Vowel duration variation induced by coda consonant voicing and perceptual short/long vowel categorization in Czech

Václav Jonáš Podlipský, Kateřína Chládková

Department of English and American Studies, Palacký University, Olomouc, Czech Rep.
vaclavjonaspodlipsky@centrum.cz, katerina.chladkova@seznam.cz

Abstract: It is possible that in Czech a vowel is consistently longer before a tautosyllabic voiced as opposed to voiceless obstruent. The main purpose of this study was to determine if this hypothesized variation interacts with the perception of vowel quantity. Two experiments were conducted that examined the effect of a voiceless vs. devoiced coda context and of a voiceless vs. voiced coda context on perceptual short/long vowel categorization. It was found that quantitatively ambiguous vowels were more likely to be perceived as short before a voiced coda than before a voiceless coda. It is concluded that vowel duration is indeed affected by coda voicing in Czech and that listeners are sensitive to this variation because they adjust the perception of vowel quantity accordingly.

1. Introduction

There is a tendency across languages for vowels to be longer when followed by voiced obstruents than when followed by voiceless obstruents. It has been proposed that this tendency has an articulatory motivation and that it is universal [1]. Moreover, this vowel-duration variation can be encoded in the phonology of a language. For instance in English, as is well documented, the effect is relatively large and several authors have suggested that the variation was phonologized to enhance the voicing contrast of the following obstruents (e.g. [2]). Others have even pointed out that in English underlying coda-obstruent voicing is primarily manifested by nuclear vowel duration rather than by phonetic voicing of the coda itself [3]. An explanation of how nuclear vowel duration could have replaced phonetic voicing in marking underlying voicing of codas in English was proposed e.g. by Nearey [4]. Other evidence that the coda-voicing-induced vowel-duration differences can be a part of the phonology of a language comes from what has been termed ‘incomplete neutralization’. It has been shown that many languages which display final devoicing do not neutralize final voicing contrasts completely because these contrasts surface in the form of durational differences of the vowels they follow; to our knowledge incomplete neutralization was observed in German, Polish, Catalan, Dutch, and Russian [5, 6].

Few studies have examined whether or not vowel duration is influenced by voicing of the following obstruent in Czech, if yes to what extent, and how this process interacts with final (or other) devoicing. Machač and Skarnitzl [7] showed that only some vowels were significantly longer before (phonetically) voiced stops and that only applied to some places of articulation. From their results it therefore appears that the effect is not very strong in Czech. In their experimental material however, vowel-consonant (VC) sequences where V and C were heterosyllabic (i.e. V and C were separated by a syllable boundary, V.C) were not treated separately from tautosyllabic VC sequences (i.e. sequences where the V and C belong to the same syllable). In addition, the authors do not indicate what the proportion of tautosyllabic versus heterosyllabic VC sequences was in their material. Likewise, in their study Machač and

Skarnitzl did not isolate devoiced codas from other contexts and no conclusions about possible incomplete neutralization in Czech can thus be made from their data.

We hypothesize that in tautosyllabic VC sequences the effect consonantal voicing exerts on the duration of the preceding vowel is stronger and/or more consistent than in heterosyllabic sequences. What is more, it is only when examining tautosyllabic VC sequences that devoicing can be considered. Before pursuing the main interest of our study (that is whether the phenomenon at hand is reflected in perception), we will report on the preliminary production measurements that we performed in order to assess our hypothesis and also the possibility of incomplete neutralization of coda voicing contrasts in Czech.

Ideally, what would have been most informative in this respect is to examine productions of a set of 3 meaningful and comparably frequent Czech words with the syllable structure CVC.CV, minimally contrasted by the coda of the first syllable: one word with a voiceless coda, one word with a voiced coda, and one word with an underlyingly voiced coda that is devoiced on the surface because it assimilates with the following segment. To the best of our knowledge such a minimal set does not exist in Czech. Our solution was twofold: we chose (1) a non-minimal set *kapky* ('drops'), *babky* ('grandmas'), ragby ('rugby'), where the coda of the first syllable differs in the desired way but there are also unwanted differences (in the initial consonants above all), and (2) a minimal set of 3 nonsense words *tapka*, *tabka*, and *tabga*, where all irrelevant variation is controlled for. These words and non-words were selected so that ambisyllabicity of the differing coda was very unlikely.

We elicited and analyzed productions of the non-minimal set by 19 Czech speakers and of the nonsense set by different 9 Czech speakers. Each word from both sets, and a number of fillers, was read in the carrier phrase *Slovo* __ *neznám* 'I don't know the word __'. There were 11 repetitions for each speaker for the non-minimal set (thus giving 627 tokens to be analyzed) and 10 repetitions for the nonsense set (270 tokens analyzed). In each token the duration of the V in the first syllable was measured.

