University of Groningen

Phonetic Realisation and Phonemic Categorisation of the Final Reduced Corner Vowels in the Finnic Languages of Ingria

Kuznetsova, Natalia; Verkhodanova, Vasilisa

Published in:
Phonetica

DOI:
[10.1159/000494927](https://doi.org/10.1159/000494927)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2019

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Kuznetsova, N., & Verkhodanova, V. (2019). Phonetic Realisation and Phonemic Categorisation of the Final Reduced Corner Vowels in the Finnic Languages of Ingria. *Phonetica*, 76(2-3), 201-233.
<https://doi.org/10.1159/000494927>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Phonetic Realisation and Phonemic Categorisation of the Final Reduced Corner Vowels in the Finnic Languages of Ingria

Natalia Kuznetsova^{a, b} Vasilisa Verkhodanova^c

^aInstitute for Linguistic Studies, Department of the Languages of Russia, Russian Academy of Sciences, St. Petersburg, Russia; ^bDipartimento di Lingue e Letterature straniere e Culture moderne, Università degli Studi di Torino, Turin, Italy; ^cRijksuniversiteit Groningen, Campus Fryslân, Leeuwarden, The Netherlands

Abstract

Individual variability in sound change was explored at three stages of final vowel reduction and loss in the endangered Finnic varieties of Ingria (subdialects of Ingrian, Votic and Ingrian Finnish). The correlation between the realisation of reduced vowels and their phonemic categorisation by speakers was studied. The correlated results showed that if V was pronounced >70%, its starting loss was not yet perceived, apart from certain frequent elements, but after >70% loss, V was not perceived any more. A split of 50/50 between V and loss in production correlated with the same split in categorisation. At the beginning of a sound change, production is, therefore, more innovative, but after reanalysis, categorisation becomes more innovative and leads the change. The vowel *a* was the most innovative in terms of loss, *u/o* were the most conservative, and *i* was in the middle, while consonantal palatalisation was more salient than labialisation. These differences are based on acoustics, articulation and perception.

© 2019 S. Karger AG, Basel

1. Introduction

1.1 Synchronic Variation and Sound Change

Systematic individual variation in speech perception and production produces a pool of variation which becomes the source of language change (Kruszewski, 1883; Baudouin de Courtenay, 1895; Ohala, 1989; Labov, 1999; Baker et

al., 2011; Yu, 2013; Stevens & Harrington, 2014; Bybee, 2015). Language change is propagated through the repeated exposure of several generations of speakers to a gradually changing variable pool of realisations. Learning theories, placed on a continuum between rational Bayesian approaches and associationist models inspired by biological discoveries, are now at the core of psychophysical sound change models.

Associative learning in phonology implies constant bidirectional updating of the connection weights in mappings between acoustic cues and phonological/subphonemic categories. Learning is distributional in that the learner acquires knowledge of the frequency distribution of various phonetic stimuli and builds a mental phonological model of the language on this. Frequency distribution is even suggested to be a more important factor in the formation of phonemic categories than minimal pairs (Maye & Gerken, 2000; Vallabha et al., 2007; Wanrooij et al., 2013; Olejarczuk et al., 2018). This approach also explains the puzzling cases of near-mergers, when speakers already categorise items into the same phonemic class when there is still a phonetic difference in the realisation of two former classes (Labov et al., 1991; Barnes, 2006; Roettger et al., 2014). The exact structure of such mental constructions is still, however, under debate between prototype, exemplar and other models (Mompeán-Gonzalez, 2004; Gureckis & Goldstone, 2008; Johnson, 2015; Davis & Poldrack, 2014; Kapatsinski, 2018).

Sound change, as any language change, follows the S-curve path, where a weighting jumps to a different value at some point during the change (Hyman, 1976; Kirby, 2010, p. 148; Blythe & Croft, 2012, p. 293). Its actuation is discussed (Baker et al., 2011; Stevens & Harrington, 2014; Priva, 2017), as well as the exact mechanism of the jump between values. The latter might be linked to properties of the articulatory/acoustic relation, when the manipulation of an articulator can result in non-monotonic variability in an acoustic parameter (Stevens, 2004). The usage-based approaches have also followed the hypothesis that lexical frequency prompts sound change. The latter starts from frequent words and morphemes due to a higher level of automatisisation in their production and can later spread throughout the entire phonological system (van Bergem, 1995; Bybee, 2001, pp. 11–12; Bybee et al., 2016; Hay & Foulkes, 2016; Hall et al., 2018; Kapatsinski, 2018).

Sound change implies two connected processes: (1) a change in the structure of the pool of phonetic realisations and (2) a categorical reanalysis in the mind of the speakers/listeners. The temporal and causal correlations between these two processes are still unclear (Bybee, 2001, p. 55). Modern phonology still has to reconcile the data on the continuous and variable nature of the phonetic signal and on the behaviour of symbolic processes in a consistent fashion (Barnes, 2006, p. 222; Kirby, 2010, p. 149). In the associative learning framework, the same question concerns the relations between typicality distributions in perception and frequency distributions in production (Kapatsinski, 2018, p. 275). The concept of attractor landscape used in non-linear dynamic systems might be of use in modelling this link between continuous and categorical variation. A dynamic system is continuous, but there are specific stable states (attractors) it moves to (Ritter et al., 2018). A change in the weighting of attractors can model the change in the frequency distributions of different realisations throughout the sound change.

The present paper explores the correlation between production and mental representation in a case study on vowel reduction and loss in several minor Finnic varieties.

1.2 Vowel Reduction: General and Particular Mechanisms

Vowel reduction and loss is a cross-linguistically frequent phenomenon. However, studies taking a cross-linguistic and general theoretical approach to it are still scarce. Lindblom (1963) suggested that vowel reduction occurs through the mechanism of formant undershoot, which is a function of decrease in vowel duration. This view was supported by Delattre (1969), Flemming (1995, 2004), Kirchner (1998) and Barnes (2006), although the causal relation between undershoot and duration was reversed by Crosswhite (2004). The matter of reduction is discussed in a number of functionalist works, where the language system is represented as a trade-off between the needs of the speaker to economise the effort and the listener to be able to decipher the message. Lindblom (1990), for example, proposed an H&H framework, where a message varies in articulatory clarity as a compromise between hypospeech, minimising articulatory effort, and hyperspeech, maximising discriminability. Reduction characterises pieces of speech with low informativity and is a manifestation of hypospeech, a “part of planned speech behaviour rather than an accidental by-product of vocal organ inertia” (Harris, 2005, p. 132).

Reduction does not affect all vowel qualities or positions in a word or phrase equally, nor does it work always in the same direction. For example, a word-final and especially a phrase-final position manifests both vowel strengthening (lengthening and strengthening of articulation) and vowel weakening (devoicing, laryngealisation, nasalisation and loss). Barnes (2006) explains the weakening effects by the perceptual weakness of final vowels, in spite of their possible articulatory strength. Vowel reduction could also have different underlying mechanisms. Kapatsinski (2018, p. 286) opposes phonetically gradual reduction produced by automatised execution in production to phonetically abrupt loss of low-salience parts left meaningless by overshadowing in perception.

Two general paths of vowel reduction are distinguished: centripetal (centralisation towards schwa) and centrifugal (dispersion towards the three corner vowels *a*, *i*, and *u*). The corner vowels are generally known to be special in various respects: the most stable and focalised, perceptually salient, the easiest for neural processing because of maximal distinction, etc. (Crosswhite, 2004; Polka & Bohn, 2003, 2011; Harris, 2005; Johnson, 2015; Manca & Grimaldi, 2016). Data on acoustic, perceptual and other differences within the corner vowels are, however, scarce. The typological studies on vowel reduction show that vowel height is affected before frontness/backness, rounding or advanced tongue root contrasts (Barnes, 2006; Flemming, 2004). Reduced speech is characterised by compression of the acoustic space between F1 and F2 through F1 raising, an effect of lessened jaw opening (Lindblom, 1963; Uchanski, 2005). The bottom-up direction of the compression suggests that high unstressed vowels would be less marked than non-high ones (Walker, 2011, p. 29). The latter require more jaw opening and a longer time to be realised. At the same time, reduction-based sonority scales presume that the vowel *a* is less marked, but that schwa is more marked than *i* and *u* (Crosswhite, 2004, p. 209; de Lacy, 2006, p. 286).

The existence of differences between *i* and *u* is not much discussed in the surveys on vowel reduction. Some argue for a role of F2-based harmony in blocking the reduction of front vowels (Pearce, 2008; Szeredi, 2010). Evidence for the disparity between *i* and *u* comes also from research on vowel perception and neuroimaging, where place of articulation and tongue height are seen as relatively simple features. They directly correspond to F1 and F2 values, which, in turn, find their straight correlates in regions and types of brain activity. The rounding feature appears more complex, as it requires higher-level information processing and is acoustically less reliable and perceived with significant help from the visual channel (Traunmüller & Öhrström, 2007; Eulitz & Obleser, 2007; Vatakis et al., 2012; Manca & Grimaldi, 2016). One might thus suggest that *u* is less perceptually robust and salient than *i* and, therefore, more prompt for reduction, especially in languages with fronting vowel harmony.

Our study offers further experimental data to explore the general mechanisms of reduction and loss, as well as vowel markedness hierarchies at different subsequent stages of reduction.

2. Methods

2.1 Aims, Data and Background of the Study

Correlations between the frequencies of various realisations of the three corner vowels in production and mental categorisation were explored in a comparative phonetic field study (2014–2016) on final vowel reduction and loss. We looked at three Finnic languages of the Lower Luga area in the west of historical Ingria (currently, the vicinity of St. Petersburg in Russia): Ingrian, Votic and Finnish (see maps in Kuznetsova et al., 2015, p. 128; Kuznetsova & Sidorkevič, 2012, p. 565). They have been in close contact for centuries and formed a Lower Luga Sprachbund (Muslimov, 2005). Besides, a group of Ingrian and Finnish speakers was expelled from this region to Western Siberia in 1803–1804 after a strike against Baron von Ungern-Sternberg. A contact Siberian Ingrian/Finnish language developed there in isolation from its sister varieties (Nirvi, 1972; Sidorkevič, 2013).

The process of reduction advances through several stages, still observed in the living varieties of these languages. The following varieties were chosen for this study: (1) the Kurkola Ingrian Finnish dialect (IF); (2) the Luutsa dialect of Votic (V); (3) the Central (CI) and (4) the Southern (SI1, SI2) varieties of the Lower Luga dialect of Ingrian; and (5) Siberian Ingrian/Finnish (S). The data were obtained from 1 speaker per variety, with the exception of South Lower Luga Ingrian, for which 2 speakers were recorded (Table 1).

