

INHERENT VOWEL DURATION IN RUSSIAN: PRODUCTION
AND PERCEPTION DATA

VLADIMIR B. KUZNETSOV

ARVO OTT

ANATOLY V. VENTSOV

Moscow State Institute
of Foreign Languages
Moscow, USSR 119034

Dept. of Computer Control
Institute of Cybernetics
Tallinn, Estonia, USSR 200108

Pavlov Institute of
Physiology
Leningrad, USSR 199164

ABSTRACT

In this paper the quantitative data concerning inherent vowel duration in Russian are presented. The established duration differences are estimated from the point of view of their perceptual significance in an experiment on lexical stress location.

INTRODUCTION

It is a well established fact that all else being equal, duration of stressed vowels is determined by the characteristics of the corresponding vowel gesture: the higher the tongue, the shorter the vowel. To be more specific, vowel duration is considered to be a result of superposition of jaw movement on the opening and closing gesture of the lips [1]. Thus, the vowel of a given phonetic identity has its own characteristic duration - inherent vowel duration (IVD).

Although in Russian IVD phenomenon has been studied by many researchers [2,3,4] the question whether the observed duration differences have any significance either at the production or perception level is still opened. This is mainly due to the fact that most results of measurements are qualitative. Among the other reasons, poor control of the experimental conditions, interpretation of the data in terms of different phonetic categories (phonemes vs. allophones), unjustified averaging over the speakers and phonetic contexts have to be mentioned.

The aim of the experiments reported in this paper was to gather statistically reliable and valid data on IVD and to assess the perceptual importance of the determined IVD differences.

EXPERIMENT 1: INHERENT VOWEL DURATION (IVD)

Vowel durations were measured in a monosyllable of CVC type, spoken as word in an identical sentence frame "Say...again".

Palatalized or nonpalatalized fricative [s] was used to form a symmetrical environment. Each of four speakers (two male ones: M1, M2 and two female ones: W1, W2) recorded a list of 330 sentences (10 vowels * 33 repetitions). To achieve constant speech rate throughout the recording session the speaker was asked to synchronize the onset of the sentences with a periodic light pulse. The phonetic identity of the test vowels was checked up by 8 listeners during an identification experiment.

Three segmentation procedures were used to measure IVD. The beginning and end of a vowel were recognized: (1) by the onset/offset of voicing; (2) by the offset/onset of high-frequency noise; (3) by sharp minimums on the amplitude curves. There was a good agreement among the three sets of measurements. Taking into consideration the fact that the third segmentation procedure produced the least scatter of measurements, only the data obtained by this procedure were subjected to further analysis.

The amplitude envelope was obtained by processing the tape recordings of the speech material through a Brüel&Kjaer graphic level recorder operated at a 100 mm/sec paper - and 1000 mm/sec writing-speeds with a high frequency preemphasis. To increase time resolution the play-back speed of the tape-recordings was reduced twice. Preliminary spectrographic analysis of the test words revealed that the minima on the amplitude envelope coincided with the onset of the voicing of the vowel, on the one hand, and the rapid energy decrease in the frequency region of the second and higher vowel formants and the onset of the frication noise, on the other. It is commonly agreed, that these acoustic cues provide reproducible and valid boundaries for the measurement of vowel duration [5]. Mean vowel duration is regarded as an estimate of IVD. The four data matrices (10*33) were submitted to the following statistical analysis: (1) to determine the effect of tempo changes a one-way analysis of variance was carried out (to test the null hypothesis that the mean vowel dura-

tion of the 33 consecutive groups of ten different vowels are the same); (2) similar technique was used to assess significance of IVD differences; (3) T-method of multiple comparisons [6] was applied to determine the critical value of IVD difference; (4) to test the significance of the proposed classification of vowels according to their IVD the duration data was subjected to S-method of multiple comparisons [6]. All statistical tests were conducted with alpha=0.05. More detailed description of the experimental method and statistical procedures is presented in [7].

Results of the analysis of variance have shown that the speaking rates were kept constant and the differences in vowel durations were statistically significant. Critical IVD difference was within 6-7 msec. By means of the method of multiple comparisons the vowels were reliably rank-ordered and subdivided according to their IVD values into the following classes: {ä,a}, {ö,ë,э,ю,ы} and {y,и,и} (cyrillic characters are used to symbolize the vowels, the dots above letters indicate the allophones in the context of palatalized consonants). Typical duration ratio of the IVD means in the classes is 1.00:0.90:0.75. The IVD patterns of the four speakers are presented in Fig.1.

