This is the accepted manuscript version of an article published by S. Karger AG in Phonetica/2019/76/2-3/201-233 (Special Issue "Constancy and variation in speech), DOI: 10.1159/000494927, and available on www.karger.com/Article/FullText/10.1159/000494927.

# Phonetic realisation and phonemic categorisation of the final reduced corner vowels in the Finnic languages of Ingria

*Natalia Kuznetsova[1,2]*, Vasilisa Verkhodanova[3]*

[1] Institute for Linguistic Studies, Dept. of the Languages of Russia, Russian Academy of Sciences, St. Petersburg, Russia; [2] Universita degli studi di Torino, Dipartimento di Lingue e Letterature straniere e Culture moderne, Turin, Italy; [3] Rijksuniversiteit Groningen, Campus Fryslân, Leeuwarden, The Netherlands

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, Ingrian Finnish). A correlation between the realisation of reduced vowels and their phonemic categorisation by speakers was studied. Correlated results showed that if V was pronounced >70%, its started loss was not yet perceived, apart for certain frequent elements, but after >70% of loss V was not perceived any more. A split of 50/50 between V and loss in production correlated to the same split in categorisation. At the beginning of sound change, the production is, therefore, more innovative, but after reanalysis the categorisation becomes more innovative and leads the change. Vowel 'a' was the most innovative in terms of loss, 'u/o' the most conservative, and 'i' in the middle, while the consonantal palatalisation was more salient than labialisation. These differences are grounded in their acoustics, articulation and perception.

## *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 and 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 and 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 in 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, etc. models (Mompéan-Gonzalez 2004, Gureckis and Goldston 2008, Johnson 2015, Davis and Poldrack 2013, 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: 148, Blythe and Croft 2012: 293). Its actuation is discussed (Baker et al. 2011, Stevens and Harrington 2014, Cohen 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 a nonmonotonic varialibity of an acoustic parameter (Stevens 2004). The usage-based approaches have also hypothesised that lexical frequency prompts sound change. The latter starts from frequent words and morphemes due to a higher level of automatisation in their production and can later spread throughout the entire phonological system (van Bergem 1995, Bybee 2001: 11-12, 2002, Bybee et al. 2016, Hay and Foulkes 2016, Kapatsinski 2018, Hall et al. 2018).

Sound change implies two connected processes: the change in the structure of the pool of phonetic realisations and the categorical reanalysis in the mind of the speakers/listeners. The temporal and causal correlation between these two processes is still unclear (Bybee 2001: 55). Modern phonology still has to reconcile the data on the continuous and variable nature of the phonetic signal and on the behavior of symbolic processes in a consistent fashion (Barnes 2006: 222, Kirby 2010: 149). In associative learning framework, the same question concerns the relations between typicality distributions in perception and frequency distributions in production (Kapatsinski 2018: 275). The concept of attractor landscape used in non-linear dynamic systems might be of use in modelling this link between continious 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), Kirschner (1998), 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 effort and the listener to be able to decipher the message. Lindblom (1990), for example, proposed a H&H framework, where a message varies in articulatory clarity being 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: 132).

Reduction does not affect all vowel qualities or positions in word or phrase equally, nor does it work always in the same direction. For example, word-final and especially phrase-final position manifests both vowel strengthening (lengthening and strengthening of articulation) and vowel weakening (devoicing, laryngealisation, nasalisation, 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: 286) opposes phonetically gradual reduction produced by automatisation of 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, 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 the maximal distinction etc. (Crosswhite 2004, Polka and Bohn 2003, 2011, Harris 2005, Johnson 2015, Manca and Grimaldi 2016). The 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 ATR contrasts (Barnes 2002, 2006, Flemming 2004). Reduced speech is characterised by the compression of the acoustic space between F1 and F2 through F1 raising, an effect of less jaw opening (Lindblom 1963, Uschanski 2005). The bottom-up direction of the compression suggests that high unstressed vowels would be less marked than non-high ones (Walker 2011: 29). The latter require more jaw opening and 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: 209, de Lacy 2006: 286).

The existence of differences between *i* and *u* is not much discussed in the surveys on vowel reduction. Some argue on the 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, is acoustically less reliable and perceived with significant help from the visual channel (Traunmüller and Öhström 2006, Eulitz and Obleser 2007, Vatakis et al. 2012, Manca and 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.

### *2.1. Aims, data and background of the study*

Correlations between the frequencies of various realisations of the three corner vowels in production and mental categorisation are explored in a comparative phonetic field study (2014-2016) on final vowel reduction and loss. We look 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: 128, Kuznetsova and Sidorkevič 2012: 565). They have been in a 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 Sourthern (**SI1**, **SI2**) variety of the Lower Luga dialect of Ingrian, and (5) Siberian Ingrian/Finnish (**S**). The data were obtained from one speaker per variety, with the exception of South Lower Luga Ingrian, for which two speakers were recorded (Table 1).

Table 1. Sociolinguistic data on the speakers

| Variety code | Language | Dialect | Subdialect | Sex | Birth year | 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 |

This is a limitation of our study, stipulated by little availability of fluent speakers able to participate in such experiments, as individual speakers even of the same language may display different reduction behavior (Hanique et al. 2015). General reduction patterns in Lower Luga and adjacent areas were, however, established prior to this experiment in Kuznetsova (2009, 2012a, 2016) on the basis of existing published sources, as well as audiodata 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 toward the Estonian language, which has been the most innovative in this respect and has completely lost reduced vowels. For example, the abovementioned varieties were ranged from the least to the 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* [ˈpysːy] > [ˈpysːy̆] > [ˈpysːẙ] > [pysʲʷː] > [pysʲː] > [pysː] ‘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).

The role of the lexical and grammatical factor at the initial stages of reduction, predicted by the usage-based approaches (see 1.1), has been also noticed previously 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: 73-74), and it occurred first of all in the most frequently used morphemes (Leskinen 1973: 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 generalized to purely phonetic (Kuznetsova 2016: 9-11).

The present experimental study was designed following the patterns established in the abovementioned works and was 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 less than ten 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, 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 : ī, ü : ǖ, u : ū, e : ē/ie, ö : ȫ/üö, õ : ȭ, o: ō/uo, ä : ǟ, a : ā*. Unrounded back vowel *õ:ȭ* is present only in Votic. The languages are characterised by significant prosodic differences between initial (stressed) and non-initial (unstressed) syllables. In certain varieties, long initial mid vowels were raised into diphthongs *ie, üö, uo*. Stems in all varieties are characterised by the fronting vowel harmony within the domain of the root plus the following derivative and inflectional suffixes, like in Standard Finnish (*a, o, u* can occur in back-vowel stems, *ä, ö, ü* in the front-vowel ones, “neutral” vowels *i* and *e* in both; for irregularities in Votic see e.g. van der Hulst 2018: 176-178).