This is a limitation of our study, precipitated by the limited availability of fluent speakers able to participate in such experiments, as individual speakers even of the same language may display different reduction behaviour (Hanique et al., 2015). General reduction patterns in Lower Luga and adjacent areas had, however, been established prior to this experiment by Kuznetsova (2009, 2012a, 2016) on the basis of existing published sources as well as audio data on several dozens of speakers. It was observed that the degree of reduction increases from the north to the south of the Lower Luga area towards the Estonian language, which has been the most innovative in this respect and has completely lost reduced vowels. For example, the abovementioned varieties were graded from least to most susceptible to reduction in the following way: Kurkola Ingrian Finnish > Votic and Central Lower Luga Ingrian > South Lower Luga Ingrian > Siberian Ingrian/Finnish. Observed processes include qualitative and quantitative reduction, devoicing and speech elision, e.g., *püssü* ['pys:y] > ['pys:ÿ] > ['pys:ÿ̥] > [pys^{jw}:] > [pys^h:] > [pys:] “rifle.” While vowels still preserve their segmental status in the Lower Luga area, they turned into the consonantal features of labialisation and palatalisation in Siberian Ingrian/Finnish (Sidorkevič, 2013).

Table 1. Sociolinguistic data on the speakers

Variety code	Language	Dialect	Subdialect	Sex	Year of birth	Birth place	Recording place
IF	Finnish	South-Western	Kurkola Ingrian Finnish	F	1933	Hakaja	Sutela
V	Votic	Western	Luutsa	F	1928	Liivčülä	Liivčülä
CI	Ingrian	Lower Luga	Central	F	1927	Ropsu	Ropsu
SI1			South	M	1924	Vanakülä	Vanakülä
SI2				F	1932	Dal'n'aja Pol'ana	Dal'n'aja Pol'ana/ Narva
S	Ingrian/ Finnish (mixed)	Siberian variety: South Lower Luga Ingrian/ Rosona Ingrian Finnish		F	1950	Ryžkovo (Omsk region)	Tallinn

The role of the lexical and grammatical factor at the initial stages of reduction, predicted by the usage-based approaches (see 1.1), has also previously been noticed in the Finnic languages of Ingria and South-Western Finnish dialects. Grammatically conditioned vowel elision is claimed to have emerged earlier than purely phonetically conditioned elision (Laanest, 1980, pp. 73–74), and it occurred first of all in the most frequently used morphemes (Leskinen, 1973, p. 218). Specifically, lexically and grammatically conditioned vowel reduction has been attested in phonologically more archaic Ingrian Finnish and Soikkola Ingrian, while in more innovative Lower Luga Ingrian the conditioning is generalised to purely phonetic (Kuznetsova, 2016, pp. 9–11).

The present experimental study was designed following the patterns established in the abovementioned works and aimed at further clarifying the results obtained mainly from auditory impressions. All these languages share the same type and drift of reduction and differ just by its degree. Therefore, in this case it is possible to transpose this geographic variability along the north-south axis into the reduction progress along the time axis. All four languages are severely endangered: the number of speakers ranges from fewer than 10 to a couple of hundred (Kuznetsova et al., 2015; Sidorkevič, 2013). Therefore, the observed differences in production and categorisation of reduced vowels can hardly be attributed to the very fact of their endangerment.

The vowel inventories of these varieties contain low, mid and high vowels, front and back vowels, and labialised and non-labialised vowels. The systems in their most archaic variant in terms of non-initial vowel reduction, which can serve as a reference point for the processes described in the study, can be summarised as follows: *i:ĩ, ü:ũ, u:ũ, e:ẽ/ie, õ:õ/üõ, õ:õ, o:õ/uo, ä:ä and a:ā*. The unrounded back vowel *õ:õ* is present only in Votic. The languages are characterised by significant prosodic differences between initial (stressed) and non-initial (unstressed) syllables. In certain varieties, the long initial mid vowels were raised to the diphthongs *ie, üõ* and *uo*. Stems in all varieties are characterised by fronting vowel harmony within the domain of the root plus the following derivative and inflectional suffixes, as in Standard Finnish (*a, o* and *u* can occur in back-vowel stems, *ä, ö* and *ü* in front-vowel ones, and the “neutral” vowels *i* and *e* in both; for irregularities in Votic see, e.g., van der Hulst, 2018, pp. 176–178).

In general, short vowels can undergo reduction and loss at the end of a final or a non-final foot and in the second syllable of a trisyllabic foot (cf. examples from South Lower Luga Ingrian: *tüttõ* [tyt:õ] “girl,” *avahtü* [ava:htü] “be_opened:PST:3SG,” *sukküris* [suk:ÿ,r’is:] “sugar:IN,” *lisähümine* [lisä:h’ÿ,mine] “be_added:NMLZ,” *ihmised* [ihm’ised] “man:PL,” *jähüteä* [jæ:h’ÿte,tä] “cool:IPS,” *kiskohumine* [kiskõhu,mine] “be_torn_off:NMLZ”; for a detailed chart of reduction and loss positions see Kuznetsova, 2011, p. 189; 2012a, pp. 59–60). Reduction in non-initial long vowels is outside the scope of this study; for a general account see Kuznetsova (2016).

2.2 Methods of Data Collection and Analysis

In the phonetic experiment, open disyllables ending in the three corner types of vowel *a*, *i* and *u* (or *o*) after both voiced (*n*, *l*, *r*, *m* or *v*) and voiceless (*t*, *k*, *p*, *s* or *h*) singleton consonants were studied in the phrase-initial and the phrase-final position (3 vowels × 2 consonants × 13 words × 4 iterations × 2 positions = 624 tokens per sample). Most types of word-final combination of these vowels with consonants were covered. Based on existing phonological descriptions of these languages (Leppik, 1975; Kuznetsova, 2009; Markus & Rožanskij, 2011), one can argue that at least consonantal palatalisation might be stronger in front-vowel stems, and that geminates could be affected by it less than singletons. We, therefore, limited ourselves to singleton consonants and to stems with back and neutral vowels. Chosen stems were mostly morphophonologically back (the few front-vowel stems, which contained only the neutral vowels *i* and *e*, are underlined in Appendix 1).

The questionnaires were nearly identical (~5% variability) for all varieties, which share a substantial part of the lexicon. Words ending in *o* were taken instead of those with *u* in about one-third of cases. First, rounded vowels are much rarer in non-initial syllables than unrounded ones. Due to the endangered state of the varieties, it proved in some cases impossible to find words ending in the required combinations of *u* and a consonant that would be familiar to the speakers. Second, in the process of vowel reduction and loss in these varieties, the mid vowels *o*, *õ* and *e* are raised to *u*, *ü* and *i* (Mägiste, 1925, p. 80; Kuznetsova, 2012a, 2012b, 2016; see, e.g., *maito/maitu* “milk,” *pudro/pudru* “porridge” and *viero/vieru* “wheel” in Appendix 1). Third, the loss of both *o* and *u* results in consonantal labialisation; thus, from this point of view, they are functionally similar.

The two phrasal positions were thought to be prosodically different enough to attest a wide range of phonetic variability in vowel realisations. Words in the phrase-initial position were pronounced in the context before the consonant *s*. The most typical position for complete vowel loss in these varieties is in sandhi before a following vowel. A position before a consonant was chosen because it allowed for subtler differences in the process of loss of different vowel qualities to be better identified. In the prevocalic context, where all vowels are nearly invariably lost in fast speech, these differences are neutralised. The data were recorded with a Zoom H4n digital recorder in the field, segmented and classified in Praat, and analysed in SPSS 11.5.0. The speakers had to translate the Finnic sentences with a carrier word from a phrase asked in Russian and repeat them 4 times. We subsequently counted the ratios of various types of vowel reflex within each pool of realisations along several scales. The most general binary scale included two main types: (1) “vowel” and (2) “loss.” These were further divided into six subtypes in the following way:

“Vowel” (= “vocalic” realisations): (1) modal, (2) partially devoiced or (3) fully devoiced vowels

“Loss” (= “consonantal” realisations): (4) heavy segmental aspiration (>30–35 ms) after the consonant, (5) palatalised or labialised consonant or (6) complete vowel loss without any traces

For the first three speakers (IF, V and CI), a more detailed scale of variability in production was also used. These speakers showed low ratios of vowel loss, but still significant reduction in vowel quality. Their “vocalic” types were additionally assessed according to three independent scales:

Presence of strong laryngealisation: (1) yes or (2) no

Vowel quality: (1) full, (2) partially reduced or (3) completely reduced (to schwa)
 Devoicing: (1) modal; (2) slightly aspirated vowels; vowels with (3) 10–30%, (4) 30–50%, (5) 50–70% or (6) 70–90% of devoicing; or (7) full devoicing

The devoicing scale is a more detailed variant of the six-type scale: type 1 of the latter includes devoicing types 1–3; type 2 includes devoicing types 4 and 5; and type 3 includes devoicing types 6 and 7.

Vowel reflexes were classified manually on the basis of the spectrographic data. Examples of main types of realisation are given in Figures 1 and 14–16. The “vocalic” types still preserved F1 and F2. If just one formant was present, the case was considered as “heavy segmental aspiration.” Being shorter than 30–35 ms, such aspiration was seen as a consonantal feature of palatalisation or labialisation.

In a separate session from the phonetic experiment, we ran a parallel psycholinguistic test on how speakers categorise the reduced vowel reflexes. They were asked to write down in any preferred orthography the carrier words from the phonetic questionnaire the way they perceived them (~78 words). The task was formulated in Russian as: “Please write down a word for ‘bird’ in your language whatever way you prefer.” If speakers noted that they did not know how to write in their language, the researcher emphasised her interest in the way how a person “feels” the word, not in the right orthography.

Three speakers used Cyrillic and three others (IF, SI1 and S) Latin letters; speakers SI1 and S, though, also sporadically used Cyrillic letters. For example, **lintu* “bird” could be typically written as *линт*у/*lintu* or *линт*/*lint*. We did not give a multiple-choice task to the speakers, so as not to attract their attention to the final vowels. However, if speakers spontaneously noted that a word could be pronounced both with a vowel and without it, we counted these cases as two separate tokens. We counted the ratios of final vowel presence to loss for each speaker (sizes of the samples: IF = 78, V = 81, CI = 76, SI1 = 78, SI2 = 81 and S = 85).

