

Perceptual Assessment of the Degree of Russian Accent

Lya Meister

Laboratory of Phonetics and Speech Technology
Institute of Cybernetics
Tallinn University of Technology
Akadeemia tee 21, Tallinn 12618, Estonia
lya@phon.ioc.ee

Abstract

This paper deals with the perceptual assessment of Russian-accented Estonian. Speech samples were recorded from 20 speakers with a Russian background; clips of about 20 seconds from each speaker were selected for this perceptual study. The accentedness was rated in two tests: first, 20 native Estonian speakers judged the samples and rated the degree of foreign accent on a six-point interval scale; secondly, two experienced phoneticians carried out a perceptual study of the same samples and compiled the list of pronunciations errors. The results of both listening tests were highly correlated – the higher the degree of accentedness given to a L2-speaker by naïve listeners, the more pronunciation errors were found by trained experts. The classification of most frequent pronunciation errors based on acoustic-phonetic features is given, as well.

1 Introduction

Native speakers/listeners can easily identify non-native speech and are able to rate the degree of foreign accent (FA). Naïve listeners' judgments of FA degree are based on their general perceptual impression rather than on conscious use of acoustic-phonetic knowledge about their own first language (L1). Accentedness ratings result in the degree of global foreign accent which is an impressionistic measure to which the speech of a second

language (L2) speaker deviates from that of L1-speakers (Southwood & Flege, 1999). On the contrary, a trained phonetician should be able to identify and classify different accent phenomena as well as describe them in terms of deviations of acoustic-phonetic features.

Following the findings and methodology presented in a recent paper (Meister, 2006; for methods employed in different studies see Jesney, 2004) on the accentedness rating of foreign-accented Estonian, two further listening experiments have been designed. The aim of these experiments is to compare the accentedness ratings given by naïve listeners, and the results of perceptual analysis of pronunciation errors carried out by experienced phoneticians. It is expected that the results of these two groups of raters harmonize quite well, i.e., the higher the accentedness ratings by naïve listeners of L2 speakers are, the more pronunciation errors are listed by experts. The study serves also a long-term goal – the development of criteria for speaking proficiency assessment, including the degree of FA.

2 Method

2.1 Speech samples and speakers

The speech material used in the study was recorded from 20 L2-speakers (14 female, 6 male) during the high-level language test at the National Examination and Qualification Centre. One of the sub-tasks the examinees have to perform is the conversation in pairs on a given topic which should demonstrate different speaking skills: expression of opinion, argumentation, turn-taking and carry-

ing on the conversation, etc. (Pajupuu et al., 2002). It is expected that a person with high-level language skills is able to communicate in written and spoken Estonian with near-native proficiency.

The recordings of the conversations were carried out using a digital recorder (sampling frequency 44.1 kHz, 16 bit, mono) and a high-quality microphone placed at a ca. 1 m distance from the speakers. With each pair of subjects, six to eight minutes of spontaneous conversation was recorded. A continuous clip of speech with the duration of ca. 20 seconds from each subject's speech was chosen for perceptual assessment. The clips were stored into an audio file in random order with an inter-stimuli interval of five seconds.

In addition to the speech recordings, each subject filled out a questionnaire concerning their linguistic background, age of L2 acquisition, use of L1, L2, etc. The summary of the speakers' information is presented in Table 1.

2.2 Listeners

Two groups of listeners were employed in the study. First, a group of naïve (non-linguist) listeners was composed of 20 native Estonians (10 female, 10 male) in the age range of 17 to 62. All of them had some knowledge of Russian and diverse exposure to foreign-accented Estonian spoken by Russians; none of them reported any hearing problems.

A second group of judges consisted of two trained phoneticians (native Estonians, one female, one male, both 49 years of age) with good knowledge of Russian and experience in experimental studies of Estonian as L2.

2.3 Experimental setup

Before the test a foreign accent scaling technique was introduced and several examples of L2 speech with different degrees of accentedness were played to the listeners. The participants were instructed to focus only on deviations in pronunciation, while grammatical and lexical errors should be ignored.

