JYVÄSKYLÄ STUDIES IN HUMANITIES
151

Riikka Ullakonoja

Da. Eto Vopros! Prosodic Development of Finnish Students' Read-aloud Russian during Study in Russia

JYVÄSKYLÄN YLIOPISTO

JYVÄSKYLÄ STUDIES IN HUMANITIES 151

Riikka Ullakonoja

Da. Eto vopros! Prosodic Development of Finnish Students' Read-aloud Russian During Study in Russia

Esitetään Jyväskylän yliopiston humanistisen tiedekunnan suostumuksella
julkisesti tarkastettavaksi yliopiston Vanhassa juhlasalissa S212
helmikuun 11. päivänä 2011 kello 12.

Academic dissertation to be publicly discussed, by permission of
the Faculty of Humanities of the University of Jyväskylä,
in building Seminarium, auditorium S212, on February 11, 2011 at 12 o'clock noon.

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2011

Da. Eto vopros! Prosodic Development of Finnish Students' Read-aloud Russian During Study in Russia

JYVÄSKYLÄ STUDIES IN HUMANITIES 151

Riikka Ullakonoja

Da. Eto vopros! Prosodic Development of Finnish Students' Read-aloud Russian During Study in Russia

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2011

Editor
Hannele Dufva
Department of Languages, University of Jyväskylä
Pekka Olsbo, Sini Tuikka
Publishing Unit, University Library of Jyväskylä

Cover picture by Riikka Ullakonoja

URN:ISBN:978-951-39-4209-0
ISBN 978-951-39-4209-0 (PDF)

ISBN 978-951-39-4199-4 (nid.)
ISSN 1459-4331

Jyväskylä University Printing House, Jyväskylä 2011

Le monde est un livre dont chaque pas nous ouvre une page.
(The world is a book; each step opens a page for us.)
Alphonse de LAMARTINE, Voyage en Orient VIII

ABSTRACT

Ullakonoja, Riikka
Da. Eto vopros! Prosodic development of Finnish students' read-aloud Russian during study in Russia
Jyväskylä: University of Jyväskylä, 2011, 159 p.
(Jyväskylä Studies in Humanities
ISSN 1459-4331; 151)
ISBN 978-951-39-4209-0 (PDF), 978-951-39-4199-4 (nid.)
Finnish summary
Diss.

This study consists of six published papers (Appendices 1-6) and their summary. A part of this research has been accepted as the author's Licentiate thesis (Ullakonoja 2009). The papers focus on L2 prosody in Russian read-aloud by Finnish university students. The aim was to determine the prosodic and perceptual correlates of fluency and intonation in L2 Russian. This experimental phonetic research aimed also at examining prosodic development during the time the subjects were studying in Russia.

The phonetic data of the studies consisted of dialogues read-aloud by six to twelve Finnish university students (aged 19–24) who were studying Russian in Finland and who participated in a semester-long study-in-Russia programme. The Russian dialogues were recorded prior to, during, and following their stay in Russia. In three studies, the same dialogues read aloud by six to seven native Russian speakers' (aged 19–28) were used for comparison. An important part of the research project was to compare the acoustic analysis of the phonetic data with perceptual evaluation. Russian teachers were asked to evaluate the students' samples for fluency and native Russian speakers to evaluate them for acceptability of question intonation.

The study showed, that for the majority of the students, their read-aloud fluency in Russian developed during their semester abroad. This was seen both in the acoustic measurements of pause frequency and speech rate and in the teachers' evaluations. With respect to intonation, the study shows that even after their semester in Russia, many students still failed to pronounce Russian yes/no-questions acceptably as evaluated by native speakers. Notwithstanding, many students improved their production as a result of semester in Russia. The acoustic analysis showed that pitch placement, sharpness and height differed between the L2 and L1 Russian speakers. However, the Finnish students' mean pitch was higher and pitch range wider in Russian than in Finnish, which indicates that they were making an effort to use a native-like pitch in their L2.

Keywords: prosody, Russian language, second language acquisition.

Author's address

Riikka Ullakonoja
Centre for Applied Language Studies
P.O. Box 35
40014 University of Jyväskylä, Finland
riikka.ullakonoja@jyu.fi

Supervisors

Professor Hannele Dufva
Department of Languages
University of Jyväskylä, Finland

PhD. Viola de Silva
Department of Languages
University of Jyväskylä, Finland

Reviewers

Professor Ineke Mennen
ESRC Centre for Research on Bilingualism in Theory and Practice
Bangor University, UK

Directeur de Recherche Daniel Hirst
CNRS Laboratoire Parole et Langage
University of Provence, France

Opponents

Professor Ineke Mennen
ESRC Centre for Research on Bilingualism in Theory and Practice
Bangor University, UK

Directeur de Recherche Daniel Hirst
CNRS Laboratoire Parole et Langage
University of Provence, France

FOREWORD

This dissertation is a collection of six published articles. One potential drawback of such a dissertation is that the papers have been published during the process of becoming a researcher. Hence, a paper that was published first may indeed end up looking like it alongside various more recent papers. However, I hope that if the first papers seem rather preliminary or incomplete to the reader, they will nevertheless be seen to deserve their place in this dissertation and will in fact make it complete.

The work was partly motivated by my own experience of participating in a 4-month-study-abroad programme in Russia some 8 years ago and finding myself becoming much more confident and fluent in using Russian in such a short time. Already then I was starting to wonder about how this had happened and what lies behind a learner's fluency development. My research interest in phonetics was aroused during my work on my Master's thesis when I had the opportunity to participate in two research projects: Spontaneous Speech of Typologically Unrelated Languages (Russian, Finnish and Dutch) (funded by INTAS) and Russian and Finnish Prosody and its Effect on Segments (funded by the Academy of Finland). Some time after completing the Master's thesis I began to work on this Doctoral dissertation "as an agreeable pastime". This pastime came increasingly to mean hard work and long hours. Now it is time to thank everyone who made it less painful.

First, my biggest thank you goes to the students, teachers and staff who participated as subjects or judges in the experiments – without them this research would not have been possible.

I also want to thank the reviewers and opponents of this dissertation, Prof. Ineke Mennen and Daniel Hirst, Directeur de Recherche CNRS, for their excellent and precise comments, which have helped me to improve the work. I also wish to thank the reviewers of my Licenciate Thesis, Prof. emer. Antti Iivonen and Pekka Lintunen, PhD. for their valuable comments.

I thank my supervisors for encouraging me to do it… and to do it "my way". To Hannele, who decided not to abandon phonetics after all, thank you for always being available for discussion and for patiently and meticulously reading, rereading, correcting and recorrecting my texts. To Viola, who got me involved in phonetics already during my undergraduate studies and the INTAS project and who introduced me to many of the wonders of Russian and Russian phonetics. Viola, thank you for your down-to-earth approach. You both definitely deserve a break from this now.

I thank the Department of Languages, Faculty of Humanities and Centre for Applied Language Studies at the University of Jyväskylä for allowing me to devote some of my working time to this research while holding different posts. Thanks, too, to all my colleagues, who have commented on my work, presentations etc. formally or informally. Thank you for proofreading: Michael Freeman, Tamra Hood, Olivia Lane, Katja Mäntylä and Piia Varis. Thank you Matti Haveri-Heikkilä for technical support. Thank you to Mietta Lennes and

Hanna Anttila from the University of Helsinki for encouragement and for teaching me to use Praat. I also wish to thank my colleagues at the St. Petersburg State University, the late Prof. Liya Bondarko, Prof. Pavel Skrelin and Docent Nina Volskaya for help and encouragement. A special thanks to Nina Volskaya for her help in conducting the perception experiments of Study V. And thank you members of the DIALUKI team for showing that there is interesting research to do after this PhD. A special thank you to Lea Nieminen, for Finnish proofreading and for tolerating me in the same room while I was putting the final touches on this work.

I have received great support from several unofficial workshops. I want to thank everybody in the "underground coffee workshop" in the building P, especially Piia Varis, Hanna Kärkkäinen, Saija Peuronen, Leila Kääntä, Sanna Lehtonen, Elina Tergujeff and Saku, and in the "corridor workshops" in the same building, especially Jean-Michel Kalmbach, Sinikka Lahtinen and Katja Mäntylä and in the after-workshop outside the building, Anna-Maria Strengell.

I also acknowledge my gratitude to the several organisations and foundations that have helped fund my research trips: the Academy of Finland, Finnish Concordia Fund, Langnet (the Finnish Graduate School in Language Studies), NGSLT (the Nordic Graduate School of Language Technology) and the Otto A. Malm Foundation. I also wish to thank the University of Jyväskylä for funding my 1.5-month-research visit to the CNRS, UMR 6057 Laboratoire Parole et Langage, Université de Provence in Aix-en-Provence, France where I finally had the possibility to concentrate on this work without phones ringing, people popping into my office and classes to teach. I honestly think this could have never been finished without the Provencal sun. In addition to my stay in Provence, no matter how mundane it sounds after Provence, the Konnevesi Grant from the University of Jyväskylä truly helped me to prepare the manuscript for pre-examination in the middle of the busy spring semester.

Finally, I want to thank the person who baked me enough pulla to get through all this and who has the most amazing capacity to cheer me up and make me happy no matter what. Mika, thank you for keeping my feet on the ground and reminding me that there is after all a much more important life outside the university (which involves such things as wandering in the wilderness of Lapland). Also, I take this opportunity to thank my parents, for raising me to believe that I am able to accomplish anything I want. This has helped me to work on this dissertation from day one as I always knew that it would be accomplished one day if I really wanted it to be. Accomplished is perhaps not the right word... with my perfectionist attitude this work can never be completely finished. But I have decided to let go of it now…

I dedicate this work to my friends and host families across the miles in Australia, Russia, Belgium and France, who have helped me, in addition to appreciating their cultures, to become fluent in their languages.

In one of my favourite places in the world, Lauhalan pirtti
25 October 2010

LIST OF ORIGINAL PUBLICATIONS

This dissertation is based on six papers, that are referred to as follows:

Study I: Ullakonoja, Riikka. 2008. Pausing as an Indicator of Fluency in the Russian of Finnish Learners. In Barbosa, Plínio A. & Madureira, Sandra & Reis, César (eds.) *Proceedings of the Speech Prosody 2008 Conference. Campinas, Brazil.* São Paolo: Editora RG/CNPq. 339–342.

Study II: Ullakonoja, Riikka. 2009. Speech Rate as an Indicator of Fluency in the Russian of Finnish Learners. In O'Dell, Michael & Nieminen, Tommi (eds.) *Fonetiikan päivät 2008 – The Phonetics Symposium 2008.* Tampere Studies in Language, Translation and Culture, Series B. 97–109.

Study III: Ullakonoja, Riikka & Dufva, Hannele. 2008. Perception of L2 Fluency in Study Abroad Context. *Academic Exchange Quarterly*, Fall 2008 (12) 3. 62–66.

Study IV: Ullakonoja, Riikka. 2007. Comparison of Pitch Range in Finnish (L1) and Russian (L2). In Trouvain, Jürgen & Barry, William J. (eds.) *Proceedings of the 16th International Congress of Phonetic Sciences, 6-10 August 2007, Saarbrücken, Germany*. Saarbrücken: Universität des Saarlandes. 1701–1704.

Study V: Ullakonoja, Riikka. 2010. How Do Native Speakers of Russian Evaluate Yes/no Questions Produced by Finnish L2 Learners? *Rice Working Papers in Linguistics,* Vol. 2. 92–105.

Study VI: Ullakonoja, Riikka. 2010. Pitch Contours in Russian Yes/no Questions by Finns. In Hasegawa-Johnson, Mark & Bradlow, Ann & Cole, Jennifer & Livescu, Karen & Pierrehumbert, Janet & Shih, Chilin (eds.) *Proceedings of the Speech Prosody 2010 conference, Chicago*.

TABLES

FIGURES

ABBREVIATIONS

AR – Articulation rate
F0 – Fundamental frequency
Hz – Hertz
IK – Intonational construction (Интонационная конструкция)
L1 – First (native) language
L2 – Second language (second, third, fourth etc. foreign language learnt)
min – Minute
ms – Millisecond
PW – Phonetic word
s – Second
SA – Study abroad
SR – Speech rate
ST - Semitone
WPM – Words per minute
YNQ – Yes/no question

CONTENTS

1 INTRODUCTION

Da. Eto vopros![1] (*Да. Это вопрос!*) This could be an example of a Finn in communication with a Russian. The Finn is intending to produce an interrogative *Da?*[2], but the use of inappropriate intonation makes the utterance sound like a statement to the Russian ear. This was the starting point of the present study, which examines the development of prosody, or more precisely fluency and intonation, in Russian as L2 (second language[3]) during SA (study abroad).

Prosodic features, such as pausing and intonation are often mentioned as important components of fluency and communicative competence. If L2 users speak disfluently, listeners may become frustrated and lose interest. Also, if L2 speakers pronounce certain sounds in a non-native way or use inappropriate intonation, some words may not be understood and communication hindered. For example, intended questions may not be understood as such by the listeners. Hence, it can be argued, that in successful communication in L2 fluency and intonation play a crucial role.

First, *fluency* can be interpreted in a number of ways (see 2.1). In this dissertation the term refers to two measures of fluency: production and perceptual. Production is measured acoustically along such dimensions as pausing and speech and articulation rates (see 3.2.3). The perceptual measure in this study is teachers' auditory evaluations of fluency with respect to samples of Russian spoken as a L2 (see 3.3). Second, *intonation* can similarly be investigated in a number of ways. The present study focuses on intonation in one utterance type only: the yes/no questions (henceforth: YNQs) which are known to be difficult to produce for Finnish learners of Russian. Intonation is investigated by acoustic measurements of pitch range and pitch contours (see 3.2.4) and listener ratings of the learner's success in expressing interrogativity (see 3.4).

---

1 In English: 'Yes. That is a question'
2 In English: 'Yes'.
3 Here I use the term second language for a second, third, fourth etc. additional language learnt after one's first language(s).

Here, I use the term intonation to refer to changes in pitch and melody in speech (see 2.2).