Neither variety has a literary standard, so such a test provided a unique opportunity to observe more or less directly speakers’ intuitions about the presence/absence of a vowel word-finally. At the same time, a classic perception test was not possible in those field conditions, given the advanced age and fragile health conditions of the subjects. The Russian language and the Finnic varieties belong to different families (Indo-European vs. Altaic), so the Russian tokens for carrier words were not expected to significantly influence the outcome of the test. Moreover, both the Cyrillic and the Latin mediating orthographies rely on the phonemic principle of encoding, and so they automatically prompted subjects to reflect in writing whether there was any vowel word-finally or not.

In some cases, a more detailed scale was used for this categorisation test:

“Vowel”: (1) full vowels or (2) reduced vowels

“Loss”: (3) retention of consonantal palatalisation or labialisation, or (4) zero

Palatalisation was coded by speakers with the use of the Russian “soft sign” “ь.” The results of the categorisation of palatalisation and labialisation should be considered tentative, as the Russian orthography does not have a corresponding sign for labialisation. The latter was depicted only by the Siberian speaker as (*o*) or (*u*) in parentheses after the consonant, while she explicitly claimed the absence of final vowels. The observed asymmetry in the depiction of the two features might be partially influenced by this orthographic disparity. Reduced vowels were rendered by some by means of the Russian “ы” (high unrounded mid vowel).

Fig. 1. Examples of reduction in vocalic quality. **a** Full modal V (**hāpa* [ha:pa] “aspen”) MO. **b** Laryngealised V (**velka* [veɫkə] “debt”) MO. **c** Schwa (**āpa* [a:pə] “aspen”) PK. **d** Slightly aspirated V (**kumpa* [kumpə^h] “which of the two”) MO. **e** Partially devoiced V (**āpa* [a:p̥] “aspen”) PK. **f** Devoiced V (**āra* [a:r̥] “branch”) PK.

(For figure see next pages.)

1

3. Results and Discussion

The results show a robust correlation between production and categorisation. The general structure of the category prototypes (Rosch, 1978) was the same in phonetic realisation and phonemic representation at each of the three stages of vowel loss observed (see Fig. 11). At stage 1 (IF, V and CI), “vocalic” realisations composed more than 90% of the sample, which correlated with their only one robust mental prototype [+SEGMENT]. At stage 2 (SI1 and SI2), there was a roughly 50/50 split both between “vocalic” and “consonantal” realisations, on the one hand, and [+SEGMENT] and [−SEGMENT] categorisations, on the other hand. At stage 3 (S), with vowel loss in >70% of the cases, only one category prototype [−SEGMENT] prevailed. These results clarify the findings by Kuznetsova (2016), where less phonetic reduction was expected for Kurkola Ingrian Finnish and more for Central Lower Luga Ingrian, respectively.

Below, we address individual features of production and categorisation and summarise the tendencies at each stage. In the General Discussion (section 4), we outline main trends in the loss of vowel quality and main differences between the six speakers and the three vowel types. Phonetic differences across positions, consonantal types and individual words and nuances concerning vowel duration largely remain outside the scope of this paper. The differences across phrasal positions and after voiced versus voiceless consonants were indeed noticeable in terms of the percentage of vowel loss, duration, and quality. Vowels expectedly manifested much more devoicing after voiceless consonants. Initial phrasal position was primarily characterised by strong qualitative reduction (apparently triggered by extremely reduced duration), while final position exhibited more devoicing. Vowel duration divided speakers into two groups: (1) stage 1 – short vowels (90–100 ms) phrase-finally and reduced vowels (<80 ms) phrase-initially, and (2) stages 2 and 3 – reduced vowels (<80 ms) in both phrasal positions.

3.1 Stage 1: Ingrian Finnish, Votic and Central Lower Luga Ingrian

The samples at stage 1 of reduction belong to three different languages: Finnish, Votic and Ingrian. Even if similar in the general structures of distribution, they exhibit slightly different configurations of vowel loss in realisation and categorisation (see the percentages of loss in Fig. 5–7). Ingrian Finnish represents the most conservative variety, and Central Lower Luga Ingrian the most innovative one, with Votic in the middle. In all three samples, the vowel **u*¹ reveals exactly the same pattern, being the most conservative of all the vowels: full preservation in mental categorisation and just 1–2% loss in production. What differs across the varieties in question is the configuration of the vowels **a* and **i*. In Ingrian Finnish, **i* is the most innovative in terms of both production and categorisation, while in Central Lower Luga Ingrian, it is the vowel **a*. In Votic, the production pattern corresponds to the one found in IF (**i* is the most innovative and **a* is as conservative as **u*), while the categorisation rather resembles that of CI, where **a* is more innovative than the other vowels.

¹ Hereafter, the original etymological vowel qualities, whose reflexes are studied in the experiments, are marked with an asterisk.

Fig. 2. IF (Ingrian Finnish speaker): vowel devoicing. pal., palatalised; lab., labialised.

Fig. 3. V (Votic speaker): vowel devoicing. For the legend explaining the types of devoicing, see Fig. 2.

A more detailed look at vowel devoicing (Fig. 2–4) and reduction of quality (Fig. 5–7), as well as lexical considerations (Table 4), clarify the possible reasons for these differences. Vocalic segments are still largely present at stage 1, but their quality is reduced along three dimensions: aspiration, centralisation and laryngealisation (cf. Klatt & Klatt, 1990; Laver, 1994, pp. 189–191; Barnes, 2006,

Fig. 4. CI (Central Lower Luga Ingrian speaker): vowel devoicing. For the legend explaining the types of devoicing, see Fig. 2.

pp. 114–150). Completely non-aspirated variants, in fact, accounted for just about half of those realisations which were considered modal according to the six-type scale (see Fig. 2–4). In total, partially or fully aspirated and devoiced allophones overwhelmingly prevailed over the “clear” modal ones even at stage 1. The percentage of non-aspirated modal allophones is in negative correlation with the percentage of complete loss in production in nearly all the cases (apart from **a* in Votic). The prototypical realisations (type 6: “zero”) of the new category [–SEGMENT] are therefore gaining strength in production first of all at the cost of the prototypical realisations (type 1: “modal non-aspirated vowel”) of the old category [+SEGMENT]. The belt of intermediate types preserves roughly the same structure for all the three vowels within each speaker and just slides down the scale.

Qualitative vowel reduction reveals quite a different picture (Fig. 5–7). Noticeable differences in the structure of phonetic variability appear between vowel types but not across speakers. Phonetic reasons for these differences are, therefore, to be sought in the articulatory and perceptual properties of vowel qualities rather than in other factors. In all three samples, *a* has undergone an extremely strong reduction to schwa (around 65% of complete schwa realisations and <30% of full vowels), and at the later stages of reduction, it is realised as schwa almost invariably. In contrast, **i* was the vowel most likely to retain its quality (around 80% of non-reduced allophones). The vowel **u* occupied an intermediate position, with about half of its realisations being non-reduced. In quality assessment, a more detailed scale for categorisation was used (see 2.2). The category of a reduced vowel rendered via Russian “ы” occurred only in the vowel **a* of the Votic speaker.

Fig. 5. IF (Ingrian Finnish speaker): qualitative reduction. pr, production; ct, categorisation. For the legend explaining the types of qualitative reduction, see Fig. 6.

Fig. 6. V (Votic speaker): qualitative reduction. pal., palatalised; lab., labialised; pr, production; ct, categorisation.

Qualitative reduction and devoicing manifested very differently, sometimes with opposing distributional patterns. We explored which of the two correlated better with the percentages of vowel loss in production and categorisation.

In Ingrian Finnish, the devoicing structure is the same for all vowel types, so it cannot be a factor conditioning the differences in their loss. Qualitative re-

Fig. 7. CI (Central Lower Luga Ingrian speaker): qualitative reduction. pr, production; ct, categorisation. For the legend explaining the types of qualitative reduction, see Fig. 6.

duction, in turn, correlates negatively with loss in production and categorisation. The only obvious phonetic factor correlating with the level of loss is, therefore, the type of vowel itself, as the level of loss in **i* is higher than in the other vowel types (for non-phonetic factors see 3.4).

In Central Lower Luga Ingrian, on the other hand, loss in production and categorisation positively correlates with the level of devoicing in all cases, and in **a*, also with qualitative reduction. One could say that in **a*, devoicing and centralisation reinforce each other as phonetic drivers for reduction, resulting in a relatively high percentage of loss in production (19%) and even more so in categorisation (35%). We will see later that in Lower Luga Ingrian, it is indeed vowel devoicing, reinforced by qualitative reduction, that is the primary driving force of loss, especially of the drastic loss of **a* from speakers' awareness and production. Devoicing with quality preservation leads to the rise of phonemic consonantal palatalisation as a trace of **i*.

The Votic speaker presented a mixed strategy between these two. As in IF, loss in the production and categorisation of **i* did not correlate with devoicing and negatively correlates with centralisation. At the same time, the configuration for **a* resembled that found in CI, although the equation was not perfect. The level of loss in categorisation positively correlated with those of qualitative reduction and devoicing. At the same time, the level of loss in production correlated with all three negatively. In other words, even if the speaker centralised and devoiced **a*, this did not lead to an increased drop of this often voiceless schwa from her production. In fact, Votic **a* was the only vowel in the stage 1 speakers which showed a clear negative correlation between the levels of loss in production and categorisation. The loss of reduced **a* had already

Fig. 8. SI1 (South Lower Luga Ingrian speaker 1): vowel devoicing, pr, production; ct, categorisation. For the legend explaining the types of vowel devoicing, see Fig. 9.

started in the mental categorisation but not yet in the production of the Votic speaker, so she manifested also a more conservative production pattern for **a* than the CI speaker. She was the only one who was aware of the qualitative reduction of **a* among the stage 1 speakers. This awareness might be related to the presence of the unrounded mid back vowel *õ* in Votic, unique among these varieties (see 2.1). The Votic speaker might have identified the schwa with this *õ*.

In IF and V, final vowels also underwent laryngealisation (27% of the IF tokens and 15% of the V tokens). In the CI speaker, this process was not attested. Summary laryngealisation patterns (IF + V) across the vowel types broadly correlated with the patterns for qualitative reduction. The vowel **i* tended to be the most conservative (17% of laryngealisation), and **a* the most innovative (26%), with **u* in the middle (19%), although these were not strong tendencies.

