Sound contraction in Russian spontaneous speech and its implications for spoken word recognition

Yulia Nigmatulina
Saint Petersburg State University
julia.nigmatic@yandex.ru

Abstract
This study, based on data from spontaneous Russian and on data from read Russian, quantifies all possible vowel as well as some consonant contractions and investigates patterns of their formation. The results reveal no strict rules that could predict contraction; however, stress presence seems to play a significant role for vowels: unstressed vowels contract easier than stressed ones. At the same time duration was not confirmed to be a reliable feature that a listener could use in the process of recognition.

Keywords*: Spontaneous speech; sound contractions; speech segmentation.*

1. Introduction

Word-external and word-internal sound contraction is a known feature of spontaneous speech (Nespor et al. 1987; Casali 1997; Smith 2008). The present paper reports on the first known study of word-external sound contraction in spontaneous Russian and shows that there are different patterns, word-externally vs word-internally, with regard to which sounds contract and where contraction occurs. Quantitative data for all vowel and some consonant word-external contractions are presented. Some observations about what can influence occurrence of contraction are also made.

The article is structured as follows: firstly, contraction is defined; then the data are presented. Then the results concerning vowel and consonant contractions in spontaneous Russian are described. Finally, contractions in spontaneous and read speech are compared in order to answer the question whether contractions are inherent for speech in general.

2. Contractions

In spontaneous continuous speech word-final and the following word-initial sounds can contract and, as a result, a new sound appears at the boundary of two words. Hence, contraction causes reorganization of syllables and diffusiveness of word boundaries. The contraction phenomenon at word boundaries is studied in order to understand how listeners manage to establish word boundaries in cases of contraction.

By sound contraction I understand the contact of two adjoining sounds which can lead to the formation of a single sound. Consider these examples from Russian:

(1) *yem**u_u**zhe*** [iˈm**u**ʐ̩] 'he-DAT_already'
(2) *n’**eo**td’el’imy* [nʲ**æ**dʲiˈlʲima] 'inseparable-PL'
(3) *ra**z_s**pas’ibo*** [ˈra**s**paˈsʲyvə] 'time-SG_thanks'
(4) *d’et’**i_o**ttuda* [ˈdʲɛtʲætuda] 'children_therefrom'

Contraction of contrasting sounds often leads to the formation of a sound qualitatively different from both original ones, exemplified in (2). In the present paper the following conditions for contractions are regarded: 1) for vowels – hiatus; 2) for consonants – a group of identical consonants.

Previous studies of consonant groups in Russian that can lead to contraction were mostly using word-internal data (Kasatkin and Choy 2005). Although hiatus at word boundaries can be observed in Russian speech (Zemskaya 1973, 1979), contractions at word boundaries were considered not typical for Russian literary language because of the lack of spontaneous data for analysis.

3. Data

The present research is based on The Corpus of Spontaneous Russian that consists of radio interviews and TV talk-shows. All recordings (115 minutes and 15106 items (words) have orthographic and phonetic transcriptions. The transcriptions were manually made by a group of trained phoneticians at Saint Petersburg State University. For phonetic transcriptions the experts used spectrogram and waveform displays, as well as auditory cues. The transcription process partly involved a verification process. In cases of disagreement the preference was given to the results of spectral analysis.

To avoid influence of grammatical and lexical knowledge on the results of phonetic transcription all the recordings were considered in 1-2 syllable fragments.

4. Method

All letter combinations that correlated with sounds that could undergo contraction were extracted from the orthographic transcription of the texts. Examples (5)-(6) show vowel combinations (hiatus) and examples (7)-(8) show consonant combinations.

(5) *byl**a** **O**l'ga* [bəˈl**au**lʲgə] 'was-FEM_Olga'
(6) *dayt'**e** **A**l'eksandru* [ˈdʲɛtʲ**e a**lʲiksandru] 'give-IMP.PRS.2PL_Alexander-DAT';
(7) *plat'at_tam* [ˈpladʲə**t**am] 'pay-PRS.3PL_there'
(8) *khot'elos'_skazat'* [xaˈtʲœ**s**kaˈzatʲ] 'would like_say-INF'

Only consonants with the same place and manner of articulation were regarded; voiced/voiceless and palatalized/non-palatalized consonants were not distinguished, as voice assimilation and palatalization assimilation can easily result in a single sound instead of two initial sounds.

