

# Speech rate and pauses in non-native Finnish

Minnaleena Toivola<sup>1</sup>, Mietta Lennes<sup>1,2</sup>, Eija Aho<sup>1</sup>

<sup>1</sup>Department of General Linguistics, University of Helsinki, Finland

<sup>2</sup>Department of Speech Sciences, University of Helsinki, Finland

Minna.Toivola@helsinki.fi, Mietta.Lennes@helsinki.fi, Eija.Aho@helsinki.fi

## Abstract

In this study, the temporal aspects of speech are compared in read-aloud Finnish produced by six native and 16 non-native speakers. It is shown that the speech and articulation rates as well as pause durations are different for native and non-native speakers. Moreover, differences exist between the groups of speakers representing four different non-native languages. Surprisingly, the native Finnish speakers tend to make longer pauses than the non-natives. The results are relevant when developing methods for assessing fluency or the strength of foreign accent.

**Index Terms:** speech rate, articulation rate, pause, second language, foreign accent, Finnish

## 1. Introduction

The temporal aspects of speech play an important role in the production of a second language. Speech rate has been shown to correlate with fluency ratings as well as with global accentedness ratings [1, 2]. According to Cucchiari, Strik and Boves [3], two factors are important for perceived fluency in read-aloud speech: the rate at which speakers articulate the sounds and the number of pauses they make. Trofimovich and Baker [2] found that pause duration and speech rate contributed to foreign accent more than pause frequency, stress timing or peak alignment.

It is well known that native speakers tend to speak significantly faster than non-native speakers. In addition, speech rate is affected by the age of second-language learning. Guion, Flege, Liu and Yeni-Komshian [4] discovered that late L2 learners tended to produce slower speech than early learners. Respectively, late L2 learners were often judged as sounding more accented than early learners.

Various elicitation methods have been used in previous studies of speech rate and pausing in second-language speech. Cucchiari et al. asked their speakers to read aloud a list of isolated sentences [3]. However, isolated sentences may not elicit many sentence-internal pauses from native speakers, and the relevance of pauses between isolated sentences is somewhat questionable. Consequently, it is difficult to reliably compare the pauses produced by native and non-native speakers. Trofimovich and Baker's material consisted of declarative sentences elicited as responses to question prompts [2]. Guion et al. [4] presented their speakers with a written list of five sentences. In their study, model pronunciations of the sentences were presented aurally in short dialogues. It is unclear whether the use of a pronunciation model may confuse the speaker or disguise the speaker's actual pronunciation skills. Moreover, few studies have reported separate results for different native languages, although it may be expected that languages differ in the temporal features of reading aloud and that this might be reflected in L2 speech.

The temporal properties of speech affect both fluency and, to some extent, the perceived strength of foreign accent especially at the early stages of L2 learning [5]. Nevertheless, there are very few studies of the temporal aspects in the speech of low-proficiency L2 speakers, as most research has been conducted with intermediate or advanced learners, e.g., university students.

Currently, there are no experimental studies of the temporal aspects of non-native Finnish speech. The aim of this study is to compare the speech rate and pauses in non-native low proficiency speech with those of native Finnish speech.

## 2. Material and methods

### 2.1. Speakers

For this study, the speech of 16 non-native speakers (NNS) of Finnish was recorded. Four different mother tongues were represented in the NNS group: six of the subjects were native speakers of Russian (five female), three were native speakers of Thai (two female), four were native speakers of Turkish (two female) and three were native speakers of Vietnamese (three female). Individual speakers will be referred to by codes, consisting of three letters for the speaker's native language (e.g., *tha* for a Thai speaker), one capital letter for gender (F=female, M=male), and a running index number.

The NNS had been living in Finland for the median duration of 12 months. Although most of the NNS had immigrated fairly recently, three speakers (*thaF1*, *turM1* and *vieF1*) had already stayed in Finland for almost three years. However, the speakers *thaF1* and *turM1* had been using English as their first language of communication before attending any Finnish lessons, as many Finns know English well and are eager to use it. Moreover, the NNS (including *vieF1*) had quite often used their own mother tongue in the beginning of their stay. The NNS had been attending beginner-level Finnish language courses for the median duration of 8 months, and they ranged in age from 18 to 48 (median 28.5 years). All of the NNS had completed either secondary school (younger speakers) or vocational school (elder speakers). According to an informal assessment by the experimenters, all speakers had distinct foreign accents. In addition, six native speakers of Finnish (NFS, three female; age 22–28 years,  $M=24.8$ ) were recorded. All speakers were living in the capital city area of Finland.

