

A Note on Native and Non-Native Accentedness Judgments

Zinny S. Bond and Verna Stockmal
Ohio University

Dace Markus
University of Latvia, Latvia

Talkers in a second language can readily be identified as speaking with a foreign accent, characterized by both specific and more general deviations from the phonology of a target language. We examined the identifications of native and non-native talkers by listeners with various amounts of knowledge of the target language. Native and non-native speakers of Latvian provided materials. All the non-native talkers spoke Russian as their first language and were long-term residents of Latvia. A listening test, consisting of identical sentences excerpted from a short recorded passage, was presented to three groups of listeners: native speakers of Latvian, Russians for whom Latvian was a second language, and Americans with no knowledge of either language. The listeners were asked to judge whether each utterance was produced by a native or non-native talker. The Latvians identified the non-native talkers very accurately, 88%. The Russians were somewhat less accurate, 83%. The American listeners were least accurate, but still identified the non-native talkers at above chance levels, 63%. Sentence durations correlated with the judgments provided by the American listeners but not with the judgments provided by the native or L2 listeners.

What speech characteristics lead listeners to judge that they are hearing a non-native speaker? This question has led to an extensive literature examining listener abilities to identify non-native speech and the information on which they base their judgments. Two related conclusions emerging from these studies are that native listeners are extremely accurate at distinguishing native from non-native speech and that they are able to take advantage of multiple sources of information in making judgments.

Listeners are extremely accurate in discriminating between native and non-native speech even from minimal information. For example, Flege (1984) reported that listeners were able to recognize accented spoken samples of their native language within 30 ms, basing their judgments on phonetic cues present in fragments of syllables. Similarly, Major (1987) found that native speakers could distinguish Brazilian Portuguese from English using the pronunciation of two contrasting vowels, /ɛ/ and /æ/. In perhaps the most extreme example of accurate judgments from minimal information, Munro, Derwing and Burgess (2003) found that listeners could

discriminate native from non-native productions of a short phrase played backwards, presented to listeners without segmental or appropriate prosodic information.

When listeners have more than minimal speech available to them, their accentedness judgments employ multiple sources of information. For example, Cunningham-Andersson and Engstrand (1989) investigated the role of segmental features in the identification of Finnish accent in spoken Swedish. They used speech in which a talker deliberately introduced one or more deviations from normal Swedish which are characteristic of Finnish accent. Although some features were more effective in suggesting an accent than others, in general, the more deviations, the more likely listeners identified the speech as accented. Magen (1989) presented computer-edited and natural versions of sentences produced in Spanish-accented English; she found that listeners were sensitive to a variety of segmental and suprasegmental factors such as syllable structure, vowel and consonant quality, and stress placement. Munro (1995) presented filtered versions of sentences produced by native and non-native talkers. He found that listeners could identify foreign-accented speech even without clear information about segmentals. Presumably, the listeners were relying on intonation patterns, timing, and speaking rate for accentedness judgments. Munro and Derwing (2001) examined the possible contribution of speaking rate, both naturally produced and in expanded or contracted sentences. They found that listener judgments of accentedness exhibited a small, significant relationship to speaking rate, although the relationship was not linear.

It is worth adding that acoustic-phonetic investigations which have compared non-native speech to native speech have found differences in all the sources of information which listeners employ in making accentedness judgments: segments (Flege & Bohn, 1989; Flege, Munro, & MacKay, 1990), suprasegmental properties such as speech rate and rhythm (Ling, Grabe, & Nolan, 2000), as well as other indices of lack of proficiency (Guion, Flege, Liu, & Yeni-Komshian, 2000).

In attempting to assess the relative contribution of various sources of information, studies have varied the listening materials but kept the knowledge of the listeners constant, in the sense that the listeners have been native speakers of the target language.

The purpose of this investigation was to vary listener knowledge of the target language in order to obtain a preliminary assessment of language-specific vs. universal indicators of accentedness. We used Latvian as the target second language and speech samples produced either by native speakers of Latvian or by ethnic Russians for whom Latvian is a second language. Three groups of listeners participated: native speakers of the target language, ethnic Russians, and monolingual Americans. The native speakers of Latvian had native-speaker knowledge of the language; the Russians had some knowledge of Latvian, both as spoken natively and by Russians who speak it as a second language; the Americans had no previous exposure to either Latvian or Russian, and little exposure to non-native English or other languages.