Word combinations as exemplified above were divided into two groups depending on whether contraction did or did not occur. (In the phonetic transcription such word combinations were written as a single word; in orthography: two words were joined with underscore). The examples with marked contraction were then checked one more time.[1] Vowel contraction was stated, if there was no sufficient change in formant frequencies. Figures 1 and 2 show spectrograms of examples with uncontracted and contracted sounds. If formant contour is uniform to the extent of the fragment that depicts two supposed adjacent sounds (no gaps, falls or rises are noticed), then contraction is fixed. If there are marks of formant frequencies change, more than one sound is considered.

[1] Spectograms were used to verify if a contraction had taken place.

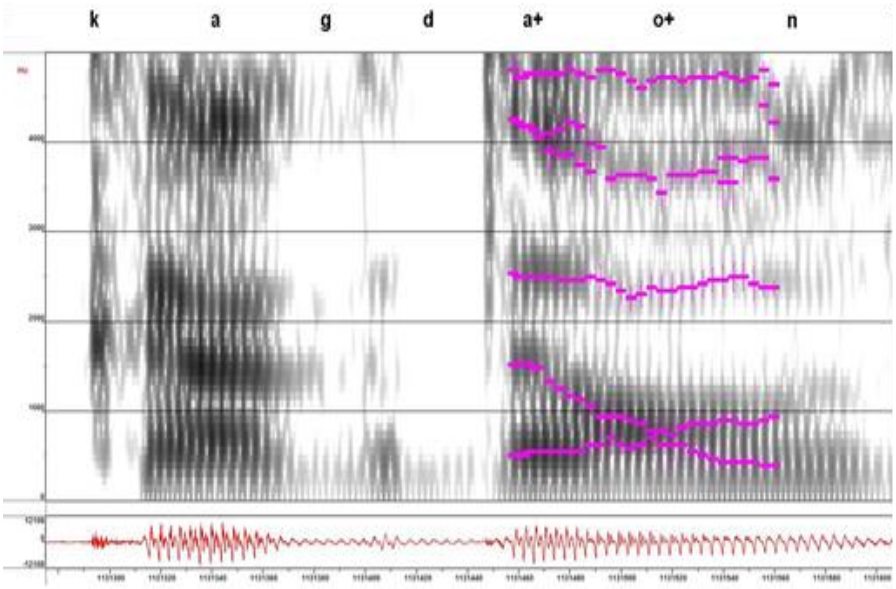

Figure 1. Spectogram of uncontracted *kogda_on* [kagˈda ˈon] ‘when_he’.

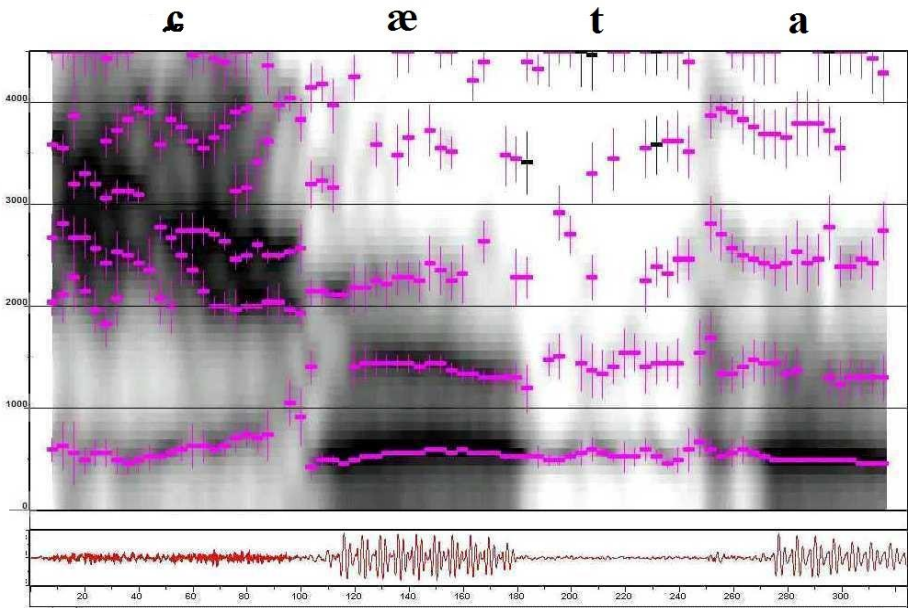

Figure 2. Spectogram of the contraction shchi_oto [ɕæta], instead from [ɕi ata] (from the word combination *v’shchi_ oto-zhd’estvl’at’* [ˈvʲɛɕætazʑdʲistˈlʲætʲ] ‘tings_identify-INF).