In the first part of the experiment the stimuli were played to subjects from a notebook computer via headphones in a quiet environment. The task of the judges was to rate the degree of foreign accent of each stimulus on an interval scale from 1 – "no foreign accent" to 6 – "very heavy foreign accent". The group of naïve listeners heard each stimulus only once; during the inter-stimulus intervals they had to write down their ratings on an answer sheet. The duration of the listening session was about nine minutes.

In the second part of the experiment, two experts carried out an exhaustive perceptual analysis of each stimulus and compiled the list of perceived pronunciation errors classified into five major groups typical to Russian-accented Estonian: (1) deviation of temporal structure, (2) location of word stress, (3) quality of vowels and diphthongs, (4) palatalization, and (5) voicing of voiceless consonants (Meister and Meister, 2005).

In the first stage the experts carried out error analysis independently from each other using repeated listening: this resulted in two different lists of pronunciation errors. Later, the disagreements in errors were discussed and analyzed together until the experts reached a common agreement.

Table 1. Summary of the background information of L2-speakers (EST = Estonia(n), RUS = Russia(n), UKR = Ukraine (Ukrainian)).

Speaker's ID	L2-speakers' data																			
	Sp1	Sp2	Sp3	Sp4	Sp5	Sp6	Sp7	Sp8	Sp9	Sp10	Sp11	Sp12	Sp13	Sp14	Sp15	Sp16	Sp17	Sp18	Sp19	Sp20
Age	52	23	19	19	16	25	26	32	19	18	20	51	19	43	20	33	18	46	45	32
Gender	F	M	F	F	F	M	F	F	F	F	M	F	M	F	M	F	F	M	F	F
Country of birth	Est	Ukr	Est	Est	Est	Est	Est	Est	Est	Est	Rus	Est	Est	Est	Est	Est	Est	Rus	Rus	Est
Age of L2 acquisition	5	9	5	7	7	1	12	12	9	5	9	5	9	8	9	20	5	30	23	9
Language(s) used at home	Rus	Rus	Rus	Rus	Rus	Est Rus	Rus	Rus	Ukr	Rus	Rus	Rus	Rus	Est	Rus	Rus	Rus	Rus	Est	Rus
Language(s) used at work	Rus Est	Est Rus	Rus	Est	Rus	Est Rus	Est Rus	Est Rus	Rus	Rus	Est	Est	Rus	Est Rus	Est	Rus Est	Rus Est	Rus	Est Rus	Rus Est
Friends include Estonians	No	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	No	No	Yes	No	No	No

Table 2. The results of the perceptual assessment in ascending order by the mean of the perceived degree of global foreign accent (L2 speakers marked as Sp1...Sp20, raters marked as R1...R20).

