Kuronen, M., P. Lintunen & T. Nieminen (toim.) 2017. Näkökulmia toisen kielen puheeseen – Insights into second language speech. AFinLA-e. Soveltavan kielitieteen tutkimuksia 2017 / n:o 10. 258–276.

Minnaleena Toivola[1] & Riikka Ullakonoja[2]

[1]University of Helsinki, [2]University of Jyväskylä

Identification of Russian accented Finnish by native and non-native listeners with and without Finnish proficiency

The study focuses on how different groups of listeners, that is native Finnish speakers (n = 18), non-native learners of Finnish (n = 12) and non-native non-learners of Finnish (n = 18), judge samples (n = 80) of Russian accented Finnish in an auditory experiment. The samples are read-aloud phrases of three Finnish speakers and three L1 Russian beginner learners of Finnish who were recorded three times at six-month intervals during their residence in Finland. The results show that the listener groups differ significantly from each other in the ability to identify the Russian accented samples. In addition, all the listener groups reported to have paid more attention to prosodic and segmental cues than speech rate. Finally, we conclude by considering implications of the study to further longitudinal studies on perception of foreign accented speech.

Keywords: Foreign accent, Finnish as a second language, non-native raters, pronunciation

1 Introduction

The small scale study reported in this paper aims to investigate how different listener groups, including naive listeners, successfully identify Russian speakers in an auditory experiment consisting of short read-aloud samples from native Finnish speakers and L1 Russian beginner learners. We explore the potential of such a non-standard experimental design in the study of foreign accent. In addition, we are interested in what phonetic aspects the listeners based their judgments on. The listener groups are: native Finnish speakers (n = 18), non-native learners of Finnish (n = 12) and non-native non-learners of Finnish (n = 18), to be further referred to as G1, G2 and G3, respectively. The study aims to contribute to the almost non-existent literature on judgments of foreign accented Finnish and to offer suggestions for further studies on this topic.

We will now briefly describe the prosody of the languages under investigation. Russian prosody has a greater role in communication than Finnish prosody, because Russian uses intonation extensively for distinguishing questions from statements whereas Finnish does not (Bondarko 1998; Iivonen 1978). Finnish intonation has often been described as rather monotonous and produced with a narrow pitch range with creaky voice occurring frequently (Iivonen 1998, 2009a). Russian intonation, on the other hand, is more variable and lively and creaky voice typically does not occur (Volskaya 2009). In Finnish, word stress is fixed on the first syllable (Iivonen 2009b), whereas in Russian it can be placed on any syllable and even change position in different forms of the same syllable (Bondarko 2009). Finnish stressed vowels do not differ as greatly from unstressed ones in quality and quantity compared to Russian ones, which differ from their unstressed counterparts a great deal (Bondarko 2009; Iivonen 1998). Previous research on Russian accented Finnish shows, for example, that non-native like prosodic chunking, stress (exaggerated stress on the word-initial syllable) and pitch variation (rapid increase in pitch in word-final syllables or utterance final positions) are typical for Russian learners of Finnish (Aho & Toivola 2008).

In this paper, we first summarize previous studies on foreign accent, focusing on the perception of foreign accent and the role of listener's background in perception. A material and methods section, describing the auditory experiment in detail, follows this introduction. We describe the listener groups as well as the three L1 Russian speakers, beginner learners of Finnish in more detail as well as explain the statistical methods used. Next, we present our main findings and finish with a discussion.

2 Perception of foreign accent

Foreign accent is the term often used for non-native pronunciation that deviates from native speech (Scovel 1969; Flege 1981; Munro 2008). As mentioned by Munro & Derwing (2015), studying foreign accent can be a key to understanding speech processing, and the results from such studies can often be applied to pronunciation teaching.

The goal for pronunciation learning is often comprehensibility (Jenkins 2000; Walker 2011) rather than complete native-likeness or speaking without any foreign accent, at least for adult learners (e.g. Abrahamsson & Hyltenstam 2009). Nonetheless, foreign accent can in some cases hinder comprehensibility, whereas in other cases accented speech can be completely comprehensible (Munro & Derwing 1999). For immigrants, foreign accent has societal relevance in integrating into the host society, for example, because native speakers may value it negatively (Lippi-Green 2012; Leinonen 2015). As Russian speakers are the largest immigrant group in Finland (n = 75,444 in 2016, Statistics Finland 2017), negative accent perception is of particular concern to this group in Finnish society. Attitudes towards Russian speakers have been rather negative; for example, over 60% of Finns reported they do not wish Russian speakers to move to Finland (Jaakkola 2006, 2009). Aho & Toivola (2008) found in their study that many Russian learners of Finnish wanted to speak Finnish without a Russian accent. For immigrants, sounding native, or at least not easily being recognized as a non-native speaker, can thus be a justifiable learning goal.

The present study focuses on factors contributing to the perception and identification of foreign accent. More precisely, the study investigates how successfully native and non-native listeners identify the foreign accent in Finnish spoken by native speakers of Russian and what phonetic aspects their judgments are based on. Non-learners of the language have been used as listeners only in a few previous studies measuring foreign accent, fluency or pronunciation (e.g. Major 2007; Weber & Pöllmann 2011; Wilkerson 2010). In the majority of previous studies the listeners have been native speakers or non-native learners of the language to be judged (Gonet & Pietroń 2004; MacKay et al. 2006; Munro et al. 2006). The studies have shown that different groups of listeners can differ in their evaluations. For example, in Weber & Pöllmann's (2011) study the judgments of non-native non-learners differed from native and non-native language learners. However, non-native non-learners are also able to judge the stimuli very reliably (Major 2007), especially the stimuli with a strong foreign accent. The studies speculate that non-native non-learner listeners may rely on their first language knowledge in their judgments. As one explanation for perception of foreign accent Major (2007) proposes the term *displaced foreign accent detection*, which refers to the fact that the listener

is paying attention to the features of their native language that can be heard in the unknown language. Pilot studies by Gilbert (1980) have indicated the strong role of prosody in recognizing the speaker's L1. Furthermore, Gupta (2005) showed that different accents of English are easier to understand if one is familiar with the accent in question.