When people spend some time abroad and speak the language of the country, they are generally considered to become more competent and/or more fluent in that language. Learning a language through staying abroad differs from classroom learning in many ways. For example, the situations of language use are authentic and there is a real need to communicate with the native speakers. Moreover, in SA contexts, the students need to use oral skills more than in an ordinary language classroom and there are also more possibilities to do so. A number of previous studies (Freed 1995; Simões 1996; Towell et al. 1996; Freed 1998; Freed et al. 2003; Collentine & Freed 2004; Freed et al. 2004; Lafford 2004; Trofimovich & Baker 2006) have found oral skills or oral fluency to improve during SA. Some studies (Walsh 1994; Harley & Hart 2002; Segalowitz & Freed 2004) have also argued that SA may be a better setting for L2 learning in general than the classroom at home. However, the present dissertation is to my knowledge the first study to focus on the development of both fluency and intonational production during SA.

Here Finnish university students' read-aloud speech is investigated, prior to, during and following a 3.5-month stay in Russia. In this study, I use the term L2 learner/speaker to refer to my subjects. All the subjects were native speakers of Finnish who had studied Russian as one of their L2s in a formal context (school/university) and participated in a 3.5-month study in Russia programme during the present study. Hence, Finnish is the L1 of the participants and Russian their L2. Finnish belongs to the Finno-Ugrian group of languages while Russian is a Slavic language. Thus, they are genetically and typologically different. They also differ in many ways in their phonological systems (and their phonetic realisations).

The Finnish sound system has 8 vowels: /i e æ y ø ɑ o u/, which can occur as either long or short vowels in all syllables; for example, *vaaraakaan* /ʋɑɑ-rɑɑ-kɑɑn/ (a partitive from the noun *vaara* 'a danger' with the particle *-kaan* 'even'). Finnish has 18 diphthongs and over 80 vowel sequences that include a syllable boundary. In Russian, on the other hand, there are only 6 vowels: /i, e, ɨ[4], ɑ, o, u/. It is a controversial issue whether Russian has diphthongs or not, but they are certainly not as common as in Finnish.

Finnish has 13 native consonants: /p t d k m n ŋ l r s ʋ j h/, and 4 that appear in loan words: /b g f ʃ/. Most consonants can occur as doubles between vowels, making a single-double contrast. (Iivonen 2009b.) Russian has 36 consonants: /b bʲ p pʲ m mʲ v vʲ f fʲ d dʲ t tʲ t͡s z zʲ s sʲ ɬ lʲ r rʲ n nʲ t͡ʃʲ ʒ ʃ ʃʲː j g gʲ k kʲ x xʲ/. However, the phonemic status of velars (g gʲ k kʲ x xʲ) and the long palatoalveolar soft sibilant /ʃʲː/ is disputable. Russian consonants can be grouped into 11 voiced/unvoiced obstruent pairs and 15 palatalized/non-palatalised consonant pairs. (Bondarko 2009). Hence, the Russian sound system

[4] Orthographically this is represented by the Cyrillic letter 'ы'.

has a voicing opposition and palatalisation opposition of consonants, which does not exist in plosives in native Finnish words[5].

On the prosodic level, Finnish word stress is always fixed on the first syllable and all the eight vowels can occur in word-initial stressed syllables and thereafter in unstressed or secondary stressed syllables (Iivonen 2009b). In Russian, on the other hand, word stress is not fixed and its position is distinctive, as in *мука́*[6] (muka) /muˈkɑ/ and *му́ка*[7] (muka) /ˈmukə/. Vowel articulation is strongly influenced by word stress both quantitatively and qualitatively and by palatalisation or non-palatalisation of the neighbouring consonants. Thus, unstressed vowels are shorter and more centralized than stressed ones. (Bondarko 2009.)

In Finnish, vowel and consonant length is a distinctive feature. This can be illustrated by examples such as *kuka*[8] /kukɑ/, *kukka*[9] /kukːɑ/and *kukkaa* /kukːɑː/ (a partitive form of *kukka)*. In Russian, length is not distinctive. There are, however, two consonants in Russian /ʃʲː/ and /ʒː/ or /ʒʲː/ that are usually analysed as long consonants. Other consonants can also occur as long consonants, e.g., on the boundary of a prefix or preposition and the stem and in loan words. For example: *поддать*[10] (poddatʼ) /pʌˈdːɑtʲ/ (prefix and stem) *из зала*[11] (iz zala) /ɪˈzːɑɬə/ (preposition and stem) *программа*[12] (programma) /prʌˈgrɑmːə/ (loan word). (de Silva et al. 2010, 104–105.)

The intonational features of Finnish, and functions of intonation are also different from those in Russian (de Silva & Ullakonoja 2009). This is especially evident in YNQs which are produced mainly by morphological means in Finnish but by prosodic means in Russian. In Russian intonation can distinguish statements from questions, whereas in Finnish this rarely occurs. These intonational differences between the two languages have been shown to influence Finns learning Russian as L2 (Lyubimova 1988; Kuosmanen & de Silva 2003; 2007). For a more detailed discussion on Finnish and Russian intonation see sections 2.2.2 and 2.2.3.

Next, I will give a general outline of the dissertation, and present the research questions.

---

5 In Finnish loan words plosives can sometimes be voiced in opposition to their unvoiced use in native Finnish words, e.g., *pussi* 'bag' – *bussi* 'bus'.
6 In English: 'flour'
7 In English: 'torment'
8 In English: 'who'
9 In English: 'a flower'
10 In English: 'to kick'
11 In English: 'from a hall'
12 In English: 'a programme'

TABLE 1 The studies comprising the present dissertation.

Study	Main Focuses	Languages	Title and reference	Type of publication
Study I	• Evaluated fluency • Pausing	Russian (L2)	R.U. 2008. Pausing as an indicator of fluency in the Russian of Finnish learners. In Barbosa, Plinio A. & Madureira, Sandra & Reis, César (eds.) *Proceedings of the Speech Prosody 2008 Conference. Campinas, Brazil.* São Paolo: Editora RG/CNPq. 339–342.	Anonymously reviewed international conference proceedings
Study II	• Speech rate • Articulation rate • Fluency	Russian (L2), Finnish (L1)	R.U. 2009. Speech rate as an indicator of fluency in the Russian of Finnish learners. In O'Dell, Michael & Nieminen, Tommi (eds.) *Fonetiikan päivät 2008 – The Phonetics Symposium 2008.* Tampere Studies in Language, Translation and Culture, Series B. 97–109.	National conference proceedings
Study III	• Normalized fluency evaluations • Self-evaluation of language skills	Russian (L2)	R.U. & H. Dufva. 2008. Perception of L2 fluency in study abroad context. *Academic Exchange Quarterly,* Fall 2008 (12) 3. 62–66.	Anonymously reviewed international journal
Study IV	• Pitch range	Russian (L1), Russian (L2), Finnish (L1)	R.U. 2007. Comparison of Pitch Range in Finnish (L1) and Russian (L2). In Trouvain, Jürgen & Barry, William J. (eds.) *Proceedings of the 16th International Congress of Phonetic Sciences, 6-10 August 2007, Saarbrücken, Germany*. Saarbrücken: Universität des Saarlandes. 1701–1704.	Anonymously reviewed international conference proceedings
Study V	• Evaluation of yes/no questions by native speakers	Russian (L2)	R.U. 2010. How do Native Speakers of Russian Evaluate Yes/no questions Produced by Finnish L2 Learners? *Rice Working Papers in Linguistics*, Vol. 2. 92–105.	Anonymously reviewed international journal
Study VI	• Pitch contours in questions • L2 performance compared to L1 evaluation	Russian (L2), Russian (L1)	R.U. 2010. Pitch Contours in Russian Yes/no Questions by Finns. In Hasegawa-Johnson, Mark, Bradlow, Ann, Cole Jennifer, Livescu, Karen, Pierrehumbert, Janet & Shih, Chilin (eds.) *Proceedings of the Speech Prosody 2010 conference, Chicago*.	Anonymously reviewed international conference proceedings

1.1 General Outline

This dissertation focuses on L2 Russian prosody in read-aloud speech. More specifically, the interest is on the fluency and intonation of Finnish university students who are learning Russian (L2). My main interest was to determine how the SA semester (3.5 months) in Russia affected the students' prosody in read-aloud speech. The material for this longitudinal research project consisted of several recordings made throughout the university studies of the subjects (see Table 7, p. 66). The purpose of the dissertation is discuss the findings overall of the six published papers (Studies I–VI in the appendices) and discuss their relevance in the prosody research field. The results of the studies along with the relevant theoretical background can be found in the papers themselves (Appendices 1–6).

Table 1 above recapitulates the studies included in this work: their foci, languages, titles and the type of publication. The first article, "Pausing as an Indicator of Fluency in the Russian of Finnish Learners" (Study I, Appendix 1), focused on the fluency development of students during their stay in Russia. More particularly, it concentrated on pausing as a temporal correlate of fluency, and on teachers' evaluations as perceptual correlates of fluency. The second article, "Speech Rate as an Indicator of Fluency in the Russian of Finnish Learners" (Study II, Appendix 2), focused on the students' speech and articulation rates, which were compared with the fluency ratings obtained in the previous study. The third article, "Perception of L2 Fluency in Study Abroad Context" (Study III, Appendix 3), summarised the results of the students' self-assessment and investigated their relationship with the fluency ratings of Study I, as well as recalculated the ratings using normalisation. The first three articles were presented as the author's Licentiate thesis (Ullakonoja 2009a).

The fourth article, "Comparison of Pitch Range in Finnish (L1) and Russian (L2)" (Study IV, Appendix 4), was published first. It was not, however, included in the Licentiate thesis as its focus was not fluency. Instead, Study IV dealt with mean pitch and pitch range in L1 and L2 and it was initially motivated by my own perceptual observation that people seemed to use a different register of voice when speaking different languages. The fifth paper, "How do Native Speakers of Russian Evaluate Yes/no questions Produced by Finnish L2 Learners?" (Study V, Appendix 5), concentrated on the successfulness of Finnish students' YNQs in Russian. There I investigated how native speakers of Russian evaluated utterances intended as questions by Finns. The last paper, "Pitch Contours in Russian Yes/no Questions by Finns" (Study VI, Appendix 6), focused on the acoustic analysis of pitch contours in YNQs in L1 (native) and L2 Russian. Furthermore, it compared the acoustic analysis to the perceptual evaluations discussed in Study V.

The outline of the dissertation is structured as follows. The first chapter is an introduction to the topic. The second chapter addresses the theoretical issues

involved and presents the key terminology and theoretical framework. The third chapter introduces the research material and methods. Finally, the fourth chapter summarises the results of the Studies (I-VI) and discusses the main findings.

1.2 Research Questions

As the present study focuses on L1 Finnish speakers' prosodic development in their L2 Russian, it stands at the cross-roads of at least four fields: 1) the study of Russian language 2) applied linguistics 3) instrumental phonetics and 4) second language acquisition. The aims of the study were, first, to investigate how pausing and speech/articulation rate function as prosodic characteristics of fluent L2 read-aloud speech; second, to detect possible improvement in the learners use of prosodic features during their SA semester using both acoustic measures and teachers' evaluations of fluency, and third, to characterize the L2 learners' use of pitch in general, as well as in Russian YNQs, in comparison with L1 speakers. The aims are met by addressing the following empirical research questions:

Fluency

1. How do Finnish L2 speakers of Russian change in their read-aloud fluency during the SA period? (Studies I and III)
2. Do the temporal/acoustic variables studied (speech and articulation rates and pausing) correspond to the fluency ratings? (Studies I and II)
3. Is there a relationship between a speaker's self-assessment and language behaviour in Russia and her fluency rating? (Study III)

Pausing

4. Are speakers perceived as more fluent in their L2 if they use fewer and shorter pauses and pause at syntactically appropriate locations? (Study I)

Speech and articulation rate

5. Are speakers evaluated as more fluent in their L2 if their speech and/or articulation rate is faster? (Study II)
6. Are speech and/or articulation rate speaker-dependent? (Study II)

Pitch range

7. Are mean pitch and pitch range different in L1 Finnish and L1 and L2 Russian for female speakers? (Study IV)
8. If so, do Finnish L2 learners of Russian develop a more native-like mean pitch and pitch range in Russian during their stay in Russia? (Study IV)

Pitch contours in YNQs

9. How do native Russian speakers perceive and judge Finnish L2 learners' utterances intended as YNQs? (Study V)
10. Are L2 learners more successful in producing YNQs perceived as questions following their stay in Russia than during it? (Study V)
11. Do the acoustic pitch measurements of L1 and L2 Russian speakers differ? (Study VI)
12. What characterises the relationship between recognition rate and intonation in L2 questions? (Study VI)

In sum, the aim of the study is to examine, 1) how acoustic measurements of pausing, speech and articulation rates contribute to fluency, 2) how the students improve in fluency and intonation during their study abroad period, and 3) how Finnish students of Russian use pitch, in particular in YNQs.

2 PROSODIC FEATURES OF FLUENCY AND INTONATION

This section introduces the relevant phonetic, mainly prosodic, research from L2 point of view, focusing on fluency and intonation. While the majority of phonetic research has recently shifted its focus from segments to prosody, in L2 phonetics the main focus continues to be on segmental phenomena, and prosodic studies are scarce. The few studies on the prosodic level have investigated such factors as rhythm, intonation, word stress, prominence, duration, fundamental frequency, pausing, speech and articulation rates, and boundaries. In many phonetic studies on L2, the main interest has been foreign accent, fluency, or teaching phonetics or pronunciation. My own focus in the present study is on prosody: pausing, speech and articulation rates, pitch range and pitch contours, and fluency.

In the L2 phonetic studies, learners' perception of L2 segments has been the main direction taken (see, e.g., Cruz-Ferreira 1989; Baker & Trofimovich 2001; Altenberg 2005; Humalajoki et al. 2006; Chen & Fon 2007; Frieda & Nozawa 2007). The theoretical justification for this approach has been the argument that perception precedes production (Lado 1961, 78; 1964, 85). However, some L2 perception studies also include prosodic phenomena. For example, Cruz-Ferreira (1989) analysed the perception of intonation patterns in L2 and Baker & Trofimovich (2001) discussed whether perception precedes production.

Currently, learners' L2 oral production seems to be studied less (see Zampini 2008, for an extensive review). Again, a lot of researchers (Flege & Hillenbrand 1986; Bohn & Flege 1990; Flege 1993; Flege et al. 1999; McAllister et al. 1999; MacKay et al. 2001) have studied segmental production in L2. However, learners' L2 prosodic production has also been studied by some. For example, Mennen and others (Mennen 1998b; 2004; 2007; Mennen et al. 2007; Chen & Fon 2008; Chen & Mennen 2008) focused on intonation and pitch range. Jilka (2000; 2007) and Holm (2008) have also studied L2 intonation, particularly from the point of view of foreign accent. Furthermore, Quené & van Delft (2010) examined durational patterns in L2 speech and their evaluation by L1 speakers.