3.2 Stage 2 (South Lower Luga Ingrian) and Stage 3 (Siberian Ingrian/Finnish)

The speakers at stages 2 and 3 of reduction (Fig. 8–10) manifested a continuation of the same tendencies, especially CI. The speakers at stage 2 belonged to the same variety (South Lower Luga Ingrian), but the male speaker SI1 was not a typical one. He used to be a community manager and a local cultural leader and has a notably higher level of linguistic awareness than others. In his notebooks, one can find texts and words in his own variety in an orthography created by himself, reasonings for choices of orthography, texts in other Finnic languages copied from published sources, and etymological comparisons between cognate Finnic words.

Fig. 9. SI2 (South Lower Luga Ingrian speaker 2): vowel devoicing. pal., palatalised; lab., labialised; pr, production; ct, categorisation.

Fig. 10. S (Siberian Ingrian/Finnish speaker): vowel devoicing. pr, production; ct, categorisation. For the legend explaining the types of vowel devoicing, see Fig. 9.

The main reasons for the differences observed between SI1 and SI2 could be attributed to these specific characteristics of SI1. He has a more innovative production pattern for **a*, and more conservative ones for **i* and **u*, than SI2. Categorisation suggests a clue to the origins of this difference. In SI1, it was the most consistent of all six speakers (apart from the Siberian speaker, where the

sound change process had already reached the terminal stage). He categorised the final **a* > ə always as zero, although he was actually not as consistent regarding schwas in non-final positions, where he often used “ɪ” (Kuznetsova, 2012b), much as the Votic speaker did. Seemingly, this is the closest perceptual Russian correlate of schwa for speakers of these local languages. Final **i* and **u*, on the other hand, were always perceived as vowels, though the Siberian speaker was aware of their reduced character and called them “half-vowels.” He seemed to target these mental categorisations in his pronunciation consciously, and his percentage of loss was correspondingly higher for **a* and lower for **i* and **u* when compared to the otherwise linguistically very close speaker SI2. He was obviously not able to gain full control over his production, though, and his pattern of loss for the three vowels still has a scalar shape similar to that of other Ingrian speakers (CI and SI2).

In SI2, the phonemic categorisation (as zero) was consistent only for the vowel **a*, which had reached the critical threshold for complete loss. Interestingly, **i* and **u* showed reverse patterns of loss in her production versus categorisation, which is apparently rooted in the acoustic and perception properties of these two vowels. In general, at stages 2 and 3, one observes a robust cluster of palatalisation for **i* in all speakers (~40% of phonetic production). The vowel **u* also manifested a visible cluster of labialisation, completely absent at stage 1, but it accounted for only 15% of its phonetic production. If one adds the clusters of strong aspiration and palatalisation/labialisation to the “vocalic” realisations of vowels, the distributional patterns in production of **i* and **u* of both SI speakers match those of categorisation much more closely. It seems that at the intermediary stage 2, the “consonantal” reflexes of the vowels **i* and **u*, which give colour to consonants, still correspond rather to [+SEGMENT] in mental categorisation. Especially in the case of **i*, one could argue that both SI1 and SI2 still perceived and targeted, more or less consciously, the full vowel. This might be linked to a robust salience of **i*-reflexes in both perception and articulation (see 4.2). The speakers succeeded in reaching a vowel in only about half of the cases, though, ending up with a more or less aspirated palatalised final consonant in the other half instead. The lower perceptual salience of **u* can be seen in its relatively innovative categorisation by SI2, which was, in general, not as systematic as that by SI1. For the even less perceptually salient **a*, however, a relatively robust cluster of consonantal aspiration did not prevent complete loss from categorisation, as **a* does not colour consonants.

In the Siberian speaker at stage 3, we saw the next step of the same processes. Here, all the vowels had already reached the critical threshold for loss in production in order to be lost from mental categorisation. Judging by all three samples from stages 2 and 3, one could estimate this threshold at about 70%. Categorisation becomes innovative for vowels which have reached it, while their production still lags behind. In the Siberian speaker, the structure of phonetic loss for **i* already closely followed that of **a*; that is, the middle step of the “ladder-like” pattern flattened. We still saw a more conservative production of **u* with respect to the other vowels, though. The phonetically conservative nature of **u*, observed in all the speakers at stages 2 and 3, cannot be explained by its categorisation properties and should apparently be attributed to general physiological factors (acoustics, articulation, perception and storage in the memory;

see 1.2 and 4.2). At the same time, while in the Siberian speaker the palatalisation cluster was still as big as the cluster of complete loss, the labialisation cluster was already twice as small as the latter. One might hypothesise that while palatalisation could still have a chance of being preserved as a phonemic feature in these languages, labialisation has already lost the historical sound change battle, even though the Siberian speaker still perceived its presence. The aspiration cluster as a reflex of **a* was especially robust in the Siberian speaker, but as at stage 2, this did not affect its perceived complete loss.

3.3 Statistical Tests on the Main Findings

The main results of the study are corroborated by statistics. We compared the means of “vowel” (= 0) versus “loss” (= 1) in the overall production and categorisation, across the vowels, the speakers, the vowels in speakers and the stages of reduction. One-way ANOVA and Levene’s *F* showed a highly significant difference ($p < 0.001$) in all cases apart from the first comparison. The overall production ($n = 3,744$, $M = 0.36$, $SD = 0.481$, $SE = 0.008$) and categorisation ($n = 479$, $M = 0.35$, $SD = 0.479$, $SE = 0.022$) did not differ ($F(1, 4,221) = 0.12$, $p = 0.729$, Levene’s $F = 0.501$, $p = 0.479$), which supports the general correlation in production and perception of vowel loss (the difference in SE can be explained by the unequal size of the groups). For the other cases, we ran two post hoc tests for pairwise within-group comparisons in big samples of unequal size and variance: Tamhane’s *T2* (*T*; more conservative) and Games-Howell (GH; more liberal) at the 95% confidence interval. The few differences between them are reported below as *T*/GH.

There was a significant effect of stage on the level of loss ($F(5, 4,217) = 654.27$, $p < 0.001$). Both post hoc tests showed no difference between production and categorisation at stage 1 ($MD = -0.02$, $SE = 0.018$, $p = 1$), a difference at stage 2 ($MD = 0.15$, $SE = 0.042$, $p = 0.004$) and a highly significant difference at stage 3 ($MD = -0.16$, $SE = 0.015$, $p < 0.001$). In other words, both production and categorisation are still conservative at the first stage, then production becomes significantly more innovative, which leads to the shift in categorial analysis: categorisation becomes significantly more innovative and drives the loss at the terminal stage.

The overall results for the three vowels showed a highly significant difference ($F(2, 4,220) = 77.49$, $p < 0.001$), as each vowel has its unique configuration of loss in production and categorisation (Table 2; Fig. 12). The post hoc tests on these two aspects analysed separately showed that both the production and the categorisation of **u*, as well as the categorisation of **i*, did not differ and were conservative. At the same time, the more innovative production of **i* manifested a relatively significant difference from these three, being closer to the even more innovative production of **a*. The latter was insignificantly more conservative than the production of **a*. In sum, **u* was conservative and **a* innovative in both aspects, while **i* was conservative in categorisation and intermediate in production.

The overall results for the six speakers lacked any difference for IF and V ($MD < 0.01$, $p = 1$), but CI differed from these two (CI-IF: $MD = 0.05$, $p = 0.003$ *T*/0.002 GH; CI-V: $MD = 0.05$, $p = 0.002$). CI manifested a slightly more advanced substage of reduction inside stage 1, where **a* overtook **i* as the leader of loss. At

Table 2. Mean differences between production and categorisation in vowels

Vowel (test)	i (pr)	u (pr)	a (ct)	i (ct)	u (ct)
a (pr)	0.07** (<i>p</i> = 0.006)	0.21***	−0.11 (<i>p</i> = 0.086 T/ 0.065 GH)	0.21***	0.22***
i (pr)		0.14***	−0.18***	0.14** (<i>p</i> = 0.002)	0.15** (<i>p</i> = 0.002 T/ 0.001 GH)
u (pr)			−0.32***	0.0	0.01
a (ct)				0.32***	0.33***
i (ct)					0.01

pr, production; ct, categorisation. ** *p* < 0.01, *** *p* < 0.001.

stage 2, SI1 did not significantly differ from SI2 (MD = −0.06, *p* = 0.262). All other differences between speakers were highly significant (*p* < 0.001). Production and categorisation are further analysed across speakers in Table 3 (see also Fig. 11, 13). At stage 1, both production and categorisation by IF and V, as well as categorisation by CI, did not differ and were conservative. Production by CI was slightly more innovative: it showed moderate-to-weak differences from production by IF and V (but not from categorisation by CI). Production and categorisation by S highly differed from everything else and from each other (her production was significantly more conservative than her production due to the categorical shift at stage 3). At the intermediate stage 2, production and categorisation by SI1 and production by SI2 did not show significant differences. However, categorisation by SI1 stands out against all effects in Table 3. The peculiarity of categorisation by SI1 is likely linked to the speaker’s unusual linguistic awareness and full systematicity in transcription (**a* as zero, **i* and **u* as vowels; see 3.2).

3.4 The Lexical Factor in Reduction at Stage 1

A lexical factor effect at the initial stage of reduction (see 1.1, 2.1) was observed also in the present data, with nuances concerning vowel type, speaker and the correlation between production and categorisation. Table 4 shows the data on the lexical distribution of the cases of loss at stage 1. In column 1, the total number of lexical words in each sample is given. Column 2 provides the number of lexemes in which at least one token of “loss” in production was attested, their percentage in the sample for each speaker, and their distribution across the **a*, **i* and **u/o* types. The vowel types are arranged in the parentheses from those with the highest number of words exhibiting loss to those with the lowest number exhibiting loss. Column 3 cites only those words in which more than half of the tokens showed loss (i.e., *n* > 4; the exact number of such tokens is given in parentheses for each word). Column 4 summarises the number of words from column 3 and their percentage in each sample. Finally, in column 5, the ratio between the numbers of words in columns 4 and 2 is calculated, providing an idea of the lexical compactness of the distribution of vowel loss.

