LARYNGEAL COARTICULATION IN TWO TYPES OF DEVOICING: AN ELECTROGLOTTOGRAPHIC STUDY OF RUSSIAN AND ENGLISH

Mayuki Matsui [1], Alexei Kochetov [2]

[1] National Institute for Japanese Language and Linguistics/ JSPS/ University of Amsterdam, [2] University of Toronto
[1] matsui-ma@ninjal.ac.jp, [2] al.kochetov@utoronto.ca

ABSTRACT

Effects of voiced/voiceless stops on voice quality of preceding vowels are known to be cross-linguistically variable, while those of voiced/voiceless fricatives are thought to be universal. Little is known, however, about whether such effects apply in a devoicing context. The current study investigated laryngeal coarticulation in two types of devoicing – phonological (albeit incomplete) devoicing in Russian and a phonetic (gradient) devoicing in English.

Electroglottographic data from 10 Russian and 10 Canadian English speakers showed that fricatives tend to boost the Open Quotient values of the vowel offset regardless of the language, suggesting that the glottal abduction gesture precedes the oral constriction of the fricatives. However, Russian devoiced obstruents had an abrupt upward trend, compared with English devoiced obstruents. Taken together, the current study documents both differences and similarities between two types of devoicing in terms of laryngeal coarticulation.

Keywords: Voicing, Voice quality, Electroglottography, Russian, English

1. INTRODUCTION

Production of speech sounds consists of multiple, continuous articulatory gestures implemented over time. As a consequence, a gesture at a point of time can be overlapped by gestures of adjacent sounds – the phenomenon referred to as coarticulation. The purpose of the current study is to investigate coarticulatory effects of voiced/voiceless obstruents on voice quality of the preceding vowel in two types of devoicing contexts as described below.

Previous acoustic studies have revealed that there are crosslinguistic differences in the effect of voiced/voiceless stops on voice quality of adjacent vowels [4, 5, 15]. These studies, which examined Germanic and Romance languages, showed that some languages (e.g. Swedish) exhibit voice quality differences at the vowel offset as a function of the voiced/voiceless status of the following stop, while other languages (e.g. French) do not show such an effect [4, 5, 15]. This suggests that coarticulatory effects of voicing on the voice quality is a language-specific, learnable process.

Unlike for stops, fricatives are thought to affect the vowel universally [5]. Specifically, vowels preceding a voiceless fricative showed an increasingly breathy voice quality, regardless of the language. This is in line with the observation that voiceless fricatives involve a larger and earlier glottal spreading [8, 10]. However, it is less clear as to whether this applies to all phonetic contexts. For example, a previous acoustic study [6] suggested that English (Scottish and Southern British standard varieties) word-final fricatives were different from Russian word-final fricatives in terms of the timing of glottal abduction. Specifically, the lag between glottal abduction and the onset of oral constriction was greater in Scottish and Southern British English, while it was smaller in Russian. It should be noted that the English varieties tested by [6] were those with robust pre-aspiration before voiceless fricatives (i.e. [ʰs] vis-à-vis [z̥]) while Russian shows such preaspiration (i.e. [s] vis-à-vis [z̥]). It is not clear, however, how stops pattern in the same context.

The current study thus examines the effect of voiced/voiceless fricatives and stops on voice quality of the preceding vowel in Russian and (Canadian) English, with special attention to two types of devoicing. Both Russian and Canadian English employ a two-way voicing contrast, and both have no robust aspiration before voiceless fricatives (as well as stops). They differ from each other, however, in terms of the status of word-final devoicing. In Russian, devoicing has been traditionally understood as a phonological, categorical process for both stops and fricatives, which results in neutralization of the contrast (e.g. [7]). However, a growing body of instrumental research suggests that this process is incomplete and not categorical as previously thought ([3, 9, 13, 14] for Russian; see also [16, 18] for German, [22] for Dutch). On the other hand, in some varieties of English including Canadian English, devoicing is thought to be a gradient process, which does not neutralize the phonological contrast (e.g. [2, 19] for American English, which is a variety close to Canadian English). The findings of incomplete devoicing in turn raise an empirical question: What are the differences and/or similarities between the incomplete phonological devoicing and gradient phonetic devoicing?