In general, short vowels can undergo reduction and loss in 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: *tütt**ŏ̈*** [ˈtytːø̥̆] ‘girl’, *avaht**ŭ*** [ˈavaˑhtŭ̥] ‘be_opened:PST:3SG’, *sukk**ŭ**riš* [ˈsukːŭ̥ˌrʲisː] ‘sugar:IN’, *lisäh**ŭ̈**mine* [ˈlisäˑhʲẙˌmineˑ] ‘be_added:NMLZ’, *ihm**ĭ**sed* [ˈihmĭ̥sed] ‘man:PL’, *jāh**ŭ̈**teťä* [ˈjæːhʲẙteˌtːä] ‘cool:IPS’, *kisk**o**humine* [ˈkiskŏ̥huˌmineˑ] ‘be_torn_off:NMLZ’ (for a detailed chart of reduction and loss positions see Kuznetsova 2011: 189, 2012a: 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 vowels, *a*, *i*, *u* (or *o*) after both voiced (*n, l, r, m, v*) and voiceless (*t, k, p, s, 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 combinations of these vowels with consonants were covered. Based on existing phonological descriptions of these languages (Leppik 1975, Kuznetsova 2009, Markus and 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 morphonologically back (the few front-vowel stems, which contained only neutral vowels *i* and *e*, are underlined in Appendix 1).

Questionnaires were nearly identical (~5% of 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 1/3 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, ö, e* are raised to *u, ü, i* (Mägiste 1925: 80, Kuznetsova 2012a, 2012b, 2016), see e.g. *maito/maitu* ‘milk’, *pudro/pudru* ‘porridge’, *viero/vieru* ‘wheel’ in Appendix 1. Third, the loss of both *o* and *u* results in consonantal labialisation, so 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. Speakers had to translate the Finnic sentences with a carrier word from a phrase asked in Russian and repeat them four times. We subsequently counted the ratios of various types of vowel reflexes within each pool of realisations along several scales. The most general *binary scale* included two main types: (1) “vowel” and (2) “loss”. The latter were further divided into *six subtypes* in the following way:

— “vowel” (=“vocalic” realisations): (1) modal, (2) partially and (3) fully devoiced vowels;

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

For the first three speakers (IF, V, 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 the *three independent scales*:

— presence of strong laryngealisation: (1) yes, (2) no;

— vowel quality: (1) full, (2) partially and (3) completely reduced (to schwa);

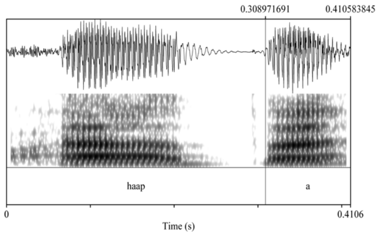
— devoicing: (1) modal, (2) slightly aspirated vowels, and vowels with (3) 10-30%, (4) 30-50%, (5) 50-70%, (6) 70-90% of devoicing, or (7) full devoicing.

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

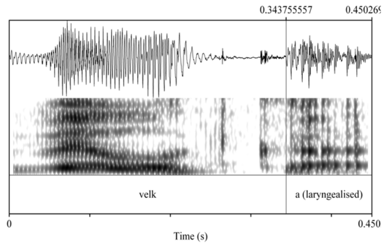
Vowel reflexes were classified manually on the basis of spectrographic data. Examples of main types of realisations are given in Figures 1 and 14-16. “Vocalic” types still preserved F1 and F2. If just one formant was present, the case was considered as “heavy segmental aspiration”. Being shorter than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

Figure 1. Examples of reduction in vocalic quality

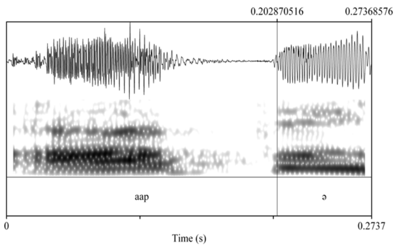
a. Full modal V (**hāpa* [haːpa] ‘aspen’) MO



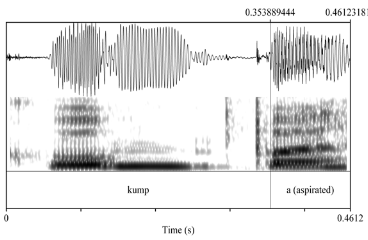

b. Laryngealised V (**velka* [vełka̰] ‘debt’) MO



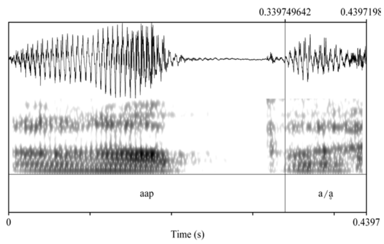

c. Schwa (**āpa* [aːpə] ‘aspen’) PK



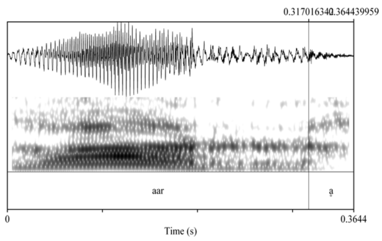

d. Slightly aspirated V (**kumpa* [kumpa$^h$] ‘which of the two’) MO




e. Partially devoiced V (**āpa* [aːpa̤] ‘aspen’) PK




f. Devoiced V (**āra* [aːrḁ] ‘branch’) PK




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

Neither variety has a literary standard, so such a test provided a unique possibility to observe more or less directly speakers’ intuitions about the presence/absence of a vowel word-finally. At the same time, a classical 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 Cyrillic and 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, (2) reduced vowels;

— “loss”: (3) retention of consonantal palatalisation or labialisation, (4) zero.

Palatalisation was coded by speakers with the use of the Russian “soft sign” ‘*ь*’ . The results on the categorisation of palalisation and labialialisation 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 with the means of Russian ‘*ы*’ (high unrounded mid vowel).