### 2.2. Recordings

The recordings were performed in a sound-treated studio using high-quality audio equipment. Each speaker's voice was recorded with a head-mounted microphone at a sample rate of 44.1 kHz and sample size of 16 bits.

The speakers were asked to read aloud a Finnish text of 125 words extracted from a library brochure. They were instructed to read the text at their normal speaking rate. Each

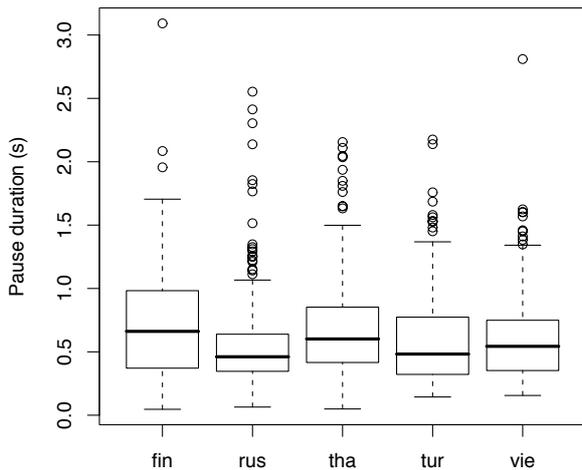


Figure 1: The distributions of pause durations in read-aloud Finnish for native Finns and native speakers of Russian, Thai, Turkish and Vietnamese.

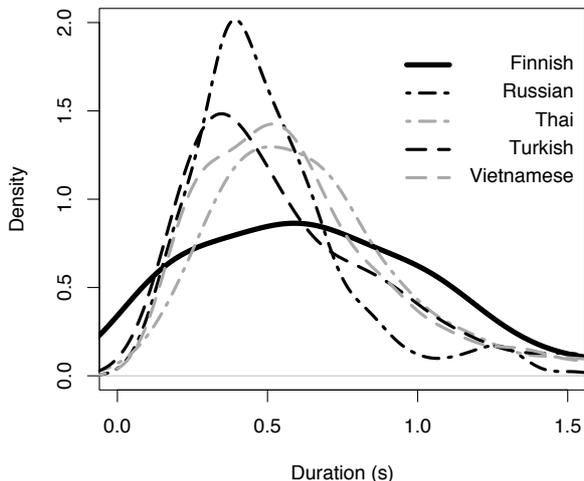


Figure 2: The probability density functions for pause durations in read-aloud Finnish for native Finns and native speakers of Russian, Thai, Turkish and Vietnamese. Each curve represents the probability density function for the measured pause durations.

recording of the complete read-aloud text was saved as one sound file for further analysis.

### 2.3. Analysis

In order to be able to perform the measurements automatically, the sound files were annotated using the Praat program [6]. First, utterance boundaries were marked for each speaker. Utterances were defined as stretches of speech during which the speaker was continuously articulating. Each utterance was then transcribed nearly orthographically. In order to be able to calculate the true number of syllables for each utterance, the words and the syllables were labeled into their separate tiers within the same annotation file. Although several Praat scripts were used in order to speed up the annotation process, the results were manually checked.

Another Praat script was used for calculating the duration of each pause as well as the duration, the number of syllables and the articulation rate for each utterance. The script was designed so as to run through all the annotation files in the speech material, to collect the relevant information for one utterance or pause at a time using the different annotation tiers, and to write this information into a table. The initial pause and the final pause of each recording were excluded from the analysis. Speech rate was calculated for each speaker as the total number of syllables divided by the total duration of the read-aloud text. Articulation rate was calculated as the number of syllables produced within the utterance, divided by the duration of the utterance, excluding pauses. Disfluencies and filled pauses were considered as part of utterances. Thus, e.g., hesitation noises would cause the articulation rate to decrease within the utterance. On the other hand, syllable boundaries were also marked for repeated or mispronounced words and false starts, and therefore such disfluencies would not necessarily affect articulation rate.

The statistical analysis of the data was performed using the R statistical programming environment [7] and SPSS.

## 3. Results

### 3.1. Pauses

During the same read-aloud text, the native Finnish speakers paused 14–28 times ( $M = 14.5$ ), whereas the native speakers of Russian had 32–75 pauses ( $M = 38$ ), the native Thai speakers 46–86 pauses ( $M = 78$ ), the native Turkish speakers 19–70 pauses ( $M = 59$ ), and the native Vietnamese speakers 78–104 pauses ( $M = 94$ ).