Previous research has also focused on rater's expertise as a background factor influencing ratings and the results have been controversial (see review by Piske et al. 2001). In a study by Cunningham-Andersson & Engstrand (1989) it did not play a significant role as naive listeners were able to identify foreign accent from rather short samples as well as expert listeners. Similarly, in Kennedy & Trofimovich (2008) English as L2 teachers' foreign accent ratings did not differ from naive listeners' ratings. However, Bongaerts, Mennen and van der Slik's (2000) study showed that language teachers and non-teachers differed in their ratings. Thus, Piske et al. (2001) recommend in their review that listeners from different backgrounds should be used in studies investigating the perception of foreign accent.

Previous studies have identified a number of factors contributing to the perception of accentedness. Research focusing on fluency and the strength of foreign accent (e.g. Major 2007; Weber & Pöllmann 2011) suggest that the utterance duration as well as slower speech rate would particularly indicate that the speaker is a non-native speaker. However, the link has not always been this clear. Derwing & Munro (2001) found that speech that was too slow or too fast did not sound native-like. Major (2007) proposes that duration differences alone do not account for the foreign accent ratings. Trofimovich & Baker (2006) found that speech rate and duration of pauses were associated with foreign accent more than stress timing and peak alignment.

Pinget et al. (2014) studied both suprasegmental and segmental features and measured pitch alternation hypothesizing that monotonous pitch contributes to the perception of foreign accent. The results showed, however, that pitch alternation and the sound segments chosen for the study explained only a small proportion of the strength of foreign accent. Additionally McCullough (2013) studied multiple acoustic properties such as VOT, vowel quality, f0 (fundamental frequency), vowel duration and sentence stress in English words produced by American English, Hindi, Korean and Mandarin speakers. In the study VOT and vowel quality were linked to the perception of foreign accent more than f0 and vowel duration. Saito et al. (2017) studied 11 linguistic factors in the perception of foreign accent of Japanese learners of English. They found that native listeners' ratings of accentedness were associated with word stress and vowel/consonant errors.

In Riney, Takagi and Inutsuka's (2005) study the listeners disagreed on which of the non-native speakers were perceived the most and least native-like. They focused on Japanese and American raters listening to Japanese L2

speakers of English and showed that all listeners were able to identify the native speakers. However, non-native listeners paid more attention to intonation, fluency and speech rate as compared to native listeners, who based their judgments more on deviances in the segmental level (especially /r/ and /l/ sounds).

Previous literature on L2 Finnish is scarce both from perception and production points of view. In previous perception studies the raters have been native Finnish speakers (e.g. Toivola 2011; Leinonen 2015; Uzal et al. 2015). The perception study by Leinonen (2015) showed that Finnish adolescents living in different parts of Finland rated foreign accented Finnish of a number of different L1s consistently. Speakers were rated differently based on their L1. Swahili and Vietnamese L1 speakers were given the strongest foreign accent ratings, whereas Arabic L1 speakers were rated the weakest and English and Russian L1 speakers in between. There is one recent study (Ahola & Tossavainen 2016) on L2 Finnish pronunciation assessment focusing on Estonian accented Finnish. It found that the similarities between Estonian and Finnish pronunciation make Estonian accented Finnish rather easy to understand and to listen to. However, the study also found that raters perceived Finnish spoken by Estonian L1 speakers too fast and not containing enough pauses.

Features related to L2 Finnish pronunciation and fluency have been studied by Toivola and colleagues. Toivola et al. (2009) compared Finnish native speakers and Finnish L2 learners (from four different L1 backgrounds: Russian, Thai, Turkish and Vietnamese) and found that pauses were longer in read-aloud speech by Finnish native speakers than L2 learners. Toivola et al. (2011) also found that articulation rate becomes faster in a longitudinal setting, as the length of residence in Finland increases, both in read-aloud and conversational speech. The number of pauses in read-aloud speech also decreases, but the mean duration of pauses increases. They conclude that learning native-like pausing in spoken Finnish requires the ability to use context-dependent pause durations in speech. Toivola (2011) also studied the role of phonetic features in perception of foreign accent and found that the number of single deviant phonetic segments, filled pauses and articulation rate contributed to the perceived degree of accentedness in the spontaneous speech of experienced Finnish L2 learners (with Russian L1 background). In addition, in the read-aloud data the number of single deviant phonetic segments and their quality explained the majority of the strength of the foreign accent by Russian speakers (Toivola 2011).

Small scale studies on L2 Finnish spoken by Russians have also been conducted by Ullakonoja and colleagues. These studies have focused on how adult Russian native speakers with no prior knowledge of Finnish imitate short Finnish utterances (Ullakonoja et al. 2014a,b) and how young Russian immigrants produce Finnish segmental duration and length in a read-aloud task

(Ullakonoja & Kuronen 2015). The results showed, firstly, that subject's working memory is correlated with the comprehensibility of imitated utterances, and secondly, that it is challenging to imitate/to learn the Finnish durational contrasts. All studies revealed great interspeaker variation in the success of production.

3 Aims, data and methods

The main aim of the study is to investigate the identification of Russian accented Finnish by three different listener groups. In particular, the present study sought to answer the following questions:

1. How well were non-native, that is Russian L1, speakers identified by different listener groups (that is native Finnish speakers, native Russian speakers who are learning Finnish and native Russian speakers with no knowledge of Finnish)?

2. Are samples from consequent recording sessions rated differently, and if so, is the change due to perceived changes in pronunciation?