*3. Results and discussion*

The results showed 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 observed three stages of vowel loss (see Figure 11). At Stage 1, (IF, V, CI), the "vocalic" realisations comprised more than 90% of the sample, which correlated to their only one robust mental prototype [+SEGMENT]. At Stage 2 (SI1, SI2), there was a roughly 50/50 split both between the "vocalic" and "consonantal" realisations, on the one hand, and the [+SEGMENT] and [–SEGMENT] categorisations, on the other hand. At Stage 3 (S), with vowel loss in >70% of cases, only one [-SEGMENT] category prototype prevailed. These results clarify 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 a general discussion, 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 vs. 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; (2) Stage 2 and 3: reduced vowels (<80 ms) in both phrasal positions.

*3.1. Stage 1: Ingrian Finnish, Votic, Central Lower Luga Ingrian*

Samples at Stage 1 of reduction belong to three different languages: Finnish, Votic and Ingrian. Even if similar in the general structures of distributions, they exhibit slightly different configurations of vowel loss in realisation and categorisation (see the percetages of loss in Figures 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% of loss in production. What differs across the varieties in question, is the configuration of 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 other vowels.

A more detailed look at the vowel devoicing (Figures 2-4) and the reduction of quality (Figures 5-7), as well as lexical considerations (Table 4), clarify 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 and Klatt 1990, Laver 1994: 189–191, Barnes 2006: 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 Figures 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 (6: 'zero') of the new category [-SEGMENT] are therefore gaining strength in production first of all at the cost of the prototypical realisations (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.

Figure 2. IF: vowel devoicing

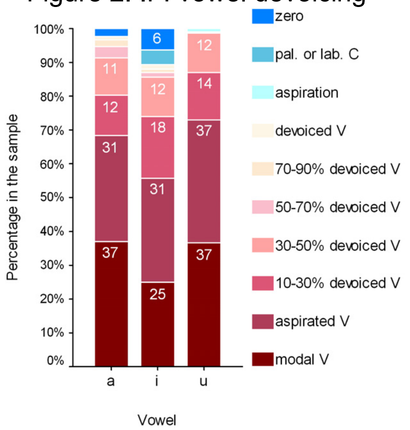



Figure 3. V: vowel devoicing

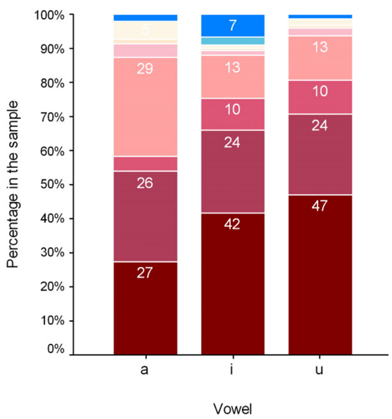
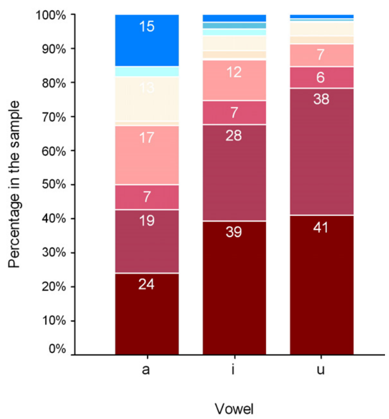



Figure 4. CI: vowel devoicing




Qualitative vowel reduction reveals quite a different picture (Figures 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 less than 30% of full vowels), and at the later stages on reduction it is realised as schwa almost invariably. On the contrary, **i* was the most liable to retain its quality

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

(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.

Figure 5. IF: qualitative reduction

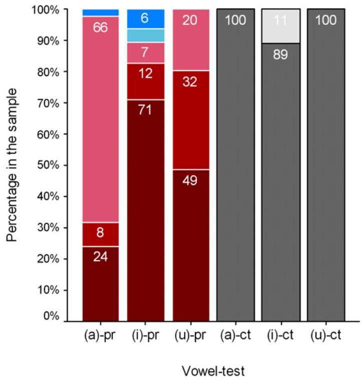



Figure 6. V: qualitative reduction

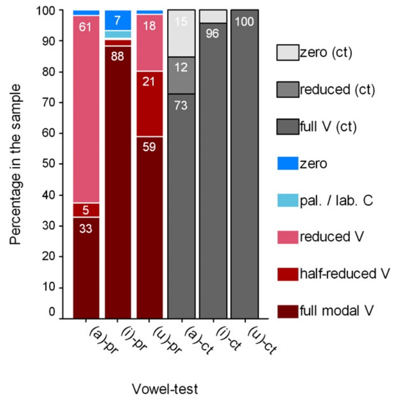
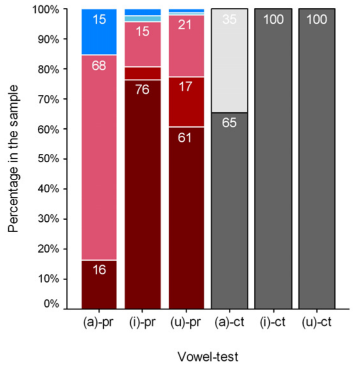



Figure 7. CI: qualitative reduction




Qualitative reduction and devoicing manifest very differently, sometimes opposing distributional patterns. We explored which of the two correlates best to 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 reduction, 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 in itself, as the level of loss in **i* is higher than in other vowel types (for non-phonetic factors see 3.4).

In Central Lower Luga Ingrian, on the contrary, 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 presents a mixed strategy between these two. As in IF, loss in the production and categorisation of **i* does not correlate with devoicing and negatively correlates with centralisation. At the same time, the configuration for **a* resembles that found in CI, although the equation is not perfect. The level of loss in categorisation positively correlates with those of qualitative reduction and devoicing. At the same time, the level of loss in production correlates with all three negatively. In other words, even if the speaker centralises and devoices **a*, this does not lead to an increased drop of this often voiceless schwa from her production. In fact, Votic **a* is the only vowel in the Stage 1 speakers which shows a clear negative correlation between the levels of loss in production and categorisation. The loss of reduced **a* has already started in the mental categorisation but not yet in the production of the Votic speaker, so she manifests also a more conservative production pattern for **a* than the CI speaker. She is 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 an unrounded mid back vowel *õ* in Votic, uniquely among these varieties (see 2.1). The Votic speaker might identify the schwa with this *õ*.

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

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

Speakers at Stages 2 and 3 of reduction (Figures 8-10) manifest the continuation of the same tendencies, especially those observed in CI. Speakers at Stage 2 belong to the same variety (South Lower Luga Ingrian), but the male speaker SI1 is 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 compared to 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.