Contrary to what one might expect, the native Finnish speakers had generally longer pauses than the non-natives (see figure 1). The median duration of pauses was 0.6612 s for the native Finnish speakers, 0.4612 s for the native Russian speakers, 0.6018 s for the native speakers of Thai, 0.4826 s for Turkish and 0.5434 s for Vietnamese. The native speakers of Thai tended to have the most similar pause durations in comparison to the native Finns, whereas the speakers of Turkish and Russian tended to keep their pauses shorter than 0.5 s.

The shapes of the distributions of pause durations are shown as probability density functions in figure 2. Visual inspection confirms that there are some differences in pause durations among the native and non-native groups. Since the number of measurements was quite small and differed across speakers and groups, and since the pause durations were not normally distributed, there were few statistical tests available. Nevertheless, the medians for all five native languages were significantly different (median test,  $p = 0.000$ ). According to pairwise comparisons using the median test, there were significant differences between the speakers of Finnish and Russian as well as Russian and Thai ( $p < 0.001$  for both pairs), Finnish and Turkish, Russian and Vietnamese, Thai and Turkish ( $p < 0.005$ ), and Finnish and Vietnamese ( $p < 0.05$ ).

The distribution of pause durations was relatively flat for the native Finns, whereas the non-native speakers, especially the native speakers of Russian, were less variable in their pause durations. Indeed, the standard deviation of pause durations was 0.49 s for NFS and only 0.37 s for NNS (native Russian 0.36 s, Thai 0.40 s, Turkish 0.38 s and Vietnamese 0.34 s).

### 3.2. Speech rate and articulation rate

The speech rates and overall articulation rates for each speaker are shown in table 1. These two measures were well correlated in this data (Spearman's  $\rho = 0.95$ ). However, speech rate is probably sensitive to variation in pause durations and it cannot be separately calculated for each utterance.

With regard to both speech rate and articulation rate, the native Finnish speakers were systematically faster than the non-native speakers (see figure 3). The native speakers of Turkish tended to have higher articulation rates than the other non-natives, whereas the native speakers of Vietnamese were the slowest in both their articulation rates and their speech rates. The standard deviation of the articulation rate was 0.93 for native Finns, 0.89 for native speakers of Russian, 0.84 for Thai, 0.86 for Turkish and 0.65 for Vietnamese.

### 3.3. Phonation-time ratio

In order to describe the relationship between speech and pauses, the phonation-time ratio (PTR) was calculated for each speaker as the percentage of the sum of the durations of all utterances from the total duration of the recording (excluding initial and final pauses). The resulting PTRs ranged from 77.8 % to 88.45 % for the NFS and from 63.04 % to 88.53 % for the NNS (see table 1). Most of the native speakers of Russian and a few speakers of the other languages had PTRs similar to those for the native Finns. Thus, many non-native speakers spent the same overall proportion of time producing speech as the native speakers. A weak correlation between PTR and articulation rate was found in the present data (Spearman's  $\rho = 0.51$ ).

## 4. Discussion

As expected, the native Finnish speakers were clearly faster speakers in comparison to the non-natives. Since the speakers were reading aloud a longer passage and not only a list of sentences as in many previous studies, pauses occurred both within and between sentences for both native and non-native speakers. The native speakers tended to have longer and fewer pauses in their read-aloud speech. The difference in pause durations remained even when comparing only those pauses that occurred at the clause boundaries in the text. Thus, the pausing frequency does not appear to affect the duration of pauses. There was no correlation between median pause duration and pause frequency for individual speakers (Spearman's  $\rho = -0.17$ ).

The results are important for the definition of "fluent" speech. For instance, it is important to note that individual speakers differ in their pausing habits, and all speakers may not be equally "fluent" readers. One of the native Finnish speakers (*finF3*) in this data exhibited clearly smaller pause durations than any of the other natives, rushing from one sentence to the next. Another Finnish speaker (*finF8*) spoke very calmly and clearly and exhibited a large variance in her pause durations, even overlapping the distributions of most non-natives. Therefore, the language-related differences in pause durations can be considered as tendencies only. Pause duration cannot be used as an index of fluency, since it probably depends on many factors such as speech style, and since it is prone to individual variation. As long as 'fluency' is not a clearly and consistently defined concept, it cannot be reliably measured with any single variable or even a combination of variables.