To compare these two types of devoicing with each other, an electroglottographic (EGG) study was conducted. EGG is a non-invasive method used to examine the direct signal caused by the glottal activity (opening and closing), apart from the effects of supralaryngeal resonance. By observing the change in the glottal activity as a function of time, the current study documents some cross-linguistic similarities and differences in laryngeal coarticulation. Given the previous findings [4, 5, 6, 15], it is predicted that (i) if glottal abduction gesture starts before oral constriction, at least for voiceless fricatives, the opening phase of the glottal activity in the preceding vowel would increase as a function of time; (ii) voiceless fricatives would show a similar pattern regardless of the language, since both Russian and Canadian English involve no preaspiration; (iii) the devoiced series might show some differences, since Russian and Canadian English involve different types of devoicing.

2. METHODS

Acoustic, EGG, and ultrasound imaging data were collected simultaneously. Given the purpose of the present study, only the EGG data are reported here.

2.1. Participants

The participants were 10 native speakers of Russian (mean age 23, SD 4.6) and 10 native speakers of Canadian English (mean age 23, SD 3.0). All of them reported to speak respective languages as their first language, to have grown up in a monolingual household, and to have no history of speech or language disorders. The Russian participants had lived in Canada for less than 5 years.

2.2. Speech materials

The speech materials consisted of monosyllabic ($C_1C_2V_1C_3$) nonce words containing word-final voiced or voiceless obstruents (C_3) preceded by a low vowel (/a/ for Russian, /ɑ/ for English), as summarized in Table 1. The target words were embedded in carrier sentences ("*Oksana skazala* [target]*.*" 'Oksana (female name) said [target]' for Russian; "*I saw* [target]*.*" for English).

Table 1: Word list.

Russian	/flad/, /flat/, /flas/, /flaz/
English	/flɑd/, /flɑt/, /flɑs/, /flɑz/

2.3. Recording procedures

The participants were tested individually in the University of Toronto phonetics lab. They were asked to read randomized sentences based on orthographic prompts. 5-6 repetitions were elicited. In total, there were 440 intended tokens (4 Cs x 5 or 6 repetitions x 10 speakers x 2 languages). Three speakers' data (1 Russian speaker and 2 English speakers) were systematically excluded from the analysis due to an EGG recording error.

The EGG signal was generated using *EG2-PCX2* (*Glottal Enterprises* Inc.) and sent to *Audacity®*. The electrodes for EGG were placed on the participant's neck, using a *Velcro®* strap. The audio signals were sent to Audacity via a preamplifier to synchronize with the EGG signals.

2.4. Analysis procedures

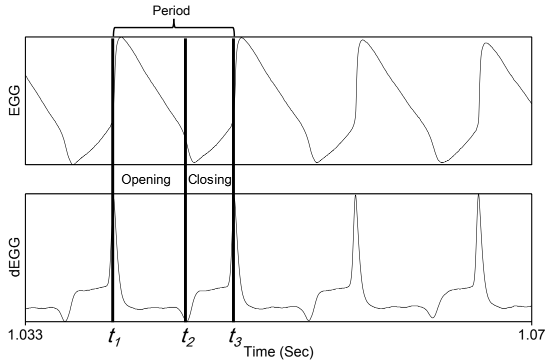
Analysis procedures consisted of three steps. First, to detect regions of interest in the EGG signal, the interval of the preceding vowel was annotated on the basis of the audio signal, using Praat [1]. Second, Open Quotient (OQ) values in the preceding vowel were measured based on the time-derivative (dEGG) of the raw EGG signal. Finally, the values were fitted to the Generalized Additive Mixed-Effects Models (GAMM, [20, 24]) to evaluate the change in glottal gestures as a function of time.

OQ in the preceding vowel of the final obstruent was semi-automatically computed using the *Praatdet* tools [11]. OQ reflects how much time the glottis is open during a period (e.g. [12]). Greater OQ implies a longer glottal opening, in comparison to the the closing time during a period, implying breathier or more lax phonation.

Before the computation, the raw EGG waveforms were high-pass filtered at 75 Hz to eliminate the noise due to exhalation interfering with periodicity, and the closing peaks were set on the top of the waveform (and the opening peaks were on the down). The OQ computation was done on the basis of the raw EGG waveform and the time-derivative of it (dEGG, see Figure 1): The closing peaks were detected from a time-derivative of the dEGG waveform, and the opening peaks were detected using "Howard's method" featured in Praatdet. The OQ values were then automatically calculated as (1). After the automatic calculation, the obtained OQ plots were visually checked before submitting statistical analyses.

(1) OQ = opening phase/period
opening phase = $t_2 - t_1$
period = $t_3 - t_1$

Figure 1: An example of the OQ measurement (produced by a Russian male speaker).