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*, compared to SI2. Categorisation suggests a clue to the origins of this difference. In SI1, it is the most consistent of all six speakers (apart for the Siberian one, where the sound change process has already reached the terminal stage). He categorised the final **a*>*ə* always as zero, although he is actually not as consistent regarding schwas in non-final positions, where he often uses 'ы' (Kuznetsova 2012b), much as the Votic speaker does. Seemingly it is the closest perceptual Russian correlate for schwa for the speakers of local languages. Final **i* and **u*, on the contrary, are always perceived as vowels, though he is aware of their reduced character and calls them "half-vowels". He seemed to target these mental categorisations in his pronunciation consciously, and his percentage of loss is correspondingly higher for **a* and lower for **i* and **u* than in the otherwise linguistically very close SI2. He was obviously not able to attain full control over his production, though, and his pattern of loss for the three vowels still has a scalar shape similar to other Ingrian speakers (CI and SI2).

Figure 8. SI1: vowel devoicing

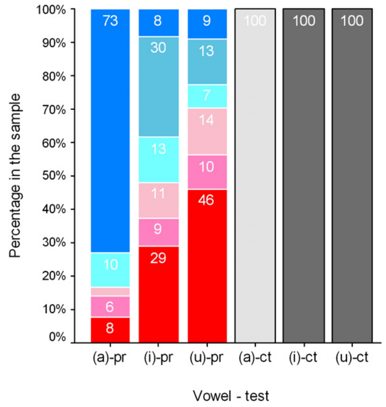



Figure 9. SI2: vowel devoicing

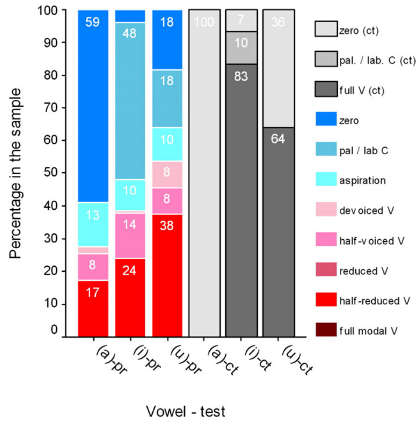
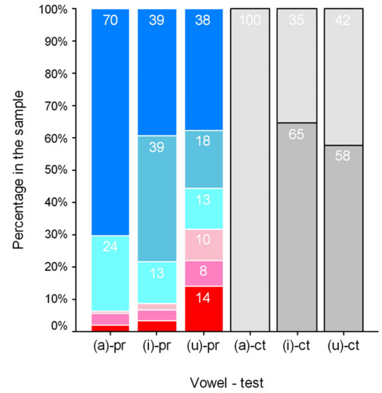



Figure 10. S: vowel devoicing




In SI2, the phonemic categorisation (as zero) is consistent only for the vowel *a, which has reached the critical threshold for complete loss. Interestingly, *i and *u show reverse patterns of loss in her production vs. 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). Vowel *u also manifests a visible cluster of labialisation, completely absent at Stage 1, but it accounts 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 the *i and *u of both SI speakers match those of categorisation much closer. It seems that an 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 perceive and target, more or less conciously, the full vowel. This might be linked to a robust salience of *i-reflexes both in perception and articulation (see 4.2). Speakers succeed in reaching a vowel in only about half of cases, though, ending up with a more or less aspirated palatalised final consonant in the other half instead. Less perceptual salience of *u can be seen in its relatively innovative categorisation by SI2, which is in general is not as systematic as that by SI1. In an even less perceptually salient *a, however, a relatively robust cluster of consonantal aspiration does not prevent a complete loss from categorisation, as *a does not colour consonants.

In the Siberian speaker at Stage 3, we see the next step of the same processes. Here, all the vowels have already reached the critical threshold for loss in production in order to be lost from mental categorisation. Judging by all the three samples at Stages 2 and 3, one could estimate this threshold at about 70%. Categorisation became 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 follows the one of *a, i.e. the middle step of the "ladder-like" pattern flattened. We still see a more conservative production for *u with respect to 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, storage in the memory, see 1.2, 4.2). At the same time, while in the Siberian speaker the palatalisation cluster is yet as big as the cluster of complete loss, the labialisation cluster is already twice as small as the latter. One might hypothesise that while palatalisation could still have a chance to be preserved as a phonemic feature in these languages, labialisation has already lost the historical sound change battle, even though the Siberian speaker still perceives its presence. The aspiration cluster as a reflex of *a is especially robust in the Siberian speaker, but as at Stage 2, this does not affect any more 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) vs. "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 for the first comparison. The overall production ($N$=3744, $M$=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, 4221)=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 SE difference can be explained by the unequal size of the groups). For other cases, we ran two post-hoc tests for pairwise within-group comparisons in big samples of unequal size and variance: *Tamhane's T2* (more conservative) and *Games-Howell* (more liberal) at the 95% confidence interval. Few differences between them are reported below as *T / GH*.

There was a significant effect of the stage on the level of loss: $F$(5, 4217)=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 a categorial analysis: categorisation becomes significantly more innovative and drives the loss at the terminal stage.

Overall results on the three vowels showed a highly significant difference ($F$(2, 4220)=77.49, $p$<0.001), as each vowel has its unique congiruration of loss in production and categorisation (Table 2, Figure 12). Post-hoc tests on these two aspects analysed separately showed that both the production and 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.

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** (p=0.006) | 0.21*** | -0.11 (p=0.086 T / 0.065 GH) | 0.21*** | 0.22*** |
| i (pr) | | 0.14*** | -0.18*** | 0.14** (p=0.002) | 0.15** (p=0.002 T / 0.001 GH) |
| u (pr) | | | -0.32*** | 0.0 | 0.01 |
| a (ct) | | | | 0.32*** | 0.33*** |
| i (ct) | ***p<0.001, **p<0.01 | | | | 0.01 |

Overall results on the six speakers lacked of any difference in IF and V (*MD*<0.01, *p*=1), but CI differed from both (CI–IF: *MD*=0.05, *p*=0.003 *T* / 0.002 *GH*; CI–V: *MD*=0.05, *p*=0.002). CI manifests a little more advanced substage of reduction inside Stage 1, where **a* takes over **i* as the leader of loss. At 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). Reduction and categorisation are further analysed across speakers in Table 8 (see also Figures 11, 13). At Stage 1, both the production and categorisation by IF and V, as well as the categorisation by CI did not differ and were conservative. The production by CI was slightly more innovative: it showed moderate-to-weak difference from the production by IF and V (but not from the categorisation by CI). The production and categorisation by S highly differed from everything else and from each other (her production is significantly more conservative than production due to the categorial shift at Stage 3). At the intermediate Stage 2, the production and categorisation by SI1 and the production by SI2 did not show significant differences. However, the categorisation by SI1 stands out of all effects in Table 8. The peculiarity of SI1 categorisation is likely linked to his unusual linguistic awareness and full systematicity in transcription (**a* as zero, **i* and **u* as vowels, see 3.2).