	Perceptual ratings given by 20 raters (R1 - R20)																			
	Sp14	Sp12	Sp4	Sp6	Sp5	Sp13	Sp1	Sp7	Sp10	Sp16	Sp2	Sp9	Sp8	Sp19	Sp20	Sp15	Sp3	Sp11	Sp17	Sp18
R1	2	1	1	1	2	3	3	2	3	4	4	3	3	5	6	5	5	5	5	6
R2	1	2	1	1	2	3	2	3	4	4	4	3	4	4	3	5	4	5	5	6
R3	1	1	1	3	2	2	2	3	3	4	4	3	4	4	5	5	5	5	5	6
R4	1	2	1	1	1	2	2	2	2	3	3	[2]	4	3	4	3	4	[3]	4	4
R5	1	1	1	1	2	2	2	3	3	4	5	5	5	4	5	5	5	5	6	6
R6	1	1	2	3	2	3	4	4	4	5	4	4	[6]	5	6	6	5	6	6	6
R7	1	1	1	1	1	2	2	2	2	2	5	3	5	3	4	4	6	6	5	6
R8	1	1	2	1	2	2	3	2	[6]	5	5	5	3	4	3	6	6	6	6	6
R9	2	2	1	1	2	2	3	3	2	4	3	5	5	5	5	6	5	6	6	6
R10	1	1	1	2	2	3	3	4	5	5	3	5	5	5	5	4	5	6	6	6
R11	1	1	2	1	3	2	2	3	3	4	3	4	4	5	5	4	4	5	5	6
R12	1	2	2	3	3	3	2	4	3	4	4	3	5	3	4	5	5	6	6	6
R13	1	1	2	1	1	2	4	2	3	3	3	4	4	4	6	4	5	5	5	6
R14	1	1	2	1	2	2	3	3	4	3	4	4	5	5	5	5	5	6	6	6
R15	1	1	2	1	2	3	4	4	3	2	4	3	5	4	4	5	[3]	5	5	5
R16	2	2	3	3	3	3	4	4	4	3	4	4	4	4	4	4	4	5	5	6
R17	1	1	1	2	2	2	3	3	5	4	4	5	5	5	5	5	5	6	6	6
R18	1	1	1	2	2	2	2	2	3	2	3	3	3	3	3	3	4	4	[3]	4
R19	1	1	1	1	1	2	2	2	4	2	3	4	3	5	5	4	5	6	6	6
R20	1	1	1	1	2	2	2	3	5	5	3	4	3	4	3	6	5	6	6	6
Mean	1,2	1,3	1,5	1,6	2,0	2,4	2,7	2,9	3,4	3,6	3,8	3,9	4,2	4,2	4,5	4,7	4,8	5,5	5,5	5,8
Stdev	0,37	0,44	0,60	0,83	0,60	0,49	0,80	0,79	0,96	1,05	0,72	0,81	0,83	0,77	1,00	0,92	0,60	0,61	0,61	0,64
CI 95%	0,16	0,19	0,27	0,36	0,27	0,21	0,35	0,35	0,42	0,46	0,31	0,35	0,37	0,34	0,44	0,40	0,26	0,27	0,27	0,28

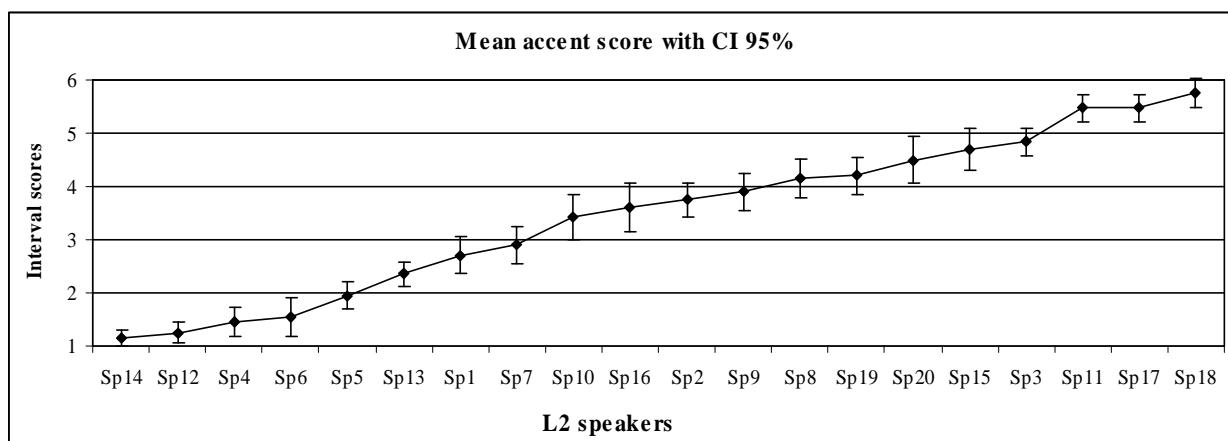


Figure 1. The mean scores of global foreign accent with a confidence interval of 95 %. L2 speakers (Sp1 ... Sp20) ordered by the mean accent score in ascending order.

