SHORT VS. LONG CATEGORY PERCEPTION AFFECTED BY VOWEL QUALITY

Einar Meister[a], *Stefan Werner*[b] *& Lya Meister*[a]

[a]Laboratory of Phonetics and Speech Technology, Institute of Cybernetics, Tallinn University of Technology, Estonia;
[b]Dept. of General Linguistics and Language Technology, University of Eastern Finland, Finland
einar@ioc.ee; stefan.werner@uef.fi; lya@phon.ioc.ee

ABSTRACT

The paper studies the impact of intrinsic vowel duration on category perception in boundary conditions by examining the short vs. long category boundary perception in Estonian. Since the intrinsic duration of a close vowel /i/ is about 10-15 ms shorter than that of an open vowel /a/, we hypothesize that the short vs. long category boundary in /i/ occurs at a shorter duration than in the case of /a/.

Twelve native Estonian subjects participated in the perception tests involving binary category decision in CV(:)CV stimuli where the duration of the primary stressed vowel was manipulated in a range from 100 ms to 190 ms embracing the short vs. long category boundary. The test results support our hypothesis and show that the differences in vowel quality, hence in intrinsic duration of vowels play a role in short vs. long category perception.

Keywords: Estonian**,** intrinsic duration, category boundary

1. INTRODUCTION

In quantity languages such as Estonian and Finnish short and long phonological oppositions exist in both vowels and consonants. Short vs. long category discrimination is crucial in perception since vowel and consonant duration serves to distinguish lexical minimal pairs in Estonian and Finnish. Typically, in different perception studies aiming at defining the perceptual boundary of the binary quantity oppositions in vowels, stimuli involving only one vowel in different consonantal contexts are used, most frequently the vowel /a/ (e.g. [8, 18]). The studies on microprosodic features have established systematic differences in the intrinsic duration of high and low vowels (e.g. [4, 19, 22]). It has been reported that in non-quantity languages in boundary conditions when primary features do not provide sufficient information for category discrimination, the intrinsic duration of vowels acts as a secondary feature facilitating the perceptual decision [11], [23]. Recently, a similar role of microprosodic variations in vowel quality perception has been reported for quantity languages (Estonian and Finnish), as well [20].

The aim of the current study is to find more evidence for the impact of intrinsic vowel duration on category perception in boundary conditions by examining the categorical distinction short vs. long. Taking into account that the intrinsic duration of a high vowel /i/ is about 10-15 ms shorter than that of a low vowel /a/, we hypothesize that the short vs. long category boundary in /i/ occurs at a shorter duration than in the case of /a/. The hypothesis is not trivial since in quantity languages duration has to be intentionally controlled by a speaker to signal quantity contrasts and this "higher order" control can "override" the intrinsic features attributed to physiological properties of the human vocal tract. To our knowledge no experimental study has been carried out to investigate the role of vowel quality (with respect to intrinsic duration) in the perception of the short vs. long category in Estonian.

2. ESTONIAN QUANTITY SYSTEM

Estonian is well-known as a language with a three-way quantity system involving contrastive prosodic patterns traditionally referred to as short (Q1), long (Q2) and overlong (Q3) quantity degrees. The early treatments of Estonian quantity system (e.g. [1]) postulated a three-way contrast based on segmental duration of the stressed syllable (i.e. all vowels and consonants occur in three contrastive quantities in this position).

However, a large number of subsequent studies have shown that the system is much more sophisticated – nowadays there is a general agreement that the ternary phonological contrast in Estonian is characterized by a complex interaction

of durational and tonal cues within a disyllabic foot and only a binary opposition (short vs. long) exists at the phonemic level (e.g. [7, 8, 13, 15, 16]). Although the duration of the vowel in the unstressed syllable varies to a great extent, there is no phonological length contrast in the unstressed syllable; vowels in unstressed syllables are classified as phonologically short.

2.1. Duration of Estonian short and long vowels: acoustic data

Duration data of speech segments obtained from acoustic analysis of Estonian speech has been reported in a large number of studies (mainly in the context of word prosody), starting with seminal works such as [5, 14, 17] to more recent studies [2, 6, 8, 12, 21] among others.

Pooling duration data from the different studies above shows that in Estonian phonologically long vowels are 1.6-2.0 times longer than phonologically short vowels when comparing stressed vowels of Q2 and Q1, and 2.1–2.9 times longer when comparing stressed vowels of Q3 and Q1. In a stressed syllable the absolute duration of a phonologically short vowel ranges from 60 ms to 129 ms and the duration of phonologically long vowel spans from 124 ms to 371 ms; in an unstressed syllable the duration of a vowel varies from 62 ms to 227 ms.

Similar duration data has been provided for other quantity languages such as Finnish and Japanese [9, 10, 24, 25].