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

| Sp | 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(*) (p=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* (p=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** (p=0.001) | 0.4*** | 0.91*** |
| SI1(pr) | | | | 0.05 | 0.29*** | 0.51*** | 0.48*** | 0.43*** | *0.2 (p=0.042 T / 0.03 GH) | 0.06 | 0.45*** |
| SI2(pr) | | | | | 0.24*** | 0.56*** | 0.53*** | 0.48*** | **0.25 (p=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 (p=0.001 T / <0.001 GH) | -0.42*** | -0.93*** |
| CI(ct) | | | | | | | | | *0.23 (p=0.047 T / 0.034 GH) | -0.38*** | -0.88*** |
| SI1(ct) | | | | | | | | | | -0.15 | -0.65*** |
| SI2(ct) | ***p<0.001, **p<0.01, *p<0.05 | | | | | | | | | | -0.51*** |

### *3.4. Lexical factor in reduction at Stage 1*

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 correlation between production and categorisation. Table 4 gives data on 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 a number of lexemes in which at least one token of the "loss" in production is attested, their percentage in the sample for each speaker, and their distribution across **a*, **i* and **u/o* types. The vowel types are arranged in parentheses from those with the highest number of words exhibiting loss to those with the lowest number. Column 3 cites only those words in which more than half the tokens show 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 4. Lexical specification of vowel loss at Stage 1

| Sp (N) | N of lexemes exhibiting phonetic loss | Lexemes with phonetic loss *n>4* (>50%) | N of lexemes with phonetic loss >50% | Lexical compactness of loss | Lexemes which exhibited loss in categorisation (by glosses) |
|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 |
| IF (78) | **12** (15,3%) (*i, a*) | *pien**i*** 'small' (6), *ūs**i*** 'new' (8) | **2** (2,6%) | **0,17** | **i**: small, child (*i*/Ø), elk |
| V (81) | **10** (12,3%) (*i, a, u/o*) | *sūr**i*** 'big' (6), *pēn**i*** 'small' (6) | **2** (2,5%) | **0,2** | **i**: big (*i*/Ø); **a:** dog, change:IMP, barley, nail, bath broom |
| CI (76) | **21** (27,6%) (*a, i, u/o*) | *nāgl**a*** 'nail' (6) | **1** (1,3%) | **0,05** | **a:** nail, leg, which (of the two), black (*a*/Ø), change:IMP, debit, twig, bath broom, broom, class, skinny |

In IF and V, the level of the lexical compactness of loss is four times higher than in CI. The vowel loss in these speakers is concentrated in very few frequent basic words, while in CI the lexical dispersion of loss is much higher. It is remarkable that in IF and V, the loss in frequent words concerns 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', *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 more innovative variety where reduction is already conditioned phonetically, **a* becomes the reduction leader.

These differences in production find a parallel in categorisation (see Column 6; words encoded with V/Ø were cited by the respective speakers as having both a variant with a vowel and without it). In Ingrian Finnish, the loss was perceived only in *i*-words, in Votic, both in *i*- and in *a*-words, and in CI, only in *a*-words. The number of lexical items with perceived loss also increases from IF to CI. It is worth noting that the match between production and categorisation is close in a statistical sense, but not in the lexical one. 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 the 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 vowels, and only one speaker was taken for each variety apart 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 the category (Gureckis and Goldston 2008, Kirby 2010: 34-37).

The main vectors summarizing 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 (mean values of each test on the scale between 0="vowel" and 1="loss" and the SE bars are given; red stands for production and grey for categorisation). The dotted horizontal lines are drawn at 70% of loss and at 70% of preservation of segments, which appeared to 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 on Figure 11. At Stage 1, production and categorisation are closely matched. As discussed in 3.4, vowel reduction and especially loss at this initial stage (speakers IF, 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 a 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, i.e. 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 automatisation 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 vowel (Stage 2). When the new category arrived at more than 70% of realisations (Stage 3 and some vowels at Stage 2), the crucial jump in categorisation happened. The pattern drastically reversed: the old value was 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 automatisation-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* is the most advanced in terms of loss, and here, in turn, categorisation is more innovative than realisation. This is obviously linked to the fact of its extremely strong qualitative reduction and that it does not leave any colour to the consonants. In general, the more the vowels were reduced to schwa, the less their presense 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 a sharp distinction made in Crosswhite (2004). In the course of vowel reduction and loss in the Finnic languages of Ingria we observed the elements of both patterns. Mid vowels *o, ö, e* are raised to high vowels *u, ü, i*, which can be seen in variants *maito/maitu* 'milk', *pudro/pudru* 'porridge', *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 individual speakers. Both the speakers and the vowels are placed starting from the most conservative to the most innovative ones, which gives an idea of an S-curve of the sound change. The chart shows that the reverse in a 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, sped forward by its innovative categorisation. 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.

Figure 11. Summary: speakers

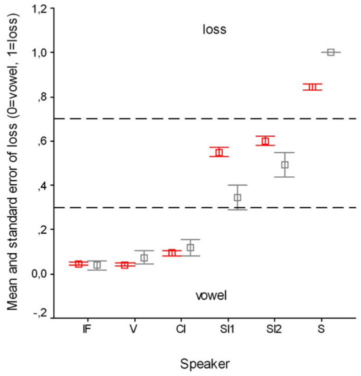



Figure 12. Summary: vowels

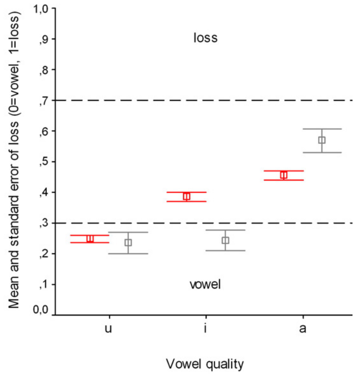
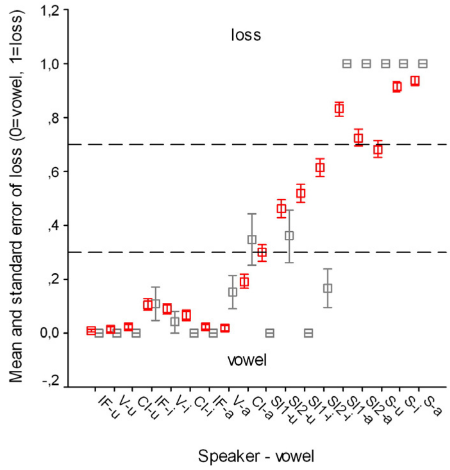



