PERCEPTION OF CZECH VOWEL QUANTITY BY ENGLISH LEARNERS OF CZECH

Václav Jonáš Podlipský

Dept. of English and American Studies, Palacký University, Olomouc, Czech Republic
vaclavjonaspodlipsky@centrum.cz

ABSTRACT

Acquiring L2 vowel quantity can be difficult for native speakers of languages like English where vowel duration cues stress. This study tested whether English learners of Czech would categorize short and long vowels in a stressed or in an unstressed syllable differently than native listeners. The role of L2 experience was also explored. Results showed that the native and non-native listeners did not differ in category boundary locations in either syllable, although non-native perception was less categorical in the unstressed syllable. No effect of experience was found. It is concluded that the L2 learners redefined the value of vowel duration as a cue.

Keywords: Czech, vowel quantity, stress, L2 acquisition.

1. INTRODUCTION

Czech belongs to languages that exhibit phonemic quantity distinctions. The role of quantity in Czech phonology is relatively important: all 5 short Czech vowels contrast with their long counterparts and there is no limitation in the distribution of short/long (V/V:) oppositions across positions in a word. Vowel duration cues seem to be reserved exclusively for quantity. This is because within 4 V/V: pairs duration represents the primary acoustic difference (the last pair, /ɪ/ and /iː/, differs critically also in quality) (see [13] for a review). Czech stress, which is non-contrastive and is fixed to the first syllable of a word, is cued by the overall F0 contour and possibly intensity [14] rather than by vowel duration. Therefore, stress and vowel quantity are independent (cf. minimal pairs like /ráda/ 'advice' vs. /ráːda/ 'glad' and /hóle/ 'sticks' vs. /hóleː/ 'bare'; see also [9]).

For L2 learners, Czech vowel quantity may pose problems, especially if vowel duration serves as a cue for stress in their L1. Such problems seem to be noticeable even to lay public. This can be illustrated by a recent popular commercial in which foreign accent is imitated by producing the word /bóbiːka/ as [bóːbika].

A number of studies have examined if L2 users of a quantity language perceive vowel length differently than native speakers do. For example, McAllister *et al.* [10] showed that native English and Spanish advanced learners of Swedish were less successful in distinguishing correct and incorrect renderings of Swedish words containing short and long vowels and consonants. In this study the role of stress was not investigated: all target vowels were in a stressed syllable. Minagawa-Kawai *et al.* [11] found that English listeners tended to misidentify Japanese accented vowels as long and unaccented vowels as short. Finally, Nenonen *et al.* [12] tested vowel length categorization in advanced Russian learners of Finnish. Native and non-native responses differed in the second, unstressed syllable but not in the first, stressed one.

Learners of a quantity language who have an L1 like English (where durational cues are used allophonically [8] and only secondarily for vowel identification [4]) are faced with the challenge of altering their cue weighting. It can be expected that their chances to succeed in this task will improve as they gain experience. However, in the study by McAllister *et al.* [10] L2 experience did not predict L2 perceptual performance for the English and the Spanish listeners. On the other hand, Hirata [6] showed that English listeners' perception of Japanese quantity improved with training. There is also evidence from non-speech categorization that cue weighting can change when input is manipulated [7].

The focus of the present study was to investigate perception of Czech vowel quantity by L2 speakers of Czech, who had English as their L1. The specific questions were as follows: (1) Do native and non-native listeners differ with respect to category boundary locations in a stressed and in an unstressed syllable? (2) Do they differ with respect to boundary sharpness? (3) Do their

reaction times differ? (4) Does categorization become more native-like as learners gain L2 experience?

2. METHOD

To address the above questions, a perceptual experiment, inspired by the one described in Nenonen *et al.* [12], was conducted.

2.1. Subjects

Ten native Czech speakers were the control group. Sixteen American English speakers formed two experimental groups of 8: advanced and less advanced learners of Czech. These two groups differed significantly ($p<.01$) in years of learning and years of residence (see Table 1). They did not differ in self-estimated usage of Czech. An unwanted difference ($p<.05$) was that the advanced learners had a higher age of onset of learning, which could be a confounding factor. All the advanced (and some less advanced) learners spoke Czech fluently. No subjects reported hearing impairments or language dysfunctions.