Method

Talkers

Ten ethnic Russian talkers and ten Latvians were recorded reading a short passage. Both groups of talkers included males and females and varied in age, from high school students to retirees. The Russian talkers represented different levels of proficiency in Latvian, according to self-report. Details about the talkers are given in Table 1. For ease of reference, the talkers were assigned an identifying code based on age. The code for Russian talkers indicates their own assessment of proficiency in Latvian.

Table 1

Talkers Providing Language Samples

<i>Talker</i>	<i>Year of Birth</i>	<i>Sex</i>	<i>Latvian proficiency_(a)</i>	<i>Russian proficiency_(a)</i>	<i>% Latvian_(b)</i>
L0A	1985	F	4	2	70
L0B	1985	F	7	1	100
L1A	1980	F	7	6	90
L1B	1980	F	7	7	50
L2A	1975	F	7	5	95
L2B	1969	F	7	7	60
L3A	1962	F	7	4	93
L3B	1963	F	7	7	50
L4A	1923	F	7	4	100
L4B	1942	F	7	-	70
R0L	1985	F	4	7	10
R0H	1985	F	5	7	60
R1H	1981	F	7	7	80
R1L	1981	F	1	7	38
R2H	1974	F	7	7	50
R2L	1970	F	3	7	30
R3H	1953	M	7	7	70
R3L	1955	M	2	7	10
R4L	1941	M	2	6	10
R4H	1942	M	6	7	60

- a. Latvian and Russian proficiency represents self-report on a 7-point scale. Some talkers failed to respond.
- b. Percent of time using Latvian represents self-report.

Materials

The same two sentences for each talker were excerpted from the recorded passage. Because the text was identical, talkers could not be distinguished by lexical selection or morphological and syntactic errors. The forty sentences were arranged in random order on a test recording. Before the test sentences, three practice items were included.

Procedure

The participants were asked to provide two different kinds of judgments—identification of speech samples as produced by native vs. non-native speakers and evaluation of speech samples. Thirty-one monolingual Americans, students at Ohio University, listened to the recordings and judged each speech sample as produced by a native or non-native talker. Nineteen different monolingual Americans drawn from the same population evaluated each speech sample on a 7-point scale. The end-points of the scale were defined as ‘definitely a native speaker’ and ‘definitely not a native speaker’.

Twelve ethnic Russians and 28 Latvians, high school students in Riga, also evaluated the talkers. These listeners were either University or technicum students, long-term residents of Riga. For each speech sample, the listeners tested in Latvia indicated whether they thought the talker was a Latvian or a Russian and evaluated the speech samples on a 7-point scale with end points defined in the same way as for the Americans.

Ideally, the procedure for testing the Americans and the listeners with knowledge of Latvian would have been identical. Because of time constraints and difficulties in listener recruitment in Riga, the Russians and Latvians were asked to provide both kinds of judgments at the same time.

Results

In the first task, identifying native vs. non-native talkers, both Latvian and Russian listeners who knew the target language identified talkers as native or non-native at better than 90% accuracy, consistent with previous investigations (e.g., Scovel, 1995). The American listeners also identified native speakers accurately—they identified as non-native those Russians who indicated that their proficiency in Latvian was low, but failed to distinguish between native Latvians and relatively high-proficiency Russian talkers. Overall, the Americans identified native vs. non-native background correctly at 63%, significantly different from chance expectation, $t(30) = 9.28, p < .001$. Listener judgments are displayed in Figure 1.

