PERCEPTUAL ASSESSMENT OF RUSSIAN-ACCENTED ESTONIAN

Lya Meister and Einar Meister

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

ABSTRACT

In this paper experiments on perceptual assessment of Russian-accented Estonian are introduced. Speech samples were recorded from 20 speakers with a Russian background; clips of about 20 seconds from each speaker were selected for perceptual assessment. Two tests were carried out: 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 analyzed the same samples and compiled the list of pronunciation errors. The higher the degree of accentedness judged by naïve listeners the more pronunciation errors were found by the experts. This general tendency is violated by some opposite examples. Finally, a simple method for accent assessment is proposed.

Keywords: foreign accent, perceptual analysis, accent degree, interval scale, pronunciation errors.

1. INTRODUCTION

As shown in numerous studies native speakers are able to identify non-native speech as well as to give adequate ratings of the degree of foreign accent (FA) [1]. Accentedness ratings result in the degree of global foreign accent which is a measure of how much the speech of a second language (L2) speaker deviates from that of L1 speakers [1], [5]. Naïve listeners' judgment of FA degree is based on general perceptual impression rather than on conscious use of acoustic-phonetic knowledge about their own first language (L1). Contrary to naïve listeners, trained phoneticians should be able to identify and classify different accent phenomena in L2 speech, and describe the pronunciation errors in terms of deviations of acoustic-phonetic features.

Following the findings and methodology presented in a recent paper [2] on the accentedness rating of foreign-accented Estonian two further listening tests 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 two experienced phoneticians. It is expected that the results of these two groups of raters harmonize well, i.e. the higher the accentedness of a L2 speakers is rated by naïve listeners 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. One subtask the examinees have to perform is conversation in pairs on a given topic which should demonstrate different speaking skills: expression of opinion, argumentation, turn-taking and carrying on the conversation, etc. [4]. 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 ca 1 m distance from the speakers. With each pair of subjects 6 to 8 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 in audio files in random order with an inter-stimuli interval of 5 seconds.

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

Table 1: Summary of the background information of L2 speakers (Est = Estonia(n), RusS = Russia(n), Ukr = Ukraine (Ukrainian)).

	L2-speakers' data																			
Speaker's ID	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 among Estonians	No	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	No	No	Yes	No	No	No

2.2. Listeners

Two groups of listeners were employed in the study. The group of naïve listeners involved 20 native Estonians (10 female, 10 male) aged from 17 to 62. All of them have some knowledge of Russian and have had diverse exposure to foreign-accented Estonian spoken by Russians; none of them reported any hearing problems.

The second group of judges consisted of two trained phoneticians (native Estonians, 1 female, 1 male, both 49 years of age) with good knowledge of Russian and experience of experimental study of Estonian as L2.

2.3. Experimental setup

Before starting the test the FA scaling technique was introduced and several examples of L2 speech with different degree of accentedness were played to the listeners. It was also stressed that the listeners should focus only on the deviations in pronunciation and grammatical as well as lexical errors should be omitted.

In the first experiment the stimuli were played to subjects on a notebook via headphones in a quiet environment. The task of the judges was to rate the degree of FA 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 their ratings on the individual answer sheet. The duration of the listening session was about 9 minutes.

In the second test two experts carried out an exhaustive perceptual analysis of each stimulus and compiled the list of perceived pronunciation errors classified into 5 major groups typical for Russian-accented Estonian: temporal deviations, location of word stress, quality of vowels, palatalization, voicing of voiceless consonants [3].

First the experts carried out error analysis independently from each other; this resulted in two slightly different lists of pronunciation errors. Later the disagreements were discussed and analyzed together until the experts reached a common agreement. In many cases correct classification of the error type was rather difficult as often simultaneous deviations of several acoustic-phonetic features co-occurred. However, only one (major) error per word was counted.

3. RESULTS

3.1. Test 1

The summary of ratings of the first group of judges is shown in Table 2; the mean scores of the ratings are presented also in Figure 1 (upper curve), where L2 speakers are sorted by the mean of the perceived degree of FA in ascending order; error bars show the confidence interval of 95%.

The accent ratings of different listeners coincide well (mean and median inter-subject correlation is 0.85; minimum 0.7; maximum 0.96); few extreme ratings were excluded from the statistics as outliers. From six outliers two were given by one listener, the others were from different listeners; the outlying scores were distributed randomly among L2 speakers. The variations in accent ratings among listeners are natural and can be explained by different internal standards of accentedness of listeners. For all judges the duration of stimuli (20 seconds) and inter-stimuli interval (5 seconds) was sufficient for giving their perceptual ratings.

Table 2: Ratings of the global foreign accent given by naïve Estonian listeners. L2 speakers (Sp1…Sp20) in ascending order by the mean rating score.

	Summary of perceptual ratings given by 20 raters																			
	Sp14	Sp12	Sp4	Sp6	Sp5	Sp13	Sp1	Sp7	Sp10	Sp16	Sp2	Sp9	Sp8	Sp19	Sp20	Sp15	Sp3	Sp11	Sp17	Sp18
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

3.2. Test 2

The findings of two phoneticians are presented in Table 3. The analysis of pronunciation errors shows that the most frequent errors are related to deviations in temporal structure; the voicing of voiceless consonants (mainly short plosives) is frequent, too. Indistinct vowel quality of some vowels and diphthongs (which are missing in the Russian sound system) as well as errors of palatalization are typical for L2 speakers with Russian background; incorrect location of word stress was observed just in a few cases. The results of error analysis confirm the earlier findings on Russian-accented Estonian [3].