Figure 13. Summary: speakers and vowels




*4.2. Formation and loss of consonantal palatalisation, labialisation, and aspiration*

Already at Stage 1, 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 is yet lexically and grammatically stipulated to a large extend (see 3.4). Later on, as loss is generalised throughout the entire phonetic system, *a takes over as a 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 contrary, 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 percentage first of all at the cost of the prototypical variants (full clear modal vowels) of the old category [+SEGMENT]. 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 from 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 vowel, but labialisation already rather as loss. This is likely linked to a higher acoustic, articulatory, and therefore perceptual salience of *i and palatisation over *u and labalisation.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 lost from the language slower. The possible impact of the front vowel harmony characteristic of the Finnic languages of Ingria in 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 the phonemic consonantal labialisation over the loss of the segmental vowel at the later stages of reduction. The vowel *u is the most conservative among the three vowels in terms of vowel loss, as it retains the largest “vocalic” cluster of realisations in all speakers, but it manifests gradual transitions between the stages in all aspects: qualitative reduction, devoicing and loss. No robust cluster of consonantal labialisation as the trace of *u is formed, the segmental vowel is rather directly lost. 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.

A difference in the size of consonantal palatalisation and labialisation clusters at 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 a 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 Figures 14 and 15). The labialised aspiration is then “eaten” away by reduction, and the consonant remains plain.

Consonantal palatalisation (Figure 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 and Urek 2016). Our preliminary observations show that the number of palatalised consonants and the degree of their 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 are, the bigger number of palatalised consonantal phonemes could be distinguished for any particular variety and the stronger the palatalisation is from the phonetic point of view. Ingrian Finnish and Soikkola Ingrian have dental palatalised 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 the 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 in pronounced (viz. Leppik 1975: 116-117, Kuznetsova 2009: 195-235, Markus and Rožanskij 2011: 17-18).

These observations are to be further verified, but one could hypothesise this trend to be a result of re-phonologisation. Front high vowel quality, which originally stipulated 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 a well-formed consonantal palatalisation. Consonants *p, k, n* manifested palatalisation in 30-60% of cases, and consonants *s, h, r, n, m, v* only sporadically. In these two groups, especially in the last one, the tendency towards complete depalatalisation was observed.

Figure 14. Aspirated labialised C (**lastu* [łast$^{\circ h}$] 'chip') PS

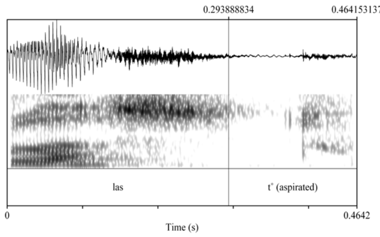



Figure 15. Aspirated C (**lasta* [łast$^{h}$] 'child:PRT') PS

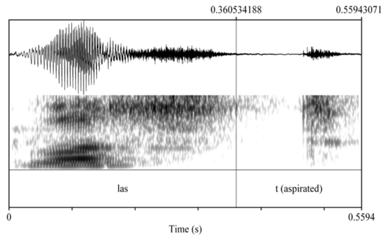
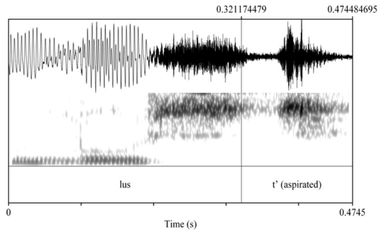



Figure 16. Aspirated palatalised C (**lusti* [łust$^{jh}$] 'beautiful') AU




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 neighboring 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 and 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: 15-16, Kiparsky 1963) or Irish (Greene 1973, Anderson 2016). Labialised consonants are much rarer cross-linguistically than palatalised ones. Blevins (2004: 204) explains the rarity of certain phonological contrasts through the uncommon occurrence of sound changes giving rise to them. Consonantal palatalisation accounted for 145 (3.18%) cases in the PBase (Mielke 2008, Brohan and Mielke 2018: 218-219), being one of the most frequent types of sound change, while labialisation included just 38 entries.

We do not know of cases of the emergence of phonemic consonantal aspiration as a result of the loss of the following plain schwa, a reflex of **a* (nor the P-base gives such examples). This is just an additional indication of a very low perceptual saliency of this schwa, especially in the final positions. In our case study, this leads to its fast disappearance both from the mind and the production of speakers. One might still wonder if, at Stages 2 and 3, it is still possible to distinguish synchronically 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 the speakers at Stages 2 and 3 still perceive reflexes of **i* and **u*, but none perceives 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: 223-238, see also 1.1).

## *5. Conclusion*

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 usually involve 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 request was rather 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 (only if the intermediary orthography has a segmental principle of encoding).

At the initial stage, vowel reduction and loss are linked to the automatisation 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 speakers pronounced a vowel in more than 70% of cases, they typically perceived its presence. A decisive qualitative shift seemed to happen in categorisation after the new realisational types have gained more than 70%. 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 now meaningless parts, and this last stage can contain near-mergers. Automatisation of execution is likely still be at work here, only now the production of the new category is being automatised. The comparison of several stages of vowel reduction and loss revealed no irreconcilable contradiction between the two main patterns of reduction, centrifugal and centripetal. Both were observed in our data: the rise of mid 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 suggested two types of the markedness hierarchies of these vowels. As regards the process of reduction and loss itself, the hierarchy of vowels (from the most to the least innovative) is *a>i>u*. The phonological saliency of secondary consonantal localisations emerging in the process of loss of these vowels, in turn, would suggest the hierarchy *a>u>i* (from the least to the most salient secondary localisations). Some possible physiological features stipulating both hierarchies were outlined in the paper.