In contrast to most previous studies on L2 pronunciation, the present study focuses on pronunciation in a longitudinal research setting and includes a listener group with no knowledge of the language under investigation. We focus on three adult beginner learners of Finnish, with Russian as their L1, recorded three times within 6-month intervals. Extracts of the recordings were subjected to a perception experiment, through which we investigated the differences between rater groups in identification and if the perceptual judgments change over time and if so, whether the pronunciation also changes as reported by the raters. Seeing how non-native non-learners of Finnish judge the speech stimuli in different recording sessions is of particular interest, as such a listener group is very rarely included in L2 perception studies. The underlying hypotheses are that, firstly, the degree of proficiency in Finnish is linked with the ability to identify the Russian accented speech (the better the raters' proficiency in Finnish, the better their ability to identify Russian speakers' samples is) and, secondly, that all the different groups of raters judge the L2 speakers more often as native speakers of Finnish in the final recording session.

3.1 Raters

Three different groups of raters (n = 48), listened to the read-aloud stimuli by L1 and L2 speakers in the auditory experiment: L1 Finnish speakers (G1), Russian speakers with L2 Finnish (G2), L1 Russian speakers with no previous

knowledge or proficiency in Finnish (G3). None of the listeners reported hearing difficulties and they were all students. G1, the L1 Finnish speakers (n = 18, all female) were from the Helsinki or Jyväskylä regions aged 20–33 (mean age 23) with no knowledge or previous study of Russian. Their only L1 was Finnish, but they had studied at least English and Swedish at school (as is typical in Finland). All Finnish speakers can be considered at least somewhat familiar with Russian-accented Finnish, as Russian is one of largest immigrant minority languages in Finland. G2, the L1 Russian speakers (n = 12, all female) with L2 Finnish, aged 20–35 (mean age 28) were living in the Helsinki or Jyväskylä regions at the time of the experiment. They had a self-rated minimum of B1-level oral proficiency of Finnish on the CEFR scale (CEFR 2001), i.e. were at least at an intermediate level. They had studied Finnish 2–13 years (mean 6) and had lived in Finland from 1 to 13 years (mean 6). G3, the L1 Russian speakers (n = 18, 2 male, 16 female) with no Finnish proficiency came from the St. Petersburg region and were between 17 and 30 years of age (mean age 20).

3.2 Speakers

In the auditory experiment there were samples from two groups of speakers: non-native speakers of Finnish with Russian as their L1 and native speakers of Finnish. The non-native speakers were three female speakers (further Rus1, Rus2 and Rus3) from the ProoF-project corpus (ProoF 2012; see Aho et al. 2016, for a description of the corpus) aged 25–29 (mean age 27). They were recorded three times with 6-month intervals (further T1, T2 and T3) and they had lived in Finland for 8–12 months (mean 11 months) and had studied Finnish for 4–9 months (mean 7 months) before the first recording session. Two speakers continued their Finnish studies throughout the experiment and one (Rus3) interrupted her Finnish lessons after T2. Two speakers (Rus1 and Rus3) reported that they mostly spoke Finnish outside the home, and Rus3 also at home, whereas Rus1 spoke Russian at home. Rus2 said that she spoke some Finnish, but mostly English both at home and outside the home. The native speakers of Finnish were three female speakers from the Helsinki region aged 21–24 (mean 22), as well as contributors to the ProoF-project corpus.

3.3 Speech samples

The read-aloud speech samples were recorded in a sound-proof studio, using high quality audio equipment within the ProoF-project. Every speaker was recorded with a head-mounted microphone at a sample rate of 44.1 kHz and sample size of 16 bits. Speakers were asked to read aloud short sentences using an ordinary speech rate. As opposed to previous studies, which have used longer stimuli, we chose short extracts of read-aloud speech for the auditory

experiment. The reason for this is that, firstly, we hypothesized that producing Finnish intonation, which has a high beginning followed by declination, would be difficult for Russian speakers to produce. Secondly, longer samples would have contained more pauses and hesitation as well as a potentially slower speech rate in the L2 speech, which would have increased the likelihood of being identified as "Russian". The following seven phrases were extracted from the beginning of longer read-aloud sentences from each speaker (for L2 speakers from three consequent recording sessions): *viime yönä* (last night), *kuten tiedät* (as you know), *hän sanoi* (s/he said), *en tiedä* (I don't know), *eilen illalla* (last night), *mielestäni* (in my opinion), *viime kesänä* (last summer).

3.4 Listening test

The auditory experiment consisted of a total of 80 samples, 63 of which were obtained from three Russian native speakers from the three recording sessions over time, and the remaining 17 samples from three Finnish native speakers (5–6 samples per speaker). Thus, the total number of all ratings was 3840, of which 3024 were ratings of Russian speakers' samples.

In the experiment (23 min), the listeners first heard seven practice stimuli, containing both L2 and L1 samples. Then, the actual stimuli were presented in two different randomized orders so that all listeners listened to the stimuli only once and had 10 seconds to respond after each stimulus. The experiments were all administered in a quiet room and the stimuli were presented either through headphones or loudspeakers (where the use of headphones was impossible for practical reasons). The listeners chose the language (Finnish/Russian) of the rating sheet. First, the listeners were asked to decide whether the speaker's mother tongue was Russian. Second, the two listener groups (L1 Finnish speakers and L1 Russian speakers with L2 Finnish proficiency) who knew Finnish were asked to define the basis for their judgments for each stimulus in a multiple-choice question, (allowing for multiple responses per stimulus) where the alternatives were prosody (referring to intonation, stress and rhythm), speech rate and vowel and consonant quality/quantity. As we considered such a multiple-choice question too demanding for the participants for whom Finnish was an unknown language, the non-learners of Finnish group were asked to express the basis of their judgments across all the stimuli freely in their own words. This was a good decision, as we observed that for some listeners the fact that they had to participate in an experiment in a language they have never heard was enough to make them question their willingness and capability to participate.