Also, Möhle (1984) and Trofimovich et al. (2006) have studied temporal variables of L2 speech during SA. A few scholars (Wenk 1985; Tortel 2009; Tortel & Hirst 2010) have been interested in the production of rhythm in L2.

Other studies dealing with prosodic aspects of L2 speech have commonly focused on the evaluation of the foreign accent, intelligibility or comprehensibility of the L2 speaker (Flege et al. 1995; Munro & Derwing 1995; Munro 1995; Magen 1998; Munro & Derwing 1998; Piske & MacKay 1999; Guion et al. 2000; Piske et al. 2001; Derwing et al. 2006; Flege et al. 2006; Trofimovich & Baker 2006; Bent et al. 2007; Meister & Meister 2007; Aoyama et al. 2008; Munro 2008 Trofimovich et al. 2009). Most of these studies have concentrated on the age of acquisition, i.e., the age of the learners when starting to learn L2, or length of residence in an L2 country. They have attempted to establish if "a critical period" exists and if so, at what age and also what influence length of residence has. Another direction taken in the field of L2 prosody has been the contrastive approach. A number of studies (see, e.g., Lehtonen et al. 1977; Grosjean 1980b; Keijsper 1983; Nevalainen 1990) have investigated the phonology/phonetics of two languages with the aim of applying the results to L2 learning. For example, Keijsper (1983) studied Russian and Dutch intonation contrastively by applying the IPO[13] approach.

Turning now to studies with Finnish as L1, it has to be pointed out that with very few exceptions the L2 has been English and the subjects Finnish university/polytechnic students. The present study is different in this respect as the L2 of the subjects is Russian. The first studies in which Finnish was the L1 were conducted in the 1970s, when Lehtonen and colleagues (Lehtonen et al. 1977; Lehtonen 1987) attempted to find ways of automatically (and acoustically) evaluating the fluency of learners' speech. Hirvonen (1967; 1970) and Toivanen (1998; 1999; Toivanen & Waaramaa 2005; Toivanen 2006b; 2009) on the other hand, concentrated on how Finns acquire English intonation. Hirvonen (1967; 1970) took a contrastive approach to studying intonation: his works describe the Finnish and English intonational systems and include both perception and production experiments with Finnish students of English. Toivanen's (1999; 2003; 2004; 2005; 2006a; 2006b; 2007; 2009) studies provide a detailed description of English intonation and compare English intonation of Finnish students to that of L1 speakers. The most recent dissertation (Paananen-Porkka 2007) in this field concentrated on the acquisition of certain rhythmic parameters in the speech of Finnish high school students.

There are some recent studies on prosody with Finnish as L1 and Russian the L2. For example, de Silva & Shcherbakova (1998) and de Silva (1999) studied the rhythmic structure of Finnish and Russian words and the perception of word stress in Russian by Finns. Kuosmanen & de Silva (2003; 2007) investigated Finnish university students' (n = 10) question intonation in Russian, including both production and perception experiments. Their results indicated

---

13 The IPO (Institute for Perception Research) approach derives from the Dutch school of intonation description on the perceptual level (see 't Hart 1990 et al. for a thorough description).

that it was indeed very difficult for Finns to produce intonation in Russian YNQs comprehensibly, even at the advanced level, and motivated my Studies V and VI. Also Shcherbakova's (2001; 2002) studies focused on intonation. She studied Finnish students' intonation in spontaneous Russian and the perception of Russian intonation by Finns and Finnish intonation by Russians. Kärkkäinen et al. (2006; 2007) and Kärkkäinen (2009) focused on the role of the fundamental frequency in dividing speech into intonation units in Russian and Finnish, and in producing prominence in L2 Russian. Ullakonoja et al. (2007) presented preliminary results on the learning of intonational phrasing as well as pitch contours in L2 Russian.

Some studies have also been done on learning L2 Finnish prosody by Russian speakers. Ylinen (formerly Nenonen) and colleagues (Nenonen 2001a; 2001b; Nenonen et al. 2003; 2005; 2005b; Ylinen et al. 2005a; 2006) investigated L2 Finnish speakers' perception of Finnish phonological length in comparison with that of L1 Finnish speakers. One of their main results was that learning to perceive Finnish vowel quantity distinction is difficult even for Russians who have resided in Finland for years. Aho, Toivola and colleagues (Aho & Toivola 2008; Toivola et al. 2009a; 2009b) also recently studied L2 prosody in Finnish of immigrants of different L1s, including Russian.

The present study is based on methods of acoustic analysis of the speech corpus, its statistical analysis as well as two perception experiments and their analysis. My main interest is in studying pauses, speech rate and pitch by using acoustic analysis and comparing the acoustic measurements to the results from the two perception experiments (fluency evaluation and evaluation of interrogativity). Studies that combine both acoustic and perceptive fluency measurements and focus on development of fluency during the stay abroad are relatively scarce. To my knowledge this is one of the first studies to focus on the acoustic analysis of pauses, speech/articulation rate, pitch, L2 teachers' fluency evaluations and L1 speakers' evaluation of interrogativity in the SA context. A study somewhat similar to the present one was conducted by Towell et al. (1996), whose focus was on L2 learning in a SA setting, and who analysed both quantitative (speech rate, articulation rate, mean length of run and phonation time ratio) and qualitative aspects of speech. However, their study on spontaneous speech did not include perceptual ratings of fluency by teachers but was based solely on the perceptions of the researchers themselves and on the assumption that faster speech and articulation rate are "automatically" more fluent. The experimental design was in other ways very similar to the present study: there was no control group or no comparison of different learning environments such as can be found in, for example, Freed and colleagues' numerous studies (Collentine & Freed 2004; Freed et al. 2004; Lafford 2004; Segalowitz & Freed 2004; Segalowitz et al. 2004).

Furthermore, although intonation in L2 speech has been studied previously, the acoustic comparison of pitch contours with native speaker evaluations has not been the most widely used method. Instead, phonological descriptions (introduced more in detail in section 2.2.4) have often been used.

Moreover, investigating pitch range in L1 and L2, which is done here, has been a subject of only a few previous studies. Also, this is one of the first studies to investigate the L2 prosody of languages other than English or Mandarin.

All in all, previous research on L2 prosody is scarce, especially from the point of view of Finns learning Russian. Also, many previous studies have focused on the perception of prosodic features by learners of L2. Very few have focused on learners' production of L2 intonation, or attempted to compare prosodic phenomena to fluency. Also, most studies have discussed the learning of English as L2 and many have studied prosody from the point of view of foreign accent only. Hence, more research is needed, both in studying other L2s than English and investigating the influence of the SA context longitudinally.

Above, I have sketched a rough outline of the previous research on L2 prosody. Now, I turn first to fluency, which is often defined in terms of prosody. Specifically, pausing and speech rate as correlates of fluency will be explored here. Then, the research addressing pitch features, namely pitch range and pitch contours in yes/no questions, is discussed.

2.1 Fluency

Fluency is a term that is widely used among both specialists and researchers but also in every-day conversations. There is great variation in its usage. In both L2 teaching and evaluation, fluency is considered an important goal. In every-day discussions, L2 learners themselves often express a wish to become fluent in the L2. Furthermore, it is generally believed and has also been shown in some studies that when L2 learners spend some time in the L2 country, their speech becomes more fluent (Freed 1995; Towell et al. 1996; Freed et al. 2003; 2004; Segalowitz & Freed 2004). Below, I discuss the term fluency, its relevance in L2 learning and in the SA context, and give my own definition of it, as used in the present study.

There is a large and still growing body of literature where the focus is on fluency from different points of view. However, there is no agreement on the definition of fluency, and often rather vague ones are used (see, e.g., Hieke 1985; Hedge 1993; Freed & Ferguson 1995; Moore & Korpijaakko-Huuhka 1996; Cucchiarini et al. 2000; Lauranto 2005 for a review). Scholars from a number of domains, linguistics (phonology & phonetics, syntax, semantics, lexicology), psychology (absence of phonological distortion, pauses and hesitations) and sociolinguistics, have been trying to capture the notion of fluency. Some experimental research on fluency has also been carried out in the field of speech pathology and logopaedics where the aim has been to develop efficient techniques of evaluating the fluency of patients (Korpijaakko-Huuhka 1996; Moore & Korpijaakko-Huuhka 1996). In L2 acquisition studies fluency has been a more infrequent research topic in recent decades. Most of the previous studies have concentrated either on the acquisition of fluency (Segalowitz & Freed 2004; Segalowitz 2007), effect of experience on fluency (Freed 1995; Lapkin et al. 1995;

Freed et al. 2004), or such acoustic parameters of fluent speech as pausing, speech rate, articulation rate and intonation (Lehtonen 1978; 1981; Simões 1996; Paananen 1998; Wennerstrom 2000).

As Segalowitz (2010, 38) formulates it, different researchers see fluency from different stand points and earlier research has not been consistent in what it has studied as fluency:

> "What really needs to be decided is whether fluency should be considered first and foremost something that resides in the ear of the beholder (Freed, 2000) and in the mind of the listener ("an impression on the listener's part"; Lennon, 1990, p. 391) or, on the contrary, whether fluency refers first and foremost to particular characteristics of oral production, regardless of how that production is actually perceived and judged by listeners."

Riggenbach (1991, 439) has pondered upon the difficulty of defining fluency in the following way:

> "We might speculate that fluent speakers resemble each other, but there may be a number of ways to identify nonfluent speakers… In order for there to be fluency, then, it appears that many different conditions have to be met – some proficiency in grammar, pronunciation, and vocabulary, to mention a few. […] Nonfluency, on the other hand, can arise from a deficiency in any one of these areas: the inability to produce a given grammatical structure may be the first link in a chain of disfluencies that may as easily have begun with a comprehension lapse, a pronunciation problem, or a motivation for precision in word choice."

According to Riggenbach (1991, 432), the *chunking together* of disfluencies (several disfluencies in a three-word sequence) may be a specially important indicator of fluency. *Clusters of disfluency* have also been mentioned as one of the important correlates of fluency by Freed et al. (2003). In Study I, I have called these *disfluency clusters*. There I studied repetitions, repairs and disfluent pauses, and calculated the number of disfluency clusters, i.e., places where there were at least two disfluencies within a three-word sequence.

Lennon (1990, 389-391; 2000, 25–26) distinguishes two uses of the term fluency: in its broad sense it means nearly the same as oral proficiency, whereas the narrow sense is often used to refer only to a part of oral proficiency, such as, correctness and native-like rapidity. As Lennon (2000, 26, 40) explains, fluency is

> "the rapid, smooth, accurate, lucid, and efficient translation of thought or communicative intention into language under the temporal constraints of online processing. […] In principle, then, performance may be fluent but erroneous. In practice, however, error will often be associated with uncertainty on the speaker's part, which will adversely affect fluency."

This corresponds exactly to my understanding of the term fluency: also fluent speech can contain errors.

Lehtonen's (1981, 331) definition of fluency is a combination of communicative acceptability and smooth continuation of speech:

> "To be fluent in the right way, one has to know how to hesitate, how to be silent, how to self-correct, how to interrupt, and how to complete one's expression. According to this definition of fluency, one must speak in a way that is expected by the linguistic community and that represents normal, acceptable and relaxed linguistic behaviour."

Here fluency is looked at from three different points of view: 1. linguistic acceptability, 2. smooth continuation of speech, and 3. communicative acceptability. Therefore, according to Lehtonen (1977), the speaker's fluency depends also on the communicative situation and the (spoken or written) text, not only on the features of his/her speech. However, as I investigate read-aloud data, Lehtonen's *linguistic acceptability* or *communicative competence* cannot be taken into account when evaluating fluency. Therefore, the only criterion I find suitable from his argumentation is the *smooth continuation of speech*.

Segalowitz (2010, 48–50, 163–165) distinguishes between cognitive fluency, utterance fluency and perceived fluency. He argues that cognitive fluency is the underlying mechanism that contributes to utterance fluency (i.e., oral fluency). Using that term, it is possible to make a distinction between one's knowledge and the implementation of that knowledge. Cognitive fluency includes parameters such as speed and efficiency of lexical access, attention control, working memory etc. Utterance fluency includes speech rate, pausing and hesitation phenomena. Perceived fluency, on the other hand, means the listener's inferences about the speaker's fluency. Segalowitz (2010: 163–165) broadens the concept of fluency into motivation-, context- and experience-related issues that all influence L2 speech and hence also fluency.

The most common parameters that have been used to define and evaluate L2 oral fluency include calculating and measuring the number of pauses (overall or per minute), their place and duration, syllable duration, hesitation phenomena, linking, rhythm, mean length of run[14], speech rate, articulation rate, phonation-time ratio, phonological grouping and intonational features (Sajavaara & Lehtonen 1980, 69; Hieke 1981; 1984, 352; Riggenbach 1991; Walsh 1994; Moore & Korpijaakko-Huuhka 1996; Perales & Cenoz 1996, 82; Towell et al. 1996; Cucchiarini et al. 2000; Temple 2000; Riggenbach 2001, 253; Cucchiarini et al. 2002; Freed et al. 2004; Segalowitz & Freed 2004; Kormos 2006, 162–164; Trofimovich & Baker 2006; Paananen-Porkka 2007). Hence, from the phonetic point of view, fluency is all about prosody (see, e.g., Cutler 1983). For some scholars, such as Segalowitz (2007) and Raupach (1980), fluency means reading at an appropriate rate, without too many hesitations and with a small number of relatively short pauses. In my view, all of these measures reveal something about non-struggle with the language, and thus pausing, speech and articulation rate are the factors that will be seen as the parameters of fluency in the present study.

Segalowitz (1986, 4) refers to oral fluency as "rapid and accurate ability to use the vocabulary and syntax of the second language", being "generally skilled

---

14 Mean length of run is the average duration of a continuous sequence of speech not interrupted by pauses.

at reading the second language" and doing so with a habitual speech rate. Lehtonen (1977, 22; 1978) points out, that speech with only few pauses is not necessarily always perceived as fluent. In fact, when the learner's speech is too fast and there are only few pauses, it can be incomprehensible (Lehtonen 1979, 35). As stated by Lehtonen (1981, 331), "There is no single "normal" speech rate, nor a "correct" number of pauses typical of fluent speech". However, I believe that it is possible to define an acceptable variation in speech rate or number of pauses in a certain speaking situation as criteria for fluent speech.