During the work with the material determining elisions from contractions turned out to be difficult. Therefore, the process was simplified by considering only two groups: a group with one pronounced sound, which includes both elision and contraction types of instances, and a group with two pronounced sounds.

For comparison, in 66 minutes of the data, contractions inside the words were also considered. The results are presented in Table 1.

Table 1. Number of contractions and places for potential contractions.

	Vowels		Consonants	
	Total number of instances	Contrac-tions	Total number of instances	Contrac-tions
At word boundary (115 min analyzed)	559	310 (55.4%)	143	106 (74.1%)
Inside words (66 min analyzed)	317	173 (54.6%)	182	173 (95.1%)

Table 1 shows that in general the vowel contractions at word boundaries as well as inside words occur with approximately the same rate (54-55%). As for consonants, contraction is more frequent word-internally than at word boundaries. Comparison of the percentage of contractions for vowels and consonants according to Table 1 indicates that consonants are more inclined to contract and that they appear more seldom than vowels. However, it is important to remember that in the study all vowel combinations were considered, i.e. including qualitatively opposite vowels. At the same time for consonants only qualitatively close sounds were considered. Therefore, at the present stage of analysis, a direct comparison of contraction tendencies for vowels and consonants is not possible.

4.1. Vowels

Many languages tend to avoid vowel hiatus (Nespor 1987; Casali 1997; Cabré and Prieto 2005; Alba 2006; Smith 2008). Vowel combinations in the languages reported on in these studies often cause modification or elision. In Spanish, elision, diphthongization or new vowel formation occurs in 94% of all potential instances of hiatus (Smith 2008). Regular elision of one of two vowels takes place at word boundaries according to Casali (1997: 493) (for example in many Niger-Congo languages): "although elision of the first of two adjacent vowels is more common cross-linguistically, elision of the second vowel is also attested". At the present stage of the study I concentrate mostly on the number of pronounced sounds.

For the investigation of contractions in spontaneous Russian, all possible word-external vowel combinations were considered. The most frequent vowel combinations are: a_a (12.0%), a_i (11.4%), a_e (6.4%), a_u (6.1%), i_i (5.9%), a_o (3.9%), o_u (3.8%), i_a (3.6%).

In 68.4% of all instances of vowel contractions the pronounced sound was equal to the one of the sounds of combination. A new sound different from the two original ones appears as the result of contraction in 31.6% of all instances, for example:

(9) *sht**o_ye**vo* [ˈʂt**ɨ**vu] that_he-ACC instead of [ˈʂt**o i**vu];
(10) *posl’**e_o**konchan’iya* [ˈposlʲ**æ**kænˈʧænʲə] after_termination-GEN instead of [ˈposlʲ**i a**kænˈʧænʲə]

Such instances seem to be the most interesting for speech segmentation modeling because along with quantitative reduction of sounds (one is pronounced instead of two), the change of sound quality takes place, in this way the sound of contraction cannot be unambiguously associated with one of the two words. Either the boundary between the words can be identified inside the sound of contraction or the sound of contraction can be referred to one of two words.

Stress

The effect of the presence or absence of stress on contraction was examined. It has been shown for Catalan (Cabré and Prieto 2005) that the first vowel, V1 (the last vowel of the first word), does not undergo any change when followed by a vowel, V2, bearing nuclear stress (or phrasal stress). On the other hand, if V2 (first vowel of the second word) is not stressed, contraction is likely to occur.

In order to investigate the contraction-stress correlation all instances of two vowel combinations were divided into 4 groups: 1) 0 stress (when two vowels were unstressed); 2) stressed V1; 3) stressed V2; 4) two stressed vowels.

Table 2. Dependence of contraction at word boundaries on the presence and absence of stress.[2]

Type of stress	Total number	Contraction	No contraction
0 (both unstressed)	211 (39.5%)	148 (70.1%)	63 (29.9%)
1 (V1 stressed)	133 (24.9%)	59 (44.4%)	74 (55.6%)
2 (V2 stressed)	106 (16.7%)	57 (53.8%)	49 (46.2%)
3 (both stressed)	84 (15.7%)	38 (45.2%)	47 (54.8%)

The combinations with two unstressed vowels turned to be the most frequent and in 70.1% of these instances vowel contraction occurred. Contraction in combination with at least one stressed vowel is less expected.