3.5 Data analysis

The perceptual judgments were analyzed and processed quantitatively using MS Excel 14.0 and IBM SPSS Statistics 24.0 software. For investigating the differences between the listener groups, we used one-way between-groups ANOVA analysis of variance and reported the values of the Welch test as the assumption of the homogeneity of variance was violated. Further, we used the Tukey HSD post-hoc comparisons. In the further analysis of the data, we used non-parametric tests, as the parametric ones require normal distribution of the data, which was not the case here for the smaller sub-groups. When comparing listener groups we used the Kruskal-Wallis test, a non-parametric equivalent to one-way variance analysis allowing for comparison of three groups. In comparing the three time points, we used the Friedman Test, an alternative to the one-way ANOVA with repeated measures.

4 Results

4.1 Identification by the listener groups

First, we present the results considering the comparison of the three listener groups in the success of identification samples spoken by a native Russian speaker and by a native Finnish speaker. Figure 1 shows the differences in identification between the listener groups. Not surprisingly, the non-Russian, i.e. Finnish speakers' stimuli (n = 17) were almost perfectly recognized by (G1) Finnish listeners (n = 18) in the experiment, i.e. in only 2% of these cases Finnish listeners responded that the speaker's mother tongue was Russian. Russian listeners with Finnish proficiency (G2) (n = 12) judged Finnish stimuli as native speakers of Russian in 7% of the cases. Non-learner listeners (G3) (n = 18) falsely identified Finnish stimuli as native speakers of Russian in 15% of the cases. A one-way between-groups analysis of variance was conducted to explore the impact of listener group on identification. As the Levene's test indicated that the homogeneity of variance assumption of the ANOVA was violated, we are reporting the values of the Robust test of equality of means, the Welch test. There was a statistically significant difference at the $p < .0001$ level in the three listener groups: $F(2, 23.78) = 9.85$, $p = .001$. Post-hoc comparisons using the Tukey HSD test indicated that the mean score for G3 ($M = 14.44$, $SD = 2.15$) was significantly different from G1 ($M = 16.61$, $SD = .61$) and G2 ($M = 15.91$, $SD = .90$). However, G1 did not differ from G2. This indicates that proficiency in Finnish (whether native or language learner) helps to identify the Finnish native speakers' samples better than listeners with no knowledge of Finnish.

However, it is more interesting to investigate how successfully the Russian

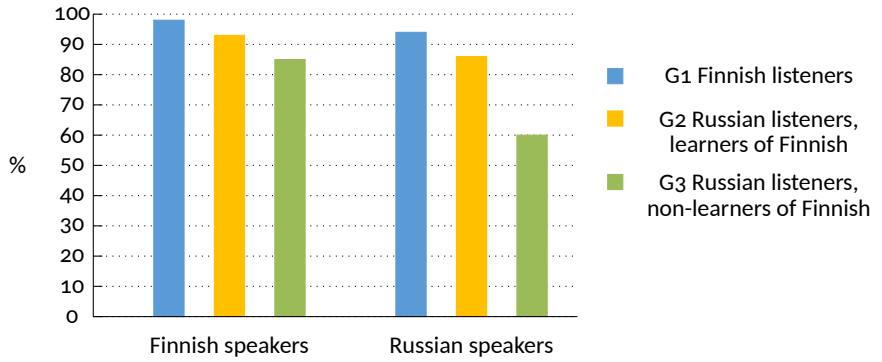

FIGURE 1. Identification rates of Finnish and Russian speakers' samples by the three listener groups.

speakers' stimuli (n = 63) were identified. As Figure 1 shows, Finnish native listeners were in general better than Russian listeners at recognizing speakers with a Russian language background (94% recognition rate for all speakers). Also Russian listeners with Finnish proficiency (G2) were better than those with no proficiency in Finnish (G3) in recognizing the Russian speaker samples. G2 recognized 86% and G3 60% of the Russian stimuli. A one-way between-groups analysis of variance was conducted to explore the impact of listener group on identification. As the Levene's test indicated that the homogeneity of variance assumption of the ANOVA was violated, we are again reporting the values of the Robust test of equality of means, the Welch test. There was a statistically significant difference at the $p < 0.0001$ level in the three listener groups: $F(2, 23.43) = 126.24$, $p = 0.0001$. Post-hoc comparisons using the Tukey HSD test indicated that the mean score for G3 ($M = 36.83$, $SD = 4.89$) was significantly different from G1 ($M = 59.28$, $SD = 3.32$) and G2 ($M = 53.17$, $SD = 7.18$). G1 also differed significantly from G2 in mean score. Thus, the degree of proficiency in Finnish seems to be linked to the ability to identify Russian speakers' stimuli from the native Finnish ones.

4.2 Identification across speakers and time

Differences between the listener groups for all the three speakers at three time points (recording sessions T1, T2 and T3) are shown in Figure 2. Kruskal-Wallis test $X^2(2, 48) = 28.76$ (T1), 32.34 (T2) and 34.62 (T3), $p = 0.001$ shows a statistically significant difference in ratings across time. Judgments by the native Finnish listeners (G1) (95%, 93% and 94%) and Russian listeners with Finnish proficiency (G2) (83%, 85% and 88%) were significantly different from Russian listeners with no proficiency in Finnish (G3) (67%, 52% and 66%) in

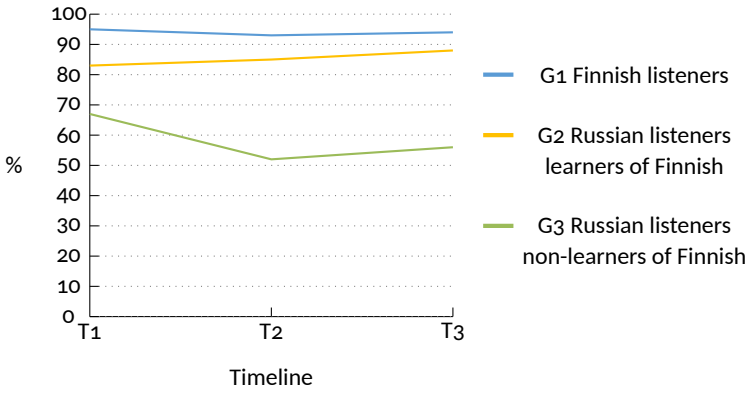

FIGURE 2. Judgments of all the Russian speakers at different recording sessions (T1, T2, T3) by all rater groups.

