ] is an allophone of /n/ before velar consonants (in the case of Finnish, only /k/). However, similarly to /d/ (cf. Section 4.2), /ŋ/ emerged in the weak grade of stems containing the /nk/ ([ŋk]) cluster in the strong grade: actually, /ŋ/ always emerged as a geminate /ŋ:/ there. Its phonemic status is the most salient when compared to the weak grade forms of stems with a cluster /nk/ in the strong grade: /van̄in/ ([van̄in]) ‘prisoner:SG.GEN’ vs. /vankin/ [van̄kin] ‘robust:SUP.SG.NOM’ (cf. the unsuffixed forms /vanki/ [van̄ki] ‘prisoner:SG.NOM’ in the former and /vank:a/ [van̄k:a] ‘robust:SG.NOM’ in the latter case). However, the position of /ŋ/ is strengthened by loanwords, where it also appears as a short consonant: /tan̄o/ ‘tango’, /aŋlismi/ ‘anglicism’, /maŋnet:i/ ‘magnet’, /dōp̄(:)in/ ‘doping’ etc. (Hakulinen et al. 2004, § 3).

The case of Erzya /ŋ/ is also worth mentioning. In Erzya, [ŋ] occurs before /k/ or /g/ (although not always, only morpheme-internally). However, Rueter (2010, 58–59) also mentions the word /ŋot/ ‘so’, in which [ŋ] occurs before a vowel (in a minimal pair of the Russian loanword /not/ ‘note’). It may seem that the emergence of the phoneme /ŋ/ could have been supported by the fact that it had already existed as an allophone. (For a further discussion of Erzya /ŋ/, see Sections 2 and 7.) Nonetheless, allophones are not likely to phonematise without a reason (e.g. changes of their environments). Hungarian also has an allophonic [ŋ] which occurs before /k/ or /g/, but speakers are not aware of pronouncing a different sound in this

¹⁴Alternatively, it would be appropriate to analyse [i] after palatalised coronals as the allophone of /i/. Nonetheless, as [i] occurs very rarely after non-coronals and word-initially, it would be strange to suppose that /i/ occurs after coronals but the otherwise frequent [i] does not.

environment, and they struggle with pronouncing it in other positions (e.g. the English ending *-ing* is often pronounced as *[-iŋg]* even by Hungarians speaking English fluently).

Sygva Mansi also presents an interesting case. While */c/* is a marginal phoneme occurring mostly, but not exclusively, before */i/* in Sosva and Ob Mansi (cf. 6), *[c]* is much more frequent in Sygva Mansi. This is a result of a palatalisation process: *[k]* changed into *[c]* before */i(:)/* and */e(:)/* (cf. Kálmán 1976, 23). As a result, most *[c]*s can be analysed as the allophones of */k/* – it also holds for words in which *[c]* precedes *[i]* also in Sosva and Ob Mansi. As a consequence, */c/* seems to be even more marginal in Sygva Mansi than in Sosva and Ob Mansi. However, the change *[k] > [c]* had some exceptions, e.g. there was no palatalisation in the word *[aki]* ‘uncle’. Thus, palatalisation is not an automatic phonological process; consequently, *[c]* cannot be simply analysed as an allophone of */k/*, but, even when it alternates with *[k]*, it might be analysed as the phoneme */c/*. If we accept this analysis, */c/* is less marginal in Sygva Mansi than in Sosva and Ob Mansi; still, its marginality remains, as it only occurs before */i(:)/*, although much more frequently.

6. PHONOLOGICAL ENVIRONMENT

In Livonian (at least in some variants of Livonian in the 20th century), the vowels */ɤ/* and */ɤ:/* emerge only in initial syllables following stem-initial labial consonants (Viitso 2008, 311). These sounds *[ɤ]* and *[ɤ:]* developed by delabialisation (and monophthongisation) from *[u̯o]* and *[u̯o:]*, and this delabialisation process is clearly connected to the labiality of the consonants preceding them. However, as *[o]* and *[o:]*, and *[u̯o]* etc. occur in similar positions (e.g. *[potʰa:]* ‘clay pot’, *[podʰ]* ‘porridge for children’, *[pu̯oŋa]* ‘son’; Ernštreits et al. 2024), *[ɤ]* and *[ɤ:]* must be considered the realisation of distinct phonemes; evidently, of */ɤ/* and */ɤ:/*.