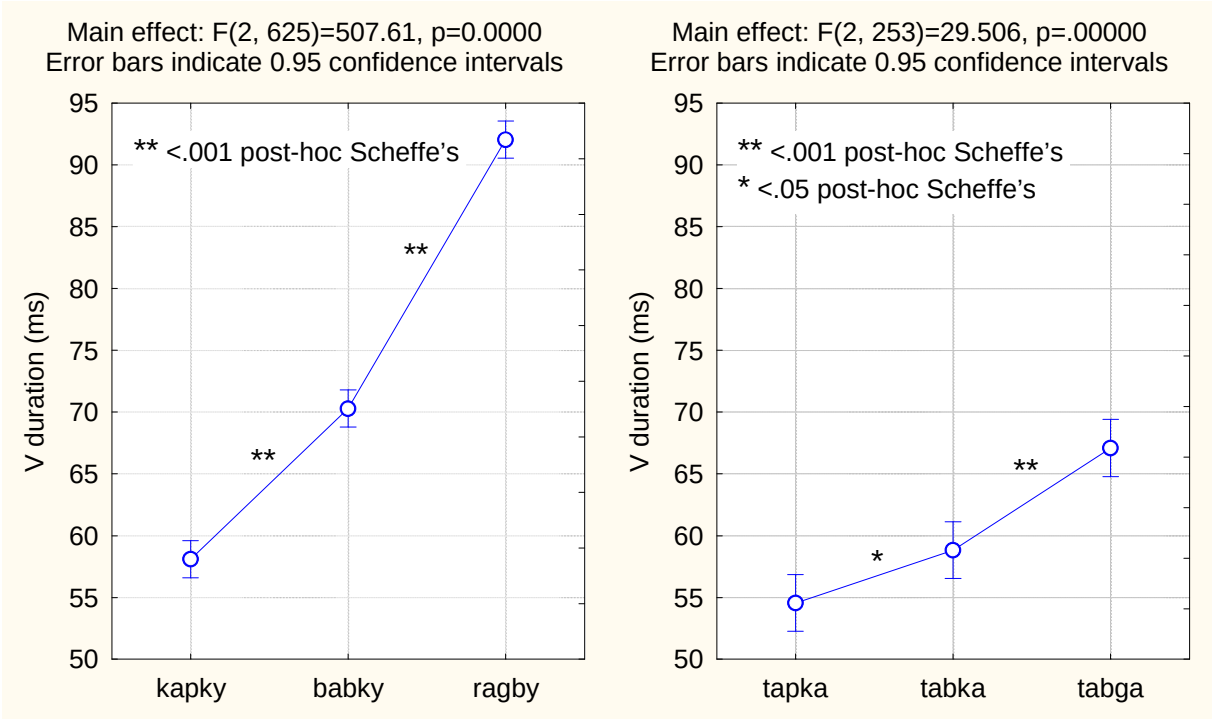

Figure 1: ANOVA results and means of the durations of [a] in the first syllable of two sets of words (real words on the left, and nonsense words on the right). See text for more comments.

The results are presented in Fig. 1. For the non-minimal set of 3 existing words, there was a clear-cut difference in the durations of the vowel [a] in the first syllable across the three phonetic contexts. It is very likely however, that at least part of the effect is due to the undesired differences, especially the differences in the initial consonants. For instance, [k], unlike [b], has a positive VOT and that is bound to affect the duration of the following vowel. What is probably more interesting therefore are the results for the nonsense set in which there is no irrelevant variation between the words and all the effect that is observed can only be attributed to the differences in the codas. The differences in the duration of the target vowel [a] are not as large as in the real-words set, however ANOVA showed that the main effect is highly significant ($p<0.00001$) and post-hoc Scheffe's tests revealed significant differences between all three contexts.

From these data, we tentatively conclude that in Czech a vowel is relatively short when followed by a voiceless obstruent, longer when followed by a devoiced (underlyingly voiced but phonetically voiceless) obstruent (in other words that incomplete neutralization takes place), and even longer when followed by a phonetically voiced obstruent. It should be stressed again that this is a tentative conclusion. It is possible that the lengthening effect was observed partly because the words are nonsensical (although they are phonotactically possible) and speakers do not treat them naturally. In addition, the process may function differently for different vowel qualities and different consonantal places of articulation, as is suggested by Machač and Skarnitzl's data [7].