As shown above, both word-internal and word-external contractions appear with approximately the same rate (about 55%). Unstressed vowels, however, undergo contraction inside the words more often than at word boundaries:

Table 3. Dependence of contraction inside words on the presence and absence of stress.

Type of stress	Total number	Contraction	No contraction
0 (both unstressed)	166 (52.4%)	138 (83.1%)	25 (15.1%)
1 (V1 stressed)	84 (26.5%)	27 (32.1%)	57 (67.9%)
2 (V2 stressed)	71 (22.4%)	19 (26.8%)	52 (73.2%)

When both vowels are unstressed contraction is more likely to occur inside words (83.1%) than at words boundaries (70.1%). Closer connection of sounds inside words rather than at their boundaries may explain the different figures for word-external and word-internal contractions. However, if at least one of the sounds is stressed, word-internal contraction occurs more seldom (32.1% and 26.8%) than word-external contraction (44.4%, 53.8% and 45.2%). A possible explanation for this mismatch could be the fact that combination of the same sounds (that undergo contraction easily) is not typical inside words. At word boundaries the same-vowel combinations are

[2] The sum of contraction and non-contraction instances does not give 100% because few examples, when zero sound was pronounced, were not counted.

more frequent,[3] which would lead to an increase of the whole percentage of word-external contractions. Alternatively, one may assume that in Russian there is a rule according to which word-internal vowel clusters tend not to be contracted (in certain stress contexts), while word-external vowel clusters do, at least to a higher extent.

4.2. Consonants

For consonants, formation of a new sound different from the source sounds does not usually occur. Yet, adjacency of two consonants can lead to considerable assimilation and deletion of one of the sounds. In the present study, only combinations of consonants with the same place and manner of articulation were considered (consonants that differ only in features of voiced/voiceless and palatalized/non-palatalized were included in possible combinations).

According to the data the most frequent consonants in the C1 position (the final consonant of the first word) in Russian are the consonants /t/ and /tʲ/ (52.5% of all instances), the percentage for C2 consonants /d/ and /dʲ/ is 23.1%. Also combinations involving the phonemes /s/, /sʲ/, /z/ and /zʲ/ are quite frequent – 21.0%; other consonant combinations included /k/g/, /m/mʲ/, /n/nʲ/, /p/ and /v/ (see Appendix). In all combinations of voiced and voiceless consonants regressive place assimilation took place, for example t+d=d: *vo**t_d’**et’i* [voˈ**dʲ**ɛtʲi] 'here children'; k+g=g: *ka**k_g**ovor’its’a* [kɛ**g**əvoˈrʲɛsa] 'as the saying goes'

Duration

In order to investigate how (or if) duration may indicate the number of (underlying) sounds, duration of all consonant contractions was measured. There is some difference in duration between contractions inside words and at word boundaries: word-external sounds are generally longer than word-internal sounds.

[3] The assumption was checked for 66 minutes of the material. The combination /aa/ was found 6 times inside words and 36 times at word boundaries.