Based on the statistics of detected errors a simple measure of pronunciation accuracy (phonetic error rate) was calculated by dividing the total number of errors by the number of words produced by the speaker during the 20 second clip. The error rate values are presented in Table 3 and also in the Figure 1 (lower curve).

Due to varying lexical (and thus phonemic) content of the rated speech samples the possible occurrence of different error types is rather uneven, therefore the error rate value is just a rough estimate of the pronunciation accuracy (only verbatim speech would provide a correct basis for speakers' comparison).

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
Sp18	34	11		5	8	6	30	0,88

4. DISCUSSION

The results of the two listeners' groups are highly correlated – the correlation between the mean accent score and the error rate is 0.942 ($p < 0.001$). Nevertheless, the mean scores of the ratings presented in Figure 1 (upper curve) and the error rates (lower curve) show slightly different behavior. The general tendency – the higher the accentedness rated by naïve listeners the more pronunciation errors found by the experts – is obvious, but it is not a one-to-one correspondence.

There are several exceptions where higher accent score is accompanied by lower error rate (Figure 1: cf. Sp16-Sp2, Sp9-Sp8, etc) and substantial increase of phonetic error rate is associated with minor increase on accent score (Figure 1: cf. Sp8-Sp19). However, matching the subjective accent score to the more objective error rate seems plausible. This has been performed by grouping L2 speakers into 5 accentedness groups according to the mean accent score and calculating the average error measure within each group. The error measure can be interpreted as a rough estimate of accentedness. The accentedness groups and the corresponding error measures are presented in

Table 4. The proposed classification should be tested in further experiments and could be implemented in accent assessment during language tests or adapted for computer-based pronunciation training systems.

Table 4: Accentedness groups and corresponding error measures.

Accentedness group	Group description	Error measure: words per error
1	Very weak accent	10-20
2	Weak accent	4-10
3	Moderate accent	2-4
4	Strong accent	1,5-2
5	Very strong accent	1-1,5

Figure 1: The mean scores of global foreign accent with confidence interval of 95% (upper curve, scale on the left) and error rate (lower curve, scale on the right). L2 speakers are sorted by the mean of the perceived degree of global foreign accent in ascending order.
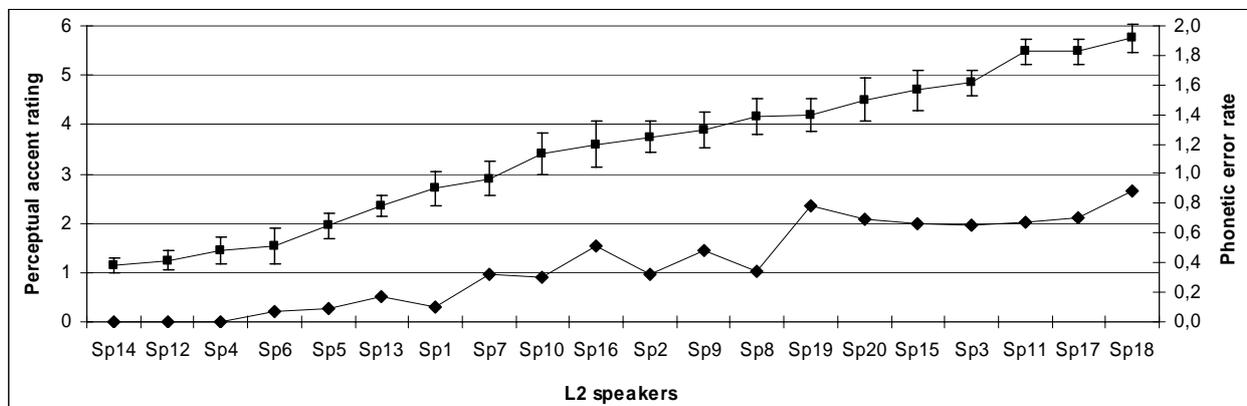
5. SUMMARY

The results of perceptual assessment of Russian-accented Estonian speech by naïve L1 listeners and by experienced phoneticians are presented and compared. The results of the two tests show good correlation – in general, higher degree of accentedness is associated with higher amount of pronunciation errors, but some exceptions exists, too. A simple method for accent assessment is proposed which could be applied during language tests or adapted to computer-based systems for pronunciation training.

The further studies will involve deeper analysis of different pronunciation errors and their role in the perception of global accent (e.g. data in Table 3 shows that errors of temporal structure and vowel quality contribute more to the error rate than the other types of error). Also the role of deviations of different acoustic features in the perception of accentedness deserves further analysis.

6. ACKNOWLEDGEMENTS

7. REFERENCES

[1] Jesney, K. 2004. The Use of Global Foreign Accent Rating in Studies of L2 Acquisition. Calgary, AB: University of Calgary Language Research Centre Reports.

[2] Meister, L. 2006. Assessment of the degree of foreign accent: a pilot study. *The Phonetics Symposium 2006*, Helsinki, 113 - 119.

[3] Meister, L., Meister, E. 2005. Acoustic correlates of Russian accent in Estonian. *Proc. SPECOM 2005,* Patras, 437 - 440.

[4] Pajupuu, H., Reins, P., Kerge, K. 2002. Eesti keele kõrgtaseme test. Käsiraamat. Tallinn: Riiklik Eksami- ja Kvalifikatsioonikeskus.

[5] Southwood, H., Flege, J. 1999. Scaling foreign accent: direct magnitude estimation versus interval scaling. Clinical Linguistics & Phonetics, Vol. 13, No. 5, 335-349.