For the purpose of the present study however, we assume that in Czech V duration is in reality (slightly but consistently) affected by the phonological and phonetic voicing of the following C if the V and C are tautosyllabic. Our second premise is that this slight variation is accessible to perceivers. This is very well possible. Warner *et al.* [5] even showed that Dutch listeners perform above chance when discriminating syllables with a voiceless coda from syllables with a devoiced coda.

With these provisions, and considering that in Czech vowel length is contrastive, we formulated the main question that our study was designed to address: Does the vowel-duration variation induced by the voicing of coda consonants in Czech affect perceptual vowel quantity categorization? In other words, do Czech listeners adjust their short/long vowel distinctions on the basis of coda voicing? A similar shift of the perceptual short/long boundary has been observed e.g. in open as opposed to closed syllables (see [8]). We conducted two experiments to answer the present question.

2. Experiment I.

In the first experiment the perception of vowel length, where the vowel was in a voiceless-coda context, was compared with the perception of vowel length, where the vowel was in a devoiced-coda context. This experiment was therefore intended to determine whether or not 'incomplete neutralization' of the voicing contrast in coda (which we presume to exist in Czech) is reflected in perception by shifting the location of the short/long vowel boundary.

2.1. Method

There were 7 different stimuli in this experiment which formed a continuum. The first stimulus was a naturally produced word [tap] (a nonsense[1] word in Czech) and in the

[1] A nonsense word was used rather than an existing word to avoid lexical effects, word-frequency effects for instance.

subsequent stimuli the vowel was lengthened (by multiplying fundamental cycles) so that the final 7th stimulus in the continuum was the word [taːp]. The vowel in the 1st stimulus lasted 106 ms and in each subsequent stimulus the vowel was 21.9 ms longer, so that in the final stimulus the duration of the vowel was 237 ms. Each stimulus was embedded in the carrier phrase *Koupila ___ konečně.* 'She bought a ___ finally.'.

Two 2-alternative forced-choice categorization tests were conducted in this experiment. The tests did not differ in the stimuli (both used the [tap] – [taːp] continuum) but they differed in the response buttons that were displayed on the screen of a computer: in one test the buttons were labeled 'tap' and 'táp'[2], while in the other they were labeled 'tab' and táb'. Within each test, each stimulus was repeated 10 times and the order of stimuli was randomized. Subjects responded by clicking on one of the two buttons using a mouse.

Forty-nine Czech listeners, all in their twenties and coming from various regions of the Czech Republic, participated in the experiment. Each person completed both tests, with at least a 20-minute break between tests, but the order of the tests was counterbalanced: 25 subjects completed the 'tap' test before the 'tab' test, and 24 subjects completed the 'tab' test first.

2.2. Results and discussion

No differences between the two tests in this experiment were found. The average categorization curves were relatively categorical, the 4th stimulus (where the vowel had 172 ms) being closest to the short/long boundary (an approximately 40% 'short' labeling) in both tests. Importantly, the location of the short/long boundary did not differ between the two tests: repeated-measures ANOVA showed no significant interaction between test type and vowel duration ($p>0.3$).

This experiment thus found no reflection of the presumed vowel lengthening before a devoiced as opposed to voiceless obstruent in perception of vowel quantity. It would be a hasty conclusion to say that perception of vowel quantity does not interact with incomplete neutralization in Czech though. It is possible that the design of the experiment, where underlying voicing was evoked by spelling, was not good enough to discover the interaction. Also, different results could be obtained with stimuli containing a vowel of a different quality and/or an obstruent with a different place of articulation. Thirdly, other devoicing than final (such as the assimilatrory devoicing in *tabka*) may be found to have an effect on perceptual vowel quantity categorization. Before these possibilities are evaluated in future studies however, it may be reasonable to take one step back and test if incomplete neutralization is actually perceptible in Czech, that is, if listeners perform above chance when discriminating voiceless from devoiced codas.

3. Experiment II.

This experiment was designed to determine whether perception of vowel length would be different before a voiceless coda as opposed to a (both phonemically and phonetically) voiced coda. A perceptual shifting of the short/long boundary was more likely here because our preliminary measurements (reported above) suggest that the duration difference between a V in a voiceless-coda context and a V in a voiced-coda context is larger than it is between a V in a voiceless-coda context and a V in a devoiced-coda context.

[2] The stroke above a vowel symbol indicates a long vowel in standard Czech orthography.