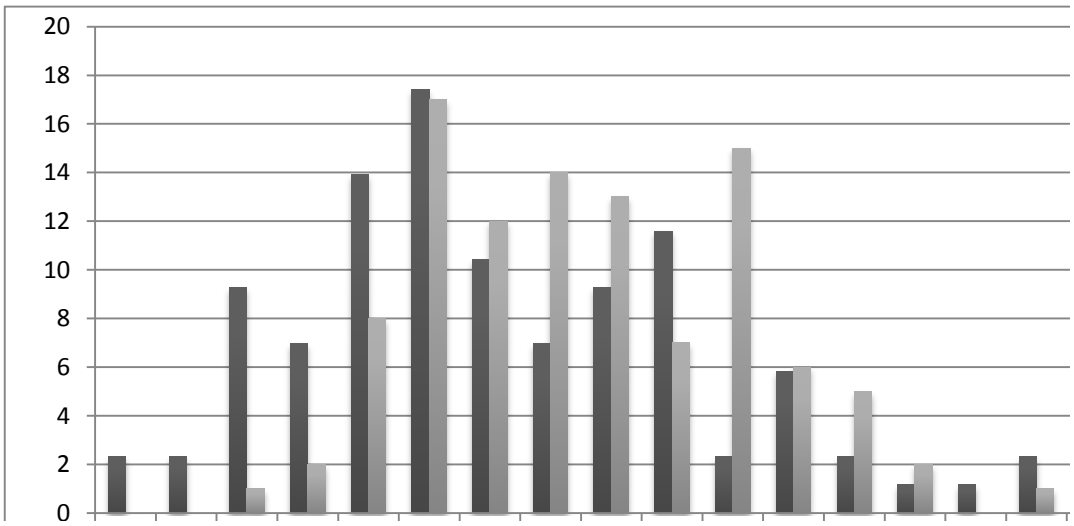

Figure 4. Duration of identical consonant contractions inside the words and at word boundaries. X-axis – consonant duration, ms; y-axis – relative number of instances.

Closure duration of plosive consonants is believed to be significant to make a distinction between geminate and non-geminate sounds (Pickett and Decker 1960; Hankamer et al. 1989; Esposito and Benedetto 1999). According to Dmitrieva (2001), the threshold of perception of double plosive consonants in Russian is 163-193 ms. So, the duration of single plosive consonants (/t/, /t^{j}/, /d/, /d^{j}/) and the same contracted ones, taken from the same interval between two pauses to diminish the speech rate influence, was measured. Figure 5 demonstrates that contracted consonants are generally longer than single ones, but they never (except for one instance – 170 ms) exceed the threshold of double sound perception.

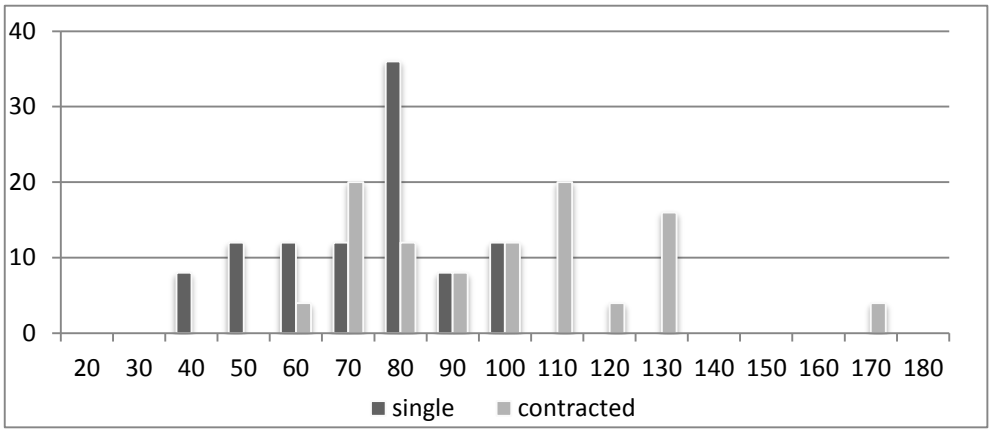

Figure 5. Durations of single and contracted consonants. X-axis – consonant duration, ms; y-axis – percentage of instances.

At the present stage of analysis, sound contraction in spontaneous Russian shows no regular rule for formation of contraction in the place of sound combination (stress or lack of stress has some influence the occurrence of vowel contraction but it is not systematical). In this way, any vowel combination (including unstressed vowel clusters) and any combination of "identical" consonants can be regarded as a potential contraction. So far, not all possible situations are examined: for example the behaviour of consonants in stressed vs unstressed syllables has not been examined.

5. Contractions in read speech

For the sake of comparison with the observations reported above, the contractions in texts read aloud were checked. 46 minutes of read texts[4] were analyzed in order to identify all possible vowel and consonant combinations that could lead to sound contraction.

Table 4. Word-external contractions in spontaneous and prepared speech

	Vowels		Consonants	
	Total	Contractions	Total	Contractions
Spontaneous speech (115 m)	559	310 (55.5%)	143	106 (74.1%)
Read texts (46 m)	303	191 (63.0%)	39	38 (97.4%)

As follows from Table 4, contractions are even more likely to occur in read speech than in spontaneous speech (63.0% and 97.4% vs 55.5% and 74.1%). Although I admit that at the moment the dataset is not large enough to make a reliable conclusion about read speech contractions, the high rate of contractions in read speech indicates inherence of contractions for Russian speech. I share the opinion that spontaneous speech and read speech are significantly different but contraction phenomena do not seem to be a particular feature of spontaneous speech as it was observed in read speech too, revealing that contraction is a common characteristic of spoken language.