Table 1: Average years of learning (YL), years of residence (YR), self-estimated usage of Czech (Cz %), and age of onset of learning (AOL) for the non-native subjects. SDs are in parentheses. The last column gives results of ANOVA between-group comparisons.

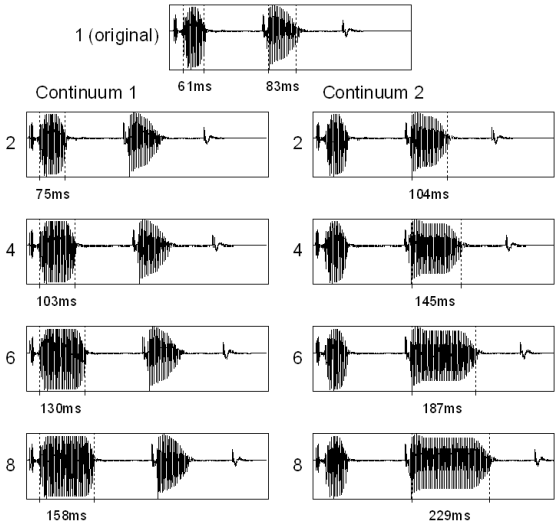
	Advanced learners	Less advanced learners	*p*
YL	8.1 (3.4)	0.9 (0.4)	<.01
YR	8.0 (3.5)	1.3 (0.8)	<.01
Cz %	38.4 (12.7)	27.5 (10.9)	-
AOL	27.0 (2.4)	21.5 (4.4)	<.05

2.2. Stimuli

Two 8-token continua were created by editing the nonsense word /kátap/ produced naturally by a female Czech speaker. The CVCVC structure was chosen, rather than CVCV, so that the lengthening effect of the final pre-pausal position (see e.g. [5]) did not coincide with the lengthening effect of an open syllable [5, 13], which would result in a large difference between the short-vowel durations in the first and second syllable.

In continuum 1, durations of the vowel in the first, stressed syllable varied (cf. audio file 1). In continuum 2, durations of the vowel in the second, unstressed syllable varied (cf. audio file 2). Vowel length was manipulated by selecting a fundamental cycle in the steady state of each vowel and pasting its copies after it. By trial and error, the optimal cycle was selected so that high-frequency noise and distortion occurring due to the concatenation were minimized. Fig. 1 shows waveforms of example stimuli. Vowel durations of all stimuli are given in Fig. 2 (horizontal axes).

Figure 1: Examples of the stimuli used in this study. In Continuum 1 the step was ca. 14ms; in Continuum 2 the step was ca. 21ms.



The specific pseudo-word was chosen because /kátap/, /ká:tap/, and /káta:p/ are all nonsense and at least 2 phonemes would need to be replaced in each to obtain an existing Czech word. Potential lexical effects (see e.g. [2]) biasing listeners' decisions were thus avoided.

The mean intensity of the stressed vowel was 2 dB higher and the mean pitch was 23 Hz higher than that of the unstressed vowel.

2.3. Procedure

Listeners participated in two separate single-interval two-alternative forced-choice tasks, one for testing V/V: categorization in each syllable. Both conditions were completed in one session but their order was counterbalanced across subjects. Within each task, the 8 respective stimuli were presented 10 times in isolation and in random order, thus giving 80 trials per tasks. Subjects listened to stimuli via headphones and indicated their decisions by clicking one of two buttons using a mouse. The buttons were marked 'katap' and 'kátap' or 'kataːp', i.e. using normal Czech orthography where a stroke indicates a long vowel. Prior to each condition, a training task was completed at least once with a subset of the stimuli (the endpoints were included). Testing took place in a quiet room.

2.4. Data processing

For each categorization curve, the location of the category boundary and curve steepness were estimated using the following procedure. It was assumed that the 'short' response probability *y* of a vowel of length *x* follows a step-like function:

$$y \approx \frac{1}{1+e^{a-bx}} \quad (1)$$

Standard log-link transformation of the *y* probabilities was used and linear regression then applied to calculate durations that would receive 75%, 50% and 25% 'short' labeling within each categorization curve, or *P*75, *P*50, and *P*25 respectively. Besides the boundary values where responses were mixed, only the longest stimulus with 100% 'short' labeling and the shortest stimulus with 0% 'short' labeling entered the computation. Steepness *s* was expressed as:

$$s = \frac{P25 - P75}{P50} \quad (2)$$

The category boundary (*P*50) and steepness (*s*) scores were then submitted to separate ANOVAs along with the overall categorization data and reaction times.