3.1. Method

This experiment, like the first one, involved two tests. Two different vowel-duration continua (one for each test) were created from a naturally produced nonsense word. Both continua consisted of 8 stimuli in which the duration of the first vowel was manipulated. The continua differed in the voicing of the stop following the manipulated vowel ([tapka] – [taːpka] vs. [tabga] – [taːbga]) but importantly, they didn't differ in vowel durations (in both continua the vowel in the 1st stimulus lasted 64 ms and for subsequent stimuli the increment was 12.1 ms so that the last stimulus lasted 148 ms). This was achieved by creating the voiceless-coda continuum first (by multiplying fundamental cycles) and then splicing the initial CV(:) of each stimulus and the sequence [bga] excised from a naturally-produced [tabga]. Again, the nonsense words were selected to avoid ambisyllabicity of the C following the manipulated V. Each stimulus was embedded in the carrier phrase *Slyším* ___ *jasně.* 'I can hear ___ clearly.'.

Two 2-alternative forced-choice categorization tests were conducted, one for each continuum. Within each test, each stimulus was presented 10 times and the order of stimuli was randomized. Subjects responded by clicking on one of the two buttons displayed on the screen (*tapka* or *tápka* in the first test, and *tabga* or *tábga* in the second test) using a mouse.

Fifty-four Czech listeners, all in their twenties and coming from various regions of the Czech Republic, participated in the experiment. Each person completed both tests, with at least a 20-minute break between the tests, but the order of the tests was counterbalanced.

3.2. Results and discussion

The identification curves averaged across subjects are shown in Fig. 2. Repeated-measures ANOVA revealed a highly significant (p<.00001) interaction between vowel duration and test

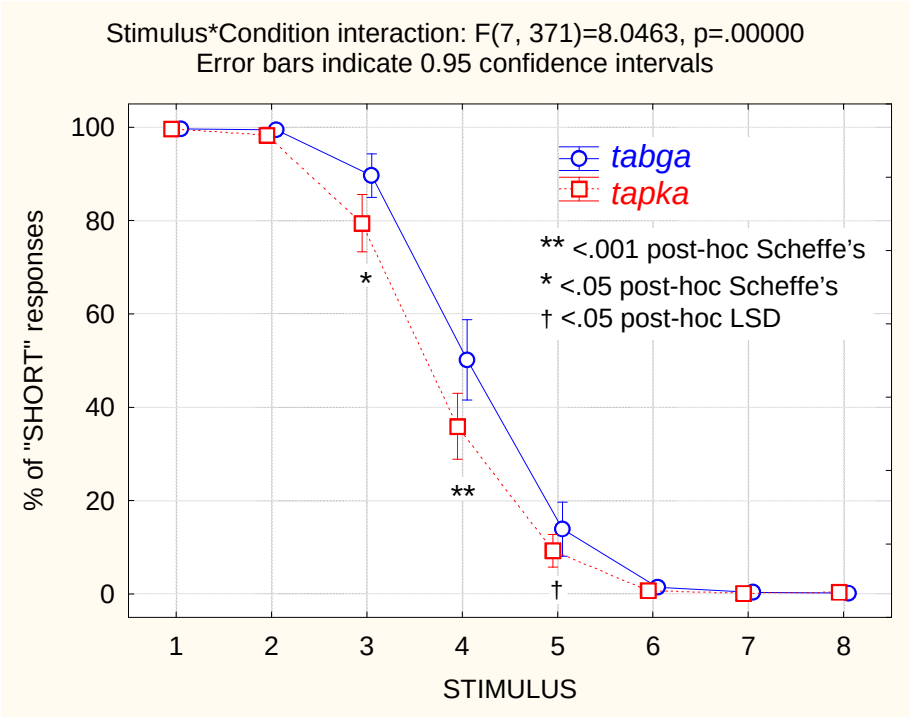

Figure 2: Average V/V: identification curves for the voiceless (dotted) and voiced (solid) coda context. Repeated-measures ANOVA found a significant interaction between vowel duration and coda voicing. Post-hoc tests showed differences in responses to near-boundary vowel durations.

condition (i.e. coda voice value). Post-hoc tests compared mean responses to individual stimuli and showed that stimulus 4, which was the closest to the category boundary, differed most significantly ($p<.001$ Scheffe's test) between coda contexts, and the two neighboring stimuli differed too although not so significantly (see Fig. 2). In order to assess the magnitude of the boundary shift, the location of the category boundary for each curve (i.e. the vowel duration corresponding to a 50% labeling) was computed using the procedure described in [9]. The difference in the boundary location was approximately 3.3 ms.

These findings suggest that listeners shift the location of the V/V: boundary so that a vowel needs to last slightly longer when followed by a voiced coda to be perceived as long than it does when followed by a voiceless coda. Although the magnitude of the shift was very small in the present experiment (3.3 ms), listeners seem to be adjusting their perception of V quantity very consistently given the high significance of the V duration × Coda voicing interaction.