Table 3. Mean differences between production and categorisation by each speaker

Speaker	V (pr)	CI (pr)	SI1 (pr)	SI2 (pr)	S (pr)	IF (ct)	V (ct)	CI (ct)	SI1 (ct)	SI2 (ct)	S (ct)
IF (pr)	0.0	-0.05(*) (<i>p</i> = 0.058 T/0.042 GH)	0.5***	0.55***	0.8***	0.01	0.03	0.07	0.3***	0.45***	0.95***
V (pr)		-0.05* (<i>p</i> = 0.013 T/0.011 GH)	0.51***	0.56***	0.8***	0.0	0.03	0.08	0.3***	0.45***	0.96***
CI (pr)			0.46***	0.4***	0.75***	0.06	0.02	0.02	-0.25** (<i>p</i> = 0.001)	0.4***	0.91***
SI1 (pr)				0.05	0.29***	0.51***	0.48***	0.43***	*0.2 (<i>p</i> = 0.042 T/0.03 GH)	0.06	0.45***
SI2 (pr)					0.24***	0.56***	0.53***	0.48***	**0.25 (<i>p</i> = 0.002 T/0.001 GH)	0.11	-0.4***
S (pr)						0.81***	0.77***	0.73***	0.5***	0.35***	-0.16***
IF (ct)							-0.04	-0.08	-0.31***	-0.46***	-0.96***
V (ct)								-0.04	**(*)0.27 (<i>p</i> = 0.001 T/<0.001 GH)	-0.42***	-0.93***
CI (ct)									*0.23 (<i>p</i> = 0.047 T/0.034 GH)	-0.38***	-0.88***
SI1 (ct)										-0.15	-0.65***
SI2 (ct)											-0.51***

IF, Ingrian Finnish; V, Votic; CI, Central Lower Luga Ingrian; SI, South Lower Luga Ingrian; S, Siberian Ingrian/Finnish; pr, production; ct, categorisation. * *p* < 0.05, ** *p* < 0.01, *** *p* < 0.001.

In IF and V, the level of the lexical compactness of loss was 4 times higher than in CI. The vowel loss in these speakers was concentrated in very few frequent basic words, while in CI the lexical dispersion of loss was much higher. It is remarkable that in IF and V, the loss in frequent words concerned only the vowel **i*. In Soikkola Ingrian, the other still existing Ingrian dialect which is about as archaic as IF from the point of view of reduction (Kuznetsova, 2016), the same type of **i*-loss in frequent words became lexicalised. For example, the following words in our questionnaire do not have the final **i* in Soikkola Ingrian: *pēn/pīn* “small,” *ūs* “new,” *sūr* “big,” *laps* “child” and *nōr/nūr* “young.” Grammatical morphemes (even more frequent elements of the language) of Ingrian Finnish, Votic and Soikkola Ingrian manifest such grammaticalised loss for both **i* and **a*. In CI, a speaker using a more innovative variety where reduction is already conditioned phonetically, **a* became the reduction leader.

These differences in production find a parallel in categorisation (see Table 4, column 6; words encoded with V/Ø were cited by the respective speakers as having both a variant with a vowel and one without it). In IF the loss was perceived only in **i*-words, in V it was perceived both in **i*- and in **a*-words, and in CI it was perceived only in **a*-words. The number of lexical items with perceived

Table 4. Lexical specification of vowel loss at stage 1

Speaker (N)	Lexemes exhibiting phonetic loss, <i>n</i>	Lexemes with phonetic loss <i>n</i> > 4 (>50%)	Lexemes with phonetic loss >50%, <i>n</i>	Lexical compactness of loss	Lexemes which exhibited loss in categorisation (by glosses)
IF (78)	12 (15.3%) (<i>i, a</i>)	<i>pieni</i> “small” (6), <i>ūsi</i> “new” (8)	2 (2.6%)	0.17	<i>i</i> : small, child (<i>i/∅</i>), elk
V (81)	10 (12.3%) (<i>i, a, u/o</i>)	<i>sūri</i> “big” (6), <i>pēni</i> “small” (6)	2 (2.5%)	0.2	<i>i</i> : big (<i>i/∅</i>); <i>a</i> : dog, change:IMP, barley, nail, bath broom
CI (76)	21 (27.6%) (<i>a, i, u/o</i>)	<i>nāgla</i> “nail” (6)	1 (1.3%)	0.05	<i>a</i> : nail, leg, which (of the two), black (<i>a/∅</i>), change:IMP, debit, twig, bath broom, broom, class, skinny

IF, Ingrian Finnish; V, Votic; CI, Central Lower Luga Ingrian.

loss also increased from IF to CI. It is worth noting that the match between production and categorisation was close in the statistical sense, but not in the lexical one. The examples in Table 3 show that in each speaker’s production and perception, the trends for reduction and loss correlate better across the vowel types than across the concrete lexical items. This might provide support for distributional learning of phonological categories, which happens relatively independently of individual lexemes and minimal pairs (see 1.1).

4. General Discussion

4.1 General Course of Vowel Reduction and Loss

Our study was restricted to two phrasal contexts (in the production part) and three types of vowel, and only one speaker was taken for each variety apart from one. Even if limited by these and other methodological restrictions, the results revealed a stable correlation of frequency distributions in production and perception across all six speakers. The latter represented three main stages of vowel reduction and loss in the Finnic languages of Ingria. This correlation of the internal structure of categorical representation to the structure of production is probably best explained by the adaptive hybrid models of mental storage, which suggest the internal clusterisation of exemplars within a category (Gureckis & Goldstone, 2008; Kirby, 2010, pp. 34–37).

The main vectors summarising the general course of vowel reduction and loss in the Finnic languages of Ingria and the differences across speakers and vowel types are represented in Figures 11–13 (the mean values for each test on a scale between 0 = “vowel” and 1 = “loss” and the SE bars are given; red stands for production and grey for categorisation). The dashed horizontal lines are drawn at 70% loss and at 70% preservation of the segments, which appeared to

Fig. 11. Summary: speakers. Red stands for production and grey for categorisation. IF, Ingrian Finnish; V, Votic; CI, Central Lower Luga Ingrian; SI, South Lower Luga Ingrian; S, Siberian Ingrian/Finnish.

Fig. 12. Summary: vowels. Red stands for production and grey for categorisation.

be important thresholds for the stages of reduction and loss and changes in categorisation.

Indeed, one can observe the three stages of reduction described in the paper divided by these thresholds in Figure 11. At stage 1, production and categorisation are closely matched. As discussed in 3.4, vowel reduction and especially loss

Fig. 13. Summary: speakers and vowels. Red stands for production and grey for categorisation. IF, Ingrian Finnish; V, Votic; CI, Central Lower Luga Ingrian; SI, South Lower Luga Ingrian; S, Siberian Ingrian/Finnish.

at this initial stage (speakers IF and V) is linked to a large extent to certain frequent elements (frequent lexemes, grammatical markers). At the same time, the correlation between production and perception in each vowel of each speaker is in general closer in the statistical sense than across concrete lexemes. The learned phonological category looks more like a sum of the distributional properties of phonetic variants in production, abstracted from particular lexical words, in line with the distributive learning hypotheses (see 1.1).

By the middle phase of loss, the novel stimuli have been accumulated and spread through the phonetic system, so reduction and loss are conditioned phonetically rather than lexically or grammatically. Categorisation remains more conservative at the first two stages, that is, more of the new category is produced than perceived. Speakers at stage 2 still often seem to target the old category in pronunciation, especially in the vowels which give colour to consonants, but achieve it only partially. The mechanism of reduction with a more conservative categorisation than realisation is linked to the automatisisation of execution in production, is phonetically gradual and can likely take a considerable amount of time. If the old category ([+SEGMENT] in our case) still keeps 70% or more of realisation, the formation of the new category [-SEGMENT] is not yet perceived by speakers (stage 1). If both categories are pronounced in about 50% of cases, the categorisation is also split about 50/50 between the perceived presence and absence of a vowel (stage 2). When the new category has arrived at more than 70% of realisations (stage 3 and some vowels at stage 2), the crucial jump in categorisation has happened. The pattern is drastically reversed: the old value is no longer perceived, while it is still partially maintained in production. Reduction and loss at the stage of a more innovative categorisation imply a drop of low-salience meaningless parts, which is sometimes distinguished from the automatisisation-

based mechanism (see 1.2). At the same time, it might still mean automatisation in production, now of the new category rather than of the old one.

Major differences in production and categorisation observed between the three corner vowels are summarised in Figure 12. Each of the three vowels manifested its own combination of production and categorisation values. The vowel **u* turned out to be conservative both in production and categorisation, and in total the most conservative among these vowels. The vowel **i* had an intermediate position, with an overall categorisation as conservative, as in **u*, but production nearly as innovative as **a*. This innovativeness in loss is actually accompanied by the formation of a robust cluster of consonantal palatalisation (see 4.2). The vowel **a* was the most advanced in terms of loss, and here, in turn, categorisation was more innovative than realisation. This is obviously linked to the its extremely strong qualitative reduction and due to the fact that it does not leave any colour to the consonants. In general, the more the vowels were reduced to schwa, the less their presence was perceived.

Our results actually showed that the two main patterns of vowel reduction (centrifugal and centripetal) do not exclude each other, in line with Harris (2005) and unlike the sharp distinction made by Crosswhite (2004). In the course of vowel reduction and loss in the Finnic languages of Ingria, we observed the elements of both patterns. The mid vowels *o*, *ö* and *e* are raised to the high vowels *u*, *ü* and *i*, which can be seen in the variants *maito/maitu* “milk,” *pudro/pudru* “porridge” and *viero/vieru* “wheel” from our questionnaire (Appendix 1), occurring even in the same speaker. At the same time, all vowels could lose their quality completely at later stages and centralise to schwa.

Figure 13 gives a chart of the loss across the vowels of the individual speakers. Both the speakers and the vowels are placed starting with the most conservative and ending with the most innovative ones, which gives the idea of an S-curve of the sound change. The chart shows that the reverse in the ratio between production and categorisation happened in **a* at a much earlier point than in **i* and **u*. Only in the most conservative speaker (IF), the pattern of **a* matches those of the other vowels. The vowel **a* basically jumps over the transitory middle zone with a 50/50 split in production between the old and new values, driven forward by its innovative categorisation. The processes of loss in **i* and **u* run more smoothly. In these two consonant-colouring vowels, in turn, it is mostly an innovative realisation that drives forward the process of change.