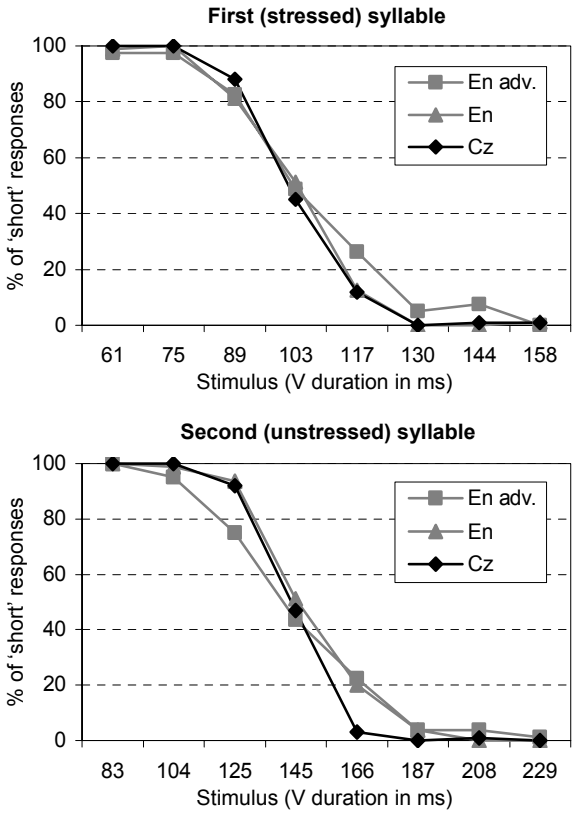
3. RESULTS

Fig. 2 shows group-averaged categorization curves for the first/stressed (top panel) and the second/ unstressed (bottom panel) syllable condition. When the categorization data were analyzed by 3 (Group) × 7 (Stimulus, rep. measure) ANOVA, there were no significant differences in the first-syllable condition. In the second-syllable condition however, a significant Group × Stimulus interaction was found ($F(12, 138)=3.1$, $p<.001$). A *post hoc* LSD test revealed that both English groups differed ($p<.01$) from the Czechs at the 166-ms stimulus and the advanced learners also at the 125-ms stimulus (cf. Fig. 2, bottom pan.). There was no difference in responses to the 145-ms stimulus which was the nearest to the boundary.

Table 2: Estimated V/V: category boundaries (ms) for the Czech (Cz), less advanced English (En), and advanced English (En adv.) listeners. SD are in parentheses. No differences reached significance.

	1st (stressed) syllable	2nd (unstressed) syllable
Cz	102.8 (3.7)	143.7 (4.0)
En	102.9 (4.6)	152.4 (7.0)
En adv.	106.9 (7.8)	144.8 (11.3)

Figure 2: Categorization curves of the native Czech (Cz), less advanced English (En), and advanced English (En adv.) listeners in the 1st-syllable (top panel) and the 2nd-syllable (bottom panel) conditions.



The latter finding is in agreement with one-way ANOVA analyses of the category boundary scores that are shown in Table 2. No differences between groups in either condition reached significance.

The significant Group × Stimulus interaction in the 2nd-syllable condition is reflected in the curve steepness scores. A main effect of Group was found in the 2nd-syllable condition ($F(2, 23)=4.43$, $p=.024$); a *post hoc* Scheffe's test revealed a significant ($p<.05$) difference between the advanced learners' and the native scores. A main effect of Group exceeded significance in the 1st-syllable condition too but it was not as strong ($F(2, 23)=3.67$, $p=.041$) and a *post hoc* Scheffe's test found no significant differences between groups.

When the reaction times (RTs) were analyzed in one-way ANOVAs a significant ($p<.0001$) main effect of Group was observed in both conditions. *Post hoc* tests showed that in both conditions, native listeners were quicker than the L2 listeners; differences between the two L2 were not significant. For all groups and both conditions, the RTs peaked at the near-boundary stimulus.