all the recording sessions (T1, T2 and T3). This was shown in pairwise comparisons at $p = 0.001$ level between G1 and G3 ($X^2 = 12.54$) and at $p = 0.045$ level between G2 and G3 ($X^2 = 24.75$) for T1. For the second recording session (T2) this was shown in pairwise comparisons at $p = 0.001$ level between G1 and G3 ($X^2 = 20.19$) and between G2 and G3 ($X^2 = 25.28$) and for the third recording session (T3) at $p = 0.001$ level between G1 and G3 ($X^2 = 26.39$) and between G2 and G3 ($X^2 = 19.75$).

As Figure 2 indicates, there is most variation across time in the judgments of the Russian listeners with no proficiency in Finnish (G3), and thus we will next focus on them and the differences in their judgments between the recording sessions. The results of the Friedman Test indicated that there was a statistically significant difference in the identification rates across the three time points for Russian listeners with no proficiency in Finnish (G3) (based on Friedman Test $X^2(2) = 11.35$, $p < .01$). For the other two listener groups, such a difference was not observed. Post-hoc testing was done using Wilcoxon Signed Ranks Test, which revealed that T1 differed from T2 significantly ($z = -3.23$, $p = .005$), but T2 did not differ from T3 in the ratings of G3.

Figure 3 shows only the ratings ($n = 1134$) of G3, that is Russian listeners with no Finnish proficiency ($n = 18$) across time and speakers. All Russian speakers were judged less often as Russian between the first two recording sessions (from T1 to T2) and Rus1 and Rus2 between the first and last recording sessions (from T1 to T3) by the Russian listeners with no proficiency in Finnish (G3) (Figure 3). For speakers Rus1 and Rus2, there was a tendency to be less often identified as Russian as the time they spent in Finland increased. However, there was a different tendency for speaker Rus3, who was identified as Russian in 74% of the samples at T1 and 48% at T2, but the identification

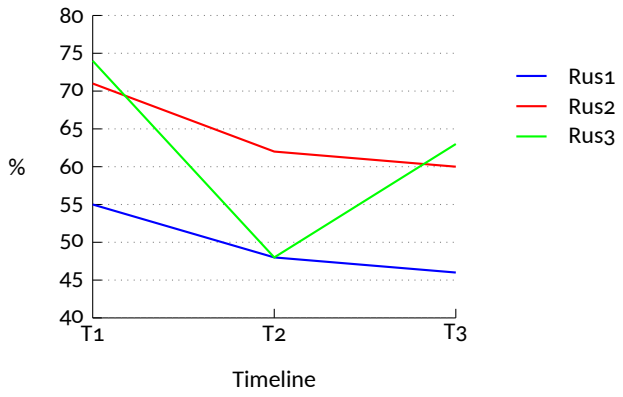

FIGURE 3. Success in identifying the three speakers' (Rus1, Rus2, Rus3) mother tongue as Russian at three time points (T1, T2, T3) by the Russian naive listener group (G3) only.

rate increased at T3 to 63%. This could be due to the fact that this speaker interrupted her Finnish course between T2 and T3.

4.3 Pronunciation features behind the judgments

The listeners also reported the basis of their judgments. Russian listeners with no Finnish proficiency (G3) answered an open-ended question for the whole task, whereas Finnish listeners (G1) and Russian listeners with Finnish proficiency (G2) were asked to choose between prosody, segmental features and speech rate for each stimulus they heard. A majority of the Russian listeners with no proficiency in Finnish (G3) mentioned one or two prosodic features as the basis of their judgments. Of these, intonation was mentioned most often, but also rhythm, stress and segment duration were reported. The Finnish listeners reported that they paid most attention to segmental features (48%), followed by prosody (37%), whereas speech rate was mentioned in only 15% of the cases.

Figure 4 shows the judgments of only one group, that is Russian listeners with Finnish proficiency (G2) at different recording sessions. This group mostly paid attention to segmental quality and duration at all the time points: 66% of the listeners mentioned this at T1, 65% at T2, and 66% at T3 66%. At the first recording session (T1) prosody (62% of the listeners) and speech rate (44% of the listeners) were also more frequently mentioned than in the other recording sessions, where prosody was mentioned by 52% of the listeners both at T2 and T3 and speech rate by 28% at T2 and 26% at T3.

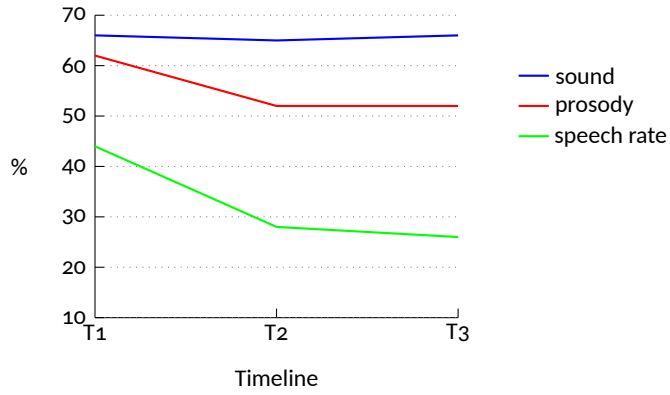

FIGURE 4. Responses of the Russian listeners with Finnish proficiency (G2) about the features of speech they based their judgments on at different recording sessions (T1, T2 and T3).