What do people then *perceive* as fluent speech? Or in Segalowitz's (2010, 163–165) terminology: what influences perceived fluency? This is how, for example, Freed (2000) measures fluency. People tend to listen to other factors than just pauses and speech rate when trying to decide whether their interlocutor's speech is fluent or not. These factors include, e.g., vocabulary size, grammar, accent, speech rhythm, confidence in speaking, voice quality and "tone of voice". (Freed 1995, 143.) The features of speech that were mentioned in my study as disfluent by teachers involve similar criteria (see Ullakonoja 2009a, 40–41 and Study I). However, we might think that each listener has different criteria according to which s/he judges fluency and that it is not possible to empirically measure fluency at all or to say omnisciently whether the speech of a speaker is fluent or not. However, some studies (Cucchiarini et al. 2000; Derwing et al. 2004) have shown that different groups of judges (phoneticians, teachers, untrained listeners and speech therapists) have rated the fluency of speech samples fairly similarly and the inter-judge reliability of these ratings has been good. Hence, people, especially experts, seem to have similar criteria for what is fluent.

There have been some studies on fluency in relation to length of residence in the L2 country. The setting in these studies is either immigration to the L2 country, or L2 students spending some time in the L2 country (SA context). The studies examining subjects who have immigrated to the L2 country have been interested in the impact of age of arrival on their L2 speech (e.g., Guion et al. 2000). The studies (Walsh 1994; Freed 1995; Simões 1996; Towell et al. 1996; Freed 1998; Freed et al. 2003; Collentine & Freed 2004; Freed et al. 2004; Lafford 2004; Segalowitz & Freed 2004; Trofimovich & Baker 2006) concentrating on the *SA context* on the other hand have yielded the result one might expect: the L2 context is advantageous to learners in improving their oral skills and their becoming more fluent. These studies have investigated longer SA periods, from a semester onwards, but statistically significant fluency improvement has been found also during a shorter stay (3–4 weeks) (Llanes & Muñoz 2009). The research covers different countries as SA contexts. Sometimes a similar context can be created in the home country, for example, when L1 English students are studying in French-speaking Canada.

A positive effect of SA on fluency has been shown, e.g., by Segalowitz & Freed (2004), Valls-Ferrer (2008) and Serrano et al. (2010). Segalowitz & Freed (2004) studied English adults learning Spanish in at home and SA contexts and found that the latter seems to help learners to improve their fluency in

spontaneous speech significantly on several measures, in particular in speech rate, mean length of run without filled pauses, and longest speech run without silent and filled pauses. To conclude, they claim that despite the many communication possibilities during the SA semester, not all the learners always improved their oral performance. Valls-Ferrer (2008, 70–71) showed that during SA, there was significant improvement in all fluency measures (including speech and articulation rates, mean length of run, phonation time ration, pause frequency, pause duration and number of disfluencies per minute). Serrano et al. (2010) on the other hand, concluded that in oral production there are significant gains in fluency even during a semester abroad.

Also, other studies have suggested that SA is not an equally beneficial learning context for all learners. For example, Simões (1996) found that even though SA helps learners in general to improve their fluency significantly, there are great interspeaker differences in using and benefiting from the opportunities afforded by SA. The pre-existing language skills of the learner seem to have an effect on the benefits of SA. For example Freed (1995, 135) and Llanes & Muños (2009) found that weaker students ameliorated their fluency more during SA than more proficient students. As Wilkinson (1998a; 1998b) argues, not all students benefit from the SA context as much as one would think a priori. For different reasons, they do not use all the opportunities available to them to use the L2, and they are in a way left out of the L2 speech community.

Furthermore, it is controversial whether the SA context is the best context for fluency development. In a study by Freed et al. (2004), which compared English students of French in three different contexts (at home, immersion and SA), it was found that the students in intensive domestic immersion gained most in terms of fluency (when fluency was understood as smooth, fast and continuous speech). Furthermore, the research indicated that the students in the SA context reported using less out-of-class time on L2 than those in the domestic immersion. However, it has to be pointed out that the students who studied abroad also improved their oral fluency compared to the regular at-home group.

Of course, other factors influence fluency development than merely residence abroad. For example, the activities of the learner in the host country as well as his/her ability to acquire the L2 and the age of onset of learning the L2 in question have been shown to have an effect on fluency development (Flege et al. 1995). Finally, Freed (1998, 50) outlines language proficiency development during SA as follows: "Those who have been abroad appear to speak with greater ease and confidence, expressed in part by a greater abundance of speech, spoken at a faster rate and characterized by fewer dysfluent-sounding pauses." This definition is consistent with the concepts of fluency I have adopted in this study, except that my focus is on read-aloud speech rather than spontaneous interaction.

A large-scale study on Russia as a SA context has been conducted by Brecht and colleagues (Brecht & Robinson 1995; Brecht et al. 1995). A total of 658 major students of Russian in an American university were tested before and

after their 4-month stay either in St. Petersburg (Leningrad) or Moscow over a period of six years (1984–1990). The great number of participants makes reliable statistical analysis possible, but the large-scale questionnaires are interesting also from the qualitative point of view. The tests done before and after the stay in Russia consisted of speaking, listening, reading, personal data and learning variables[15]. Their main results showed that the higher the pre-test score in listening and reading, the less the gain, i.e., the most advanced students were not able to benefit as much from the SA context. Also, the students who had learnt Russian already in high school gained less than those who had not. However, reading proficiency before SA was strongly associated with the improvement in listening skills during SA. In the test about 13 % of the students received "advanced" in the pre-test whereas almost 40 % did so in the post-test. This was seen as an indicator of improvement in their functional level of competence. Brecht et al. (Brecht & Robinson 1995; Brecht et al. 1995) concluded that in the SA context men improved their listening skills more than women, younger speakers more than older, students who had learnt other L2s more than those who had not, and people who had been to the L2 country before more than those who had not. On the basis of these studies, it can be concluded that the SA context is beneficial for learning different L2 skills.

The largest and newest contributions focusing on Russia as a SA context is, however, without doubt those of Davidson (2007; 2010). They are the most comprehensive of all the studies concerning linguistic development during SA, as they investigated a total of 1 881 American students who studied in Russia for 2, 4 or 9 months. In many aspects Davidson's results resemble those of Brecht et al. (Brecht & Robinson 1995; Brecht et al. 1995); for example, the initial proficiency level of the student, as well as studying Russian already in high-school, affected the amount of gain during SA. Unlike Brecht et al., Davidson (2010) compared students who resided in Russia for different periods of time. He showed, for instance, that students gained more in oral skills, the longer they stayed in Russia (a semester vs. a year).

In their extensive studies, Coleman (1998) and Freed (1995; 1998; 2004) have encapsulated the research on the SA context. This research has expanded into various areas including sociology, psychology and the educational sciences, as well as linguistics. SA has been found to have positive effects on most areas of L2 competence: listening comprehension, vocabulary recognition skills, vocabulary production skills, oral communication skills, sociolinguistic competence and communicative competence in general (Harjula & Manninen 1994; Huhta 1994; Marriott 1995; Reagan 1995; Lennon 2000; Harley & Hart 2002; Isabelli-García 2003; Segalowitz et al. 2004). Also studies (e.g., Walsh 1994) measuring language proficiency and comparing SA and at-home students have

---

15 Brecht and other American researchers (e.g., Freed et al. 2004; Lafford 2004; Trofimovich & Baker 2006), who have studied language development in the SA context, have based their studies on OPI (Oral Proficiency Interview) results or data collected in OPI. The OPI is a standardised test used in the USA to evaluate the overall speaking proficiency of the speaker (Language Testing International 2004) and is available for many languages.

found that students sojourning abroad are prone to attain higher levels of L2 proficiency than those staying at home.

Most of the above-mentioned studies have focused on fluency in spontaneous speech. In read-aloud speech fluency is a somewhat different phenomenon (Nuttall 1982, 2–18, 23). Grabe (1991, 378) defines fluent reading as follows: "fluent reading is rapid; the reader needs to maintain the flow of information at a sufficient rate to make connections and inferences vital to comprehension". Continuing with the particularities of read-aloud speech Adams (1979, 131) and Koponen (1992, 134-136) argue that the reasons for disfluencies in learner's speech include the reader's unfamiliarity with the language's lexical items, typically with long and low-frequency words. Thus, one reason for pausing in L2 reading aloud may be unfamiliarity with the lexical items of the text (Guion et al. 2000, 209). Similarly, Lehtonen & Heikkinen (1981, 329–336) showed that disfluency in reading aloud in one's L1 can be caused by a single lexical item (e.g., foreign word), and that readers tend to have increased pause duration in a text containing multiple foreign words in comparison with other assumedly more familiar texts. They also found that the longer the word, the more disfluencies it created.

In sum, most previous studies about fluency in the SA context have found an increase in the students' fluency during the time spent in an L2 country. There are also several dimensions to the term fluency. On one hand fluency may refer to the accuracy of grammar or pronunciation, while on the other hand it may indicate the speed of delivery. Hence, the common features most definitions of fluency share are perceived ease of articulation and appropriate rapidity. Many of the definitions also underline the absence or at least scarceness of hesitation phenomena in fluent speech. Although some scholars (e.g., Stahl & Heubach 2006, 190) define fluent reading both as fast and phonetically accurate, as pointed out above, I have not taken phonemic accuracy into account because my focus is on the prosodic characteristics of fluent speech. For example, a learner might have problems pronouncing all the segments of the language correctly or using appropriate intonation, but still be perceived as fluent (see, e.g., Hammerly 1991). Here fluent speech means reading aloud smoothly and at an appropriate rate and with pauses in the correct places (see, e.g., Lennon 1990; 2000).

2.1.1 Pausing

In the present study pausing was considered as an important element of fluency. This was the topic of Study I. As many studies have shown, perceived fluency is affected by the number of pauses, their place and duration (Riggenbach 1991; Walsh 1994; Riggenbach 2001, 253–256). From the point of view of this study, the most important pause classification is that of fluent and disfluent pauses (Riggenbach 1991, 426–427; Perales & Cenoz 1996, 79; Segalowitz & Freed 2004) which is based on the place (position) of the pause. Riggenbach (1991, 426–427) uses the terms fluent-sounding pause and disfluent-sounding pause. Fluent-sounding pauses are those that occur at "predictable places" at clause or phrase

boundaries, and disfluent-sounding those occurring elsewhere. I have categorized pauses on the basis of how they sound, i.e., how they are perceived (fluent or disfluent). Perales & Cenoz (1996, 79) defined fluent and disfluent pauses in spontaneous speech as follows:

> "Fluent pauses correspond to breathing and planning pauses which mainly occur at grammatical junctures and are, therefore, natural and expected. Disfluent pauses are those which are not natural in Basque [here L2] and can be either the result of transfer from the first language [Spanish] or part of the learner's specific interlanguage."

Table 2 recapitulates different terms for pauses that in my opinion support my adoption of the classification of pauses into fluent and disfluent pauses. I shall next introduce each of the definitions in more detail.

Fluent pause		Disfluent pause
juncture pause	◄►	non-juncture pause
syntactic pause	◄►	non-syntactic pause
grammatical pause	◄►	non-grammatical pause
functional pause	◄►	hesitation pause

TABLE 2 Different terms corresponding to fluent and disfluent pauses.

Kenny (1996, 36–38) applies the term juncture pauses for fluent pauses, meaning pauses that mark syntactic boundaries and term non-juncture pauses for disfluent pauses indicating hesitation and abnormality. The justification for using this classification is that syntactic boundary pauses are perceived as more adequate by native listeners (Butcher 1980). Drommel (1980) uses the terms syntactic (positioned at syntactic/constituent boundaries) or non-syntactic (within noun or verb phrases) pauses. Another justification for the categorisation is that syntactically in L1 a majority of pauses are placed at clause or sentence boundaries independent of the reading rate or type of text spoken (Hawkins 1971; Lane & Grosjean 1973; Grosjean & Deschamps 1975; Grosjean 1980b, 44; Riazantseva 2001).

Goldman Eisler (1968, 13–14) defines pause in a similar manner as in the above-mentioned studies: grammatical pauses are those occurring at grammatical junctures and those that are semantically motivated. Non-grammatical pauses, on the other hand, are pauses occurring, e.g., in the phrase medial or final position, before repetition or a false start. In Strangert's (1991) study, paragraph and sentence boundaries were found to be almost obligatory places for a pause whereas pausing at clause or phrasal boundaries depended on speech rate and the length and complexity of the clause. These findings also support the pause classification adopted in this study. The term disfluent pause coincides with the term non-grammatical pause whereas fluent pause is a grammatical pause and occurs often at clause and sentence boundaries. Also Deese's (1980, 72–75) categorisation of pauses into *functional* and *hesitation pauses* accords well with the categorisation of fluent and disfluent pauses.

Functional pauses (meaning other than hesitation pauses) are defined as pauses with a grammatical function.

Next, I shall briefly look at other pause definitions that I do not consider completely coincide with my fluent-disfluent classification. Table 3 below lists the different approaches to pause classification and introduces the terminology used. These I will now describe in more detail.

Basis of the classification	Pause types
Intensity of airflow	silent pause, low voice pause, filled pause
Reason for pausing	intentional pause, unintentional pause hesitation pause pause for repair or repetition breathing pause, non-breathing pause emotional pause emphasis pause
Perception/ production	speaker relevant pause, communication relevant pause, hearer relevant pause auditory pause, acoustic pause perceived pause, physical pause psychological pause = zero pause = virtual pause
Place of the pause	intersegmental pause, intrasegmental pause syllable pause, word pause, constituent pause

TABLE 3 Other approaches to pause classification.

First, a traditional pause categorisation is that between *silent* and *filled* pauses. A silent pause is a silent interval in speech whereas filled pauses have been associated with hesitations and involve some sound (elongated vowel, laugher, coughing, etc.). Ballmer (1980) provides a thorough threefold classification of pauses. First, pauses can be defined by the intensity of airflow into empty (silent) pauses, *low voice pauses* and filled pauses. Most other researchers do not distinguish between low voice pauses and filled pauses. Second, he classifies pauses according to their controllability into *unintentional* and *intentional pauses*. Third, he characterises pauses by the concern of the interlocutors into *speaker relevant, communicative* and *hearer relevant pauses*.