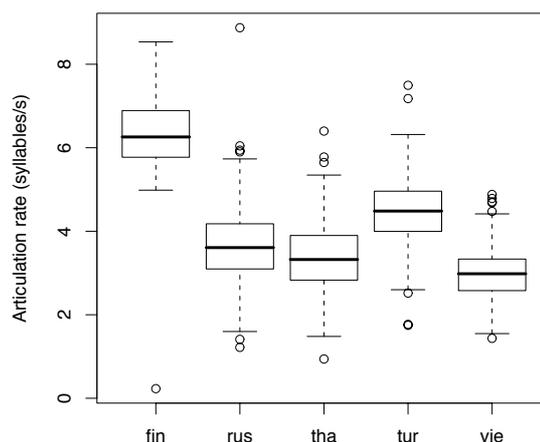


Figure 3: The distributions of articulation rates for utterances in read-aloud Finnish by native Finns and native speakers of Russian, Thai, Turkish and Vietnamese.

Table 1. Phonation-time ratio (PTR, the total duration of utterances divided by total duration of the speech sample), articulation rate (syllables per second within all utterances) and speech rate (syllables per second within the whole speech sample) for six native Finnish speakers and native speakers of Russian, Thai, Turkish and Vietnamese. Female speakers are indicated with the letter F, male speakers with an M.

Speaker	PTR (%)	Articulation rate (syll/s)	Speech rate (syll/s)
finF3	86.76	6.64	5.76
finF12	83.92	6.52	5.47
finM5	88.45	6.50	5.75
finM3	83.29	6.25	5.20
finM4	84.01	5.94	4.99
finF8	77.85	5.85	4.56
rusF1	85.00	4.46	3.79
rusF4	84.79	4.05	3.43
rusF6	81.98	3.66	3.00
rusM1	73.62	3.56	2.62
rusF3	88.53	3.54	3.13
rusF2	75.36	3.11	2.34
thaM1	63.04	3.72	2.35
thaF2	82.94	3.24	2.69
thaF1	70.61	2.96	2.09
turF4	69.87	4.93	3.44
turM5	83.33	4.42	3.68
turM1	71.72	4.18	3.00
turF1	76.66	4.17	3.20
vieF1	78.95	3.04	2.40
vieF2	67.12	3.00	2.01
vieF3	72.12	2.68	1.93

## 5. Conclusions

In this study, the read-aloud speech of six native and 16 second-language speakers of Finnish was compared. The beginner-level learners of Finnish were shown to have smaller articulation rates and more frequent pauses than the native Finnish speakers. Contrary to previous results from other languages, native Finnish speakers exhibited clearly longer pauses in their read-aloud speech than non-native speakers. It also appears that different native languages have different effects on articulation rate and pausing in L2 Finnish, at least in the early stages of L2 learning. However, more data are required in order to be able to generalize the findings.

This study presents the very first results based on the corpus of second-language Finnish speech that is being collected at the University of Helsinki, Finland. The corpus will include speech samples from speakers of over ten different native languages and represent various speech styles, including spontaneous speech. The present data do not yet indicate how speech rate, pausing and other pronunciation features in second-language Finnish may develop over longer periods of time. In order to reveal such long-term changes, follow-up recordings will be performed for a number of non-native Finnish speakers.

## 6. Acknowledgements

This study is part of the larger research project *Proof – Pronunciation of Finnish by immigrants in Finland*, <http://www.ling.helsinki.fi/ykiel/research/Proof/>, funded by the Academy of Finland (project number 122624).

## 7. References

- [1] Munro, M. and Derwing, T., "Modelling perceptions of the comprehensibility and accentedness of L2 speech: The role of speaking rate", *Studies in Second Language Acquisition*, 23:451-468, 2001.
- [2] Trofimovich, P. and Baker, W., "Learning second-language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech", 28:1-30, 2006.
- [3] Cucchiari C., Strik, H. and Boves, L., "Quantitative assessment of second language learners' fluency by means of automatic speech recognition technology." *Journal of the Acoustical Society of America*, 107(2):989-999, 2000.
- [4] Guion, S.G, Flege, J.E, Liu, S.H. and Yeni-Komshian, G.H., "Age of learning effects on the duration of sentence produced in a second language." *Applied Psycholinguistics*, 21(2):205-228, 2000.
- [5] Derwing, T., Rossiter, M., Munro, M. and Thomson, R., "Second language fluency: Judgments on different tasks." *Language Learning*, 54(5):655-679, 2004.
- [6] Boersma, P. and Weenink, D. "Praat: doing phonetics by computer" (Version 5.1.04) [Computer program]. Online: <http://www.praat.org/>, retrieved on 4 April, 2009.
- [7] R Development Core Team. "R: A Language and Environment for Statistical Computing." R Foundation for Statistical Computing, Vienna, Austria, 2006.