The acoustic data shows that in quantity languages short and long vowel categories are well distinguished in speech production – the durational ratio of short and long vowels around 2.0 is quite stable even at different speaking rates and should guarantee reliable perceptual distinction of these phonological categories.

2.2. Perception data: short vs. long category boundary

The perceptual boundary of short and long vowel categories of native Estonian subjects was studied in one-syllable synthetic CV(:)C nonsense words (*sa(:)s*) by varying the duration of the vowel (from 40 ms to 250 ms in 10 ms steps) at three speaking rates, whereas the duration of the word-initial consonant (70, 100 and 130 ms) was assumed to play the role of a cue for speaking rate. It was found that the short/long category boundary is a function of speaking rate – the mean category boundary occurred at a vowel duration of 100 ms in the case of C1 = 70 ms (fast speech), at 120 ms in the case of C1 = 100 ms (moderate speech), and at 140 ms for C1 = 130 ms (slow speech) [8].

A study on the perception of quantity degrees in Estonian using re-synthesized two-syllable CV(:)CV words with manipulated duration of the vowel /a/ in the first syllable provides the short vs. long boundary at 115-130 ms depending on the regional dialectal background of the subjects [1].

Thus, the perception studies reporting results on the short vs. long category boundary in Estonian vowels are actually obtained from experiments involving the open vowel /a/ only.

In Finnish, the short/long category boundary has been studied [26] in CV(:)CV(:) words by manipulating the duration of vowels in the first and second syllable (in contrast to Estonian, in Finnish short/long oppositions occur in both stressed and unstressed syllables) resulting in two continua: (1) from /tu:ku/ to /tuku/, and (2) from /tuku:/ to /tuku/. The perception tests showed that for native Finns the short vs. long category boundary is at 116 ms in the stressed first-syllable position and at 162 ms in the unstressed second-syllable position.

3. EXPERIMENTAL SETUP

3.1. Stimulus corpus

The experiment is designed to discover short vs. long category boundaries in close and open vowels in different plosive context of CV(:)CV nonsense words /kaka/-/kiki/, /papa/-/pipi/, /tata/-/titi/. The words were pronounced in isolation by a native male speaker as Q1 words (i.e. with a short primary-stressed vowel). In all words the duration of the first vowel was manipulated from 100 ms to 190 ms in 10 ms steps which consequently resulted in six stimulus sets from CVCV to CV:CV – /kaka/ vs. /ka:ka/, /papa/ vs. /pa:pa/, /tata/ vs. /ta:ta/, /kiki/ vs. /ki:ki/, /pipi/ vs. /pi:pi/, /titi/ vs. /ti:ti/. The durations of the other segments were kept constant (C1(burst) = 25 ms for /k/, 15 ms for /p/ and /t/; C2 = 75 ms; V2 = 240 ms); the F0 was set to a constant value of 100 Hz in both vowels. The number of different stimuli in all sets was 10. The manipulation of stimuli was done with Praat [3].

3.2. Subjects

Twelve native Estonian adults (6 male and 6 female) participated in the experiment; none of the subjects were phonetically trained or reported any hearing problems.

3.3. Procedure

All testing was conducted in a sound-isolated room and stimuli were presented to subjects via high-quality headphones. The test was administered with Praat's multiple forced-choice test facility and organized in six blocks (according to six stimulus sets). Each word was repeated three times in random order. In the test listeners had to decide on vowel quantity in a binary identification task and answer by clicking in one of two response boxes on the screen. The boxes were labeled "CVCV" and "CVVCV". In total 180 stimuli (10 stimuli x 6 sets x 3 repetitions) were presented to subjects. The duration of the test was 15-20 minutes.

4. RESULTS

The results of each test block were presented as plots of the proportion of "short" responses against stimulus duration. The categorization functions were obtained by interpolating the response data with a logistic function (figure 2) and the locations of short vs. long category boundaries (in milliseconds) were calculated as the 50% cross-over point of the fitted curve (figure 1).

Pairwise comparison of words showed that the difference of boundary means was highly significant in all pairs (/kaka/ vs. /kiki/: t = 4.026, df = 11, p-value = 0.002; /papa/ vs. /pipi/: t = 5.386, df = 11, p-value = 0.0002; /tata/ vs. /titi/: t = 5.233, df = 11, p-value = 0.0003).

Figure 1: Box-plots of short/long category boundaries in all stimulus sets.