Second, pauses have been classified according to the possible reasons for pausing. One obvious reason for pausing is the need to inhale, and hence *breathing pause* (or respiratory pause) and *non-breathing pause* have been used to categorise pauses (Grosjean 1980a; Vaissière 1983). Other commonly mentioned reasons for pausing are *hesitation, repair, reformulation*, gaining time for planning or finding suitable words (see, e.g., Paananen-Porkka 2007, 271). Zinder (1979, 277), in his summary of the different functions of a pause, talks also about the *emotional* function of a pause: by pausing at a particular place a speaker can express emotions, e.g., surprise. In Russian it is also possible to pause in the middle of a word when the speaker wants to *emphasise* a certain syllable or articulate very clearly (Nikolayeva 1977, 15; Zinder 1979, 277).

Third, pauses can also be looked at either from the acoustic (physical) or auditory (perceptive) point of view. According to Zinder (1979, 277), an *acoustic pause* is a silence in the sound signal, whereas from the physiological point of view, an *auditory pause* is a break in articulation. However, Zinder (1979, 277) points out that when perceiving pauses neither one of these characteristics have to be met: in the Russian research tradition a pause that can be perceived but that is not identifiable acoustically is called a *psychological pause*. Volskaya (2002; 2004; 2009a) and Skrelin & Volskaya (2004) continue that psychological pauses are below 200 ms in duration and are not perceived only by temporal cues, but also with the help of other prosodic (e.g., tonal) means. They also refer to them as *zero* or *virtual pauses* because there necessarily does not need to be a silent interval in speech, but listeners interpret, e.g., intonation unit boundaries as pauses. This argumentation corresponds very well with my understanding of a pause. As described above, I have defined pauses into fluent and disfluent ones based on perception, which in turn is affected by the place of the pause. Also Strangert (1990; 1991) has studied perceived pauses instead of physical ones and showed that even very short pauses (1–200 ms) can be perceived as pauses and in fact they constitute about 7–26% of all pauses.

Fourth, pauses can be defined depending on the place of the pause. For instance, Drommel (1980) states that acoustic pauses can be *intrasegmental* (e.g., in the middle of a plosive) or *intersegmental*, only the latter being auditory or audible. Furthermore, Pilon (1981) has identified three pause types according to their place in the sentence: 1) *constituent pauses*, 2) *word pauses* and 3) *syllable pauses*. Constituent pauses are situated at the constituent boundary, word pauses at the word boundary (that is not a constituent boundary), and syllable pauses in the middle of a word at the syllable boundary. Herman's (1985) study also implies a similar distinction.

However, quite often the only criterion for defining a pause has been the specific duration of a silent interval in speech. Although researchers have used different durational thresholds of silence (1–400 ms) for defining a pause, a commonly used one has been 200–250 ms (Grosjean & Deschamps 1975; Lehtonen 1979; Lennon 1984; Moore 1990; Cenoz 2000; Guion et al. 2000; Volskaya 2002). This definition has been justified by the possibility of automatically detecting pauses without regarding, e.g., the closure phase of plosives and other silent intervals belonging to articulation as pauses. Also longer and shorter pause thresholds than 200–250 ms have been applied. For instance, Derwing (2004) defined a pause threshold as 400 ms, Raupach (1980) used a cut-off point of 300 ms, while Paananen (1998; Paananen-Porkka 2007), Riazantseva (2001) and Trofimovich et al. (2006) set the limit at 100 ms. Furthermore, Adams (1979) defined a pause threshold as only 50 ms of silence, because in her study that was the shortest silent pause duration used by L1 speakers at the phrasal boundary. In his dissertation Kendall (2009, 104–105) settled on 60 ms and was still able to measure pauses automatically.

Next, I shall briefly look at studies on *pause duration* itself to understand the factors influencing it, as it was also measured in Study I. One could assume

that the most important factor influencing pause duration is speech rate. However, the picture is not that simple. Grosjean & Lane (1974) argue that when speakers modify their reading rate they mainly alter the pause frequency, i.e., insert or delete pauses at strategic syntactic places, and do not change their pause duration or articulation rate. Furthermore, Strangert (1991) claims that pauses are more frequent the slower the speech. Hence, in slow speech pauses are not necessarily longer than in fast speech, but pause frequency is higher. Also Toivola et al. (2009b) showed that pause frequency does not influence pause duration. One of the factors influencing pause duration is the place of the pause (Fant et al. 2003). Pauses at the end of sentences are generally longer than at other locations (Grosjean 1980a). In addition, Strangert (1991) found that pause duration varied among speakers in L1 Swedish, but still followed a regular pattern depending on the place of the pause. Pauses at paragraph boundaries were the longest, while at sentence boundaries they were about 60 % and at clause boundaries on average about 20 % of the mean pause duration at paragraph boundaries. Also in Volskaya's (2003) study pauses were longest at the end of a paragraph, and most sentence boundaries were marked by a silent pause. Clause and phrasal boundaries were mostly marked by virtual (perceived) pauses[16].

Next, studies focusing especially on *pausing in L2* will be summarised. Researchers seem to agree that extensive pausing is typical of non-native speech and that pauses occur in connection with hesitation phenomena such as repetition or repair (see, e.g., Raupach 1980; Riggenbach 1991; Cenoz 2000; Guion et al. 2000; Paananen-Porkka 2007). A commonly obtained result is that pausing differs between L1 and L2 speakers (see, e.g., Hieke 1987, 52–53; Temple 2000). Previous L2 studies have mainly discussed the relationship between pausing and fluency and between pausing and L2 proficiency. For example, fluent speakers have been found to hesitate less and hence also to produce fewer pauses than disfluent speakers (Riggenbach 1991). Similarly it has been found that as L2 proficiency increases, pausing becomes more native-like (Cenoz 2000; Riazantseva 2001). Also, according Trofimovich et al. (2006, 17–19), the age of starting to learn L2 has a greater impact on pause frequency and duration (the younger started, the more fluent the speech) than, for example, L2 experience (residence in an L2 speaking country). This would mean that for learning the appropriate pausing in an L2, the earlier you start learning that L2, the better.

However, the frequency of pauses cannot always be looked at as an indicator of disfluency, but can also be interpreted as transfer from L1 (see, e.g., Lehtonen 1981). For example, Olynyk et al. (1987) and Raupach (1980) found that learners pause similarly in their L1 as in their L2. Some researchers (e.g., Grosjean 1980b; Paananen 1998; Riazantseva 2001) have also suggested that pausing is culturally determined. Perales & Cenoz (1996, 75) suggest that, in every language, pausing behaviour in spontaneous conversation is determined by turn-taking strategies and the function of silence in the corresponding

16 See definition above on p. 32.

culture. Lehtonen and Sajavaara (Lehtonen 1979; Sajavaara & Lehtonen 1980) found that Finnish learners of English pause more often, for a longer period of time and at different (incorrect) places than Swedish-speaking Finns and Swedes learning English or native English speakers in spontaneous speech. It is possible that when hesitating, Finns tend to use unfilled pauses, whereas Swedes and Swedish-speaking Finns use filled pauses. (Lehtonen 1981, 325).

Paananen (1998; Paananen-Porkka 2007, 234–239, 246–253) did not find any significant differences between Finnish pupils of English and English L1 speakers in pause duration or the percentage of pauses out of speaking time. However, consistent with Lehtonen's (1979) study she found that Finnish pupils pause more often and at different (incorrect) places than L1 speakers, and also that pauses were longer in L2 English than L1 Finnish. Adams (1979, 22) found that L2 speakers paused more and for longer than L1 speakers, and that their pauses also occurred in erroneous places. This was consistent with the findings of Paananen-Porkka (2007). Learners, for example, failed to respect the constituent structure of the utterance, which resulted in grouping lexical items inappropriately and disturbing the rhythmical pattern of the sentence.

There is however, at least one controversial study on L2 pausing, which indicates that L1 speakers can have longer pauses than L2 speakers. Toivola et al. (2009b) demonstrated that L1 Finnish speakers had on average greater pause duration than L2 speakers of Finnish. The difference was statistically significant between L1 Finnish speakers and L2 speakers, whose L1 was Russian. Furthermore, L1 speakers had a lot of variation in their pause duration, whereas the L2 speakers showed less variability. This seems to support earlier findings, where pause duration has been claimed to be culturally determined. As Tannen (1985, 109) puts it: "A pause becomes a silence, and a silence is negatively valued, when it is too long or appears at what seems like the wrong time and the wrong place".

This subsection introduced the different terminologies used to characterise pauses. It also presented the pause terminology used in this study and introduced references supporting it. Furthermore, it showed that some researchers simply rely on a durational silence threshold for determining a pause without further classifying pauses, whereas others have developed different functional categorisations. To summarise, pauses occur in speech for many different reasons: planning, hesitation, repair, reformulation, breathing, emphasizing etc. Pause frequency is affected by the speech rate. The temporal characteristics of a pause were shown to depend on the speaker, text, language and place of the pause. Last, it was pointed out that in L2, pausing and fluency are linked and that L2 pausing differs from L1 pausing.

2.1.2 Speech and Articulation Rate

Apart from pausing, other important acoustic correlates of fluency are speech and articulation rates, which were investigated in Study II. In this subsection, I shall discuss the factors influencing speech and articulation rate and look at studies focusing on them in the L2 and SA contexts.

Speech rate (tempo) is a term used to indicate the number of units per total time a speaker is uttering his/her speech, including pauses, whereas the term articulation rate refers to the speech rate excluding pauses (Grosjean 1980b; Kenny 1996, 50; Tsao et al. 2006; Paananen-Porkka 2007, 123). Sometimes the term phonation rate is used synonymously with articulation rate (see, e.g., Moore 1990). In the present study the terms speech and articulation rate are used in the context of reading aloud and defined in terms of reading rate. Speech and articulation rate are defined as the "number of output units per unit of time", e.g., sounds/second, syllables/second or words/minute (Kodzasov & Krivnova 2001, 72; Tsao et al. 2006, 1156).

Hence, there is a relationship between speech rate and pauses. For example, Goldman Eisler (1968, 24) defines the speech rate and articulation rate as follows: "The longer and more frequent the pauses, the slower is the total rate of speech production. […] The articulation rate (AR) on the other hand, plays no significant part in the rate at which speed is produced over a period of time (SR)." According to Ivanova-Lyukyanova (2003, 142–143), speech rate tells us about the rate at which the speaker pronounces words, and it depends greatly on the frequency and type of pauses. One might think that it is the frequency of pauses that most influences speech rate when comparing samples of the same text produced by different speakers. However, as Crystal & House (1990, 106) argue, in addition to the higher frequency of pauses, slow speakers also need more time to utter each syllable.

Multiple factors are known to affect the speech rate of a speaker (see Trouvain 2004 for a review). Individuals are able to vary their speaking rate in different situations (Goldman Eisler 1968, 19; Trouvain 2004), from time to time (Abercrombie 1967, 96), in different text styles (Grosjean 1980b), or in different parts of the sentence (Deese 1980, 74–76). For example, before a hesitation pause L2 learners have been found to slow down their speech rate, and then make it faster after the pause (Shcherbakova 2002, 272). Speech rate is also affected by word length and word frequency (Perfetti 1985, 15). Furthermore, it has been found that clause type influences speech rate so that, e.g., declarative questions[17] are spoken more rapidly than corresponding statements (van Heuven & van Zanten 2005) and questions faster than statements (Nikolayeva 1977, 84–85). Also, the length of an utterance influences speech rate variation (Goldman Eisler 1968, 19–23; Kendall 2009, 149–152). The type of information conveyed by the phrase can also influence speech rate: the parts the speaker thinks are more important are spoken at a slower rate than, e.g., parts offering some specifying information (Nikolayeva 1977, 15). Furthermore, Paananen-Porkka (2007) and Kendall (2009, 140–142) found evidence of influence of on speech rate: in their studies women spoke faster than men.

Furthermore, the type or genre of the text is known to affect speech rate so that the more linguistically complex the text, the slower is its reading aloud. Also, for example, jokes are spoken faster than fairy tales (Sallinen-Kuparinen

---

17 Van Heuven & van Zanten (2005) use the term "declarative question" to refer to questions that fully correspond to a statement in their morphology and syntax.

1979). Also speech rate in spontaneous speech has been found to be faster than in reading aloud (Lennes 2009). Individuals have been found to differ in their habitual speaking rate in L1: some are slow speakers on a biological basis while some are fast (Tsao et al. 2006). Interestingly, it has also been suggested that the personality and ethnicity of a speaker might influence his/her speech rate (Crown & Feldstein 1985). In fact, Kendall (2009, 143) found statistically significant differences in the speech rate of speakers from different regional and ethnic backgrounds.

Speech rate in L1 Finnish has been under scrutiny in some studies (e.g., Lehtonen 1979; Sallinen-Kuparinen 1979; Iivonen et al. 1995; Moore & Korpijaakko-Huuhka 1996; Suomi 2007). In Russian, the focus recently has been on the difference between read-aloud and spontaneous speech, and, on pausing and its influence on prosodic phrasing and speech rate (see, e.g., Volskaya 2009a). The results of these studies will, however, be discussed later in more detail in section 4 together with my results.

It is not surprising that a number of studies have shown a tendency for *L2 speakers* to speak at a slower speech rate than native speakers (Riggenbach 1991; Munro 1995; Munro & Derwing 1998; Cenoz 2000; Guion et al. 2000; Trofimovich & Baker 2006; Paananen-Porkka 2007; Toivola et al. 2009b). Furthermore, the same speakers have been found to speak significantly slower or much slower in their L2 than L1 (Raupach 1980; Möhle 1984). However, comparison of the speech rate of the same speakers in different languages is often, as Lehtonen (1981) has shown, problematic. For example, the comparison of Finnish and English is difficult because if the measuring unit syllables/minute is used, then Finnish is spoken faster than English, but if words/minute is used, English is spoken faster[18]. It is also possible that when speakers become more proficient or fluent in L2, their spontaneous speech becomes in fact slower, because their ability to monitor their speech develops (Segalowitz & Freed 2004, 195).

When looking at speech rate and articulation rate from the point of view of the L2 and *SA contexts*, several studies (Möhle 1984; Lennon 1990; Freed 1995; Towell et al. 1996; Segalowitz & Freed 2004) have found that L2 learners speak faster after their SA than before it (exact values found in these studies are shown in Table 15, p. 92). For example, Möhle (1984) found that German university students who were studying French as an L2 and spent a semester in France increased their speech rate and articulation rate considerably during their stay. In her study, however, L1 French speakers who were studying German and spending a semester in Germany did not show such a great change; only their articulation rate increased a little. Möhle argues that the increase in the speech rate may be due to enlargement of the vocabulary of the students. However, in a more recent study by Freed et al. (2004) the students in the immersion context increased their speech rate more than their colleagues in the SA context.