In addition, voiced stops were classified based on the EGG signal as phonetically devoiced and voiced. If the signal ceased before reaching 50% of the constriction interval (i.e. an assumed center of the steady-state of the obstruent), the token was classified as "phonetically devoiced". The other instances were classified as "phonetically voiced". Not surprisingly, the majority of final voiced obstruents were judged as "devoiced". However, some instances of the voiced stops with no devoicing were observed in both languages. These results are presented separately in Section 3. Voiced fricatives were consistently devoiced.

The obtained OQ values were plotted onto a two-dimensional space with the normalized time as the x-axis, and the normalized OQ as the y-axis. OQ values were z-transformed for each speaker to minimize possible deviations caused by individual differences in voice quality.

Having generated the normalized space, the data were fitted onto the Generalized Additive Mixed Models (GAMM; see [20] for a practical introduction) in order to estimate a non-linear function relating the x-axis (time) with the y-axis (OQ), i.e. a change in OQ value as a function of time. In the regression models predicting OQ value, the (normalized) time in the preceding vowel was specified as the numeric dependent variable. The (normalized) time was then nested by context. By-trajectory (or by-token) random smooth as well as by-speaker random smooth were added to reduce autocorrelation [20]. The assumption here is that, if the voicing in the following obstruent affects the voice quality of the preceding vowel, the vowel would show different OQ trajectories, depending on the following voicing context, across individual speakers. All the statistical analyses were implemented using R ([17]). GAMMs were implemented with *mgcv* [23] and *itsadug* [21] packages in R. The statistical significance was assessed by visual inspection of the 95% confidence interval and model comparison with the null model.

3. RESULTS

3.1. Within-language results

Figures 2 and 3 shows OQ trajectories for Russian and English speakers respectively. For Russian (Fig. 2), the phonetically voiced stops show lower OQ, compared with voiceless stops and fricatives at the end of the vowel. However, the devoiced stops and fricatives were not significantly different from the voiceless counterparts, in terms of the OQ trajectory. English shows similar trends (Fig. 3), but the effect is somewhat smaller, compared with Russian.

Figure 2: OQ trajectories for Russian speakers (non-linear regression by GAMM). Lines indicate the mean while bands indicate the 95% confidence interval of the mean.

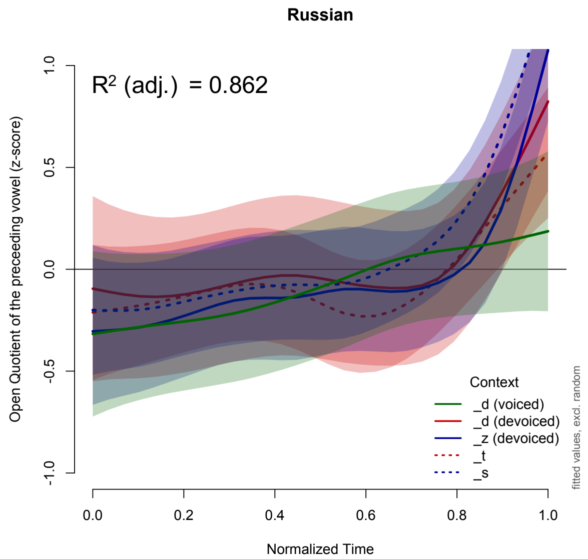
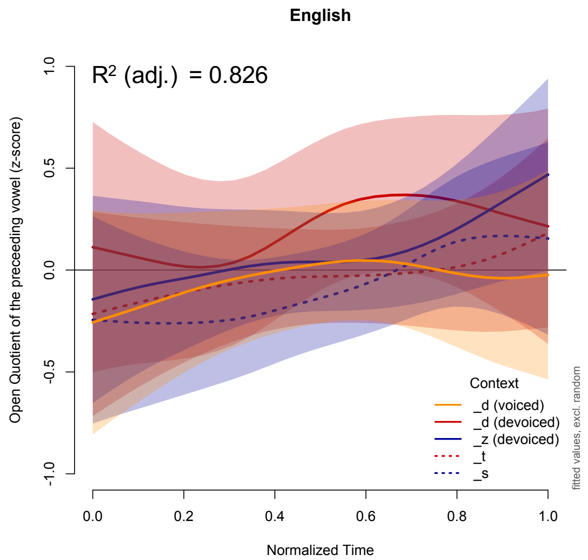


Figure 3: OQ trajectories for English speakers (non-linear regression by GAMM).