**Appendix 1: List of the most typical carrier words** (morphonologically front-vowel stems are underlined):

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

**Abbreviations**: 1 — 1 person; 3 — 3 person; ct — categorisation; IN — inessive; IMP — imperative; IPS — impersonal; NMLZ — nominalisation; PL — plural; pr — production; PRS — non-past; PRT — partitive; PST — past; SG — singular.

**Author contributions**: Natalia Kuznetsova conceived and performed the study, apart for the half of the acoustic analysis of the data and the PRAAT script for data extraction performed by Vasilisa Verkhodanova.

**Acknowledgements**: Project was supported by the Grant of the President of Russian Federation for young PhDs, no. MK-4291.2015.6 “Non-initial vowel reduction in the vanishing Finnic varieties of Ingria”. I am grateful to Maxim Fedotov for his help with a preliminary segmentation of a part of the data, to Cormac Anderson, Brittany Blankinship, four anonymous reviewers and the associate editors of this Special Issue for their valuable comments.

**References**

Anderson C (2016): Consonant colour and vocalism in the history of Irish. PhD diss. Poznań: Adam Mickiewicz University in Poznań.

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):347–74.

Barnes J (2006): *Strength and weakness at the interface: Positional neutralisation in phonetics and phonology*. Berlin; New York: Mouton de Gruyter.

Baudouin de Courtenay JN (1895): *Versuch einer Theorie phonetischer Alternationen: ein Capitel aus der Psychophonetik*. Strassburg: K.J. Trübner.

van Bergem DR (1995): Perceptual and acoustic aspects of lexical vowel reduction, a sound change in progress. *Speech Communication* 16(4):329–58.

Blevins J (2004): *Evolutionary phonology: The emergence of sound patterns*. Cambridge University Press.

Blythe RA, Croft W (2012): S-curves and the mechanisms of propagation in language change. *Language* 88(2):269–304.

Brohan A, Mielke J (2018): Frequent segmental alternations in P-base 3. In: Hyman L, Plank F (eds.) *Phonological typology.* Walter de Gruyter, pp 196-228.

Bybee J (2001): *Phonology and language use*. Cambridge University Press.

Bybee J (2015): *Language change*. Cambridge University Press.

Bybee J, File-Muriel RJ, de Souza RN (2016): Special reduction: a usage-based approach. *Language and Cognition* 8(3):421–46.

Cohen Priva U (2017): Informativity and the actuation of lenition. *Language* 93(3):569–97.

Crosswhite K (2004): Vowel reduction. In: Hayes B, Kirchner R, Steriade D (eds). *Phonetically-based phonology*. Cambridge University Press, pp 191-231.

Davis T, Poldrack RA (2013): Quantifying the internal structure of categories using a neural typicality measure. *Cerebral Cortex* 24(7):1720–37.

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.

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.

Flemming E (1995): Auditory representations in phonology. PhD diss. UCLA.

Flemming E (2004): Contrast and perceptual distinctiveness. In: Hayes B, Kirchner R, Steriade D (eds). *Phonetically-based phonology*. Cambridge University Press, pp 232–76.

Greene D (1973): The growth of palatalisation in Irish. *Transactions of the Philological Society* 72(1):127–36.

Gureckis TM, Goldstone RL (2008): The effect of the internal structure of categories on perception. *Proceedings of the 30th annual conference of the cognitive science society*. Austin, TX: Cognitive Science Society, 1876–81.

Hall KC, Hume E, Jaeger FT, 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):161–85.

Harris J (2005): Vowel reduction as information loss. In: Carr P, Durand J, Ewen CJ (eds.). *Headhood, elements, specification and contrastivity: Phonological papers in honour of John Anderson*. John Benjamins Publishing, pp 119–32.

Hay J, Foulkes P (2016): The evolution of medial /t/ over real and remembered time. *Language* 92(2):298–330.

Hulst H van der (2018): Asymmetries in vowel harmony: A representational account. Oxford University Press.

Hyman LM (1976): Phonologisation. In Juilland A (ed). *Linguistic studies presented to Joseph H. Greenberg*. Saratoga: Anma Libri, pp 407–18.

Johnson K (2015): Vowel perception asymmetry in auditory and phonemic listening. UC Berkeley Phonology Lab Annual Report 2015. Available online at https://escholarship.org/uc/item/21t337g.

Kapatsinski V (2018): *Changing minds changing tools: From learning theory to language acquisition to language change*. MIT Press.

Kiparsky V (1963): *Russische historische Grammatik. Band. I: Die Entwicklung des Lautsystems*. Heidelberg.

Kirby JP (2010): Cue selection and category restructuring in sound change. PhD diss. University of Chicago.

Kirchner RM (1998): An effort-based approach to consonant lenition. PhD diss. UCLA.

Klatt DH, Klatt LC (1990): Analysis, synthesis, and perception of voice quality variations among female and male talkers. *The Journal of the Acoustical Society of America* 87(2):820–57.

Kochetov A (2011): Palatalisation. In: van Oostendorp M, Ewen CJ, Rice K, Hume EV (eds). *The Blackwell companion to phonology. Vol. 3*. John Wiley and Sons, pp 1666–90.

Krämer M, Urek O (2016): Perspectives on palatalisation. *Glossa* 1(1):31.1–17.

Kruszewski NV (1883): *Očerk nauki o jazyke*. Kazanj.

Kuznetsova N (2009): Fonologičeskie sistemy ižorskih dialektov. PhD diss. St. Peterburg: Institut lingvističeskih issledovanij RAN.

Kuznetsova N (2011): Typology of prosodic systems in Low Luga Izhorian varieties. In: Austin PK, Bond O, Charette M, Nathan D, Sells P (eds). *Proceedings of Conference on Language Documentation and Linguistic Theory 2*. London: SOAS, pp 185–192.

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. *'Folk linguistics': Language from speakers' perspective: Paper abstracts for the conference, St. Petersburg, November 19–21*. St. Petersburg, 85–7.

Kuznetsova N (2015): Two phonological rarities in Ingrian dialects. In: Hilpert M, Duke J, Mertzlufft Ch, Östman J-O, Rießler M (eds). *New trends in Nordic and general linguistics*. Berlin: De Gruyter, pp 91–117.

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.

Kuznetsova N, Markus E, Muslimov M (2015): Finnic minorities of Ingria: The current sociolinguistic situation and its background. In: Marten H, Riessler M, Saarikivi J, Toivanen R (eds). *Cultural and linguistic minorities in the Russian Federation and the European Union*. Berlin: Springer, pp 127–167.