5 Discussion and conclusions

The purpose of the study was to investigate the potential differences between three different listener groups in evaluation of Russian accented Finnish. The study shows that proficiency in Finnish seems to be connected with the success of differentiation of Russian speakers' stimuli from the native Finnish ones in an auditory experiment where the listeners are asked to define whether each Finnish stimulus is spoken by a native Russian speaker or not. Finnish listeners (G1) identified both Finnish speakers' and Russian speakers' stimuli with very high accuracy and Russian listeners with Finnish proficiency (G2) were also rather accurate in their identification of both speaker groups. Similarly as in Weber & Pöllmann (2011) the ratings of non-native non-learners differed from ratings of other listener groups. The listeners with no proficiency in Finnish (G3) were the least successful in identifying the Russian speakers. The differences in the identification rates between the three listener groups were significant for Russian speakers' samples, but only between G1 and G2 for Finnish speakers' samples. Thus, we can conclude that some or native proficiency in Finnish makes a difference in identification of the Finnish speakers' stimuli as compared to the listeners without Finnish proficiency. For the main research question, that is, how well non-native speakers were identified by different listener groups, we can conclude that the ability to identify Russian accented samples is linked with the listener's degree of proficiency in Finnish.

When the listeners were asked to define the basis of their judgments overall, many mentioned prosodic features and segmental duration. Thus, as Russian and Finnish prosody are known to differ greatly, it is very likely that this

is explained by displaced foreign accent detection proposed by Major (2007), meaning that the listeners are making their judgments based on their knowledge of their mother tongue or language(s) they are learning. In addition, it is possible that the listener ratings were based on how proficient they were in Finnish, as could have been the case for G2 in our study. Our findings are in line with Riney et al. (2005), as native listeners (G1) paid the most attention to segmental features. Both Russian listener groups (G2 and G3), responded similarly to each other: they paid attention to prosody, but also to segmental quality and vowel /consonant duration. However, based on our findings it seems that both prosody and segments contributed equally to the judging of whether the speaker's mother tongue was Russian or not. Speech rate also received some attention at the first recording session by the listeners who were Finnish L2 learners.

As the study included only three speakers that were followed during one year, we can only draw tentative conclusions about pronunciation learning. We can say that we identified individual differences in "learning to sound less Russian" as perceived by naive listeners. Speakers Rus1 and Rus2 showed a tendency to have less Russian accent over the whole observation period, whereas Rus3 did so during the first six months (from T1 to T2), but then declined during the last six months (from T2 to T3). The decline may be explained by the fact that Rus3 interrupted her Finnish studies after T2 for family reasons and that her Finnish use changed at the same time, from rather frequent to a lot less frequent, and she started using more Russian at home.

From the pedagogic point of view, the tentative results of this study imply that participating in a language course while residing in the country where the language to be learnt is spoken can diminish features of foreign accent in speech. In this limited sample (n = 3) of Russian learners of Finnish, formal instruction in Finnish seemed to make a difference in how successfully listeners recognized Russian speakers. As was mentioned above, speaker Rus3, who dropped out of the language course during the data collection, was more often recognized as Russian after dropping out than before.

There are a number of possible directions for the future longitudinal studies on assessment of foreign accented speech, as the extant literature covers only limited issues. First, it would be important to record a great number of speakers during the first phase of data collection, as it is impossible to control for background variables, e.g. participation in language courses, interaction with native speakers and it can be challenging to record the same speakers multiple times over time without having to sacrifice ideal conditions for phonetic data collection. In the present study only three out of the ten Russian speakers originally recorded at the first recording session were available for two other consequent recording sessions conducted in the sound proof studio over one year.

Second, acoustic measurements would complete the results from the auditory perception task. It would be interesting, for example, to acoustically measure the perceived pronunciation features reported by the raters that contribute to the perception of foreign accent. Similarly as for the auditory perception task, the acoustic analysis also requires more speakers in order to be able to draw generalizable conclusions. Third, more raters with no knowledge of Finnish could be recruited for such a study. Fourth, in the experiment, we did not ask the listeners with no proficiency in Finnish (G3) to determine the basis of their judgments for each stimulus, as we considered this too demanding a task in a language completely unknown to them. In future studies, however, it would be interesting to give the same multiple-choice questionnaire to all listener groups. Such an experiment would allow a more detailed comparison between the listener groups.

There is still a great deal of research to be undertaken in the longitudinal study of L2 pronunciation to improve our understanding of the potential general trends in pronunciation learning, as much of the scarce previous longitudinal studies (e.g. Ullakonoja 2011; Derwing & Munro 2013) have concluded that individual differences in learning paths are great.

Acknowledgements

We wish to thank the ProoF-project (funded by the Academy of Finland) for providing the data for the auditory experiment as well as the Tell Me project (funded by Tekes) and FOKUS-project (funded by Society of Swedish Literature in Finland) for enabling the authors to work on this paper. Moreover, we wish to acknowledge the help of the late docent Nina Volskaya (from St. Petersburg State University) for her help with data collection.

References

Abrahamsson, N. & K. Hyltenstam 2009. Age of onset and nativelikeness in a second language: listener perception versus linguistic scrutiny. *Language Learning*, *59* (2), 249–306. DOI: 10.1111/j.1467-9922.2009.00507.x.

Aho, E. & M. Toivola 2008. Venäläisten maahanmuuttajien suomen prosodiasta. *Virittäjä*, *112* (1), 3–23. https://journal.fi/virittaja/article/view/40638.

Aho, E., M. Toivola, F. Karlsson & M. Lennes 2016. Aikuisten maahanmuuttajien suomen ääntämisestä. *Puhe ja kieli*, *36* (2), 77–96. http://hdl.handle.net/10138/168735.

Ahola, S. & H. Tossavainen 2016. Helppoa vai vaikeaa arvioida? Yleisten kielitutkintojen suomen kielen arvioijien käsityksiä vironkielisten arvioinnista. *Lähivõrdlusi. Lähivertailuja*, *26*, 27–53. DOI: 10.5128/LV26.01.

