

# USING LANGUAGE TESTING CORPORA TO INVESTIGATE L2 PRONUNCIATION

Riikka Ullakonoja, Hannele Dufva, Mikko Kuronen, Maria Kautonen & Elina Tergujeff  
Department of Languages, University of Jyväskylä, Finland  
riikka.ullakonoja | hannele.t.m.dufva | mikko.j.kuronen | maria.e.e.kautonen | elina.tergujeff@jyu.fi

**Keywords:** L2 learning, prosody, Swedish

## 1. BACKGROUND

First, we will shortly present our project *Focus on learning pronunciation: Swedish as L1/L2* ([www.jyu.fi/fokus](http://www.jyu.fi/fokus)). Second, we focus on the possibilities of using language testing corpora to investigate L2 pronunciation. The project focuses 1) on studying the phonetic aspects of L2 Swedish with speakers of Finnish, Russian and English as L1, and 2) studying aspects of Swedish L1 speakers' production of these languages as their L2. The variant we focus on is Finland Swedish, but also Swedish as spoken in Sweden will be considered (for Swedish phonology, see, e.g. [1], [3]; for differences between Finland Swedish and Swedish spoken in Sweden, see [14], [16]).

The cross-language design of the project was chosen for studying 1) the effect of different L1-L2 interrelationships (e.g. genetic relationship; writing system/orthography; phonological distance) on learning pronunciation and transfer phenomena between each L1-L2 pair and 2) the influence of various sociophonetic factors (e.g. language attitudes attached to different foreign accents) (see [13]). Hence, by using methods of acoustic and auditory analysis we will analyse features of learner language with Swedish either as L1 or L2 with the purpose of highlighting the potential language-specific difficulties, but also potential language-independent factors such as inter-individual differences (for second language pronunciation learning, see, e.g. [6], [18], [20]). While the particular focus of analysis is on prosody (see, e.g. [23]), we are also interested in segmental features typical for each variant of L2 speech (e.g. [15]), and further, in learners' own views on learning pronunciation (e.g. [12]).

In addition to its theoretical aims of gaining new information on the L2 oral performance, the project aims at a contribution to language teaching practices (see, e.g. [5], [7], [8], [19]) and design of teaching materials (see, e.g. [4], [15]). Finnish and Swedish, as two national languages of Finland, are mandatory subjects in formal education. English, while being a very widely used language of media, youth culture and business, is also the most frequent choice for the first foreign language at Finnish schools. Finally,

while Russian speakers are the largest linguistic group in Finland after the national language communities, the status of Russian at school is one of the lesser studied languages. Research-based knowledge is now particularly needed for developing teaching of oral skills because of the new oral proficiency exam to be launched (for the current practices, see, e.g. [22], [21]).

## 2. RESEARCH QUESTIONS

The main research questions of the project team are:

1. What phonetic/prosodic features are typical of Swedish as spoken by Finnish, Russian and English speakers? What are the most central problems for each L1 group? How can these problems be addressed in teaching pronunciation?
2. What phonetic/prosodic features are typical of English, Russian and Finnish as spoken by Swedish speakers? What are the most central problems in each L2? How can these problems be addressed in teaching pronunciation?
3. What features seem to be associated with L1 effects? What aspects may be language-independent? How do findings help to develop teaching pronunciation?
4. What is the added value of imitation experiments for understanding how pronunciation is learned?
5. How are different foreign accents evaluated in the society? How can this knowledge be used in teaching pronunciation?
6. What kind of pronunciation teaching methods seem to be effective (e.g. transcription, listening, visual feedback; teacher feedback)? How should this knowledge be used in designing teaching materials?

## 3. DATA

New data sets will be collected, but there are also existing corpora that will be used for analysis.

### 3.1 Corpora to be collected

The data sets to be collected ("FOKUS data") include: 1) Speech samples consisting of reading aloud (sentences, a short text) and spontaneous speaking (a short oral narrative task) to be analysed

acoustically and by perception tests. Data includes a) L2 Swedish as produced by Finnish, Russian and English speakers and b) L2 Finnish, English and Russian as produced by Swedish speakers. 2) Imitation tasks based on pilot studies by the authors [24, 25]. Task involves imitation of utterances by learners and non-learners (i.e. persons who have not studied the language) to be analysed acoustically and by perception tests. 3) Data on Finnish university students learning Swedish. The data includes both the students' oral performance, its analysis and the students' interviews. Data (both read-aloud and spontaneous speech) are being collected on Swedish pronunciation courses given at University of Jyväskylä.