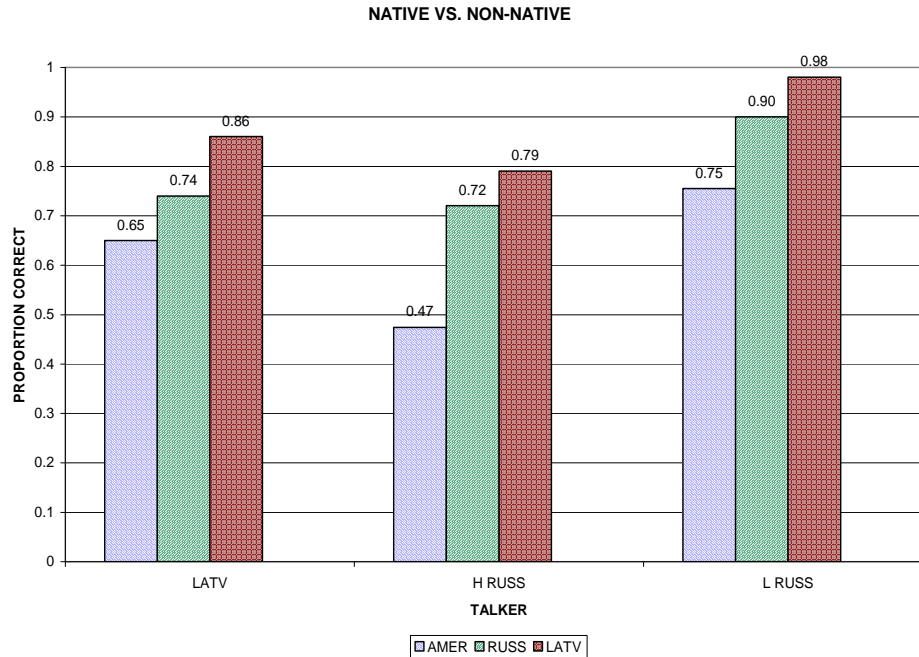


Figure 1. American, Russian and Latvian listener identifications of three groups of talkers— native Latvians, Russians with relatively high proficiency in Latvian, and Russians with relatively low proficiency in Latvian.

In the second task, evaluating the speech samples, the overall performance of the three groups was different, $F(2, 53) = 6.8, p < .005$. The average evaluation provided by the American listeners was 3.8, significantly lower than the evaluations offered by the other two groups. The Russian overall evaluation of 4.2 and the Latvian evaluation of 4.5 were not significantly different. Apparently, the Americans were reluctant to judge a large proportion of the speech samples as truly produced by native speakers.

In terms of response patterns, the Latvian listeners gave native Latvians the highest evaluation (5.8) and distinguished between low-proficiency and high-proficiency Russian talkers. Although the American responses included fewer confident native speaker judgments, they gave the highest evaluations to native Latvians (4.5) and distinguished between low and high-proficiency Russians, just as the Latvian listeners did.

The Russian listeners gave high evaluations to Russians and relatively low evaluations to Latvians. Because they could readily identify native from non-native talkers, the Russian judgments may reflect ethnic solidarity rather than linguistic judgments. Alternatively, the Russians may be basing their judgments on imperfect knowledge of Latvian phonology because of limited contact with native speakers. The evaluations of the talkers are given in Figure 2.

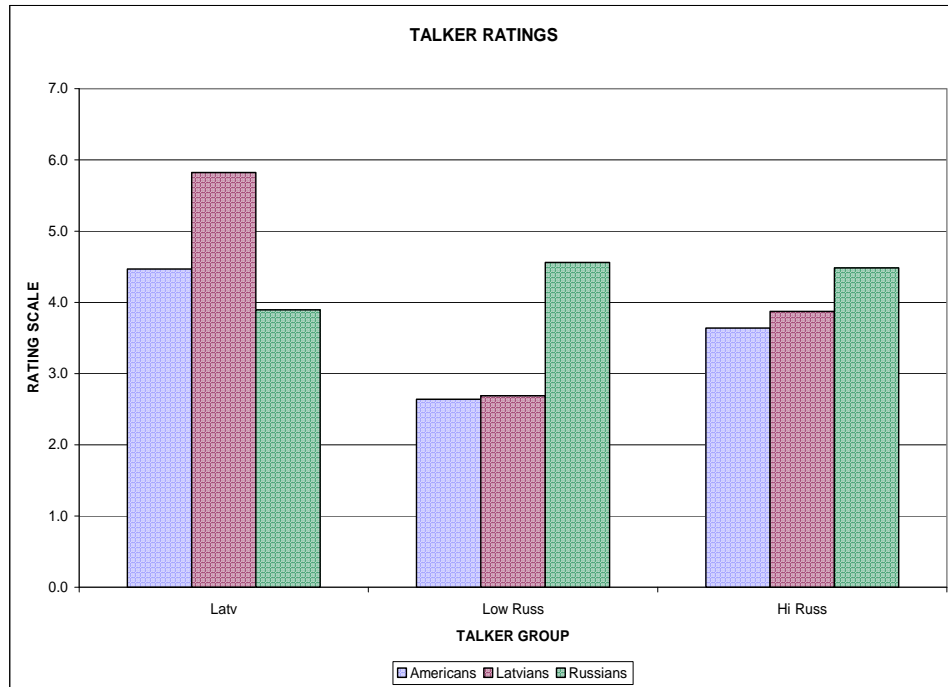


Figure 2. American, Russian and Latvian listener identifications.