---

18 Finnish has less complex syllables than English, but longer and more complex words than English because of the morphological differences between the languages.

As the empirical data of the present study comes from read-aloud speech the factors influencing *reading rate* are also relevant. Hence, before moving on to describing pitch, I will briefly discuss reading. There is not a lot of research on oral reading rates in L2. However, it is clear that cognitively reading in L2 is a complex process that can be further complicated by the different orthographies or alphabets of the L1 and L2. Also, the text can be "hard to read" if, e.g., it contains a great number of foreign, loan or low-frequency words, has unusual syntax, or if it is not coherent (see, e.g., Glushko 1981; Lehtonen & Heikkinen 1981, 328–329; Akamatsu 2005, Grabe 2009: 289–293).

Individuals differ in their reading skill: people read at different rates, can interpret the meaning of a text in different ways, have dissimilar reading aloud abilities and can read unfamiliar words differently from each other (Baron & Strawson 1976; Graesser et al. 1980; Perfetti & Roth 1981; Perfetti 1985, 15; Daneman 1991). Individual differences in reading rates have been said to result from differences in word recognition, word encoding and lexical access skills (Just & Carpenter 1987, 454–455). Reading rate differences are also claimed to reflect the lexical and syntactic knowledge of the speaker rather than, e.g., semantic or conceptual understanding of the text (Graesser et al. 1980). However, reading the same text several times (repeated reading) and hearing an L1 speaker read the text is shown to help L2 learners improve their reading rate (Taguchi & Gorsuch 2002; Taguchi et al. 2004; Gorsuch & Taguchi 2008).

The connection between reading rate and reading skills has been defined by Just & Carpenter (1987, 455) as follows: "the speed with which a reader can pronounce written words is correlated with his reading skill". In L1 reading speed depends word frequency and length: less frequent words are read more slowly than more frequent ones, and longer words more slowly than short ones (Just & Carpenter 1987, 46–47). Structural features of the text can also slow down the reading rate (Just & Carpenter 1987, 443–444). When reading aloud, a speaker can vary his/her reading rate on the basis of what s/he wants and does not want to emphasise in the text (Gut et al. 2007, 10). Segalowitz and colleagues argued in several studies (Favreau & Segalowitz 1982; Segalowitz 1986; Segalowitz et al. 1991; Segalowitz 2000) that even highly proficient bilinguals read slower in their L2 than L1, which, according to them, may be a result of reduced automaticity of word recognition, deficient activation of semantic representations of a single word and insufficient use of phonological information in memory. Whereas reading rates in L1 were about 320 WPM, reading rates of the same adult bilinguals in L2 were about 30 % slower (Segalowitz et al. 1991). Grabe (2009, 289–290) reports that L2 students reading comprehension (in academic settings) is rather good, but their reading rate is 80-120 WPM, which is about one-half or one-third of that of an L1 student.

Learners' L2 reading skills can affect their motivation to read independently out of class. Slow readers are less likely to engage in a reading activity in L2 in their spare time, because they find it laborious and not enjoyable. (Nuttall 1982, 167.) Therefore, skilled and fast readers, on the other hand, would be more likely to read independently. In the present study, all the

students were offered approximately the same time to read aloud in class, but reading in their spare time could not, of course, be controlled.

In sum, the literature review in this section showed that phonetic analysis provides explanations regarding the acoustic features of speech, such as pausing and speech rate, that make it sound fluent or disfluent. L2 acquisition studies indicate that factors such as the length of stay and age of acquisition can be predictors of improvement in language skills. Psycholinguistic approaches, in turn, show what factors influence reading rate.

Thus, I will formulate here a working definition of fluency to be tested in this study: read-aloud speech is considered fluent if it is spoken at a regular rate and if it has perceived pauses mostly at phrasal, clause, sentence or paragraph boundaries. Fluent speech does not contain excessive amounts of pauses, and the reading rate can sometimes slow down and become faster again, but the listener perceives it as having a somewhat regular rhythm.

2.2 Pitch as an Intonational Feature

In addition to fluency, the present study also deals with intonation and pitch. Pitch is the auditory perception of the acoustically measurable fundamental frequency (F0)[19] (Cruttenden 1986, 4). It can be measured from the complex waveform by determining its components in sinusoidal waveforms (Lieberman & Blumstein 1988, 24–25). Intonation can be described as variation in pitch movement (Brazil 1997, 1–6). More precisely, intonation can also be defined as "the salient pitch changes in the course of an utterance" ('t Hart et al. 1990, 69). This section sets out to introduce how pitch functions as a correlate of intonation. In this study, pitch in L1 Finnish and L1 Russian is described from the point of view of the pitch range and pitch contours in YNQs. Only YNQs were chosen from among all utterance types, to be investigated in the present study as a pilot study (Ullakonoja et al. 2007) showed them to be the most difficult to produce for Finnish learners of Russian. Below, I also summarise previous studies on L2 intonation in general. Furthermore, I present theoretical grounding to the choices made in the three papers (Studies IV-VI) that concern pitch range and pitch contours.

In the present study intonation is understood as variations in the pitch of the voice (Brosnahan & Malmberg 1970, 148), although it is clear that intensity, duration, speech rate and rhythm can also play an important role in defining intonation (Crystal 1969, 108, 195; 1975, 127). Here I limited my empirical analysis to pitch, which I consider to be the most important acoustic correlate of intonation. Thus, I will mostly talk about pitch contours instead of intonation.

The choice of pitch as the main intonational cue is also advocated by several other authors. First, it is known that F0 contours contribute to the

---

[19] Also, jitter, aperiodic perturbation of F0, is known to contribute to the auditory perception of pitch (Lieberman 1963).

perception of intonation ('t Hart et al. 1990, 97–98). Bolinger (1986, 24) and 't Hart (1990, 96) argue that it is pitch movements of a particular kind that contribute to the perception of the intonation (or melody) of the utterance. Further, and particularly from the Russian point of view, Svetozarova (1975, 500) claims that:

> "The main intonational means in Russian […] is speech melody, i.e., the change of fundamental frequency in time which has a minimal use at the segmental level and a limited use at lower prosodic levels in Russian."

However, it is impossible to talk about intonation without mentioning *accent*, which refers to the most important words in the sentence (Bolinger 1986, 10–11). Others use the term *nucleus*[20] or *prominence* or *sentence stress*. The Russian research tradition refers to this as the *intonation centre*[21] (Bryzgunova 1977), which aptly describes its function: the place where intonational information is centred. Another basic concept in intonation research is the *intonation unit*, which has been defined in several ways. Other terms used include *prosodic unit*, *prosodic phrase, intonation phrase, intonation group, tone-unit* and *breath group.* In the Russian research tradition this unit is called the *syntagma* (see p. 43 for a definition). Many definitions see the intonation unit as having a single continuous pitch contour and a syntactically coherent structure (see, e.g., Shcherba 1955, 84–88; Aho & Yli-Luukko 2005 for a review of the intonation unit terminology).

During the past four decades or so, intonation research has been a focus of interest for many researchers, and hence a number of different theories and schools of intonation research have emerged (for an extensive review see Botinis et al. 2001, 280–286). I have not followed any particular theoretical framework, but I have selectively employed the works of different authors from different approaches to be able to measure the way intonation is realized in the pitch contours of L2 speakers. Similarly, I do not use any phonological descriptions or transcriptions of intonation, because the focus here is on the phonetic realisation of pitch. Although multiple ways of transcribing intonation exist, many do not rely on acoustic analysis of a signal, but on auditory perception of pitch by the authors. The tools of transcription include dots, arrows, musical scales and lines (see, e.g., Lehtonen et al. 1977, 63–64). After the 1970s more sophisticated ways of transcribing intonation, e.g., the INTSINT transcription (see Hirst & Di Cristo 1998 for a detailed description) and the ToBI-transcription[22] (Silverman et al. 1992), based on autosegmental metrical phonology by Pierrehumbert (1987), have been proposed. Adopting ToBI to languages other than English is demanding as it was originally designed for English. Nevertheless, the ToBI transcription has been applied to Russian by Odé (2003) and Yokoyama (2001). Odé (2003) has suggested a ToRI (ToBI for

---

20 The term nucleus is used mainly in the British tradition, see, e.g., Wells (2006).
21 See p. 43 for a definition of intonation centre.
22 ToBI is short for Tone and Break Index, a short description available online at http://www.ling.ohio-state.edu/~tobi/

Russian) framework for the phonological investigation of Russian intonation. In studying L2 intonation ToBI has been the most popular transcription system used.

Multiple individual factors, for instance, the speaker's age, gender, sociocultural background and emotional state are known to influence pitch (Rossi 1981, 24–29). For example, women and children are known to speak with a higher pitch than men and older people with lower pitch than younger. In addition to individual factors, there are also microprosodic features (typical intrinsic pitch of sounds) and textual features that may influence pitch.

Thus, pitch as a correlate of intonation is a complex issue and has been studied using several different approaches. The present study focuses only on two aspects of pitch, namely pitch range and pitch contours in YNQs.

2.2.1 Pitch Range

First, I discuss pitch range, as it provides an overall vocal characteristic of speakers. Following Ladd (1996, 260), when studying pitch range I also consider it important to define the speaker's mean pitch. He proposes that pitch range should be looked at from two different dimensions: overall level and span. He uses *overall level* to refer to the overall mean pitch of a speaker and *span* to the ratio between the pitch maximum and minimum. Issues that relate to pitch contours as parameters of intonation in YNQs are addressed in the following subsection. I will now first look at global pitch range, and then *pitch range in different utterance types*.

Global pitch range refers to the mean pitch and pitch range across the entire recording of a speaker. There is a general belief that languages differ in mean pitch (see, e.g., Mennen 2007). Languages are also claimed to differ in pitch range in two ways: width and placement (Laver 1994, 457). Width refers to the size of the pitch range and placement to the place of the pitch range within the speaker's physically possible maximum range. Some cross-linguistics studies have focused on differences in mean pitch between languages. However, quite a few (Bezooijen 1995; Mennen et al. 2007) of them have found no statistically significant differences in mean pitch between two languages. For example, van Bezooijen (1995) found, somewhat unexpectedly, that there was no statistically significant difference in the mean pitch of Japanese and Dutch women, but instead she established that Japanese listeners preferred a higher pitch for female speakers whereas Dutch listeners preferred medium or low pitch. Ohara (2001), in turn, found no difference in mean pitch between Japanese male speakers when they spoke Japanese or English, but found that women spoke L1 Japanese with a higher pitch than L2 English. Similarly Mennen et al. (2007) showed in their systematic cross-linguistic comparison that the mean pitch did not differ statistically significantly between the two groups of German and English speakers.

In contrast, some studies have indicated that languages may differ in mean pitch. Altenberg & Ferrand (2006) found that English/Russian female bilinguals used a higher mean pitch in Russian than in English, but there was

no statistically significant difference in mean pitch between English and Cantonese for English/Cantonese bilinguals. Scharff-Rethfeldt et al. (2008) showed that highly proficient female bilingual speakers of English and German differentiated the two languages by mean F0 (English was higher). Interestingly, their bilingual speakers differed significantly from monolingual English and monolingual German speakers. De Silva et al. (2003) found that mean pitch was significantly higher in L1 Russian than in L1 Finnish. Thus, some languages seem to differ in mean pitch, while some do not. Perhaps measuring the median instead of the mean, as proposed by Lennes (2009), would better describe the pitch level of a speaker and language than mean and make the cross-language comparison more fruitful.

In languages in general, the mean pitch for women is, according to Cruttenden (1986, 4), 225 Hz, ranging from 180 Hz to 400 Hz. However, it has been suggested that Finnish women speak with a slightly lower mean pitch than female speakers of many other languages. A typical mean pitch for Finnish female speakers has been said to be around 200 Hz (Laukkanen & Leino 1999, 41). The claim can also be supported by the finding that Finnish listeners are known to prefer a low mean pitch level in Finnish spoken by both men and women (Valo 1994, 119). Also, people with a low pitch are perceived as reliable and competent in Finland (Laukkanen et al. 1999). In Russian, the mean pitch has not been measured in many studies, but it has been shown to be higher than that of Finnish female speakers (for more exact values and comparison with my results in Study IV, see Table 18 p. 102).

There are not many cross-linguistic studies on pitch range either. Mennen et al. (2007) indicate an obvious reason for this: the lack of a consensus on ways of measuring it. In addition to comparing different methods of pitch range measurement, Mennen et al. (2007) showed that female speakers of English used a wider pitch range than female speakers of German. A similar result is reported by Mennen (2007) in another study. Consistently, Jilka (2000, 110) reported a significantly wider pitch range in American English than German. In comparing English/Russian and English/Cantonese bilinguals, the only significant difference Altenberg & Ferrand (2006) found was the wider pitch range of Cantonese as compared to English. Therefore, it is possible that the use of pitch range differs in different languages.

Next, I discuss *pitch range in different utterance types*. In Study IV, I called this *local pitch range*, which may be slightly confusing, as many other researchers (e.g., Yuen 2007; Mennen et al. 2008) use the term local pitch range to refer, e.g., to pitch in a certain position in the utterance. The reason for investigating pitch range in different utterance types is mainly based on Ohala's (1983; 1984) Frequency Code, which has been held to explain the general tendency for questions being spoken with a higher pitch than statements. According to the Code, low pitch is associated with physically large, confident and authoritative individuals and high pitch with the opposite. This in turn implies that in asking a question the speaker feels less confident and is relying on the interlocutor for information, and in consequence adopts a higher pitch than usual. Therefore, it

is possible that utterance type affects pitch range. For example, an experimental study by Mennen et al. (2007) showed a significant effect of sentence length on pitch range in English but not in German. The values are given more precisely and compared with my results, obtained in Study IV, later (section 4.2.2).

Earlier studies have reported contradictory results in discussing whether utterance types differ in pitch range in Finnish. Iivonen (2005, 119–120; 2009a) found a difference between statements, questions and irritated answers in Finnish. Anttila (2009), on the other hand, did not find a statistically significant difference in pitch range between any utterance types. According to Iivonen (2009a), pitch range in Finnish is much wider in questions than statements. Anttila (2009) also found a similar distinction between questions and statements, but found no statistical significance. Iivonen (2005, 119–120; 2009a) studied samples of three speakers using a formal style (radio speech), whereas Anttila's (2009) results were based on a larger (eight speakers) corpus of informal spontaneous speech. In read-aloud speech the results are likely to be different, as Lennes's (2009) study showed that pitch is less varied and pitch range is narrower in read-aloud speech than spontaneous speech. One explanation for the different results might be wide interspeaker differences in pitch range (as reported by, e.g., Carlson et al. 2004) or the use of different measurements for pitch range in different studies.