Similarly, Standard Udmurt */w/* appears only in initial syllables following stem-initial */k/* and preceding */a/* at the same time (except for */kwin/*, where it is followed by */i/*) – cf. Csúcs (1990, 24); Csúcs (1998, 279); Csúcs (2005, 111); Kel'makov (1998, 79, 84); Bartens (2000, 51–52) etc.¹⁵ In most cases, the string */kwa/* goes back to **/ko/*, but stem-initial *[ko]* (additionally */ka/*, */ku/* etc.) also occurs.

The occurrence of marginal phonemes in a restricted set of environments can be explained by historical and phonetic reasons. On the one hand, these sounds could emerge as allophones in certain phonological contexts; therefore, they could become contrastive in that context. Contrastivity can emerge due to the inconsistency of the change: e.g. according to Csúcs (2005, 346–347, 351), some Proto-Permic words reconstructed with a stem-initial **/ko-/* did not go through the **/ko-/ > /kwa-/* change and begin with */ko-/* even today. Additionally, new morphemes containing the sequence eliminated by the change can emerge. The reason can be sound change (contemporary stem-initial */ko-/* can go back to several vowels following a stem initial */k-/*, cf. Csúcs 2005, 346–347) or borrowing (e.g. */koc/* < Russian */xotʲ/* ‘although; even if’).

Nonetheless, there are cases where some vowel qualities are extremely rare, but they occur in various phonological contexts, and they are not restricted to specific layers of the vocabulary. According to Idrisov (2012), in Beserman Udmurt, usually */ʌ/* corresponds to Standard Udmurt */i/*; still, */i/* emerges in various environments (although attested only in 17 words, including

¹⁵For a different analysis, see Fejes (2026).

Russian loanwords and onomatopoeic words). The only restriction seems to be that /i/ is only attested in initial syllables in non-onomatopoeic native words.

Similarly, /c/ is marginal in Sosva and Ob Mansi. E.g. in the texts published by Kálmán (1976, 104–155), it occurs in a handful of words, although in so frequent ones as /a:ci/ ‘no’, /a:ci:m/ ‘there is no’, /e:ci/ ‘at night’ /wa:ci/ ‘for a short time; not far’, /kocA/ ‘middle’, /nosa:ci/ ‘obviously’, /picmi/ ‘lip’, or the animal names /kaci/ ‘cat’ and /ku:cuw/ ‘puppy’ (cf. also Section 5.2 on Sygva Mansi). Although it is striking that /c/ occurs dominantly before /i/, it is not always the case. Additionally, it may also occur in Russian loanwords as /fakuʎcet/ ‘faculty’, /inscitur/ ‘college, institute’, /ka:cer/ ‘ship’, etc. or in proper names as /pe:ca/ ‘Pete’ or /cumeɲ/ ‘Tyumen’. However, the transcription may be misleading here. In reality, the realisations of palatal consonants in native Mansi words and palatalised consonants in Russian loanwords might be phonetically different or similar as well, especially due to the general bilingualism of Mansi speakers. If these sounds are realised differently, it would imply that they belong to separate phonemes. In that case, Russian loanwords would not contribute to the number of words with /c/ but /tʲ/. In the latter case, all palatalised consonants must be considered marginal, occurring only in loanwords.

7. THE EMERGENCE OF MARGINAL PHONEMES

Some remarks on the emergence of marginal phonemes have already been made above, but it may be worth giving an overview of the possibilities. Not surprisingly, the two main sources of marginal phonemes, as of any elements of language, are internal development and borrowing.