Bondarko, L. V. 1998. *Fonetika sovremennogo russkogo âzyka*. St Petersburg: Izdatel'stvo Sankt-Peterburgskogo universiteta.

Bondarko, L. V. 2009. Short description of Russian sound system. In V. de Silva & R. Ullakonoja (eds) *Phonetics of Russian and Finnish: general description of phonetic systems, experimental studies on spontaneous and read-aloud speech*. Frankfurt am Main: Peter Lang, 23–35.

Bongaerts, T., S. Mennen & F. van der Slik 2000. Authenticity of pronunciation in naturalistic second language acquisition: the case of very advanced late learners of Dutch as a second language. *Studia Linguistica*, *54* (2), 298–308. DOI: 10.1111/1467-9582.00069.

CEFR 2001. *Common European framework of reference for languages: learning, teaching, assessment*. Cambridge: Cambridge University Press.

Cunningham-Andersson, U. & O. Engstrand 1989. Perceived strength and identity of foreign accent in Swedish. *Phonetica*, *46* (4), 138–154.

Derwing, T. M. & M. J. Munro 2001. What speaking rates do non-native listeners prefer? *Applied Linguistics*, *22* (3), 324–337. DOI: 10.1093/applin/22.3.324.

— 2013. The development of L2 oral language skills in two L1 groups: a 7-year study. *Language Learning*, *63* (2), 163–185. DOI: 10.1111/lang.12000.

Flege, J. E. 1981. The phonological basis of foreign accent: a hypothesis. *TESOL Quarterly*, *15* (4), 443–455. DOI: 10.2307/3586485.

Gilbert, J. 1980. Prosodic development: some pilot studies. In R. C. Scarcella & S. D. Krashen (eds) *Research in second language acquisition: selected papers of the Los Angeles second language acquisition research forum*. Rowley (Mass.): Newbury House, 110–117.

Gonet, W. & G. Pietroń 2004. The Polish tongue in the English ear. In E. Waniek-Klimczak & W. Sobkowiak (eds) *Zeszyty naukowe państwowej wyższej szkoły zawodowej w Koninie*. Konin: Wydawnictwo PWSZ, 56–65.

Gupta, A. 2005. Inter-accent and inter-cultural intelligibility: a study of listeners in Singapore and Britain. In D. Deterding, A. Brown & E. Low (eds) *English in Singapore: phonetic research on a corpus*. New York (N. Y.): McGraw-Hill, 138–152.

Iivonen, A. 1978. Is there interrogative intonation in Finnish? In E. Gårding, G. Bruce & R. Bannert (eds) *Nordic prosody: papers from a symposium*. Travaux de l'Institut de Linguistique de Lund XIII. Lund: Department of Linguistics, University of Lund, 43–53.

— 1998. Intonation in Finnish. In D. Hirst & A. Di Cristo (eds) *Intonation systems: a survey of twenty languages*. Cambridge: Cambridge University Press, 311–327.

— 2009a. Finnish sentence accent and intonation. In V. de Silva & R. Ullakonoja (eds) *Phonetics of Russian and Finnish: general description of phonetic systems, experimental studies on spontaneous and read-aloud speech*. Frankfurt am Main: Peter Lang, 67–73.

— 2009b. Major features of standard Finnish phonetics. In V. de Silva & R. Ullakonoja (eds) *Phonetics of Russian and Finnish: general description of phonetic systems, experimental studies on spontaneous and read-aloud speech*. Frankfurt am Main: Peter Lang, 47–65.

Jaakkola, M. 2006. Suomalaisten suhtautuminen ulkomaiseen työvoimaan ja pakolaisiin. *Kuntoutus*, *29* (1), 5–21.

— 2009. *Maahanmuuttajat suomalaisten näkökulmasta: asennemuutokset 1987–2007*. Helsinki: Helsingin kaupungin tietokeskus.

Jenkins, J. 2000. *The phonology of English as an international language*. Oxford: Oxford University Press.

Kennedy, S. & P. Trofimovich 2008. Intelligibility, comprehensibility, and accentedness of L2 speech: the role of listener experience and semantic context. *The Canadian Modern Language Review*, *64* (3), 459–489. DOI: 10.3138/cmlr.64.3.459.

Leinonen, A. 2015. *"Riittää kun saa selvää": vieraalla aksentilla tuotettu suomi nuorten arvioimana*. Jyväskylä Studies in Humanities 275. Jyväskylä: University of Jyväskylä. http://urn.fi/URN:ISBN:978-951-39-6448-1.

Lippi-Green, R. 2012. *English with an accent: language, ideology, and discrimination in the United States*. 2nd ed. New York (N. Y.): Routledge.

MacKay, I. R. A., J. E. Flege & S. Imai 2006. Evaluating the effects of chronological age and sentence duration on degree of perceived foreign accent. *Applied Psycholinguistics*, *27* (2), 157–183. DOI: 10.1017.S0142716406060231.

Major, R. C. 2007. Identifying a foreign accent in an unfamiliar language. *Studies in Second Language Acquisition*, *29* (4), 539–556.

McCullough, E. 2013. Perceived foreign accent in three varieties of non-native English. *Ohio State University Working Papers in Linguistics*, *60*, 51–66.

Munro, M. J. 2008. Foreign accent and speech intelligibility. In J. G. Hansen Edwards & M. L. Zampini (eds) *Phonology and second language acquisition*. Amsterdam: John Benjamins, 193–218.

Munro, M. J. & T. M. Derwing 1999. Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, *49* (s1), 285–310. DOI: 10.1111/j.1467-1770.1995.tb00963.x.

— 2015. A prospectus for pronunciation research in the 21st century: a point of view. *Journal of Second Language Pronunciation*, *1* (1), 11–42. DOI: 10.1075/jslp.1.1.01mun.