To summarise, pitch range is likely to differ for L1 speakers of different languages. However, comparison of different studies is difficult as the ways of measuring pitch range are not fully comparable. As very few cross-linguistic studies or L2 studies on pitch range exist to date, more research using the same methods is clearly needed.

2.2.2 Intonation in Finnish and Russian

In addition to pitch range, pitch contours were investigated in the empirical part of the present study (Studies V and VI). This subsection briefly outlines the use of pitch in Finnish and Russian, before moving on to pitch contours in YNQs in particular.

Finnish is a language with a rather free word order and a rich case system expressed by morphological means (i.e., suffixes). Word stress is always fixed on the first syllable (Iivonen 2009b, 60). Typically, Finnish has a falling pitch contour, where all content words (except sometimes for finite verbs) are accented (Iivonen 1983; 1998). In Finnish, intonation does not have a grammatical function, such as distinguishing statements from questions (Välimaa-Blum 1993).

Little experimental research has been reported on the intonation of Finnish questions. Some groupings of pitch contours typical for Finnish in different sentence types have been done (see, e.g., Iivonen 1979; 1998; Anttila 2009; Iivonen 2009a), but a comprehensive account of Finnish pitch contours or an intonational grammar of Finnish remains non-existent. Also, the few existing empirical studies present somewhat controversial results and are not entirely comparable due to differences in their data. With this in mind, researchers at

least agree on one feature of Finnish intonation. This is a voice quality feature, more specifically *creaky voice*. It is known to occur frequently in Finnish, in particular at the end of utterances (Iivonen 1998, 320). It possibly also has a conversational meaning (Ogden 2003; 2004). Anttila (2008, 54–56) found that creaky voice was more typical in questions than in statements in both spontaneous and read-aloud Finnish. However, it was more frequent in question-word questions than YNQs.

In contrast to Finnish, Russian intonation research has a long tradition, starting with the earliest theoretical models of Russian intonation developed for teaching Russian to foreigners in the 1950s. One important aspect of Russian intonation research is the terminology used which differs from that used elsewhere. First, I introduce two central terms used in previous Russian intonation research, *syntagma* and *intonation centre*, which are not used in the Western tradition in the same sense.

Syntagma (синтагма 'sintagma') refers to the intonation unit, but, importantly, in the Russian definition it is also a unit of meaning (Shcherba 1955, 84). It is a combination of a word, or several words, which is intonationally undividable (Bryzgunova 1982). In other words, a sentence can consist of one or several syntagmas, but each syntagma carries a single intonational pattern (Bryzgunova 1982). According to other definitions, a syntagma is a single breath-group that has either filled, unfilled or psychological pauses at its boundaries (Zlatoustova et al. 1986, 79–80; Bondarko 1998, 234; Cheremisina-Enikolopova 1999, 159–160). Volskaya (2001) presents an interesting comparison stating that in the Western tradition syntax is more important than meaning in determining the intonation unit, whereas in Russian it is vice versa. Nevertheless, the problem of determining the boundaries of the unit exists in both the Russian and Western traditions. In general in both traditions the boundaries of the unit are determined by pauses, but the existence of virtual or zero pauses[23] in the Russian tradition complicates the division of the intonation unit in Russian (see, e.g., Svetozarova 1998, 274–275).

Intonation centre (интонационный центр 'intonacionnyy centr') is another concept in the Russian tradition, which means roughly the same as the nucleus (or prominence) in the Western terminology (Volskaya 2007). Its place is determined by the sentence accent (irrespective whether it is phrasal accent[24] or logical accent[25], which falls on the vowel of the accentuated syllable (Bogomazov 2001, 81–82). The most significant pitch movements of the utterance are realized in the intonation centre (Bryzgunova 1982). It is also the crucial point of reference in the description of different intonational constructions (IK), which are divided into a precentre (or prenuclear part), intonation centre (nucleus) and postcentre (postnuclear part) (Bryzgunova 1982). The intonation centre is not always realized by pitch peak, but with the help of

---

23 See p. 32 for a definition.

24 Phrasal accent is the sentence stress in a neutral sentence according to the rhythmical structure of the sentence.

25 Logical accent is the place of sentence stress when the speaker decides to emphasize a particular word.

other prosodic parameters such as duration and intensity (Volskaya 2009c). In Russian declaratives the intonation centre is often considered to be on the last stressed syllable of the utterance (Ivanova-Lyukyanova 2003, 4–5). In YNQs, on the other hand, the place of the intonation centre depends on the speaker's choice of what to ask (i.e., which word of the utterance to accentuate) (Bryzgunova 1975). This is illustrated in Figure 1, where it is possible also to see how the meaning of the utterance changes depending on the place of the centre. Here, simply by changing the intonation centre, different things are asked about going to the movies. Hence, in YNQs the intonation centre can be placed on any word (Fougeron-Benenson 1971, 18). In this subsection I will respect the Russian terminology and refer to the nucleus as the intonation centre when describing pitch contours in Russian.

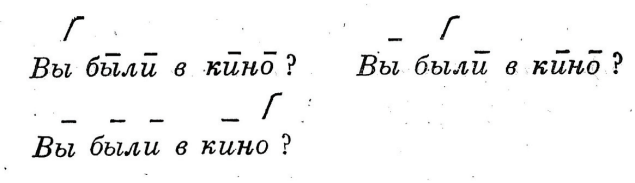


FIGURE 1 YNQ *Вы были в кино 'Vy byli v kino?'* (You went to the movies?) uttered with three different positions of the intonation centre in IK-3: 1. Was it you who went to the movies? 2. Did you go to the movies? 3. Was it the movies you went to? (Bryzgunova 1982, 97).

Russian, like Finnish, has a rather free word-order. Also there is no particular word order that would always be used in YNQs (Svetozarova 1998). In Russian one can express a question by means of prosodic cues only (Bryzgunova 1975), whereas in Finnish one cannot. In other words Russian uses pitch to signal differences in sentence type (see, e.g., Volskaya 2009a).

In many studies focusing on intonation in Russian, the communicative functions of intonation have not been described, and little acoustic data given. However, as in Finnish, so too in Russian not enough research has been conducted to enable the production of an intonational grammar of Russian (Bondarko 1998, 240). According to Svetozarova (1998) there are two different approaches to Russian intonation research. Linguists such as Nikolayeva (1977, 80–100) and Svetozarova (1982) have attempted a description of the use of intonation to differentiate different sentence types (or syntactic categories). The earlier studies by Bryzgunova (1975; 1977; 1982), along with the newer ones by Odé (1989; 2003; 2005), are focused more on determining intonational constructions through minimal pairs, i.e., through the phonological method, not through sentence types.

In Russian, no consensus has been found so far on precisely how many basic intonation patterns there are. Bryzgunova initially listed four intonational constructions (IKs) (Bryzgunova 1972; 1975) in her work, which dealt with teaching Russian intonation to foreigners, and later seven (Bryzgunova 1977; 1982). Odé (1989; 2003; 2005) on the other hand distinguished 11 patterns (or

pitch accents as she calls them). It should be pointed out that Bryzgunova's description was drawn up essentially for teaching purposes, and, like many intonation courses at the time, was based mainly on the auditory observations of the author ('t Hart et al. 1990, 176). However, there is a strong tradition of describing Russian intonation through Bryzgunova's (1977) seven IKs. According to Bryzgunova (1977), one utterance in a certain context can only be pronounced with a particular IK. For example, declaratives are usually pronounced with IK-1 and IK-2, whereas interrogatives use IK-2, IK-3 and IK-4. IK-5, IK-6 and IK-7 are reserved for rarer functions, for example, exclamations.

Bryzgunova's (1975) theory has been subjected to some criticism during the past few decades. It is true that the theory can be somewhat confusing; for example, IK-3, which is most often used in YNQs, can also be used in statements expressing continuity. According to Nikolayeva (1977, 84), the peak is a lot higher in IK-3 when it expresses interrogativity than when it expresses continuity. However, as Keijsper (1983, 129) remarks, the distinction between the different IKs is not always clear. As an example he mentions the distinction between IK-3 and IK-6, which can be neutralized if the intonation centre is in the utterance final position.

2.2.3 Pitch Contours in Finnish and Russian YNQs

The aim of this subsection is to illustrate and discuss the differences between Finnish (L1 of the subjects of the present study) and Russian (L2) YNQs in pitch. Pitch contours were chosen to be investigated in YNQs, because this has been said to be the most difficult Russian utterance type for foreigners to produce (see, e.g., Nikolayeva 1977, 84). Also, a pilot study by the author (Ullakonoja et al. 2007) showed that it was the most difficult utterance type for Finns to produce.

Following t'Hart et al. (1990, 82–84) a *pitch contour* is defined as consisting of independent movements of pitch. *Intonation pattern* on the other hand is regarded as an abstract category on the phonological level, and hence related to the intonational grammar of the language ('t Hart et al. 1990, 87-88). Therefore, a language has an unlimited number of possible pitch contours but a limited number of (linguistically significant) intonation patterns ('t Hart et al. 1990, 82-84). The present study, thus, focuses on pitch contours instead of intonational patterns.

The most "natural" way of producing an utterance is by making the pitch fall either towards or at the end of the sentence, owing to the decrease in the pulmonary air pressure (subglottal pressure). This regular fall in pitch towards the end of the utterance is called *declination* (Cohen et al. 1982). It has been observed in most languages (Lieberman & Blumstein 1988, 203; 't Hart et al. 1990, 121.) and it has even been modelled (Cooper 1981, 28–100). Hence, because the articulatorily "natural" way is to produce an utterance with a declination towards the end of the sentence, producing an utterance with a rising contour (for example in YNQs) requires extra effort on the part of the speaker (Cohen et al. 1982; Lieberman & Blumstein 1988, 200–201.). Cruttenden

(1986, 162) claims that frequently in languages YNQs differ from the corresponding statements as having "a 'terminal rise' or in some way a higher pitch than the corresponding statement pattern". This claim is supported by the studies of Bolinger (1978) and Ultan (1978).

YNQs in *Finnish* are formed by grammatical, rather than intonational means. They are marked by an interrogative particle *-ko/-kö*. In spontaneous speech, YNQs are possible without the lexical marker. However, YNQs expressed only by intonational means are perceived as very unnatural by native speakers of Finnish (Mixdorff et al. 2002). Iivonen (2001a) concludes that Finnish speakers can use intonation for expressing interrogation but that there is no single way to do it: the speaker can have a high initial peak, high pitch level in general or even final rise. Also, speakers can express interrogativity purely syntactically without intonational markers. As for pitch contours, it has been found that sentence type affects pitch so that the initial pitch is higher in questions than in statements (Hirvonen 1970, 39–40; Mixdorff et al. 2002). On the contrary, Itkonen (1972, 14) reported an overall higher pitch in questions than in statements.

According to Hirvonen (1970, 31–35), Finnish YNQs[26] differ from question-word questions[27] in their pitch contour: question-word questions have a high initial pitch followed by a fall whereas in YNQs the fall comes later, on the nucleus. In YNQs a similar contour has been observed by Iivonen (2001a). However, Iivonen (2001a) found that question-word questions are marked with a high initial pitch, whereas YNQs do not have as high a pitch in the beginning. Nevertheless, Iivonen (2001a) claims that YNQs can be characterized as having an overall high pitch level. Anttila's (2008, 84) experimental study on Finnish interrogative intonation sheds more light on interrogative pitch contours in Finnish, but is unable to determine a specific interrogative intonation contour that would be typical in Finnish. Anttila (2008, 76–77) determined that in Finnish read-aloud speech, the most typical pitch contour in YNQs was a fall or a rise-fall. However, she concluded that speakers can mark interrogatives prosodically, but the ways can be different for different speakers. In sum, it can be concluded that Finnish YNQs have an overall higher pitch than statements and that they are distinguished from question-word questions by lower initial pitch and a later fall.

A more controversial issue concerns the possibility of a rising pitch contour in Finnish questions. Traditionally, it has been considered that this does not occur (Iivonen 1987, 241). For example, Iivonen (1998) argues that interrogative intonation may not exist in Finnish, and that final rises occur in Finnish only occasionally. The fact that most of Iivonen's data is read-aloud speech may explain the absence of final rising contours. However, nowadays there is much evidence that rising pitch contours indeed exist in Finnish. For example, in a more recent study Iivonen (2001a) found final rises, e.g., in echo-questions and tag-questions. Furthermore, Routarinne (2003; 2008) and

---

26 Hirvonen (1970) uses the term general question for YNQs.
27 Hirvonen (1970) uses the term particular question for question word questions.

Routarinne & Ogden (2005) found final rises in YNQs in teenage girls' spontaneous Finnish. However, they emphasise that the existence of a final rise does not automatically mean that we are dealing with an interrogative. In fact, in their data a rising terminal pitch contour was rather common in general: they report about two rising contours per minute and their studies show, interestingly, the ways in which a rising pitch contour can function in conversation. Mixdorff et al. (2002) also found the much debated final-rises in echo-questions along with non-terminal rises in Finnish conversation speech as well as in a small data set of Finnish read-aloud speech. Furthermore, Anttila (2008, 76; 2009) reported a total of 33 final rises in the 277 questions of her spontaneous dialogue data. As most of her question data is either YNQs or question-word questions, the proportion of final rising contours is also almost the same in these question types. Consequently, in the light of these recent studies there is no doubt that a rising pitch contour is possible in Finnish and perhaps even frequent in Finnish questions. However, it is likely that is not compulsory in Finnish but rather an expressive possibility.

In fact, it can be concluded that native Finnish speakers who are learning Russian are familiar with and can produce the rising pitch contour that is needed for expressing interrogative intonation in Russian solely on the basis of their L1 alone. However, the contour with a sharp peak on the nuclear syllable that is typical in Russian YNQs (as described below) may not exist as such in Finnish. The final rising pitch contour very likely exists in the everyday Finnish of the subjects of the present study (young adult females). Of course, the subjects have also studied other L2s and might be familiar with the use of pitch in them. However, as there are no clear rules on when and how to use pitch in Finnish to convey an interrogative meaning, the students might struggle to learn the more precise rules of Russian. These rules and a description of the Russian intonational system are presented next.