Among the cases of internal development, the general case seems to be the phonematisation of allophones. Finnish /ð/ (> /d/) emerged as an allophone of /t/ and phonematised when the environment that had triggered the different realisation changed. Similarly, Livonian /ʁ/ and /ʁ:/ must have emerged as the allophones of /u̯o/ and /u̯o:/, respectively. The case of Standard Udmurt /w/ is a bit more complicated. According to Csúcs (2005, 112), in Proto-Permic, it was also a marginal phoneme, which occurred only stem-initially, preceding /ɔ/. Only one word, **/kwac/* ‘6’, can be reconstructed with a /w/ following a word-initial /k/ (and preceding /a/). The number of words beginning with /kwa/ had grown by the Proto-Udmurt period, in which /w/ was still marginal and occurred only in these two positions. The stem-initial /w/ merged with /v/, but /w/ following /k/ remained in most of the Udmurt dialects, and this is also the case in Standard Udmurt. That is, /w/ first became less marginalised but then more marginalised.

Some marginal phonemes may occur in onomatopoeic words or interjections, as [b] in Hungarian, Erzya and Komi-Permyak (cf. Gálffy 2005, 85; Juhász 2012, 102; Rueter 2010, 61; Wichmann & Uotila 1942, XII, 362 – discussed in Section 1). However, it is questionable whether these can really be identified as phonemes, or should rather be analysed as emphatic-affective realisations of central phonemes. In any case, the phonematisation of a sound emerging this way is possible.

A central phoneme may also be marginalised. In all probability, this is the case in Beserman. In this dialect of Udmurt, /i/ has changed to /ʌ/. However, there are also exceptions: while /kʌz/ ‘spruce’ is attested instead of Common Udmurt /kiz/ ‘id.’, /kiz/ ‘thick’ corresponds to Common Udmurt /kiz/ ‘id.’. It is unlikely that Beserman borrowed /kiz/ ‘thick’ from other Udmurt dialects (Idrisov 2012, 595, cf. Usačëva et al. 2017, 165, 358). Among native words containing /i/, Idrisov

(2012, 355) also mentions /ti/ ‘lung’ (not included in Usačëva et al. 2017; Usačëva et al. 2017, 201, 225, 227, 240, 243; 350 also contains the native words, /li/ ‘bone, gristle’ (and compounds containing it) and /kiz/ ‘20’.¹⁶ In addition to these, only the onomatopoeic word /iʃ-iʃ/ ‘(indicating or warning about cold)’ (Usačëva et al. 2017, 32) and clear cases of recent Russian loanwords contain /i/. The most probable scenario is that /i/ was on the verge of extinction and occurred (maybe vacillating) just in a few words when its position was strengthened by loanwords. As a consequence, it was preserved in some native words but was changed in most of them.

Another main source of marginal phonemes is borrowing: Finnish /b, g/, Czech /o:/ and Latvian /o, o:/ can be the main examples. In these cases, new phonemes become established through loanwords. Interdialectal borrowing also belongs here. In such cases, the source of the loanwords with the new phonemes is not a distinct language but a dialect of the same language (keeping in mind that sometimes the two cases are not easy to distinguish). Erzya /ɲ/ can be a good example: the only argument for the phonemic status of /ɲ/ in Erzya is the existence of the word /ɲot/ ‘so’ (Rueter 2010, 58–59). Otherwise /ɲ/ is considered to be a phoneme in some dialects (cf. Keresztes 1990, 25; Bartens 1999, 27; Keresztes 2011, 18). The word /ɲot/ ‘so’ does not occur in the normative dictionaries of Erzya (e.g. Serebrennikov, Buzakova, and Mosin 1993), and it is used in the spoken language according to Rueter as well. Based on that, the question arises whether the word belongs to Standard Erzya at all – in any case, it seems that the word /ɲot/ ‘so’ is used in a wider range of language variants than other words with /ɲ/. It seems that the most probable scenario is that the word is borrowed from a dialect where /ɲ/ is established, and it has spread to other dialects, including spoken standard, as well. However, there seems to be no conclusive evidence for this development. It is also possible that /ɲot/ ‘so’ is a relic from the times when /ɲ/ was a phoneme in all variants of Erzya. In both cases, it is confusing that /ɲ/ is/was typical in word-internal and word-final positions, but not word-initially (cf. Keresztes 1990, 25; Bartens 1999, 27, 47–50; Keresztes 2011, 18, 40).

In some cases, the causes of the emergence of the marginal phoneme are complex. Finnish /d/ is a poster child of such cases. Although [ð] emerged due to internal phonetic changes, both its change into [d] in the variant the standard is based on and its survival in the standard (while it was lost or merged with other phonemes in the different varieties) are due to foreign influence. Additionally, the emergence of the voicing contrast between the plosives /t/ and /d/ could serve as a basis for the borrowing of /b/ and /g/.