4.2 Formation and Loss of Consonantal Palatalisation, Labialisation and Aspiration

Already at stage 1, the vowel qualities revealed significant differences in their patterns of reduction and loss both in production and perception. It seems that vowel loss starts from **i*, and yet is lexically and grammatically stimulated to a large extent (see 3.4). Later on, as loss is generalised throughout the entire phonetic system, **a* takes over as the leader because its reduction and loss proceed faster (it can be seen in speakers starting from CI). This is presumably linked to the strongest qualitative reduction and devoicing of the latter, and also to the fact that **a* does not give any colour (i.e., secondary localisation) to the consonants, unlike **i* and **u*. In **i*, on the other hand, the qualitative reduction is

the weakest. For the vowel **u*, no reduction or loss is yet perceived at stage 1, even if qualitative reduction is already very frequent.

An overall trend observed across all vowels and speakers is that during vowel loss the prototypical variants (zero vowel reflexes) of the new category [−SEGMENT] are gaining ground first of all at the cost of the prototypical variants (full clear modal vowels) of the old category [+SEGMENT]. The intermediate variants form a belt which in total accounts for about 20–30% of each sample and slides down the scale. The only major exception to this is a significant cluster of consonantal palatalisation, which replaces a substantial part of the full non-reduced modal vowel **i* between stages 1 and 2. The vowel **i* is the first vowel type to exhibit loss and is nearly as fast in phonetic loss as schwa (<**a*). However, even the stage 2 speakers typically did not yet categorise the reflexes of **i* and **u* as consonantal features. SI1 was consistent in perceiving both as vowels, while SI2 perceived palatalisation also rather as a vowel, but labialisation already rather as loss. This is likely linked to a greater acoustic, articulatory and therefore perceptual salience of **i* and palatalisation compared to **u* and labialisation. Matthies et al. (2001) report the same tendency for quality preservation in *i* even in fast speech. Phonological consonantal palatalisation is apparently formed earlier than labialisation, at least in the history of these varieties, but is lost from the language more slowly. The possible impact of the front vowel harmony characteristic of the Finnic languages of Ingria on this asymmetry is yet to be investigated.

The high level of susceptibility of **u* to qualitative reduction and concomitant loss of rounding might be one of the factors impeding the formation of phonemic consonantal labialisation over loss of the segmental vowel at the later stages of reduction. The vowel **u* is the most conservative of the three vowels in terms of vowel loss, as it retained the largest “vocalic” cluster of realisations in all speakers, but it manifested gradual transitions between the stages in all aspects: qualitative reduction, devoicing and loss. No robust cluster of consonantal labialisation as a trace of **u* is formed; the segmental vowel is rather lost directly. Evidence from vowel perception and neuroimaging (see 1.2) also suggests that *u* is a more complex unit than *i* for brain processing, and acoustically and perceptually less salient.

The difference in size of the consonantal palatalisation and labialisation clusters between stages 2 and 3 could also be attributed to the articulatory properties of these two features. In the Finnic languages of Ingria, consonants do not typically undergo coarticulatory labialisation along their whole length; only the very last portion of the segment is regressively affected. Labialised consonants are often aspirated consonants where just the aspiration portion is labialised rather than the consonant itself (cf. labialised vs. plain aspiration on Fig. 14, 15). The labialised aspiration is then “eaten” away by reduction, and the consonant remains plain.

Consonantal palatalisation (Fig. 16) has a much more powerful impact on the articulation of consonants in these languages. Especially for dentals (and specifically for *l* and *t*), it is a full rather than secondary palatalisation, with a shift of the primary articulation towards the palatal region of the vowel tract (cf. surveys in Kochetov, 2011; Krämer & Urek, 2016). Our preliminary observations show that the number of palatalised consonants and the degree of their

Fig. 14. Aspirated labialised C (**lastu* [lastʰ] “chip”) PS.

Fig. 15. Aspirated C (**lasta* [lastʰ] “child:PRT”) PS.

palatalisation manifest a positive correlation with the degree of vowel reduction in the Finnic languages of Ingria. The more advanced the vowel reduction and loss, the bigger the number of palatalised consonantal phonemes to be distinguished for any particular variety and the stronger the palatalisation from the phonetic point of view. Ingrian Finnish and Soikkola Ingrian have dental pala-

Fig. 16. Aspirated palatalised C (**lusti* [lustʰj] “beautiful”) AU.

talised phonemes, but in Votic their inventory is significantly larger. At the same time, in Ingrian Finnish, Soikkola Ingrian, Votic and partially Central Lower Luga Ingrian, the consonant *t* is just secondarily palatalised and can be easily realised as a plain one. In most other Lower Luga Ingrian varieties (North, South and partially Central) and in Siberian Ingrian Finnish, in turn, this consonant is always fully palatalised in the context before the high front vowels *i* and *ü*. This palatalisation is so strong that sometimes a palatal stop is pronounced (viz. Leppik, 1975, pp. 116–117; Kuznetsova, 2009, pp. 195–235; Markus & Rožanskij, 2011, pp. 17–18).

These observations are to be further verified, but one could hypothesise that this trend is a result of re-phonologisation. High front vowel quality, which originally stimulated palatalisation, ceases to do so as vowels are progressively reduced and lost. Palatalisation starts to be perceived as a distinct property of consonants and becomes phonetically reinforced. Subsequently, even if the aspiration after consonants disappears, the palatalised articulation in those consonants where it has emerged as a stable property is preserved. Indeed, in the Siberian speaker, consonantal palatalisation was unevenly distributed across consonants. The consonants *l* and *t* were palatalised as a trace of **i* in all cases, and here we can speak about well-formed consonantal palatalisation. The consonants *p*, *k* and *n* manifested palatalisation in 30–60% of the cases, and the consonants *s*, *h*, *r*, *n*, *m* and *v* only sporadically. In these two groups, especially in the last one, a tendency towards complete depalatalisation was observed.

Palatalisation might still, therefore, survive as a phonemic feature, at least for some consonantal types. Labialisation is likely to be lost without any reflexes. Indeed, in neighbouring Estonian, which represents an even more advanced state of the same type of reduction and has passed through the stage of devoiced

vowels, consonantal palatalisation as a trace of **i* still exists (only for dental consonants and with a trend towards further loss), but no traces of consonantal labialisation were preserved (Teras & Pajusalu, 2014; Kuznetsova, 2016).

A similar disparity in the trajectories of loss of **i* and **u* is observed in the history of other languages, such as Russian (Šahmatov, 1915, pp. 15–16; Kiparsky, 1963) or Irish (Greene, 1973; Anderson, 2016). Cross-linguistically, labialised consonants are much rarer than palatalised ones. Blevins (2004, p. 204) explains the rarity of certain phonological contrasts through the uncommon occurrence of sound changes giving rise to them. Consonantal palatalisation accounted for 145 cases (3.18%) in PBase (Mielke, 2008; Brohan & Mielke, 2018, pp. 218–219), being one of the most frequent types of sound change, while labialisation included just 38 entries.

We do not know of any cases of emergence of phonemic consonantal aspiration as a result of loss of the following plain schwa, a reflex of **a* (nor does PBase give any such examples). This is just an additional indication for the very low perceptual saliency of this schwa, especially in the final position. In our case study, this has led to its fast disappearance both from the mind and the production of the speakers. Nonetheless, one might wonder if at stages 2 and 3 it is still possible to synchronically distinguish a separate phonemic series of aspirated consonants, as opposed to the non-aspirated ones (e.g., *vīn* ‘bring away:PRS:1SG’ vs. *vīn^h* [*<*vīna*] ‘vodka’). At stage 3, this series would be an addition to the plain, palatalised, labialised and labiopalatalised series (Sidorkevič, 2013; Kuznetsova, 2015, 2016). However, all speakers at stages 2 and 3 still perceived reflexes of **i* and **u*, even if none perceived plain aspiration as a reflex of **a*. If the difference between aspirated and plain consonants were still consistently maintained in production, this case would represent an example of a near-merger, the next-to-last step of phonologisation (Barnes, 2006, pp. 223–238; see also 1.1).

5. Conclusions

The novelty of our approach to the problem of correspondence in production and perception at different stages of sound change was that real varieties were studied and speakers assessed words in their own languages. Similar studies have usually involved miniature artificial languages or cross-linguistic assessments, with their own methodological restrictions. The most typical question asked is which of the vowels is perceived, while our enquiry was rather into whether any vowel is perceived at all. If a variety has no literary standard, the latter question is much easier to answer by means of the categorisation test proposed in this study (but only if the intermediary orthography has a segmental principle of encoding).

At the initial stage, vowel reduction and loss are linked to automatisisation of execution in production of the old category. Categorisation remains more conservative than production, and the phonetic loss is likely to be concentrated in a few frequent words and grammatical morphemes. At later stages, loss spreads throughout the system and its conditioning becomes purely phonetic. When our speakers pronounced a vowel in more than 70% of the cases, they typically perceived its presence. A decisive qualitative shift seemed to happen in categorisa-

tion after the new realisational types had gained more than 70%. The speakers stopped perceiving any segment and categorisation became innovative, while production still showed a certain percentage of vowel preservation and was lagging behind. Reduction is based on a loss of newly meaningless parts, and this last stage may contain near-mergers. Automatisations of execution is likely still at work here, only now the production of the new category is being automatised. Our comparison of the several stages of vowel reduction and loss revealed no irreconcilable contradiction between the two main patterns of reduction, i.e., centrifugal and centripetal. Both were observed in our data: the rise of mid vowels to high vowels and eventual centralisation of all vowels to schwa.

We studied the three basic corner vowels, which are known to share a specific set of properties with regard to reduction and loss, and observed asymmetry in their production and perception. The results suggest two markedness hierarchies among these vowels. As regards the process of reduction and loss itself, the hierarchy of vowels (from the most innovative to the least innovative) is $a > i > u$. The phonological saliency of the secondary consonantal localisations emerging in the process of loss of these vowels, in turn, would suggest a hierarchy $a > u > i$ (from the least salient to the most salient secondary localisations). Some possible physiological features stimulating both hierarchies were outlined in the paper.

Acknowledgements

I am grateful to Maxim Fedotov for his help with the preliminary segmentation of a part of the data, as well as to Cormac Anderson, Brittany Blankinship, four anonymous reviewers and the associate editors of this Special Issue for their valuable comments.

Disclosure Statement

The authors have no conflicts of interest to declare.

Funding Sources

Data collection and segmentation for this research (2014–2016) were supported by a Grant of the President of the Russian Federation for young PhDs (No. MK-4291.2015.6; “Non-initial vowel reduction in the vanishing Finnic varieties of Ingria”).

Author Contributions

Natalia Kuznetsova conceived and performed the study, apart from half of the acoustic analysis of the data and the Praat script for data extraction, which were performed by Vasilisa Verkhodanova.