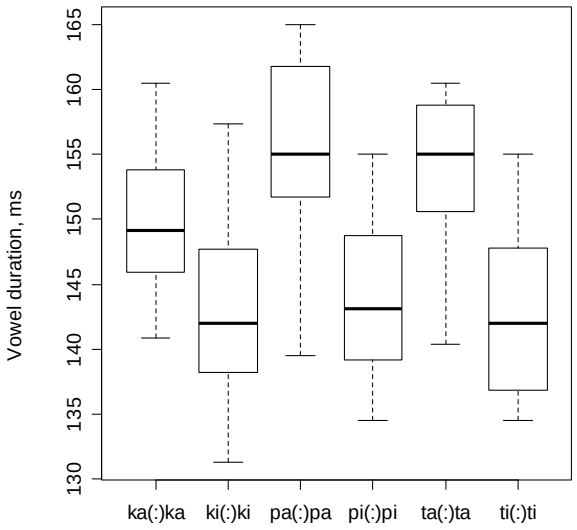
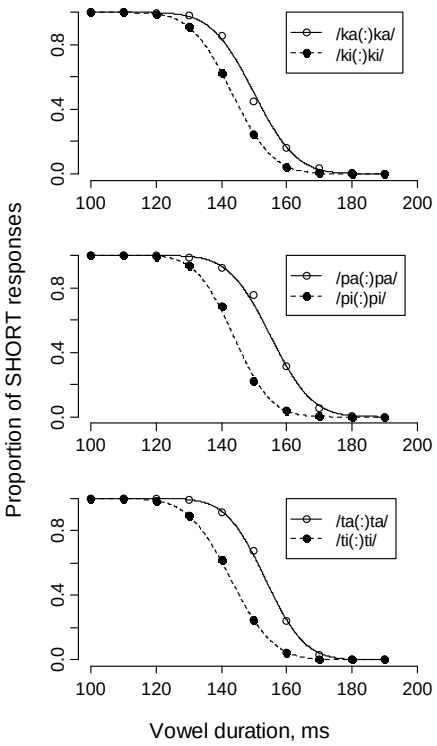


Figure 2: The averaged perception scores of each stimulus step ("•" for close vowel, "○" for open vowel) and the categorization functions in all stimulus sets.



The ANOVA analysis confirmed that consonant context had no effect on boundary means (F = 1.198, df = 2, p = 0.3), nor was the consonant * vowel interaction significant (F = 0.833, df = 2, p = 0.44), while the expected vowel type showed a highly significant effect (F = 35.469, df = 1, p < 0.001).

Table 1: Mean values and standard deviations (in ms) of short/long category boundaries in all stimulus sets.

	/kaka/	/kiki/	/papa/	/pipi/	/tata/	/titi/
Mean	149,8	143,1	155,2	143,8	153,7	142,8
SD	6,51	7,35	7,97	6,38	6,40	6,60

5. DISCUSSION

The CV(:)CV stimulus sets were designed to elicit short/long judgments at the phonemic level by manipulation of the duration of the primary-stressed vowel. For the Estonian subjects the CVCV vs. CV:CV opposition is not just a short vs. long opposition of the primary-stressed vowel, but it also unavoidably evokes the language-specific foot-level contrast of quantity degrees Q1 vs. Q2.

The results confirm our hypothesis that the intrinsic duration of a vowel plays a significant role in the discrimination of Estonian short vs. long phonological category at boundary conditions. In

our opinion, our results do not lend support to the claim that the intrinsic features are under control of the speaker [23]. On the contrary, they tend to support the idea that the intrinsic features are attributed to physiological properties of the human vocal tract and cannot be overridden by the speaker's articulatory program.

6. SUMMARY

The experiment confirmed our hypothesis that the intrinsic microdurational variations of high vs. low vowels do affect the perception of duration-based phonological oppositions in a quantity language. The results support the standpoint that intrinsic features are determined by physiological constraints of the human articulatory system and cannot be intentionally controlled by the speaker.

7. ACKNOWLEDGEMENTS

This research was partially supported by the target-financed theme No.0322709s06 of the Estonian Ministry of Education and Research and by the National Program for Estonian Language Technology. All subjects who participated in the perception tests are gratefully acknowledged.

8. REFERENCES

[1] Ariste, P. 1939. A quantitative language. *Proceedings of the Third International Congress of Phonetic Sciences* Ghent, 276-280.

[2] Asu, E.L., Lippus, P., Teras, P., Tuisk, T. 2009. The realization of Estonian quantity characteristics in spontaneous speech. In Vainio, M., Aulanko, R., Aaltonen, O. (eds.), *Nordic Prosody. Proceedings of the Xth Conference*. Frankfurt: Peter Lang Verlag, 49-56.

[3] Boersma, P., Weenink, D. 2010. Praat: doing phonetics by computer (Version 5.2) [Computer program]. Retrieved November 1, 2010, *http://www.praat.org/*.

[4] di Cristo, A. 1978. *De la Microprosodie à l'Intonosyntaxe*. Thèse d'Etat. Université de Provence, Aix-en-Provence.