Order of tests	Direction of boundary shift		
	expected	no clear shift	opposite
tapka - tabga	14	7	6
tabga - tapka	19	5	3

Table 1: The number of listeners who showed or did not show a boundary shift in either direction arranged by the order in which they took the tests.

Two remarks need to be made at this point. First, only group averages have been presented so far and it is worth considering the results within individual listeners who did not always show a V/V: boundary shift similar to the average. We sorted individual patterns into three classes: a shift in the expected direction (i.e. the one the average identification curves show), a shift in the opposite direction, and no clear shift (cases where there was either very little difference between curves or it was a difference in steepness rather than in location of the cut-off point). There were 33 expected-direction cases, 9 opposite-direction cases, and 12 no-clear-shift cases. Second, the order in which listeners completed the two tests in this experiment needs to be considered as a potential factor at play. Table 1 is a tabulation of the direction of the boundary shift (using again the 3 classes 'expected', 'no clear shift', and 'opposite') against the order of tests. The number of subjects who took the *tabga* test first and who showed a shift in the expected direction is higher (and in the opposite direction, lower) than of those who took the *tapka* test first. However a chi-square test ($p=.05$) did not eliminate the null hypothesis and therefore the character of the boundary shift probably does not depend on the order of the tests. We take this as a support of the validity of the general observation that the V/V: boundary is shifted according to the voicing of the coda, with the provision that not all individuals in the present study followed this pattern.

4. Conclusion

In Czech, the voicing value of a consonant potentially influences the duration of the preceding tautosyllabic vowel, so that it is the shortest when the following consonant is voiceless, longer when the consonant is devoiced, and still longer when the consonant is voiced. The purpose of this study was to determine whether or not this vowel-duration variation interacts with the perception of vowel quantity.

Our results did not show any effect of a voiceless as opposed to devoiced coda context on the perceptual short/long vowel distinction. However, the perceptual boundary between a

short and a long vowel was shown to shift slightly (and very significantly) when a voiceless vs. voiced context was compared: quantitatively ambiguous vowels were more likely to be perceived as short before a voiceless coda obstruent than before a (phonetically) voiced coda obstruent. This finding suggests that Czech listeners adjust the perception of vowel quantity depending on the phonetic voicing value of the coda. This conclusion indirectly provides support to the hypothesis that in Czech (like in many languages, see e.g. [1]) a voiced obstruent causes a lengthening (or alternatively, that a voiceless obstruent causes clipping) of the nucleus of the syllable it closes.

It is worth noting that the shift of the perceptual short/long vowel boundary caused by the voicing difference of the coda that we found is similar to the adjustments listeners make in differing speaking rates or in open vs. closed syllables (see [8]). The effect that we observed thus fits into a larger picture of quantity perception in interplay with the perception of other phonological entities. From a broader perspective it appears that our findings can be adduced in support of the view (see e.g. [10]) that phonetic detail is encoded in the sound system of a language user and that it is used systematically to facilitate the perception of speech.

References

[1] Chen, M. 1970. Vowel Length Variation as a Function of the Voicing of the Consonant Environment. *Phonetica* 22, pp. 129–159.

[2] Klatt, D. H. 1976. Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *J. Acoust. Soc. Am.* 59(5), pp. 1208–1221.

[3] Walsh, T., Parker, F. 1981. Vowel length and ‘voicing’ in a following consonant. *J. Phonetics* 9, pp. 305–308.

[4] Nearey, T. M. 1997. Speech perception as pattern recognition. *J. Acoust. Soc. Am.* 101(6), pp. 3241–3254.

[5] Warner, N. *et al.* 2004. Incomplete neutralization and other sub-phonemic durational differences in production and perception: evidence from Dutch. *J. Phonetics* 32, pp. 251–276.

[6] Dmitrieva, O. 2005. Final devoicing in Russian: Acoustic evidence of incomplete neutralization. *J. Acoust. Soc. Am.* 117(4), pp. 2570.

[7] Machač, P., Skarnitzl, R. 2007. Temporal compensation in Czech? *Proc. 16th ICPhS* Saarbrücken, pp. 537-540.

[8] Palková, Z. 1994. *Fonetika a fonologie češtiny s obecným úvodem do problematiky oboru*. Prague: Karolinum.

[9] Podlipský, V. J. 2007. Perception of Czech vowel quantity by English learners of Czech. *Proc. 16th ICPhS* Saarbrücken, pp. 1709-1712.

[10] Hawkins, S., Local, J. 2007. Sound to sense: Introduction to the special session. *Proc. 16th ICPhS* Saarbrücken, pp. 181-183.