4. DISCUSSION

The primary aim of this study was to determine if allophonic use of vowel duration as a cue for stress in the L1 would affect acquisition of contrastive vowel quantity in L2. While it was found that perception of the English speakers of Czech in the unstressed syllable was less categorical than that of the native Czechs, the non-native and the native listeners differed with respect to the locations of the V/V: category boundary neither in the stressed syllable nor in the unstressed one. What was reported about Russian learners of Finnish [12] thus does not hold true for English learners of Czech. The L2 listeners in the present study seem to have successfully redefined the value of vowel duration as a cue and dissociated it from stress. What set the L2 listeners in the present study clearly apart from the native Czechs was the fact that their latencies were consistently longer.

The conclusion that stress did not interfere with V/V: categorization in the L2 listeners is even more interesting if we consider that no effect of L2 experience was found. There was even a trend in the opposite direction: the less advanced learners seemed to outperform the more advanced learners (the same was observed in Spanish learners of Swedish quantity [10]). Their categorization curves more closely matched the native ones and they had shorter response times on average (although the difference was not significant). It may be the case that this is because L2 experience was confounded with AOL in this study, advanced learners having a higher AOL on average (see Table 1).

Another possible explanation can be that duration is an acoustic dimension which is salient even for speakers of languages where quantity is not phonemic. It has been repeatedly reported [1, 3, 4] that Spanish learners of English tend to prefer durational cues to spectral ones when categorizing English /ɪ/ and /i/. Whether this is a superficial connection or not will be addressed in future research.

A conclusion that can be drawn with confidence from the present results is that in Czech a vowel located in a word-final pre-pausal position needs to be much (nearly 1.5 times) longer to be perceived as long than when it is in the first, stressed syllable. This shows that what is known about pre-boundary lengthening in Czech [5, 13] is reflected in perceptual strategies.

5. ACKNOWLEDGEMENTS

Many thanks are extended to E. Boudová, K. Chládková, M. Čujanová, J. Fröhlichová, K. Janíček, and B. Majstrík (students at Palacký University), who helped to collect data, to Šárka Šimáčková (Palacký University), who commented on an earlier version of this paper, and to David Boukal (Institute of Marine Research, Bergen, Norway), who helped with data processing.

6. REFERENCES

[1] Bohn, O.-S. 1995. Cross-Language Speech Perception in Adults: First Language Transfer Doesn't Tell It All. In: Strange, W. (ed), *Speech perception and linguistic experience: Issues in Cross Language Research.* Baltimore: York Press, 279-304.

[2] Burton, M. W., Baum, S. R., Blumstein, S. E. 1989. Lexical Effects on the Phonetic Categorization of Speech: The Role of Acoustic Structure. *J. Exp. Psych: Hum. Perc. Perf.* 15, 567-575.

[3] Cebrian, J. 2006. Experience and the use of non-native duration in L2 vowel categorization. *J. Phonetics* 34, 372-387.

[4] Escudero, P., Boersma, P. 2004. Bridging the gap between L2 speech perception research and phonological theory. *SSLA* 26, 551-585.

[5] Hála, B. 1962. *Uvedení do fonetiky češtiny na obecně fonetickém základě*. Prague: ČSAV.

[6] Hirata, Y. 2004. Training native English speakers to perceive Japanese length contrasts in word versus sentence contexts. *J. Acoust. Soc. Am.* 116, 2384-2394.

[7] Holt, L. L., Lotto, A. J. 2006. Cue weighting in auditory categorization: Implications for first and second language acquisition. *J. Acoust. Soc. Am.* 119, 3059-3071.

[8] Klatt, D. H. 1976. Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *J. Acoust. Soc. Am.* 59, 1208-1221.

[9] Lehiste, I. 1970. *Suprasegmentals*. Cambridge, MA: MIT Press.

[10] McAllister, R., Flege, J. E., Piske, T. 2002. The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *J. Phonetics* 30, 229-258.

[11] Minagawa-Kawai, Y., Maekawa, K., Kiritani, S. 2002. Effects of pitch accent and syllable position in identifying Japanese long and short vowels: Comparison of English and Korean speakers. *J. Phonetic Soc. Japan* 6 (2): 88-97.

[12] Nenonen, S., Shestakova, A. Alko, P. Huotilainen, M. 2003. Perception of vowel length in native speakers and second-language users of a quantity language. *Proc. 15th ICPhS* Barcelona, 531-534.

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

[14] Palková, Z., Volín, J. 2003. The role of F0 contours in determining foot boundaries in Czech. *Proc. 15th ICPhS* Barcelona, 1783-1786.