Appendix 1: List of the Most Typical Carrier Words

T	A	I	U/O	R	A	I	U/O
t	<i>musta</i> “black”	<i>lusti</i> “beautiful”	<i>lastu</i> “chip”	l	<i>suola</i> “salt”	<i>stuoli</i> “chair”	<i>škoulu</i> “school”
t	<i>vihta</i> “bath broom”	<i>risti</i> “(a) cross”	<i>lintu</i> “bird”	l	<i>naula/nāgla</i> “nail”	<i>hili</i> “coal”	<i>laulu</i> “song”
t	<i>lūta</i> “broom”	<i>puoti</i> “shop”	<i>maito/maitu</i> “milk”	l	<i>muila</i> “soap”	<i>kieli</i> “tongue”	<i>joulu</i> “Christmas”
p	<i>hāpa</i> “aspen”	<i>sīpi</i> “wing”	<i>rūpo</i> “rubbish”; (<i>ei</i>) <i>korpu</i> “(does not) dry out:3sg”	r	<i>koira</i> “dog”	<i>sūri</i> “big”	<i>puđru/puđro</i> “porridge”
p	<i>kumpa</i> “which of the two”	<i>krāpi</i> “comb (wool):IMP”	<i>urpo</i> “willow”	r	<i>nuora</i> “rope”	<i>nuori</i> “young”	<i>vieru/viero</i> “wheel”
k	<i>jalka</i> “leg, foot”	<i>panki</i> “bucket”	<i>hanko</i> “snow-bank”	r	<i>hāra</i> “branch”	<i>hīri</i> “mouse”	<i>viiru</i> “stripe”
k	<i>nahka</i> “skin”	<i>poski</i> “cheek”	<i>pehko</i> “bush”	n	<i>vīna</i> “vodka”	<i>pieni</i> “small”	<i>hieno</i> “fine”
k	<i>poika</i> “boy”	<i>hauki</i> “pike”	<i>riuku</i> “pole”	n	<i>sauna</i> “sauna”	<i>sāni</i> “sleigh”; <i>sieni</i> “mushroom”	<i>kehno</i> “worn out”
k	<i>velka</i> “debt”	<i>olki</i> “straw”	<i>halko</i> “billet”	m	<i>ilma</i> “air”	<i>tormi</i> “storm”	<i>silmu</i> “lam-prey”
s	<i>oksa</i> “twig”	<i>lapsi</i> “child”	<i>paksu</i> “thick”	m	<i>surma</i> “death”	<i>sormi</i> “finger”	<i>formu</i> “(a) form”
s	<i>klāsa</i> “class”; <i>vatsa</i> “stomach”	<i>ūsi</i> “new”	<i>haisu</i> “(a) smell”	m	<i>māma</i> “moth-er”	<i>Suomi</i> “Fin-land”	<i>solmu</i> “knot”
h	<i>vaiha</i> “change:IMP”	<i>rīhi</i> “drying barn”; <i>jouhi</i> “horsehair”	<i>jauho</i> “wheat”	v	<i>hīva</i> “yeast”	<i>talvi</i> “winter”	<i>kaivo</i> “(a) well”
h	<i>laiha</i> “lean, thin”; <i>turha</i> “vain”	<i>tuohi</i> “birchbark”	<i>karhu</i> “(a) bear”; <i>kaiho</i> “damage”	v	<i>līva</i> “sand”	<i>hīrvi</i> “elk”; <i>sarvi</i> “horn”	<i>koivu</i> “birch”

Morphophonological front-vowel stems are underlined.

References

- Anderson, C. (2016). *Consonant colour and vocalism in the history of Irish* (Unpublished PhD dissertation), Adam Mickiewicz University in Poznań, Poland.
- Baker, A., Archangeli, D., & Mielke, J. (2011). Variability in American English s-retraction suggests a solution to the actuation problem. *Language Variation and Change*, 23(3), 347–374. <https://doi.org/10.1017/S0954394511000135>
- Barnes, J. (2006). *Strength and weakness at the interface: Positional neutralisation in phonetics and phonology*. Berlin, Germany; New York, NY: Mouton de Gruyter.
- Baudouin de Courtenay, J. N. (1895). *Versuch einer Theorie phonetischer Alternationen: Ein Capitel aus der Psychophonetik*. Strassburg: K. J. Trübner.
- Blevins, J. (2004). *Evolutionary phonology: The emergence of sound patterns*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511486357>
- Blythe, R. A., & Croft, W. (2012). S-curves and the mechanisms of propagation in language change. *Language*, 88(2), 269–304. <https://doi.org/10.1353/lan.2012.0027>
- Brohan, A., & Mielke, J. (2018). Frequent segmental alternations in P-base 3. In L. Hyman & F. Plank (Eds.), *Phonological typology* (pp. 196–228). Berlin, Germany; Boston, MA: Walter de Gruyter. <https://doi.org/10.1515/9783110451931-006>

- Bybee, J. (2001). *Phonology and language use*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511612886>
- Bybee, J. (2015). *Language change*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139096768>
- Bybee, J., File-Muriel, R. J., & de Souza, R. N. (2016). Special reduction: A usage-based approach. *Language and Cognition*, 8(3), 421–446. <https://doi.org/10.1017/langcog.2016.19>
- Crosswhite, K. (2004). Vowel reduction. In B. Hayes, R. Kirchner, & D. Steriade (Eds.), *Phonetically-based phonology* (pp. 191–231). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511486401.007>
- Davis, T., & Poldrack, R. A. (2014). Quantifying the internal structure of categories using a neural typicality measure. *Cerebral Cortex (New York, N.Y.: 1991)*, 24(7), 1720–1737. <https://doi.org/10.1093/cercor/bht014>
- de Lacy, P. (2006). *Markedness: Reduction and preservation in phonology*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511486388>
- Delattre, P. (1969). An acoustic and articulatory study of vowel reduction in four languages. *International Review of Applied Linguistics in Language Teaching*, 7(4), 295–325. <https://doi.org/10.1515/iral.1969.7.4.295>
- Eulitz, C., & Obleser, J. (2007). Perception of acoustically complex phonological features in vowels is reflected in the induced brain-magnetic activity. *Behavioral and Brain Functions*, 3(1), 26. <https://doi.org/10.1186/1744-9081-3-26>
- Flemming, E. (1995). *Auditory representations in phonology* (Unpublished PhD dissertation), UCLA, USA.
- Flemming, E. (2004). Contrast and perceptual distinctiveness. In B. Hayes, R. Kirchner, & D. Steriade (Eds.), *Phonetically-based phonology* (pp. 232–276). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511486401.008>
- Greene, D. (1973). The growth of palatalisation in Irish. *Transactions of the Philological Society*, 72(1), 127–136. <https://doi.org/10.1111/j.1467-968X.1973.tb01017.x>
- Gureckis, T. M., & Goldstone, R. L. (2008). The effect of the internal structure of categories on perception. In L. Shi, N. H. Feldman, & T. L. Griffiths (Eds.), *Proceedings of the 30th annual conference of the Cognitive Science Society* (pp. 1876–1881). Austin, TX: Cognitive Science Society.
- Hall, K. C., Hume, E., Jaeger, F. T., & Wedel, A. (2018). The role of predictability in shaping phonological patterns. *Linguistics Vanguard*, 20170027.
- Hanique, I., Ernestus, M., & Boves, L. (2015). Choice and pronunciation of words: Individual differences within a homogeneous group of speakers. *Corpus Linguistics and Linguistic Theory*, 11(1), 161–185. <https://doi.org/10.1515/cllt-2014-0025>
- Harris, J. (2005). Vowel reduction as information loss. In P. Carr, J. Durand, & C. J. Ewen (Eds.), *Headhood, elements, specification and contrastivity: Phonological papers in honour of John Anderson* (pp. 119–132). Amsterdam, The Netherlands: John Benjamins Publishing. <https://doi.org/10.1075/cilt.259.10har>
- Hay, J., & Foulkes, P. (2016). The evolution of medial /t/ over real and remembered time. *Language*, 92(2), 298–330. <https://doi.org/10.1353/lan.2016.0036>
- Hyman, L. M. (1976). Phonologisation. In A. Juillard (Ed.), *Linguistic studies presented to Joseph H. Greenberg* (pp. 407–418). Saratoga: Anma Libri.
- Johnson, K. (2015). Vowel perception asymmetry in auditory and phonemic listening. *UC Berkeley Phonology Lab Annual Report 2015*. Available from <https://escholarship.org/uc/item/21t337g>
- Kapatsinski, V. (2018). *Changing minds changing tools: From learning theory to language acquisition to language change*. Cambridge, MA: MIT Press.
- Kiparsky, V. (1963). *Russische historische Grammatik. Band I: Die Entwicklung des Lautsystems*. Heidelberg, Germany: Universitätsverlag Winter Heidelberg.
- Kirby, J. P. (2010). *Cue selection and category restructuring in sound change* (Unpublished PhD dissertation), University of Chicago, USA.
- Kirchner, R. M. (1998). *An effort-based approach to consonant lenition* (Unpublished PhD dissertation), UCLA, USA.
- Klatt, D. H., & Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *The Journal of the Acoustical Society of America*, 87(2), 820–857. <https://doi.org/10.1121/1.398894>
- Kochetov, A. (2011). Palatalisation. In M. van Oostendorp, C. J. Ewen, K. Rice, & E. V. Hume (Eds.), *The Blackwell companion to phonology* (Vol. 3, pp. 1666–1690). London, UK: John Wiley & Sons.