8. SUMMARY AND CONCLUSION

Marginal phoneme is an umbrella term for extremely rare phonemes (occurring in just a handful of words, or in extreme cases, only in one word) or just rare but restricted to some groups of words (typically loanwords and/or onomatopoeic terms) or phonological environments. However, there is no clear dividing line between marginal phonemes and rare central phonemes.

Some marginal phonemes fit well in the phoneme system, while some others are outliers. Some marginal phonemes strongly vary in their pronunciation and may be realised identically to the default realisation of some other, central phoneme. Some speech sounds may usually occur

¹⁶In all probability, /bik/ ‘rafter, truss’ (Usačëva et al. 2017, 32) is related to Russian /bik/ [bik] ‘(bridge) pier’.

as the allophones of some central phoneme, while being contrastive only in some restricted contexts. Additionally, some speech sounds may occur in a restricted set of phonological contexts, and may be contrastive there: that is, these sounds must always be analysed as the realisations of a phoneme which occurs only in that restricted set of contexts. From a historical point of view, marginal phonemes can emerge both due to internal development or borrowing, or due to a complex interaction between these processes.

Even though the number of languages analysed from this point of view is very low from a typological perspective (even including languages not mentioned in this paper), it is clear that a phoneme can be marginal in very different ways. The relatively low number and the very different characteristics of these phonemes do not make it possible to make significant generalisations. Nonetheless, some remarks, which can be regarded as educated guesses rather than well-grounded conclusions, can be made. These can serve as hypotheses for further, more comprehensive research.

First of all, it seems that every language which has contact with other languages with different phoneme inventories¹⁷ is likely to have marginal phonemes which occur only in the borrowed elements of the lexicon. The closer the contact and the higher the level of bilingualism, the more likely that phonemes are borrowed, and additionally, that they will be stable (that is, consistently realised in a contrastive way, differently from the realisation of a central phoneme). It is expected that there can be differences between the groups of speakers of the languages: those who have more contact and are more bilingual will be more consistent in using the borrowed marginal phonemes. In all likelihood, the stability of a marginal phoneme also depends on how well it fits into the phoneme system: if it fills a gap, it is more likely to be accepted even by speakers who are not in contact with the donor language. In addition, in such circumstances, the borrowed phoneme is likely to be more stable. Based on the fact that languages are only rarely isolated, it can be assumed that every language has such marginal phonemes. However, there might be major differences between the variants of the language, depending on how much speakers are exposed to the other language. (Similarly, it can be supposed that many descriptions, concentrating on the “genuine” characteristics of languages or dialects, ignore these phenomena). A further problem is that these borrowed phonemes may become relatively frequent due to intensive borrowing, and then their marginal status might be questioned or unequivocally lost.

Additionally, it can be expected that languages have phonemes which occur in just a restricted set of phonological contexts for various reasons. One reason may be that the default realisation of these phonemes occurs only in that context. Another reason may be that the default realisation of that phoneme also occurs in other contexts, but it can rather be interpreted as an allophone of some other (central) phoneme. Both scenarios may result from historical changes, i.e. some phonemes emerge or remain only in some specific contexts. The problem is that the occurrence of several central phonemes is also restricted, and there are no clear quantitative criteria for demarcating the border between marginal and central phonemes.

¹⁷It is not clear whether the differences in the phoneme system or the speech sound inventory play a more significant role. E.g. if the contacting languages have no voiced obstruent phonemes, but in one of them, intervocalic obstruents are voiced, the other can borrow a word with an intervocalic voiced allophone and adopt it with a contrastive, that is, phonemic intervocalic obstruent. The possibility is given, although the question of whether such cases are attested remains open until some examples are found.

ACKNOWLEDGEMENTS

I would like to thank the participants of the workshop *Marginal phonemes* at the 57th Annual Meeting of the Societas Linguistica Europaea in Helsinki (21–24 August 2024) for their valuable comments. I am especially grateful to Natalia Kuznetsova and Fedor Rozhansky for their advice on the call, which was also used in this paper. I am similarly obliged to Jack Rueter, the co-organiser of the workshop, for our discussions on the subject of marginal phonemes. I am also thankful to Nóra Wenszky for her assistance in finding the correct English phrasing.