Munro, M. J., T. M. Derwing & S. Morton 2006. Mutual intelligibility of L2-speech: language learning. *Studies in Second Language Acquisition*, *28* (1), 111–131. DOI: 10.1017/S0272263106060049.

Pinget, A.-F., H. R. Bosker, H. Quené & N. H. De Jong 2014. Native speakers' perceptions of fluency and accent in L2 speech. *Language Testing*, *31* (3), 349–365. DOI: 0.1177/0265532214526177.

Piske, T., I. R. A. MacKay & J. E. Flege 2001. Factors affecting degree of foreign accent in an L2: a review. *Journal of Phonetics*, *29* (2), 191–215. DOI: 10.1006/jpho.2001.0134.

ProoF 2012. Pronunciation of Finnish by immigrants in Finland, version 1.0. http://urn.fi/urn:nbn:fi:lb-20140730130 [accessed 1/5/2018].

Riney, T. J., N. Takagi & K. Inutsuka 2005. Phonetic parameters and perceptual judgments of accent in English by American and Japanese listener. *TESOL Quarterly*, *39* (3), 441–466. DOI: 10.2307/3588489.

Saito, K., P. Trofimovich & T. Isaacs 2017. Using listener judgments to investigate linguistic influences on l2 comprehensibility and accentedness: a validation and generalization study. *Applied Linguistics*, *38* (4), 1–25. DOI: 10.1093/applin/amv047.

Scovel, T. 1969. Foreign accents, language acquisition, and cerebral dominance. *Language Learning*, *19* (3–4), 245–253. DOI: 10.1111/j.1467-1770.1969.tb00466.x.

Statistics Finland 2017. Statistics Finland's databases. Helsinki: Statistics Finland. http://pxnet2.stat.fi/PXWeb/pxweb/en/StatFin/StatFin__vrm__vaerak/statfin_

vaerak_pxt_010.px/?rxid=22bb0a37-029e-4d0e-ac53-69dda024ae3c [accessed 5/4/2017].

Toivola, M. 2011. *Vieraan aksentin arviointi ja mittaaminen suomessa*. Helsinki: Helsingin yliopisto. http://urn.fi/URN:ISBN:978-952-10-7217-8.

Toivola, M., M. Lennes & E. Aho 2009. Speech rate and pauses in non-native Finnish. In M. Uther, R. Moore & S. Cox (eds) *Interspeech-2009*. Brighton: ISCA, 1707–1710.

Toivola, M., M. Lennes, J. Korvala & E. Aho 2011. A longitudinal study of speech rate and pauses in non-native Finnish. In K. Dziubalska-Kołaczyk, M. Wrembel & M. Kul (eds) *New sounds 2010: proceedings of the sixth international symposium on the acquisition of second language speech*. Poznań: Adam Mickiewicz University, 499–504.

Trofimovich, P. & W. Baker 2006. Learning second language suprasegmentals: effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition*, *28* (1), 1–30. DOI: 10.1017/S0272263106060013P.

Ullakonoja, R. 2011. *Da. Eto vopros! Prosodic development of Finnish students' read-aloud Russian during study in Russia*. Jyväskylä Studies in Humanities 151. Jyväskylä: University of Jyväskylä. http://urn.fi/URN:ISBN:978-951-39-4209-0.

Ullakonoja, R., H. Dufva, M. Kuronen & P. Hurme 2014a. How to imitate an unknown language? Russians imitating Finnish. In K. Jähi & L. Taimi (eds) *XXVIII Fonetiikan päivät: Turku 25.–26. lokakuuta 2013*. Turku: University of Turku, 10–18. https://www.utu.fi/fi/sivustot/fp2013/Documents/XXVIII_Fonetiikan_p%C3%A4iv%C3%A4t_konferenssijulkaisu.pdf [accessed 3/19/2017].

Ullakonoja, R. & M. Kuronen 2015. Young Russian immigrants' segmental duration and length in Finnish. In The Scottish Consortium for ICPhS 2015 (ed.) *Proceedings of ICPhS 2015: 18th international congress of phonetic sciences, 10–14 Aug 2015 Secc Glasgow Scotland UK*. Glasgow: University of Glasgow. https://www.internationalphoneticassociation.org/icphs-proceedings/ICP.

Ullakonoja, R., M. Kuronen, P. Hurme & H. Dufva 2014b. Segment duration in Finnish as imitated by Russians. In N. Campbell, D. Gibbon & D. Hirst (eds) *Social and linguistic speech prosody: proceedings of the 7th International conference on speech prosody*. Dublin: Science Foundation Ireland & Intenational Speech Communication Association, 507–511.

Uzal, M., T. Peltonen, M. Huotilainen & O. Aaltonen 2015. Degree of perceived accent in Finnish as a second language for Turkish children born in Finland. *Language Learning*, *65* (3), 477–503. DOI: 10.1111/lang.12126.

Volskaya, N. 2009. Aspects of Russian intonation. In V. de Silva & R. Ullakonoja (eds) *Phonetics of Russian and Finnish: general description of phonetic systems, experimental studies on spontaneous and read-aloud speech*. Frankfurt am Main: Peter Lang, 34–46.

Walker, R. 2011. *Teaching the pronunciation of English as a lingua franca*. Oxford: Oxford University Press.

Weber, A. & K. Pöllmann 2011. Identifying foreign speakers with an unfamiliar accent or in an unfamiliar language: proceedings of the sixth international symposium on the acquisition of second language speech. In K. Dziubalska-Kołaczyk, M. Wrembel & M. Kul (eds) *New sounds 2010: proceedings of the sixth international symposium on the acquisition of second language speech*. Poznań: Adam Mickiewicz University, 536–541.

Wilkerson, M. E. 2010. Identifying accent in German: a comparison of native and non-native listeners. *Die Unterrichtspraxis/Teaching German*, *43* (2), 144–153.