- Krämer, M., & Urek, O. (2016). Perspectives on palatalisation. *Glossa: A Journal of General Linguistics*, 1(1):31. <http://doi.org/10.5334/gjgl.226>
- Kruszewski, N. V. (1883). *Očerki nauki o jazyke*. Kazan, Russia.
- Kuznetsova, N. (2009). *Fonoličeskie sistemy ižorskih dialektov* (Unpublished PhD dissertation), Saint Petersburg: Institut lingvističeskikh issledovanij RAN, Russia.
- Kuznetsova, N. (2011). Typology of prosodic systems in Low Luga Izhorian varieties. In P. K. Austin, O. Bond, M. Charette, D. Nathan, & P. Sells (Eds.), *Proceedings of Conference on Language Documentation and Linguistic Theory 2* (pp. 185–192). London, UK: SOAS.
- Kuznetsova, N. (2012a). Prosodika slovoformy v nižnelužskom dialekte ižorskogo jazyka. *Acta linguistica Petropolitana*, VIII(1):43–103.
- Kuznetsova, N. (2012b). Vowel reduction in Lower Luga Ingrian: Scientific description and “folk” perception. In G. V. Golovko, M. L. Lurie, M. Z. Muslimov, & S. A. Oskolskaya (Eds.), *“Folk linguistics”: Language from speakers’ perspective. Paper abstracts for the conference, St. Petersburg, November 19–21, 2012* (pp. 85–87). St. Petersburg, Russia: Nestor-Istorija.
- Kuznetsova, N., & Sidorkievič, D. (2012). Priloženie 6: Karty k statjam. *Acta linguistica Petropolitana*, VIII(1):559–567.
- Kuznetsova, N. (2015). Two phonological rarities in Ingrian dialects. In M. Hilpert, J. Duke, C. Mertzlufft, J.-O. Östman, & M. Rießler (Eds.), *New trends in Nordic and general linguistics* (pp. 91–117). Berlin, Germany: De Gruyter.
- Kuznetsova, N. (2016). Evolution of the non-initial vocalic length contrast across Finnic varieties of Ingria and the adjacent areas. *Linguistica Uralica*, 52(1), 1–25. <https://doi.org/10.3176/lu.2016.1.01>
- Kuznetsova, N., Markus, E., & Muslimov, M. (2015). Finnic minorities of Ingria: The current sociolinguistic situation and its background. In H. Marten, M. Riessler, J. Saarikivi, & R. Toivanen (Eds.), *Cultural and linguistic minorities in the Russian Federation and the European Union* (pp. 127–167). Berlin: Springer.
- Laanest, A. (1980). Inkerioismurteiston suhteista suomen kaakkoismurteisiin. *Virittäjä*, 84, 142–149.
- Labov, W. (1999). *Principles of linguistic change. Vol. 1: Internal factors*. London, UK: Blackwell.
- Labov, W., Karen, M., & Miller, C. (1991). Near-mergers and the suspension of phonemic contrast. *Language Variation and Change*, 3(1), 33–74. <https://doi.org/10.1017/S0954394500000442>
- Laver, J. (1994). *Principles of phonetics*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139166621>
- Leppik, M. (1975). *Ingerisoom kurgola murde fonoloogilise süsteemi kujunemine*. Tallinn, Estland: Eesti NSV Teaduste Akadeemia, Keele ja Kirjanduse Instituut.
- Leskinen, H. (1973). Kaakkoissuomalaisen loppuheiton yleisyydestä ja alkuperästä. *Suomalais-Ugrilaisen Seuran Aikakauskirja*, 72, 210–221.
- Lindblom, B. (1963). Spectrographic study of vowel reduction. *The Journal of the Acoustical Society of America*, 35(11), 1773–1781. <https://doi.org/10.1121/1.1918816>
- Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. J. Hardcastle & A. Marchal (Eds.), *Speech production and speech modelling* (pp. 403–439). Dordrecht, The Netherlands: Springer. https://doi.org/10.1007/978-94-009-2037-8_16
- Mägiste, J. (1925). *Rosona (Eesti Ingeri) murde pääjooned*. Tartu, Estonia.
- Manca, A. D., & Grimaldi, M. (2016). Vowels and consonants in the brain: Evidence from magnetoencephalographic studies on the N1m in normal-hearing listeners. *Frontiers in Psychology*, 7, 1413. <https://doi.org/10.3389/fpsyg.2016.01413>
- Markus, E., & Rožanskij, F. (2011). *Sovremennyj vodskij jazyk. Teksty i grammatičeskij očerk. Vol. 2*. Saint Petersburg, Russia: Nestor-Istorija.
- Matthies, M., Perrier, P., Perkell, J. S., & Zandipour, M. (2001). Variation in anticipatory coarticulation with changes in clarity and rate. *Journal of Speech, Language, and Hearing Research: JSLHR*, 44(2), 340–353. [https://doi.org/10.1044/1092-4388\(2001/028\)](https://doi.org/10.1044/1092-4388(2001/028))
- Maye, J., & Gerken, L. (2000). Learning phonemes without minimal pairs. In S. C. Howell, et al. (Eds.), *Proceedings of the Annual Boston University Conference on Language Development* (Vol. 24, pp. 522–533). Somerville, MA: Cascadilla Press.
- Mielke, J. (2008). *The emergence of distinctive features*. Oxford: Oxford University Press. PBase available from <http://pbase.phon.chass.ncsu.edu> [last accessed 01.10.2018].
- Mompeán-González, J. A. (2004). Category overlap and neutralisation: The importance of speakers’ classifications in phonology. *Cognitive Linguistics*, 15(4), 429–469. <https://doi.org/10.1515/cogl.2004.15.4.429>

- Muslimov, M. (2005). *Jazykovyje kontakty v Zapadnoj Ingermanlandii (nižnee tečenije reki Lugi)* (Unpublished PhD dissertation), Sankt-Peterburg: Institut lingvističeskikh issledovanij RAN, Russia.
- Nirvi, R. E. (1972). Siperian inkeriläisten murteesta ja alkuperästä. *Kotiseutu*, 2/3, 92–95.
- Ohala, J. J. (1989). Sound change is drawn from a pool of synchronic variation. In L. E. Breivik & E. H. Jahr (Eds.), *Language change: Contributions to the study of its causes* (pp. 173–198). Berlin, Germany: Walter de Gruyter. <https://doi.org/10.1515/9783110853063.173>
- Olejarczuk, P., Kapatsinski, V., & Baayen, R. H. (2018). Distributional learning is error-driven: The role of surprise in the acquisition of phonetic categories. *Linguistics Vanguard: Multimodal Online Journal*, 4(s2, S2). <https://doi.org/10.1515/lingvan-2017-0020>
- Pearce, M. (2008). Vowel harmony domains and vowel undershoot. *UCL Working Papers in Linguistics*, 20:115–140.
- Polka, L., & Bohn, O. S. (2003). Asymmetries in vowel perception. *Speech Communication*, 41(1), 221–231. [https://doi.org/10.1016/S0167-6393\(02\)00105-X](https://doi.org/10.1016/S0167-6393(02)00105-X)
- Polka, L., & Bohn, O. S. (2011). Natural Referent Vowel (NRV) framework: An emerging view of early phonetic development. *Journal of Phonetics*, 39(4), 467–478. <https://doi.org/10.1016/j.wocn.2010.08.007>
- Priva, U. C. (2017). Informativity and the actuation of lenition. *Language*, 93(3), 569–597. <https://doi.org/10.1353/lan.2017.0037>
- Ritter, S., Mücke, D., & Grice, M. (2018, October 2). *The dynamics of intonation: Categorical and continuous variation in an attractor-based model*. <https://doi.org/10.31234/osf.io/qkm6b>
- Roettger, T. B., Winter, B., Grawunder, S., Kirby, J., & Grice, M. (2014). Assessing incomplete neutralisation of final devoicing in German. *Journal of Phonetics*, 43, 11–25. <https://doi.org/10.1016/j.wocn.2014.01.002>
- Rosch, E. (1978). Principles of categorisation. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorisation* (pp. 27–48). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Šahmatov, A. A. (1915). *Očerki drevneišego perioda istorii russkogo jazyka*. Petrograd, Russia.
- Sidorkevič, D. V. (2013). *Jazyk ingermanlandskih pereselencev XIX v. v Sibiri (struktura, dialektnye osobennosti, kontaktnye javlenija)* (Unpublished PhD dissertation), Sankt-Peterburg: Institut lingvističeskikh issledovanij RAN, Russia.
- Stevens, K. N. (2004). Invariance and variability in speech: Interpreting acoustic evidence. In J. Slifka, S. Manuel, & M. Matthies (Eds.), *From sound to sense: 50+ years of discoveries in speech communication* (pp. B-77–B-85). Cambridge, MA: Research Laboratory of Electronics, MIT.
- Stevens, M., & Harrington, J. (2014). The individual and the actuation of sound change. *Loquens*, 1(1), e003. <https://doi.org/10.3989/loquens.2014.003>
- Szeredi, D. (2010). Vowel centralisation and vowel harmony in Hungarian. *The Odd Yearbook*, 8, 111–137.
- Teras, P., & Pajusalu, K. (2014). Palatalisatsioonist ja prepalatalisatsioonist spontaanses eesti keeles. *Keel ja Kirjandus*, 4, 257–269.
- Trautmüller, H., & Öhrström, N. (2007). Audiovisual perception of openness and lip rounding in front vowels. *Journal of Phonetics*, 35(2), 244–258. <https://doi.org/10.1016/j.wocn.2006.03.002>
- Uchanski, R. M. (2005). Clear speech. In D. B. Pisoni & R. E. Remez (Eds.), *The handbook of speech perception* (pp. 207–235). Oxford: Blackwell. <https://doi.org/10.1002/9780470757024.ch9>
- Vallabha, G. K., McClelland, J. L., Pons, F., Werker, J. F., & Amano, S. (2007). Unsupervised learning of vowel categories from infant-directed speech. *Proceedings of the National Academy of Sciences of the United States of America*, 104(33), 13273–13278. <https://doi.org/10.1073/pnas.0705369104>
- van Bergem, D. R. (1995). Perceptual and acoustic aspects of lexical vowel reduction, a sound change in progress. *Speech Communication*, 16(4), 329–358. [https://doi.org/10.1016/0167-6393\(95\)00003-7](https://doi.org/10.1016/0167-6393(95)00003-7)
- van der Hulst, H. (2018). *Asymmetries in vowel harmony: A representational account*. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780198813576.001.0001>
- Vatakis, A., Maragos, P., Rodomagoulakis, I., & Spence, C. (2012). Assessing the effect of physical differences in the articulation of consonants and vowels on audiovisual temporal perception. *Frontiers in Integrative Neuroscience*, 6, 71. <https://doi.org/10.3389/fnint.2012.00071>
- Walker, R. (2011). *Vowel patterns in language*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511973710>
- Wanrooij, K., Escudero, P., & Raijmakers, M. E. (2013). What do listeners learn from exposure to a vowel distribution? An analysis of listening strategies in distributional learning. *Journal of Phonetics*, 41(5), 307–319. <https://doi.org/10.1016/j.wocn.2013.03.005>
- Yu, A. C. (2013). Socio-cognitive processing and the actuation of sound change. In A. C. Yu (Ed), *Origins of sound change: Approaches to phonologisation* (pp. 201–227). Oxford: Oxford University Press.