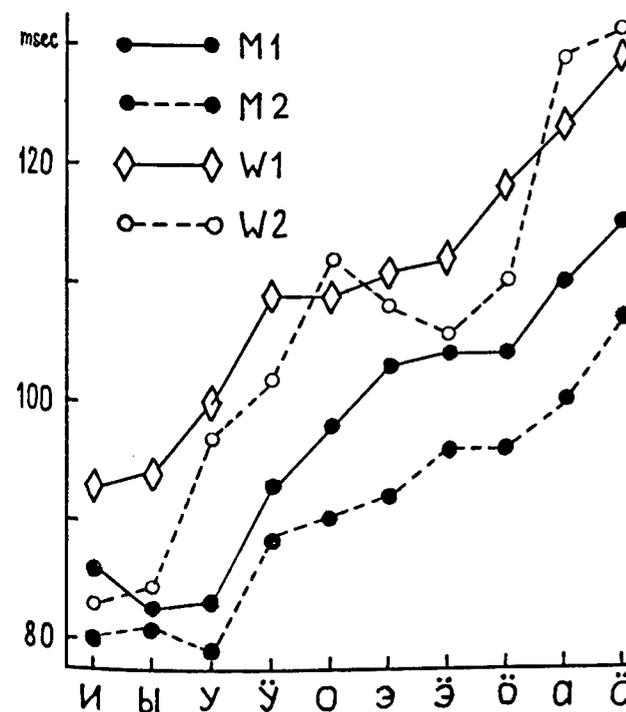


Fig.1 IVD pattern of the four speakers

EXPERIMENT 2: RANGE OF IVD VARIABILITY

To estimate the range of IVD variability two types of speech material were used: (1) the same CVC syllables as in experiment 1, but spoken in isolation, (2) five-syllable nonsense word [bist<V>tfurnaja] embedded in the carrier phrase "I was told that the oldest ... company had gone bankrupt". The immediate consonantal context of the test vowel was palatalized or non-palatalized. Speaker M1 was instructed to read the materials at his normal speech rate and not to insert any pauses into the carrier sentence. The arrangement of the speech material, recording conditions and segmentation procedure were similar to those described above. A more detailed description of the experiment may be found in [8]. Mean vowel durations measured in the two contexts are regarded as maximum and minimum IVD estimates respectively. Maximum and minimum IVD along with the normal ones averaged across vowels belonging to the same class are presented for speaker M1 in Table 1. Examination of the table reveals that our data on minimum IVD support the suggestion put forward in [9] that vowel incompressibility is relative to its inherent duration.

Table 1. Mean IVD for vowel classes (in msec)

IVD	classes			IVD ratio
	ä,a	ö,ë,э,ю,ы	y,и,и	
normal	113	101	84	1.00:0.90:0.75
maximum	167	152	142	1.00:0.91:0.85
minimum	90	75	63	1.00:0.84:0.71

Experimental data reported in this paper provide some evidence against the concept that in Russian the degree of opening is a single factor determining vowel duration. One can hardly explain, for example, within the framework of Lindblom's model of lip-jaw coordination [1] why the vowels [и] and [э] that are produced by quite similar articulatory gestures, judging from the corresponding F-patterns [13], differ so much in IVD, and the vowels [ү,ö,ä] are systematically longer than [y,о,a], though it is recognized that [ä] is closer than [a]. Further research is needed to clarify the significance of these findings. Nevertheless, it may be concluded that the acoustically-defined IVD is characteristic of Russian as well. But still there is a possibility that the auditory system of human being uses quite different segmentation criteria for the measurement of subjective duration and the differences in IVD might be already neutralized at the stage of measurements. Evidently, the

answer to this question may be obtained only by investigation of speech perception.

EXPERIMENT 3: PERCEPTUAL ROLE OF IVD

The data on the perceptual role of IVD are rather contradictory: on the one hand, there is an evidence that "naturalness" of synthetic vowels is increased if an appropriate IVD pattern is used to control the duration parameter [10], but, on the other hand, the results reported in [11] indicate that IVD is not important for the perception of stress.

Since in Russian the vowel duration is known to signal the position of word stress [12], the mechanism of stress perception has to take into account IVD which must have effect on the results of psycho-acoustic experiments with the vowel of different identity.

In the experiment described below natural russian words were used as stimuli. Most of the words were disyllabic with an open final syllable. The first vowel of the words was either [и] or [y], the second one was always [a]. It should be noted that the spectral properties of these vowels do not change appreciably in the pre- and post-stressed positions. The words formed phonetic minimally contrastive pairs, differing only in the position of stress, for example, "ти'хо-тихо'", "у'хо-уха'".

The natural vowels were replaced in the words by semi-synthetic ones of required duration. The method of stimulus generation and experimental procedure are described in full detail in [14]. Duration of the first vowel was 11 fundamental periods (one period was 8.7 msec). Duration of the second vowel varied from 7 to 23 periods. In Fig.2 and 3 the frequency of response "the second vowel is stressed" is plotted as a function of its duration for the vowel pairs [и-а] and [y-а] respectively. The lines designated by opened circles represent the data when the amplitudes of the vowels in the pair were made equal, the lines marked with crosses represent the data when the effective values of the vowels were equalized. For the sake of comparison the results of the experiment

with words comprising identical vowels are also presented (light lines without special signs).