Discussion

All three groups of listeners were able to distinguish between native speakers and non-native speakers with limited proficiency in their second language. Listeners with knowledge of the target language could also distinguish relatively proficient L2 talkers from native speakers. In making their judgments, both the native listeners and the L2 listeners were undoubtedly responding to multiple sources of information as listeners who know a language would be expected to do. The American listeners had no knowledge of the phonology of the target language, so they were making judgments only on the basis of universal, non-linguistic speech characteristics.

The most likely characteristic which the Americans employed is speech rate or fluency. Although lack of fluency may be signaled by any number of properties such as hesitations, repetitions, pauses, excessive prolongations, and so forth, these properties can be subsumed under one physical characteristic, utterance duration. That is, limited fluency increases the durations of utterances. For the American listeners, the correlation between utterance duration and accuracy in identifying talkers as native or non-native was .65. That is, the longer an utterance, the more likely American listeners judged correctly that it was produced by a non-native talker. This interpretation of listener use of duration is consistent with the suggestions of Munro et al. (2003). In contrast, there were no significant correlations between utterance duration and correct identification for Russian and Latvian listeners. Because these listeners knew the target language, they had many sources of information available to them rather than a relatively obvious characteristic such as fluency or speech rate.

Although phonological accuracy and fluency appear to be related measures in non-native speech (Burgess [2001] reported that the correlation between accuracy and fluency judgments was .40), they are separable properties of speech, in that listeners are able to make reasonably consistent and accurate accentedness judgments without knowledge of the phonology of a language. Apparently, naïve listeners have expectations about normal fluent speech which they can use as perceptual anchors in judging utterances even when listening to a language which they do not know.

Of course as an estimate of fluency, duration can reflect a number of properties of non-native speech, from excessive prolongations of segments to hesitations and pauses. Whether listeners unfamiliar with a language are able to make distinctions reflecting different sources is a matter for further investigation.

Acknowledgements

Research for this project was supported by a grant from the International Research & Exchanges Board, with funds provided by the U.S. Department of State (Title VII program) and the National Endowment for the Humanities. None of these organizations is responsible for the views expressed here. A preliminary version of this paper was presented at a meeting of the Acoustical Society of America, Nashville, TN, in May, 2003.

References

- Burgess, C. (2003). *Speaking rate, fluency, and accentedness in the speech of second language learners*. Unpublished doctoral dissertation, Simon Fraser University, Vancouver, Canada.
- Cunningham-Andersson, U., & Engstrand, O. (1989). Perceived strength and identify of foreign accent in Swedish. *Phonetica*, *46*, 138–154.
- Flege, J. E. (1984). The detection of French accent by American listeners. *Jl. of the Acoustical Society of America*, *76*, 692–707.
- Flege, J. E., & Bohn, O. (1989). An instrumental study of vowel reduction and stress placement in Spanish-accented English. *Studies in Second Language Acquisition*, *11*, 35–62.
- Flege, J. E., Munro, M. J., & MacKay, I. R. A. (1995). Effects of age of second-language learning on the production of English consonants. *Speech Communication*, *16*, 1–26.
- Guion, S. G., Flege, J. E., Liu, S. H., & Yeni-Komshian, G. H. (2001). Age of learning effects on the duration of sentences produced in a second language. *Applied Psycholinguistics*, *21*, 205–228.
- Low, E. L., Grabe, E., & Nolan, F. (2000). Quantitative characterizations of speech rhythm: ‘syllable timing’ in Singapore English. *Language and Speech*, *43*, 377–401.
- Magen, H. S. (1998). The perception of foreign-accented speech. *Journal of Phonetics*, *26*, 381–400.
- Major, R. (1987). Phonological similarity, markedness, and rate of second language acquisition. *Studies in Second Language Acquisition*, *9*, 63–82.

- Munro, M. J. (1995). Nonsegmental factors in foreign accent: Ratings of filtered speech. *Studies in Second Language Acquisition*, 17, 17–34.
- Munro, M. J., & Derwing, T. M. (2001). Modeling perceptions of the accentedness and comprehensibility of L2 speech: The role of speaking rate. *Studies in Second Language Acquisition*, 23, 451–468.
- Munro, M. J., Derwing, T.M., & Burgess, C. S. (2003, August). *The detection of foreign accent in backwards speech*. 15th International Congress of Phonetic Sciences, Barcelona, Spain.
- Scovel, T. (1995). Differentiation, recognition, and identification in the discrimination of foreign accents. In J. Archibald (Ed.), *Phonological acquisition and phonological theory* (pp. 169–181). Hillsdale, NJ: Lawrence Erlbaum.