ABBREVIATIONS

1	first person
2	second person
3	third person
ADJZ	adjectiviser
ALL	allative
CONNNEG	connegative
GEN	genitive
IMP	imperative
INF	infinitive
NOM	nominative
PL	plural
POSS	possessive
PRS	present
SG	singular
SUP	superlative
TEMP	temporal

REFERENCES

- Abondolo, Daniel. 1998a. Finnish. In D. Abondolo (ed.) *The Uralic languages*. New York, NY: Routledge. 149–183.
- Abondolo, Daniel. 1998b. Hungarian. In D. Abondolo (ed.) *The Uralic languages*. New York, NY: Routledge. 428–456.
- Abondolo, Daniel (ed.). 1998c. *The Uralic languages*. New York, NY: Routledge.
- Abondolo, Daniel and Riitta-Liisa Valijärvi. 2023a. Introduction to the Uralic languages, with special reference to Finnish and Hungarian. In D. Abondolo and R.-L. Valijärvi (eds.) *The Uralic languages*. London & New York, NY: Routledge. 1–80.
- Abondolo, Daniel and Riitta-Liisa Valijärvi (eds.). 2023b. *The Uralic languages*. London & New York, NY: Routledge.
- Anderson, Gregory D. S. 2004. The languages of Central Siberia: Introduction and overview. In E. J. Vajda (ed.) *Languages & prehistory of Central Siberia*. Amsterdam: John Benjamins, 1–119.

- Ante, Aikio (Luobbal Sámmol Sámmol Ánte) and Jussi Ylikoski. 2022. North Saami. In M. Bakró-Nagy, J. Laakso and E. Skribnik (eds.) *The Oxford guide to the Uralic languages*. Oxford: Oxford University Press. 147–177.
- Ariste, Paul. 1968. *A grammar of the Votic language*. Bloomington & The Hague: Indiana University – Mouton & Co.
- Bakró-Nagy, Marianne, Johanna Laakso and Elena Skribnik (eds.). 2022. *The Oxford guide to the Uralic languages*. Oxford: Oxford University Press.
- Bals, Berit Anne, David Odden and Curt Rice, 2007. *The phonology of gradation in North Saami*. Downloaded 1. 11. 2024. <http://web.archive.org/web/20120916094952/http://www.ling.ohio-state.edu/~odden/SaamiGradation2007.pdf>.
- Bartens, Raija. 1999. *Mordvalaiskielten rakenne ja kehitys* [The structure and the history of the Mordvinic languages]. Helsinki: Suomalais-Ugrilainen Seura.
- Bartens, Raija. 2000. *Permiläisten kielten rakenne ja kehitys* [The structure and the history of the Permic languages]. Helsinki: Suomalais-Ugrilainen Seura.
- Bergmane, Anna, Rūdolfs Grabis, Milda Lepika and Ēvalds Sokols (eds.). 1959. *Mūsudienū latviešu literārās valodas gramatika I. Fonētika un morfoloģija* [A grammar of the contemporary standard Latvian language I. Phonetics and morphology]. Riga: Latvijas PSR Zinātņu Akadēmijas Izdevniecība.
- Bolla, Kálmán. 1985. *A finn beszédhangok atlasza* [An atlas of the Finnish sounds of speech]. Budapest: MTA Nyelvtudományi Intézet.
- Bye, Patrik, 2009. Three types of marginal contrast. Paper presented at the Toronto-Tromsø Phonology Workshop, October 9–11.
- Csúcs, Sándor. 1990. *Chrestomathia votiacica*. Budapest: Tankönyvkiadó.
- Csúcs, Sándor. 1998. Udmurt. In D. Abondolo (ed.) *The Uralic languages*. London & New York, NY: Routledge. 276–304.
- Csúcs, Sándor. 2005. *Die Rekonstruktion der permischen Grundsprache* [The reconstruction of the Permic proto-language]. Budapest: Akadémiai Kiadó.
- Ernštreits, Valts, Signis Vāvere, Tiit-Rein Viitso, Pētōr Damberg, Milda Kurpniece, Gunta K]ava, Uldis Balodis, Tuuli Tuisk, Gita Kūla, Marili Tomingas, Sven-Erik Soosaar, Anna Sedláčková and Toms Jurgenovskis, 2024. *Livonian language and culture resource platform “livonian.tech”*. <https://livonian.tech/>.
- Fejes, László. 2026. An unidentified marginal phoneme: The analysis of Udmurt [w]. *Acta Linguistica Academica* 73(2). 391–405. <https://doi.org/10.1556/2062.2025.01002>
- Fennel, Trevor G. and Henry Gelsen (eds.). 1980. *A grammar of Modern Latvian, Vol. I*. The Hague, Paris & New York, NY: Mouton Publishers.
- Grünthal, Riho. 2022. Veps. In M. Bakró-Nagy, J. Laakso and E. Skribnik (eds.) *The Oxford guide to the Uralic languages*. Oxford: Oxford University Press. 291–307.
- Grünthal, Riho. 2023. The Finnic languages. In D. Abondolo and R.-L. Valijärvi (eds.) *The Uralic languages*. London & New York, NY: Routledge. 433–480.
- Gussenhoven, Carlos and Haike Jacobs. 2017. *Understanding phonology*, 4th edn. London & New York, NY: Routledge.
- Gálffy, Mózes. 2005. A magyarói nyelvjárás mássalhangzó-rendszere [The consonant system of the Magyaró dialect]. In L. Murádin (ed.) *Magyarói nyelvjárási tanulmányok* [Studies on the Magyaró dialect]. Kolozsvár: Erdélyi Múzeum-Egyesület. 83–93.
- Hakulinen, Auli, Maria Vilku, Riitta Korhonen, Vesa Koivisto, Tarja Riitta Heinonen and Irja Alho. 2004. *Iso suomen kielioppi* [A comprehensive grammar of Finnish]. Helsinki: Suomalaisen Kirjallisuuden Seura.