[4] Two types of read texts were used: 1) a constructed text read by a professional speaker in laboratory conditions; 2) several radio-reports read by one speaker.

6. Conclusion

The investigation reported on in the present paper showed that sound contraction phenomenon is typical for Russian speech in general: it appears both in spontaneous and prepared (read) speech. Vowels as well as consonants regularly contract at word boundaries leading to loss of phonetic information.

All studied sound combinations were distributed according to the number of pronounced sounds (one or two). One pronounced sound can be a result of different processes: deletion, coalescence, centralization. Therefore, further investigation should include more detailed analysis of examples from spontaneous speech. Studies of contractions may explore two issues: 1) what makes sounds contract; 2) what could indicate the contractions for listeners. Vowel stress seems to have an effect on vowel contraction but it is doubtful whether duration can indicate sound contraction. Factors such as phonetic environment, word or phrase position, articulation rate should also be included in future studies of contraction.

Acknowledegments

The work was supported be research grant number MK-3646.2013.6 from the President of the Russian Federation.

References

Alba, M. C. 2006. Accounting for Variability in the Production of Spanish Vowel Sequences. *Selected Proceedings of the 9th Hispanic Linguistics Symposium*, ed. Sagarra Nuria and Jacqueline Toribio Almeida. 273-285. Somerville, MA: Cascadilla Proceedings Project.

Cabré, T., Prieto P. 2005. Positional and metrical prominence effects on vowel sandhi in Catalan. *Prosodies*. The Hague: Mouton de Gruyter: 123–158.

Casali, R.F. 1997. Vowel Elision in Hiatus Contexts: Which Vowel Goes? *Language* 73.3: 493–533.

The Cascadilla Proceedings Project.

The Corpus of Spontaneous Russian.

Dmitrieva, O. 2011. Asymmetries between production and perception of consonant length. *Studies in the Linguistic Sciences*, Illinois Working Papers: 1-15.

Esposito, A., M.G. Di Benedetto. 1999. Acoustical and perceptual study of gemination in Italian stops. In *JASA*. Vol. 106, N 4, Pt.1.: 2051-2042.

Hankamer, J., Lahiri A., Korenan J.C. 1989. Perception of consonant length: voiceless stops in Turkish and Bengali. *Journal of Phonetics* 17: 283–298.

Kasatkin, L.L., Choy M.Ch. 2005. Dolgota / kratkost' soglasnogo na meste sochetaniy dvuh soglasnyh bukv v sovremennom russkom literaturnom yazyke. M.: Yazyki slav'anskih kul'tur.

Nespor, Marina, Ellen M. Kaisse, Arnold M. Zwicky. 1987. Vowel degemination and fast speech rules. *Phonology* 4: 61–85.

Pickett, J.M., Decker L.R. 1960. Time factors in the perception of a double consonant. *Language and Speech* 3: 11–17.

Russkaya razgovornaya rech / pod redakciei E.A. Zemskoy. 1973. M.: Nauka.

Smith, Janet M., Tanya L. Flores, Michael S. Gradoville. 2008. An Analysis of Vowels across Word Boundaries in Veracruz, Mexican Spanish. *Indiana University Bloomington Volume 8*.

Zemskaya, E. A. 2006 [1979]. Russkaya razgovornaya rech: Linguisticheskii analis i problem obucheniya. Izd. 4. M.: Flinta, Nauka.

Appendix

Number of different consonants at word boundaries

Sound	First word	Second word	Total
d	1 – 1.7%	16 – 11.2%	79 – 56.3%
d’	-	17 – 11.9%	
t	44 – 30.8%	45 – 31.5%	
t’	34 – 23.8%	1 – 1.7%	
g	-	3 – 2.1%	6 – 4.2%
k	6 – 4.2%	3 – 2.1%	
m	15 – 10.5%	9 – 6.3%	15 – 10.5%
m’	-	6 – 4.2%	
n	11 – 7.7%	4 – 2.8%	11 – 7.7%
n’	-	7 – 5.6%	
p	6 – 4.2%	6 – 4.2%	
s	14 – 9.8%	15 – 10.5%	26 – 21.0%
s’	9 – 6.3%	2 – 1.4%	
z	3 – 2.1%	8 – 5.6%	
z’	-	1 – 1.7%	
v	4 – 2.8%	4 – 2.8%	