3.2. Between-language comparison

Figures 4 and 5 show cross-linguistic comparisons of OQ trajectories for stops and fricatives respectively. For devoiced fricatives (Fig. 5), an upward trend is evident for both languages. At the same time, Russian devoiced fricatives showed an abrupt upward trend, compared to English devoiced fricatives. Such a trend is also evident for devoiced stops (Fig. 4). Additionally, while English devoiced stops show an upward trend, the rising starts somewhat earlier than Russian devoiced stops.

Figure 4: OQ trajectories for Russian and English stops (non-linear regression by GAMM).

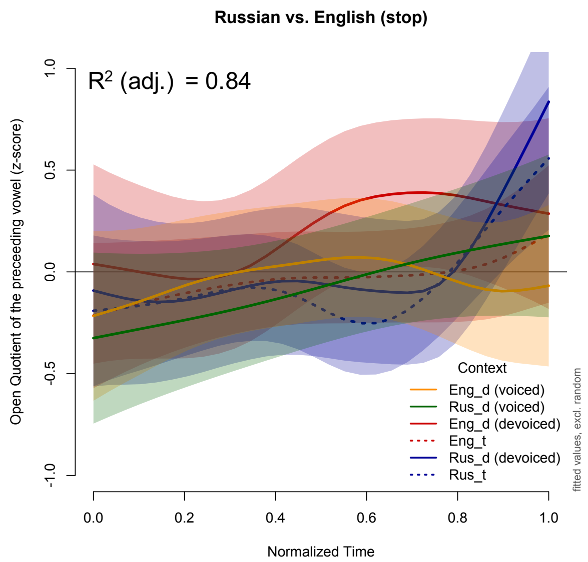
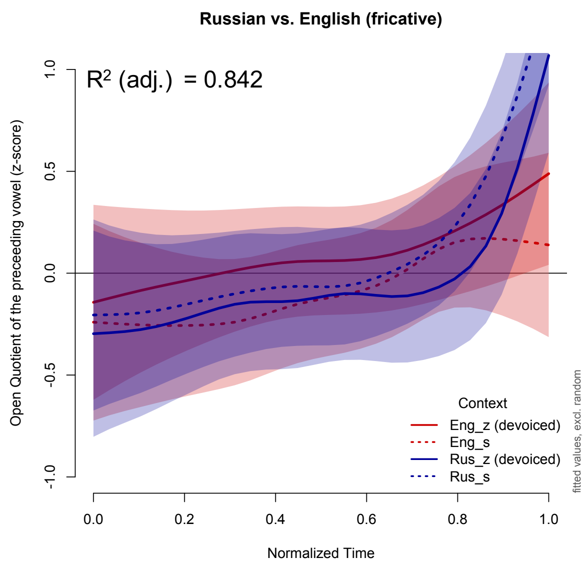


Figure 5: OQ trajectories for Russian and English fricatives (non-linear regression by GAMM).



4. DISCUSSION

The purpose of the current study was to investigate coarticulatory effects of voicing on voice quality of the preceding vowel. This was done by taking into consideration two types of devoicing – phonological (albeit incomplete) devoicing as in Russian and phonetic (gradient) devoicing as in Canadian English. Specifically, we tested three predictions stated in Section 1. Our preliminary analysis of the data revealed three main findings.

First, as a function of time, voiceless fricatives boosted the OQ values of the vowel offset in both languages, suggesting that the glottal abduction gesture starts before the oral constriction, which confirms the prediction (i). This is consistent with previous acoustic studies examining various languages [6, 15]. However, the OQ trajectories were different between Russian and Canadian English, which disconfirms the prediction (ii).

Second, for both languages, phonetically voiced stops had lower OQ trajectories, compared to voiceless stops and fricatives. However, phonetically devoiced stops and fricatives showed higher OQ trajectories, just like the underlyingly voiceless obstruents did. This suggests that in the devoicing context, effects of obstruent voicing on voice quality were affected by *surface* voicing of the following consonant, but not by underlying voicing in both languages.

Finally, the results showed that Russian devoiced obstruents have an abrupt upward trend compared with English devoiced obstruents, suggesting that two types of devoicing result in different OQ trajectories. This confirms the prediction (iii). A general implication of the current results is that, while the phonological devoicing and phonetic devoicing is proved to be closer to each other than previously thought, their difference may be due to a difference in coarticulation. Taken together, the results suggest that coarticulatory effects of voicing on the voice quality is, at least in part, a language-specific process. Future studies could further address to what extent the two types of devoicing are similar or different from each other both in terms of laryngeal and supralaryngeal coarticulation.

In conclusion, the current study documented both differences and similarities between an incomplete phonological devoicing and a gradient phonetic devoicing in terms of laryngeal coarticulation.