### 3.2 Existing language testing corpora

The existing language testing corpora include: 1) The Finnish National Board of Education Learning Performances and 2) The National Certificates of Language Proficiency. The first corpus consists of the learning performances of Finnish ninth graders in Swedish, Russian and English in a test organized by the Finnish National Board of Education [9, 10, 11]. The oral skills sub-test has been video-recorded and evaluated on the Finnish National Curriculum Scale (a more fine-tuned version of the CEFR-scale). It consisted of four different tasks in each language, one of which was a monologue and three were dialogues. In addition, extensive background questionnaires have been collected from both the pupils and their teacher. All the participants are of the same age, come from a similar language learning background (Finnish schools), but have different L1s: Finnish, Finland-Swedish, and some Russian. They have taken the test in Swedish, Russian or English and we have taken a sub-sample of their performances to be investigated more thoroughly.

The National Certificates of Language Proficiency data consist of test performances given in a national language proficiency test evaluated across CEFR (Common European Framework for Languages) proficiency levels. The test takers come from a variety of language learning backgrounds (no specific course or study unit is required to take the test) and socioeconomic groups. In the test, each speaker is recorded typically doing two monologue tasks and some "interactional" tasks, where s/he is asked to have a conversation with the person s/he hears on the tape.

## 4. METHODS

Methods will include using tools of experimental phonetics such as acoustic analysis and perception tests. Programs such as Praat [2] and KayPENTAX'

Computerized Speech Lab (<http://www.kaypentax.com>) will be used. Results will be analysed using statistical means.

## 5. DISCUSSION

Using language testing corpora as data in L2 pronunciation research has both advantages and disadvantages. By using existing data, the steps of recruiting participants and making recordings can be avoided, and the speakers' oral proficiency level is reliably assessed. However, finding suitable samples from the data bank may be laborious due to missing video or audio files and consent forms. Also, as the conditions of the recordings vary, their quality is not always optimal. Further, due to ethical principles of the language testing system, the instructions of the speaking tasks themselves are not available and the use of background information of the subjects is also limited. Similarly, as the original video file is not available to researchers, it may be difficult to distinguish the speakers in dialogue tasks.

In comparison, we have also met challenges in finding suitable designs and tasks for the purposes of the present project, especially as concerns spontaneous elicitation. For example, when piloting tasks, we found a picture telling task unsatisfactory for investigating tone accent in Swedish as it resulted in pitch patterns that are typical of reading a list. Hence, while monologue tasks might be more suitable for such purposes, they present other types of challenges, such as their potential unauthenticity.

## 6. ACKNOWLEDGEMENTS

The project *Focus on learning pronunciation: Swedish as L1/L2* is funded by The Society of Swedish Literature in Finland (SLS).