3 Results

The rating results of the first group of judges (Table 2 and Figure 1) show high inter-rater consistency. Correlation for all possible pairwise combinations of two raters was computed while a few outliers were excluded from the statistics (see Table 2 numbers in square brackets). The average correlation is $r = 0.85$ (min $r = 0.7$, max $r = 0.96$); a correlation of 0.75 is considered acceptable (Shrout and Fleiss, 1979). Good inter-rater correlation shows that the duration of stimuli of 20 sec-

onds is sufficient for reliable results (cf. (Meister, 2006), where five- and 60-seconds clips were used). Also, the narrower six-point interval scale (compared to the nine-point scale used in (Meister, 2006)) may result in less dispersed ratings.

Variability of judgments among different listeners and the occurrence of few deviating ratings suggest that listeners' internal standards of accent-ness are different. Also, it can not be excluded that grammatical and lexical errors made by L2 speakers influenced the individual accent scores.

Table 3. Classification of pronunciation errors of L2 speakers in ascending order by error rate.

Speaker ID	Number of words	Type and amount of errors					Total number of errors	Error rate
		Temporal structure	Word stress	Vowel quality	Palatalization	Voicing of consonants		
Sp14	34						0	0
Sp12	40						0	0
Sp4	40						0	0
Sp6	45	2				1	3	0,07
Sp5	33	3					3	0,09
Sp1	40	2		1		1	4	0,10
Sp13	36	3				3	6	0,17
Sp10	23	2	1	2	2		7	0,30
Sp7	41	4		2	1	6	13	0,32
Sp2	34	4	1	1	2	3	11	0,32
Sp8	56	8		7	4		19	0,34
Sp9	23	5		1	2	3	11	0,48
Sp16	33	10	2	1		4	17	0,52
Sp3	23	5		4	5	1	15	0,65
Sp15	27	9		1		8	18	0,67
Sp11	31	9		3	3	6	21	0,68
Sp20	29	9		6		5	20	0,69
Sp17	27	10		2	3	4	19	0,70
Sp19	23	8	1	2	4	3	18	0,78

The findings of two experts (Table 3) show that the most frequent errors are related to temporal structure, voicing of voiceless consonants, and quality of some vowels and diphthongs; other errors are less frequent. These results confirm earlier findings (Meister and Meister, 2005).

In order to compare different L2-speakers, a simple measure of error rate has been formed by dividing the total number of errors by the number of words produced by the speaker during a 20 second clip (see Table 3).

4 Summary

The results of the two groups of listeners are highly correlated – the correlation between the mean accent score (Table 2) and the error rate (Table 3) is 0.94. It has been shown that for L1 speakers of a non-quantity language it is difficult to acquire a contrastive temporal category of L2 as a quantity language (McAllister et al., 2000). The same seems to hold true for the case of Russian as L1 and Estonian as L2 – the errors in the temporal domain contribute most to the error rate and probably to the perceived degree of FA, as well.

Further work will focus on the analysis of relationships between the degree of global FA and the types of pronunciation errors, as well as the role of deviations of acoustic features in the perception of accentedness.

References

- Jesney, K. 2004. *The Use of Global Foreign Accent Rating in Studies of L2 Acquisition*. Calgary, AB: University of Calgary Language Research Centre Reports.
- McAllister, R., Flege, J. and Piske, T. 2000. Aspects of the Acquisition of Swedish Quantity by Native Speakers of English, Spanish and Estonian. In: *Proceedings of FONETIK 2000*. Skövde, Sweden.
- Meister, L. and Meister, E.. 2005. Acoustic correlates of Russian accent in Estonian. In: *Proceedings of SPECOM 2005*, University of Patras, 437 - 440.
- Meister, L. 2006. Assessment of the degree of foreign accent: a pilot study. In: *Fonetiikan Päivät 2006 = The Phonetics Symposium 2006*, University of Helsinki, 53:113 - 119.
- Pajupuu, H., Reins, P. and Kerge, K. 2002. *Eesti keele kõrgtaseme test. Käsiraamat*. Tallinn: Riiklik Eksami- ja Kvalifikatsioonikeskus.
- Shrout, P.E. and Fleiss, J.L. 1979. Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, 86, 420-428.
- Southwood, H. and Flege, J. 1999. Scaling foreign accent: direct magnitude estimation versus interval scaling. *Clinical Linguistics & Phonetics*, Vol. 13, No. 5, 335-349.