Let us assume that the vowels having the same subjective duration, are judged as stressed with equal probability, then from the results displayed in Fig.2 and 3 it follows that the vowel [а] must have a longer acoustic duration to be perceived as subjectively equal to the vowels [и] and [y]. Since the test vowels were produced by repetition of one fundamental

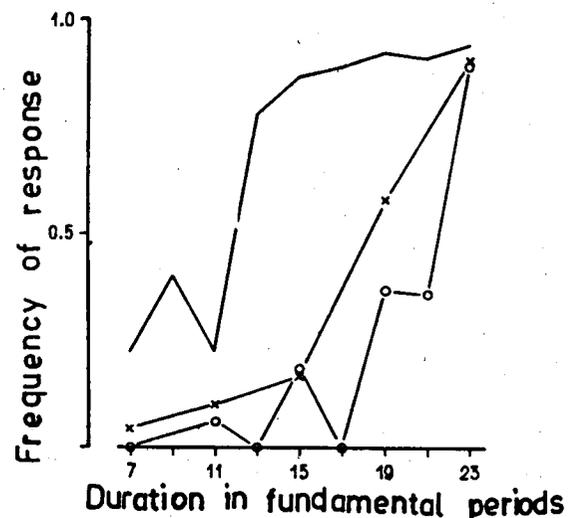


Fig. 2.

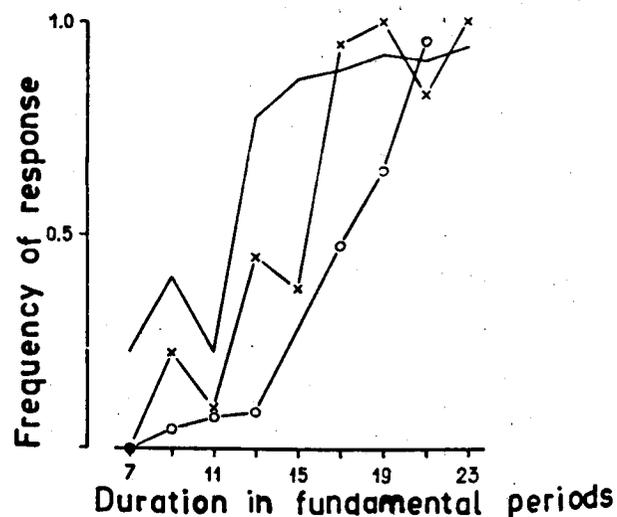


Fig. 3.

period singled out from the steady-state part of the corresponding natural vowel [14], the established discrepancy between the acoustic and subjective durations can not be ascribed to the segmentation. Consequently, this discrepancy may be considered to be a result of the IVD effect on the perception of stress when the judgement of stress is based on vowel duration. Thus, the reality of IVD in Russian has been demonstrated not only at the production but at the perception level as well.

This raises an interesting problem of establishing formal rules concerning IVD that may be used in the algorithmic descriptions of word stress perception.

REFERENCES

- [1] B.Lindblom "Vowel duration and a model of lip-mandible coordination", STL-QPSR, 4, 1967
- [2] L.V.Shcherba "Qualitative and quantitative description of russian vowels" (in russian), Peterburg, 1912
- [3] L.R.Zinder "Effect of speech rate on production of some sounds" (in russian) In: "Phonetic problems", Leningrad, 69, p.3-27, 1964
- [4] K.Bolla "Some problems of vowel duration ratio in Russian" (in russian), Bulletin of Moscow State University, Philology, Moscow, 3, p.54-62, 1968
- [5] E.Fisher-Jorgensen, B.Hutters "Aspirated stop consonants before low vowels, a problem of delimitation" Ann.Rep.Inst.Phon.Univ.CPH, Copenhagen, 15, p.77-102, 1981
- [6] G.V.Glass, J.C.Stanly "Statistical methods in education and psychology", New Jersey, 1970
- [7] V.B.Kuznetsov, A.Ott "Inherent vowel duration in Russian" Estonian Papers in Phonetics, Tallinn, p.67-95, 1987
- [8] V.B.Kuznetsov, A.Ott "Speech synthesis by rule. Algorithm of phonetic transcription and control of segmental duration" (in russian), Tallinn, 1987 (in print)
- [9] D.H.Klatt "Interaction between two factors that influence vowel duration" Journ.Acoust.Soc.Am., 54, p.1102-1104, 1973
- [10] R.Petersen "The influence of tongue height on the perception of vowel duration in Danish" Ann. Rep.Inst.Phon.Univ.CPH., Copenhagen, 8, p.1-10, 1974
- [11] E.Rosenvald "The role of intrinsic FO and duration in the perception of stress" Ann.Rep.Inst.Phon.Univ.CPH., Copenhagen, 15, p.147-166, 1981
- [12] N.I.Jinkin "Perception of stress in Russian words" (in russian) Trans. of the Academy of pedagogical sciences, Moscow, 54, p.7-83, 1954
- [13] V.B.Kuznetsov, A.Ott "Spectral properties of russian stressed vowels in the context of palatalized and non-palatalized consonants" (contribution to the present congress), Tallinn, 1987
- [14] A.V.Ventsov "What is the reference that sound durations are compared with in speech perception?", Phonetica, 40, p.135-144, 1983