- Hakulinen, Lauri. 1941. Suomen kielen rakenne ja kehitys. I. Ääne- ja muoto-oppia. [The structure and the history of the Finnish language. I. Phonology and morphology]. Helsinki: Otava.
- Hall, Kathleen Currie. 2013. A typology of intermediate phonological relationships. *The Linguistic Review* 30(2). 215–275.
- Hall, Daniel Currie and Kathleen Currie Hall. 2016. Marginal contrasts and the contrastivist hypothesis. *Glossa* 1(1). 1–23.
- Hudson, Joyce and Eirlys Richards. 1993. Walmajarri dictionary. The source is not available anymore on the url given by Macklin-Cordes and Round (aseda.aiatsis.gov.au). However, a similar material is available at ausil.org.au/Dictionary/Walmajarri/Index-en.htm. The relationship between the two dictionaries is unclear as no source dated from 1993 is referenced on the current page.
- Häkkinen, Kaisa. 1983. Suomen kielen äännerakenteen ominaispiirteistä [On some characteristics of the sound structure of Finnish]. In A. Hakulinen and P. Leino (eds.) *Nyky-suomen rakenne ja kehitys 1*. [The structure and the history of Modern Finnish]. Pieksämäki: Suomalaisen Kirjallisuuden Seura. 38–56.
- Idrisov, Rustem I. 2012. Vokalizm besermjanskogo dialekta udmurtskogo jazyka na materiale govora derevni Šamardan Jukamenskogo rajona udmurtii [The vowel system of the Beserman dialect of Udmurt based on the subdialect of the village of Shamardan in the Yukamensk Rayon of Udmurtia]. In A. I. Kuznecova (ed.) *Finno-ugorskie jazyki: fragmenty grammatičeskogo opisanija. Formal'nyj i funkcional'nyj podhody*. [Finno-ugric languages: fragments of grammatical description. Formal and functional approaches]. Moskva: Jazyki slavjanskix kul'tur. 591–607.
- Janhunen, Juha. 1998. Samoyedic. In D. Abondolo (ed.) *The Uralic languages*. New York, NY: Routledge. 457–479.
- Juhász, Dezső. 2012. The types and main characteristics of the Hungarian dialects of Moldavia. In *Language use, attitudes, strategies. Linguistic identity and ethnicity in the Moldavian Csángó villages*. Cluj-Napoca: Editura Institutului pentru Studierea Problemelor Minorităților Naționale. 97–110.
- Karlsson, Fred. 1983. Suomen yleiskielen segmentaalifoneemien paradigma [The paradigm of segmental phonemes in Colloquial Finnish]. In A. Hakulinen and P. Leino (eds.) *Nyky-suomen rakenne ja kehitys 1*. [The structure and the history of Modern Finnish]. Pieksämäki: Suomalaisen Kirjallisuuden Seura. 28–38.
- Karttunen, Kaarina. 1979. *Nykyslangin sanakirja* [A dictionary of contemporary slang]. Porvoo, Helsinki & Juva: Werner Söderström.
- Katamba, Francis. 1989. *An introduction to phonology*. London & New York: Longman.
- Kel'makov, V. K. 1998. *Kratkij kurs udmurtskoj dialektologii* [A short course of Udmurt dialectology]. Iževsk: Izdatel'stvo Udmurtskogo universiteta.
- Keresztes, László. 1990. *Chrestomathia morduinica*. Budapest: Tankönyvkiadó.
- Keresztes, László. 1993. Új magánhangzó-fonémák a magyarban? [New vowel phonemes in Hungarian?]. *Hungarologische Beiträge* 1. 45–51.
- Keresztes, László. 2011. *Bevezetés a mordvin nyelvészetbe*. Debrecen: Debreceni Egyetemi Kiadó.
- Kiparsky, Paul. 2018. Formal and empirical issues in phonological theory. In L. M. Hyman and F. Plank (eds.) *Phonological typology*. Berlin & Boston: De Gruyter – Mouton. 54–106.
- Kukhto, A. V. 2018. Fonologija [phonology]. In S. Y. Toldova and M. A. Xolodilova (eds.) *Èlementy mokšanskogo jazyka v tipologičeskom osveščanii* [Elements of the Moksha language in a typological perspective]. Moskva: Buki Vedi. 19–37.
- Kálmán, Béla. 1976. *Wogulische texte mit einem Glossar* [Mansi Texts with a Glossary]. Porvoo, Helsinki & Juva: Werner Söderström.