5. ACKNOWLEDGEMENTS

This research was supported by a KAKENHI Grant-in-Aid for Scientific Research (#17J08493 to MM) and a Social Sciences and Humanities Research Council of Canada grant (#435-2015-2013 to AK). Thanks to Kiranpreet Nara, Maida Percival, Connie Ting, and Jessica Yeung for assistance with data collection.

6. REFERENCES

[1] Boersma, P., Weenink, D. 1999–2017. Praat: doing phonetics by computer (Version 6.0.36) [Computer program] http://www.praat.org/.

[2] Davidson, L. 2016. Variability in the implementation of voicing in American English obstruents. *Journal of Phonetics*, 54, 35–50.

[3] Dmitrieva, O., Jongman, O., Sereno, J. 2010. Phonological neutralization by native and non-nativespeakers: The case of Russian final devoicing. *Journal of Phonetics,* 38, 483–492.

[4] Gobl, C., Ní Chasaide, A. 1988. The effects of adjacent voiced/voiceless consonants on the vowel voice source: a cross language study. STL-QPSR, Speech, Music and Hearing, Royal Institute of Technology, Stockholm, 23–59.

[5] Gobl C., Ní Chasaide, A. 1999. Voice source variation in the vowel as a function of consonantal context. In: Hardcastle W. J., Hewlett N. (eds.), *Coarticulation: Theory, Data, Techniques*. Cambridge: Cambridge University Press, 122–143.

[6] Gordeeva, O. B. 2007. Learnability of laryngeal abduction in voiceless fricatives: Cross-linguistic evidence. *Proc. the 16th ICPhS*, 433–436.

[7] Halle, M. 1959. *The Sound Pattern of Russian*. Hague: Mouton and Co. Printers.

[8] Hoole, P. 1999. Techniques for investigating laryngeal articulation. In Hardcastle, W. J., Hewlett, N. (eds.), *Coarticulation: Theory, Data, Techniques*. Cambridge: Cambridge University Press, 294−300.

[9] Kharlamov, V. 2014. Incomplete neutralization of the voicing contrast in word-final obstruents in Russian: Phonological, lexical, and methodological influences. *Journal of Phonetics*, 43, 47−56.

[10] Kingston, J. 1990. Articulatory binding. In Kingston, J., Beckman, M.E. (eds.), *Papers in Laboratory Phonology I*. Cambridge: Cambridge University Press, 406−434.

[11] Kirby, J. 2017. Praatdet: Praat-based tools for EGG analysis (v0.1.1). [Praat script] https://doi.org/10.5281/zenodo.1117189

[12] Klatt, D., Klatt, L. 1990. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America*, 7, 820–857.

[13] Kulikov, V. 2012. Voicing and voice assimilation in Russian stops. Doctoral dissertation, University of Iowa.

[14] Matsui, M., Igarashi, Y., Kawahara, S. 2017. Acoustic manifestation of Russian word-final devoicing in utterance-medial position. *Journal of the Phonetic Society of Japan*, 21(2), 1−17.

[15] Ní Chasaide, A., Gobl, C. 1993. Contextual variation of the vowel voice source as a function of adjacent consonants. *Language and Speech*, 36, 303–330.

[16] Port, R. F., O'Dell, M. 1985. Neutralization of syllable-final devoicing in German. *Journal of Phonetics*, 13, 455–471.

[17] R Core Team. 2016. R: A language and environment for statistical computing (Version 3.5.0). [Computer program] R Foundation for Statistical Computing, Vienna, Austria. http://www.Rproject.org/

[18] Roettger, T.B., Winter, B., Grawunder, S. 2014. Assessing incomplete neutralization of final devoicing in German. *Journal of Phonetics*, 43, 11–25.

[19] Smith, C. L. 1997. The devoicing of /z/ in American English: Effects of local and prosodic context. *Journal of Phonetics*, 25, 471–500.

[20] Sóskuthy, M. 2017. Generalised additive mixed models for dynamic analysis in linguistics: a practical introduction. arXiv:1703.05339 [stat.AP].

[21] van Rij, J., Wieling, M., Baayen, R. H., van Rijn, H. 2017. itsadug: Interpreting Time Series and Autocorrelated Data Using GAMMs. R package version 2.3.

[22] Warner, N., Jongman, A., Sereno, J. A., Kemps. R. 2004. Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch. *Journal of Phonetics*, 32, 251–276.

[23] Wood, S. N. 2006. *Generalized Additive Models: An Introduction with R*. Boca Raton: CRC Press.