## 7. REFERENCES

- [1] Bannert, R. 2004. *På väg mot svenskt uttal*. Lund: Studentlitteratur.
- [2] Boersma, P., Weenink, D. 2014. PRAAT: Doing phonetics by computer, Version 5.4, [Computer program] <http://www.praat.org/> (University of Amsterdam, Amsterdam)
- [3] Bruce, G. 2012. *Allmän och svensk prosodi*. Lund: Studentlitteratur.
- [4] de Silva, V., Mäkilä, K., Kärkkäinen, H., Ullakonoja, R. 2010. *Uusi venäjän ääntämisopas*. Helsinki: Finn Lectura.
- [5] Derwing, T., Munro, M. 2005. Second Language Accent and Pronunciation Teaching: A Research-Based Approach. *TESOL Quarterly*, 39:3, 379–397.
- [6] Flege, J. E. 1995. Second language speech learning: Theory, findings, and problems. In: W. Strange (ed). *Speech Perception and Linguistic Experience. Issues in Cross-Language Research* Timonium: York Press, 233–277.
- [7] Gilbert, J. 2008. *Teaching Pronunciation. Using the Prosody Pyramid*. Cambridge: Cambridge University Press.
- [8] Henderson, A., Frost, D., Tergujeff, E., Kautzsch, A., Murphy, D., Kirkova-Naskova, A., Waniek-Klimczak, E., Levey, D., Cunningham, U. & Curnick, L. 2012. The English Pronunciation Teaching in Europe Survey: Selected results. *Research in Language* 10.1, 5–27. <http://www.diva-portal.org/smash/get/diva2:564137/FULLTEXT02>
- [9] Hildén, R., Rautopuro, J. 2014a. Ruotsin kielen A-oppimäärän oppimistulokset perusopetuksen päättövaiheessa 2013. Helsinki: Finnish National Board of Education. Available (with summary in English) [http://oph.fi/download/160073\\_ruotsin\\_kielen\\_a\\_oppimaaran\\_oppimistulokset\\_perusopetuksen\\_paattovaiheessa\\_2.pdf](http://oph.fi/download/160073_ruotsin_kielen_a_oppimaaran_oppimistulokset_perusopetuksen_paattovaiheessa_2.pdf).
- [10] Hildén, R., Rautopuro, J. 2014b. Venäjän kielen A- ja B-oppimäärän oppimistulokset perusopetuksen päättövaiheessa 2013. Helsinki: Finnish National Board of Education. Available (with summary in English) [http://oph.fi/julkaisut/2014/venajan\\_kielen\\_a\\_ja\\_b\\_oppimaaran\\_oppimistulokset\\_perusopetuksen\\_paattovaiheessa\\_2013](http://oph.fi/julkaisut/2014/venajan_kielen_a_ja_b_oppimaaran_oppimistulokset_perusopetuksen_paattovaiheessa_2013).
- [11] Härmälä, M., Huhtanen, M., Puukko, M. 2014. *Englannin kielen A-oppimäärän oppimistulokset perusopetuksen päättövaiheessa 2013*. Helsinki: Finnish National Board of Education. Available (with summary in English) [http://oph.fi/download/160066\\_englannin\\_kielen\\_a\\_oppimaaran\\_oppimistulokset\\_perusopetuksen\\_paattovaiheessa.pdf](http://oph.fi/download/160066_englannin_kielen_a_oppimaaran_oppimistulokset_perusopetuksen_paattovaiheessa.pdf).
- [12] Kautonen, M. 2014. *Språkliga identiteter hos blivande språklärare*. Master's thesis. Department of Languages, University of Jyväskylä. <http://urn.fi/URN:NBN:fi:jyu-201410243088>
- [13] Kautonen, M., M. Kuronen, R. Ullakonoja, E. Tergujeff, Dufva, H. (2015). På väg mot bättre språkundervisning – FOKUS på uttal, *Language, Education and Society*, March 2015. <http://www.kieliverkosto.fi/article/pa-vag-mot-battre-sprakundervisning-fokus-pa-uttal/>
- [14] Kuronen, M., Leinonen, K. 2000. Fonetiska skillnader mellan finlandssvenska och rikssvenska. In: L. Jönsson et al. (eds) *Svenskans beskrivning 24. Linköpings universitet*. <http://www.ep.liu.se/ecp/006/011/ecp00611.pdf>
- [15] Kuronen, M., Leinonen, K. 2010. Svenskt uttal för finskspråkiga – teori och övningar i finlandssvenskt och rikssvenskt uttal för finskspråkiga. Tampere: Tampereen yliopistopaino.
- [16] Kuronen, M. 2000. Vokaluttalets akustik i sverigesvenska, finlandssvenska och finska. *Studia Philologica Jyväskyläensia* 49. Jyväskylä: University of Jyväskylä. <https://jyx.jyu.fi/dspace/handle/123456789/25579>
- [17] Kuronen, M. 2015. Bryter finlandssvenskarna när de talar finska? *Folkmålsstudier* 52.
- [18] Major, R. C. 2008. Transfer in second language phonology: A review. In: J. G. H. Edwards, M. L. Zampini (eds) *Phonology and Second Language Acquisition. Studies in Bilingualism* 36 Amsterdam: John Benjamins Publishing Company, 63–94.
- [19] Norris, J.M., Ortega, L. 2000. Effectiveness of L2 Instruction: A Research Synthesis and Quantitative Meta-analysis. *Language Learning*, 50:3, 417–528.
- [20] Piske, T., MacKay, I., Flege, J. 2001. Factors affecting degree of foreign accent in an L2: a review. *Journal of Phonetics*, 29, 191–215.
- [21] Tergujeff, E. 2013. *The English Pronunciation Teaching in Finland*. Jyväskylä Studies in Humanities 207. Jyväskylä: University of Jyväskylä. <http://urn.fi/URN:ISBN:978-951-39-5322-5>
- [22] Tergujeff, E., Ullakonoja, R., Dufva, H. 2011. Phonetics and Foreign Language Teaching in Finland. In: S. Werner, T. Kinnunen (eds) *XXVI Fonetikan päivät 2010*. Joensuu: UEF. 63–68. <http://urn.fi/URN:ISBN:978-952-61-0391-4>
- [23] Ullakonoja, R. 2011. *Da. Eto vopros!: prosodic development of Finnish students' read aloud Russian during study in Russia*. Jyväskylä: Jyväskylän yliopisto. Jyväskylä Studies in Humanities 151. <http://urn.fi/URN:ISBN:978-951-39-4209-0>.
- [24] Ullakonoja, R., H. Dufva, M. Kuronen, P. Hurme. 2014. How to imitate an unknown language? Russians imitating Finnish. In K. Jähi, L. Taimi (eds.) *XXVIII Fonetikan päivät*, Turku: University of Turku. 10–18. <http://www.utu.fi/fi/sivustot/fp2013/Sivut/Konferenssiulkaisu.aspx>
- [25] Ullakonoja, R., Kuronen, M., Hurme, P., Dufva, H. 2014. 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 Speech Prosody Conference*. 507-511. Dublin: Science Foundation Ireland; International Speech Communication Association. <http://fastnet.netsoe.ie/sp7/sp7book.pdf>