Kuznetsova N, Sidorkevič D (2012): Priloženie 6: Karty k statjam. *Acta linguistica Petropolitana* VIII(1):559–67.

Laanest A (1980): Inkeroismurteiston suhteista suomen kaakkoismurteisiin. *Virittäjä* 84:142–9.

Labov W (1999): *Principles of linguistic change. Vol. 1: Internal factors*. Blackwell.

Labov W, Karen M, Miller C (1991): Near-mergers and the suspension of phonemic contrast. *Language variation and change* 3(1):33–74.

de Lacy P (2006): *Markedness: Reduction and preservation in phonology*. Cambridge University Press.

Laver J (1994): *Principles of phonetics*. Cambridge University Press.

Leppik M (1975): *Ingerisoome kurgola murde fonoloogilise süsteemi kujunemine*. Tallinn: Eesti NSV Teaduste Akadeemia, Keele ja Kirjanduse Instituut.

Leskinen H (1973): Kaakkoissuomalaisen loppuheiton yleisyydestä ja alkuperästä. *Suomalais-Ugrilaisen Seuran Aikakauskirja* 72:210–21.

Lindblom B (1963): Spectrographic study of vowel reduction. *The journal of the Acoustical society of America* 35(11):1773–81.

Lindblom B (1990): Explaining phonetic variation: A sketch of the H&H theory. In: Hardcastle WJ, Marchal A (eds). *Speech production and speech modelling*. Dordrecht: Springer, pp 403–439.

Mägiste J (1925): *Rosona (Eesti Ingeri) murde pääjooned*. Tartu.

Manca AD, 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.
Markus E, Rožanskij F (2011): *Sovremennyj vodskij jazyk. Teksty i grammatičeskij očerk. Vol. 2*. Sankt-Peterburg: Nestor-Istorija.
Matthies M, Perrier P, Perkell JS, Zandipour M (2001): Variation in anticipatory coarticulation with changes in clarity and rate. *Journal of Speech, Language, and Hearing Research* 44(2):340–53.
Maye J, Gerken L (2000): Learning phonemes without minimal pairs. *Proceedings of the Annual Boston University Conference on Language Development* 24(2):522–33.
Mielke J (2008): The emergence of distinctive features. Oxford University press. PBase available from http://pbase.phon.chass.ncsu.edu, last accessed on 01.10.2018.
Mompeán-González JA (2004): Category overlap and neutralisation: The importance of speakers' classifications in phonology. *Cognitive Linguistics* 15(4):429–69.
Muslimov M (2005): Jazykovyje kontakty v Zapadnoj Ingermanlandii (nižnee tečenije reki Lugi). PhD diss. Sankt-Peterburg: Institut lingvističeskih issledovanij RAN.
Nirvi RE (1972): Siperian inkeriläisten murteesta ja alkuperästä. *Kotiseutu* 2/3:92–5.
Ohala JJ (1989): Sound change is drawn from a pool of synchronic variation. In: Breivik LE, Jahr EH (eds). *Language change: Contributions to the study of its causes*. Walter de Gruyter, pp 173–98.
Olejarczuk P, Kapatsinski V, Baayen RH (2018) Distributional learning is error-driven: The role of surprise in the acquisition of phonetic categories. *Linguistics Vanguard* 4(S2).
Pearce M (2008): Vowel harmony domains and vowel undershoot. *UCL Working Papers in Linguistics* 20:115–40.
Polka L, Bohn OS (2003): Asymmetries in vowel perception. *Speech Communication* 41:221–31.
Polka L, Bohn OS (2011): Natural Referent Vowel (NRV) framework: An emerging view of early phonetic development. *Journal of Phonetics* 39(4):467–78.
Ritter S, Mücke D, Grice M (2018): The dynamics of intonation: categorical and continuous variation in an attractor-based model (preprint). Available online at https://doi.org/10.31234/osf.io/qkm6b.
Roettger TB, Winter B, Grawunder S, Kirby J, Grice M. (2014): Assessing incomplete neutralisation of final devoicing in German. *Journal of Phonetics* 43:11–25.
Rosch E (1978): Principles of categorisation. In: Rosch E, Lloyd BB (eds). *Cognition and categorisation*. Hillsdale, NJ: Lawrence Erlbaum Associates, pp 27-48.
Šahmatov AA (1915): *Očerk drevneišego perioda istorii russkogo jazyka*. Petrograd.
Sidorkevič DV (2013): Jazyk ingermanlandskih pereselencev XIX v. v Sibiri (struktura, dialektnye osobennosti, kontaktnye javlenija. PhD diss. Sankt-Peterburg: Institut lingvističeskih issledovanij RAN.
Stevens KN (2004): Invariance and variability in speech: Interpreting acoustic evidence. In: Slifka J, Manuel S, Matthies M (eds). *From sound to sense: 50+ years of discoveries in speech communication*. Cambridge, MA: Research Laboratory of Electronics, MIT, pp B-77–B-85.
Stevens M, Harrington J (2014): The individual and the actuation of sound change. *Loquens* 1(1):e003.
Szeredi D (2010): Vowel centralisation and vowel harmony in Hungarian. *The Odd Yearbook* 8:111–37.
Teras P, Pajusalu K (2014): Palatalisatsioonist ja prepalatalisatsioonist spontaanses eesti keeles. *Keel ja Kirjandus* 4:257-69.
Traunmüller H, Öhrström N (2007): Audiovisual perception of openness and lip rounding in front vowels. *Journal of Phonetics* 35(2):244–58.
Uchanski RM (2005): Clear speech. In: Pisoni DB, Remez RE (eds). *The handbook of speech perception*. Oxford: Blackwell, pp 207–235.
Vallabha GK, McClelland JL, Pons F, Werker JF, Amano S. (2007): Unsupervised learning of vowel categories from infant-directed speech. *Proceedings of the National Academy of Sciences* 104(33):13273–8.
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. doi: 10.3389/fnint.2012.00071.
Walker R (2011): *Vowel patterns in language*. Cambridge University Press.
Wanrooij K, Escudero P, Raijmakers ME (2013): What do listeners learn from exposure to a vowel distribution? An analysis of listening strategies in distributional learning. *Journal of Phonetics* 41(5):307–19.
Yu AC (2013): Socio-cognitive processing and the actuation of Sound change. In: Yu AC (ed). *Origins of sound change: Approaches to phonologisation*. Oxford University Press, pp 201–227.