[5] Eek, A. 1974. Observations on the duration of some word structures: I. *Estonian Papers in Phonetics*, 18-31.

[6] Eek, A., Meister, E. 1998. Quality of standard Estonian vowels in stressed and unstressed syllables of the feet in three distinctive quantity degrees. *Linguistica Uralica*, 3(34), 226-233.

[7] Eek, A., Meister, E. 1999. Estonian speech in the BABEL multi-language database: Phonetic-phonological problems revealed in the text corpus. In Fujimura, O., Joseph, B.D., Palek, B. (eds.), *Proceedings of LP'98: Item Order in Language and Speech*. Prague, The Karolinum Press, 529-546.

[8] Eek, A., Meister, E. 2003. Foneetilisi katseid ja arutlusi kvantiteedi alalt (I). Häälikukestusi muutvad kontekstid ja välde. *Keel ja Kirjandus*, 11-12, 815-837, 904-918.

[9] Hirata, Y. 2004. Training native English speakers to perceive Japanese length contrasts in word versus sentence contexts. *Journal of the Acoustical Society of America* 116, 2384-2394.

[10] Isei-Jaakkola, T. 2004. *Lexical Quantity in Japanese and Finnish*. University of Helsinki.

[11] Kouznetsov, V.B., 2001. Perceptual role of inherent vowel duration as distinctive feature in Russian. *Proceedings of the XI Session of the Russian Acoustical Society*. Moscow, 443-447. *http://www.akin.ru/Rao/sess11/sect5sp.htm*

[12] Krull, D. 1997. Prepausal lengthening in Estonian: Evidence from Conversational speech. In Lehiste, I., Ross, J. (eds.), *Estonian Prosody: Papers from a Symposium.* Tallinn: Institute of Estonian Language, 136-148.

[13] Krull, D., Traunmüller, H. 2000. Perception of quantity in Estonian. *Proceedings of Fonetik 2000*, 85-88.

[14] Lehiste, I. 1960. Segmental and syllabic quantity in Estonian. *American Studies in Uralic Linguistics 1*, Bloomington: Indiana University Press, 21-82.

[15] Lehiste, I. 1997. Search for phonetic correlates in Estonian prosody. In Lehiste, I., Ross, J. (eds.), *Estonian Prosody: Papers from a Symposium.* Tallinn: Institute of Estonian Language, 11-35.

[16] Lehiste, I. 2003. Prosodic change in progress: from quantity language to accent language. In Fikkert, P., Jacobs, H. (eds.), *Development of Prosodic Systems*. Berlin, New York: Mouton de Gruyter, 47-65.

[17] Liiv, G. 1961. Eesti keele kolme vältusastme vokaalide kestus ja meloodiatüübid. *Keel ja Kirjandus*, 412-424, 480-490.

[18] Lippus, P., Pajusalu, K. 2009. Regional variation in the perception of Estonian quantity. In Vainio, M., Aulanko, R., Aaltonen, O. (eds.), *Nordic Prosody – Proceedings of the Xth Conference.* Frankfurt: Peter Lang Verlag, 151-157.

[19] Meister, E., Werner, S. 2006. Intrinsic microprosodic variations in Estonian and Finnish: Acoustic analysis. In Aulanko, R., Wahlberg, L., Vainio, M. (eds.), *The Phonetics Symposium 2006 (Fonetiikan Päivät 2006)*. Department of Speech Sciences, University of Helsinki, 53, 103-112.

[20] Meister, E., Werner, S. 2009. Duration affects vowel perception in Estonian and Finnish. *Linguistica Uralica* 45(3), 161-177.

[21] Meister, L., Meister, E. 2011. Perception of the short vs. long phonological category in Estonian by native and non-native listeners. *Journal of Phonetics* 39(2), 212-224.

[22] Peterson, G.E., Lehiste, I. 1960. Duration of syllable nuclei in English. *Journal of the Acoustical Society of America* 32(6), 693-703.

[23] Solé, M.J. 2007. Controlled and mechanical properties in speech: a review of the literature. In Solé, M.J., Beddor, P., Ohala, M. (eds.), *Experimental Approaches to Phonology*. Oxford: Oxford University Press, 302-321.

[24] Suomi, K., Toivonen, J., Ylitalo, R. 2008. Finnish sound structure: phonetics, phonology, phonotactics and prosody. *Studia humaniora ouluensia 9*, Oulu University Press.

[25] Suomi, K., Ylitalo, R. 2004. On durational correlates of word stress in Finnish. *Journal of Phonetics* 32, 35-63.

[26] Ylinen, S., Shestakova, A., Alku, P., Huotilainen, M. 2005. The perception of phonological quantity based on durational cues by native speakers, second-language users and nonspeakers of Finnish. *Language and Speech*, 48(3), 313-338.