- Laakso, Johanna. 2022. Finnish, Mäenkieli, and Kven. In M. Bakró-Nagy, J. Laakso and E. Skribnik (eds.) *The Oxford guide to the Uralic languages*. Oxford: Oxford University Press. 254–268.
- Lauerma, Petri. 1993. *Vatjan vokaalisointu* [vowel harmony in votic]. Helsinki: Suomalais-Ugrilainen Seura.
- Macklin-Cordes, Jayden L. and Erich R. Round. 2020. Re-evaluating phoneme frequencies. *Frontiers in Psychology* 11. <https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2020.570895>.
- Martindale, Colin, Sabir Medzhidovich Gusein-Zade, Dean McKenzie and Mark Yu Borodovsky. 1996. Comparison of equations describing the ranked frequency distributions of graphemes and phonemes. *Journal of Quantitative Linguistics* 3. 106–112.
- Mathiassen, Teje. [1997]. *A short grammar of Latvian*. [Oslo]: Slavica Publishers.
- Matisoff, James A. 1994. Tone, intonation, and sound symbolism in Lahu: Loading the syllable canon. In L. Hinton, J. Nichols and J. J. Ohala (eds.) *Sound symbolism*. Cambridge: Cambridge University Press. 115–129.
- Nádasdy, Ádám and Péter Siptár. 1989. Issues in Hungarian phonology: Preliminary queries to a new project. *Acta Linguistica Hungarica* 39(1–4). 3–27.
- Nuutinen, Olli. 1994. Vinskaa, venskaa, tanskaa ja ranskaa [Vinnish, vedish, tanish and Rench]. In *Hetkisen pituus ja muita kirjoituksia kielestä* [The length of the moment and other papers on language.], Vol. 128 of *Tietolipas*. Helsinki: Suomalaisen Kirjallisuuden Seura. 64–79.
- Petr, Jan (ed.). 1986. *Mluvnice češtiny (1) Fonetika. Fonologie. Morfonologie a morfemika. Tvoření slov*. [A grammar of Czech (1) Phonetics. Phonology. Morphology and morphemics. Derivation]. Praha: Academia.
- Prāuliņš, Dace. 2012. *Latvian. An essential grammar*. New York: Routledge.
- Rueter, Jack. 2010. Adnominal person in the morphological system of Erzya, Vol. 261 of *Suomalais-ugrilaisen Seuran Toimituksia*. Helsinki: Suomalais-Ugrilainen Seura.
- Salminen, Tapani. 2007. Notes on forest Nenets phonology. *Suomalais-Ugrilaisen Seuran Toimituksia / Mémoires de la Société Finno-Ougrienne* 253. 349–372.
- Sammallahti, Pekka. 1988. Historical phonology of the Uralic languages. In D. Sinor (ed.) *The Uralic languages. Description, history and foreign influences*. Leiden: E. J. Brill. 478–554.
- Schön, Zsófia and Katalin Gugán. 2022. East Khanty. In M. Bakró-Nagy, J. Laakso and E. Skribnik (eds.) *The Oxford guide to the Uralic languages*. Oxford: Oxford University Press. 608–635.
- Serebrennikov, B. A., R. N. Buzakova and M. V. Mosin. 1993. *Erzjan-ruzon' valks* [Erzya–Russian dictionary]. Moskva: «Russkij jazyk».
- Setälä, Emil Nestor. 1902². *Suomen kielioppi. Ääne- ja sanaoppi. Oppikoulua ja omin päin opiskelua varten* [Finnish grammar. Phonology and morphology. For upper grades and independent study.]. Helsinki: Otava.
- Sinor, Denis (ed.). 1988. *The Uralic languages. Description, history and foreign influences*. Leiden: E. J. Brill.
- Suomi, Kari, Juhani Toivanen and Riikka Ylitalo. 2008. *Finnish sound structure. Phonetics, phonology, phonotactics and prosody*. Oulu: University of Oulu.
- Sussex, Roland and Paul Cuddeback. 2006. *The Slavic languages*. Cambridge: Cambridge University Press.
- Szathmáry, István. 1988. La langue hongroise [The Hungarian language]. In D. Sinor (ed.) *The Uralic languages. Description, history and foreign influences*. Leiden: E. J. Brill. 478–554.
- Tamborisev, Yuri and Colin Martindale. 2007. Phoneme frequencies follow a Yule distribution. *SKASE Journal of Theoretical Linguistics* 4. 1–11. http://www.skase.sk/Volumes/JTL09/pdf_doc/1.pdf.

- Usačëva, Marija Nikolaevna, Timofej Aleksandrovič Arxangel'skij, Ol'ga Leonidovna Birjuk, Vladimir Andreevič Ivanov and Ruslan Irekovič Idrisov. 2017. Tezaurus besermjanskogo narečija: Imena i služebnye časti reči (govor derevni Šamardan) [Thesaurus of the Beserman dialect: Nominals and function words (the subdialect of the village of Shamardan)]. Moskva: Izdatel'skie rešenija. https://iling-ran.ru/usacheva/Usacheva_et_al_2017.pdf.
- Viitso, Tiit-Rein. 2008. Liivi keel ja läänemeresoome keelemaastikud [Livonian language and Finnic linguistic landscapes]. Tartu – Tallinn: Eesti Keele Sihtasutus. <http://hdl.handle.net/10062/16987>.
- Wichmann, Yrjö and Toivo Eemil Uotila. 1942. Syrjänischer Wortschatz nebst Hauptzügen der formenlehre [Komi vocabulary with the principal characteristics of morphology]. Helsinki: Suomalais-Ugri-lainen Seura.
- Ylikoski, Jussi. 2022. Lule Saami. In M. Bakró-Nagy, J. Laakso and E. Skribnik (eds.) *The Oxford guide to the Uralic languages*. Oxford: Oxford University Press. 130–146.
- Zaicz, Gábor. 1998. Mordva. In D. Abondolo (ed.) *The Uralic languages*. New York: Routledge. 184–218.