Russian YNQs (in Russian *общий вопрос 'obshchiy vopros'* = *general question*) typically have the same direct word order as statements (Volskaya 2009a) and hence a question can be differentiated from a statement solely by means of pitch variation. It is possible to mark YNQs grammatically with the particle *ли* (*li*), but it is less common than having a YNQ identical to a statement in morphology and syntax but different in prosody. Most scholars have described the distinctions between questions and statements solely reference to use of pitch. However, researchers such as Svetozarova (1982, 111–112) and Nikolayeva (1977, 84–85) argue that questions in general are spoken faster than declaratives and with a higher intensity and wider pitch range than statements.

Turning then to pitch contours in Russian, according to the pioneering work of Bryzgunova (1982), Russian YNQs are pronounced with the IK-3 pattern, which typically has a steep rise on the intonation centre. In the precentre the pitch is on the average pitch level of the speaker and in the postcentre the pitch is falling and lower than in the precentre (Bryzgunova 1977, 38). The form of the pitch contour in YNQs is illustrated in Figure 1 (p. 44) which shows that it can be different depending on the place of the intonation

centre. In addition to YNQs, IK-3 is also used to express incompleteness or surprise in statements (Bryzgunova 1977, 198–201).

Below, Figure 2 illustrates the pitch contours in the present data. A short YNQ *Соня? 'Sonya?'* is produced by three different types of pitch contour by native speakers. The context is the same: at the beginning of a telephone conversation, the speaker says *Алло. Соня?*[28] (*Allo. Sonya?)*. This is different from Bryzgunova's example in Figure 1 in that here the place of the intonation centre is always the same and the meaning of the utterance does not depend on its place.

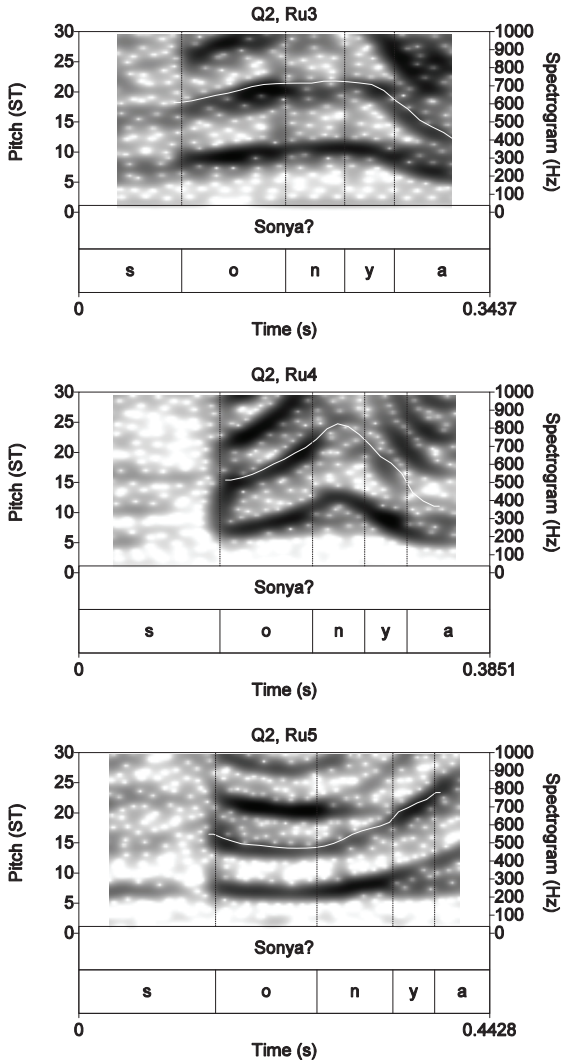


FIGURE 2 YNQ *Соня? 'Sonya?'* (Sonya, proper name) uttered by three native speakers and three different pitch contours (from my own data).

28 In English: 'Hello, is that Sonya?'

Other scholars describe the YNQ pitch contour as having a steep rise or a high peak in the intonation centre, depending on the place of the intonation centre (Svetozarova 1982, 100; Fougeron 1999; Makarova 1999; Meyer & Mleinek 2006, 1619). In general the rise-fall is placed on the intonation centre, but when the intonation centre is on the last syllable, there is a high-rise (Cruttenden 1986, 164). Volskaya (2009a) describes the typical contour as having "a pointed hat" at the intonation centre. Nikolayeva (1977, 84–85) writes that YNQs have a rise on the intonation centre followed by an F0 lower or similar to that at the beginning of the intonation centre. She adds to the description the fact that pitch range in the intonation centre is wide and that the intensity contour is either falling or very similar to the pitch contour. According to Volskaya (2001; 2009a) and Volskaya & Skrelin (2001), other pitch characteristics typical of YNQs are the absence of declination and pauses as well as no great pitch variation at the boundaries of the intonation unit or during the phonetic words[29] of the precentre.

Svetozarova's (1982, 100) description of Russian YNQs corresponds to that of Bryzgunova (1977, 38): there is a sharp rise on the stressed syllable of the "most important word" preceded by a monotonous, i.e., flat pitch on the intonation centre, followed by a declination. In Kasatkin's (2007) terminology YNQs are pronounced with the pattern TK-3.1[30], which is described as having a peak on the intonation centre, but the fall extends to the syllable following it. The precentre is pronounced with a neutral pitch and the postcentre with a lower pitch than the precentre. If the intonation centre is in the phrase final position, it is possible, according to him, only a small pitch fall will occur during the intonation centre or none at all.

Although the descriptions above seem rather similar, researchers do not completely agree on the intonation pattern used when producing Russian YNQs. Traditionally, YNQs are said to be pronounced with IK-3, or with IK-4 if they begin with the conjunction *a*[31] *(a)* (Bryzgunova 1982). The typical contour of the *a*-question is a rise-fall followed by a final rise (Fougeron-Benenson 1971; Bryzgunova 1982) (this type of YNQ was not present in my data). However, Fougeron-Benenson (1971, 63) argues that two types of pitch patterns are possible in YNQs not beginning with the conjunction *a*. The pitch pattern depends on the intention of the speaker: for a neutral question (for asking information) there is a very sharp rise-fall, but for questions that are likely to be answered affirmatively the peak is not as sharp.

Also, according to Bryzgunova (1977, 200–201), there are several other possible realisations of pitch contours in YNQs than the traditional IK-3 pattern. It is possible that the peak is produced on the last syllable of the precentre instead of the intonation centre, expressing a tone of disbelief. Also, it is possible that the peak on the intonation centre is not sharp, but flatter, and hence the peak starts already during the precentre and expresses positive

29 See p. 74 for a definition.
30 TK is short for a tonal contour.
31 In English: and what about.

surprise. Furthermore, the high pitch of the intonation centre can be continued during one or two syllables of the postcentre and thus express unexpectedness. Further, the pitch can be relatively low on the intonation centre but rising sharply during one or two syllables of the postcentre, expressing a great surprise. I would argue that this could be the same phenomenon that other researchers have referred to as *peak-delay* (see below).

Some researchers have compared the intonation of YNQs with other utterance types. In comparison with statements, according to Svetozarova (1975; 1982, 111–112; 1998, 271) Russian YNQs have a relatively monotonous pitch in the precentre and postcentre as well as an accelerated speaking rate. Declaratives, on the other hand, have more peaks. Unlike Bryzgunova (1972; 1975; 1977; 1982), Svetozarova (1975) distinguishes questions from statements not only by pitch but also by intensity, which is higher in the intonation centre of interrogatives. Moreover, Fougeron (1999) and Yanko (2008) have attempted to define the differences between statements and YNQs that are identical in lexical content. Yanko's (2008, 38, 83–85) experiments, on the other hand, showed that YNQs had a deeper and faster rise in the intonation centre than the corresponding statements. Also, the rise started lower and ended higher in YNQs than statements, which means that YNQs have a wider global pitch range. Fougeron (1999) found that in addition to differences in pitch contours YNQs often also differ from the corresponding statements by the place of the intonation centre.

In a recent study, Makarova (2007) found YNQs to differ from declaratives and exclamations in that the pitch peak was positioned significantly later in YNQs. Furthermore, she found that declaratives differ from interrogatives and exclamations in their pitch peak height. These results were confirmed in a perception task with manipulated stimuli: the later position of the peak as well as the higher peak provoked perception of the stimulus as an interrogative. This is consistent with Meyer & Mleinek (2006, 1619) who showed that in read-aloud speech the maximum pitch of the nucleus was higher in YNQs than in declaratives.

Pitch in YNQs has also been studied in comparison to question-word questions. It has traditionally been claimed that YNQs and question-word questions differ in the intonational construction they use (IK-2 for question-word questions and IK-3 for YNQs) (Bryzgunova 1982; Svetozarova 1982, 92, 100). For example, Igarashi (2006) observed that the peak was higher in YNQs than question-word questions and the initial pitch was lower (for a schematic representation of the differences, see Igarashi 2006, 184). The new finding in Igarashi's (2006) study is the difference between the two types of interrogatives in the precentre, which in contradiction to the earlier finding by Bryzgunova (1982). Igarashi (2006) claims that pitch movements in the precentre proved crucial in differentiation: in YNQs the pitch in the precentre is flat until the beginning of the intonation centre, where as in question-word questions the pitch commences to rise much earlier.

Several scholars have reported a phenomenon called *peak-delay* [32] in Russian, though not all have used this term. The phenomenon is characterised by a delayed pitch peak, i.e., the position of the maximum pitch does not fall on the stressed syllable of the intonation centre, but on the unstressed one following it. In Russian, peak-delay has been observed acoustically in 1) YNQs compared to statements (Meyer & Mleinek 2006), 2) YNQs compared to question-word questions (Igarashi 2006) and 3) interrogatives and enumerations compared to declaratives and exclamations (Makarova 1999). Volskaya (2007) observed it in an IK-3 used to express discontinuity in both spontaneous and read-aloud speech. Hence, peak delay can reasonably be adeptly argued to be characteristic of Russian YNQs.

Most researchers who have focused on determining the typical intonation pattern of YNQs might share views about the most typical shape of the contour, but do not agree on what is the *most salient part of the pattern for distinguishing different sentence types*. Some researchers, such as Bryzgunova (1975), Fougeron-Benenson (1971, 14–15, 59–60) and Odé (1989) have ignored the role of the precentre. Bryzgunova's (1982) theory is based on pitch differences in the intonation centre or postcentre, as she claims that the precentral part is usually pronounced with an average flat pitch or with very small pitch variation[33]. Similarly, according to Fougeron-Benenson (1971, 14–15, 59–60), the differences between statements and YNQs are realized in the intonation centre, (which in YNQs has a sharp rise-fall, absent in statements) and the precentre has no function in differentiating them.

In contrast, Volskaya (2001) and Kasatkin (2007), among others, claim that the most important component for distinguishing YNQs from statements is the precentre. This was demonstrated in a perception experiment, where Svetozarova (1982, 116–120) found that even in sentences where the intonation centre was in the phrase-final position and deleted, listeners were able to distinguish between statements and YNQs on the basis of the precentre only. They differentiated between statements and YNQs on the basis of the faster speech rate and more monotonous pitch of the YNQs in contrast with the statements. Thus, Svetozarova contradicts the claims made in her earlier work (1975) where she put strong emphasis on the pitch changes in the intonation centre, stating that they are great enough to differentiate different intonation patterns.

Turning now to the perception of intonation, at least one interesting Finnish-Russian contrastive study has reported on the perception of intonation. Shcherbakova (2001) studied how Russians (who had not studied Finnish) perceived Finnish utterances and Finns (without any knowledge of Russian) perceived Russian. The listeners were asked to determine what kind of contour they heard: non-final, final, interrogative or emotional. The results showed, first, that both Russians and Finns encountered difficulties in recognizing the pitch

---

32 In Russian: локализация частотного пика на заударном слоге.

33 Interestingly Bolinger (1986, 25) claims that it is the end of the sentence that is most important in distinguishing YNQs and statements in English.

contours of the other language correctly. Only 21 % of the Russian listeners rated Finnish YNQs as interrogatives, whereas the majority rated them as final statements. The Finns recognized Russian interrogative intonation (both YNQs and question-word questions) in 35 % of the cases; it was also commonly rated as emotional speech. Shcherbakova's (2001) study provides strong evidence of the differences between the two intonation systems. Although the study focused on non-learners of the target language, it suggests that merely perceiving the interrogative intonation of the other language alone demands a lot of effort. Learning to produce the interrogative intonation itself represents yet another challenge for the learners.

To summarise the works of different authors, it can be argued that the place of the intonation centre is determined by the speaker. The place of the intonation centre, then, determines the pitch contour to be used. If the intonation centre is in the phrase final position, the YNQ is pronounced with a final rising contour as the post centre does not exist. In other positions the intonation centre has a rise-fall (researchers do not agree whether it is sharp or not) and the peak is delayed (or at least realized later than in the corresponding statements). According to several recent empirical studies, the precentral part of YNQs also differs from that in statements and question-word questions by being rather flat. In sum, despite the extensive amount of research conducted on Russian intonation, disagreement remains on many issues. For example, it is difficult to draw conclusions on what constitute the most important pitch features in YNQs that a potential L2 learner should take into account. What is clear, though, is that YNQs differ from other utterance types, such as statements, in pitch.

2.2.4 Pitch Contours in L2

In the light of the literature review above it is surprising that traditionally no great importance has been attached to prosody when teaching L2 pronunciation. However, it is a well known fact that even adult L2 learners can attain a high level of proficiency in L2 prosody through phonetic training. (Neufeld & Schneiderman 1980.) However, surprisingly little experimental research has been done on L2 speech production. Some of the earlier studies have focused only on the segmental aspects of speech, but studies on L2 prosody from the production aspect are extremely rare. Hence, this subsection will provide a summary of the previous research on L2 prosody, in particular intonation, in order to outline the issues addressed in previous studies involving L1 Finnish and L2 Russian as well as some other languages.

The Common European Framework of Reference for Languages (Council for Cultural Co-operation. Education Committee, Modern Languages Division, Strasbourg & Council of Europe 2001) has underlined the role of intonation in L2 oral skills. For example, it expects the L2 learner to master intonation already on the B2 level (4th highest level on the 1–6 scale defined as "independent user, vantage") by stating that on that level the L2 speaker "has acquired a clear, natural, pronunciation and intonation" (p. 117). This rather exacting

requirement has been criticized, e.g., by de Silva & Volskaya (2005), who discussed it from the point of view of L2 Russian oral skills in Finland. They stressed that as Russian is most often learnt as the third or fourth L2, there are insufficient possibilities to focus on the prosodic features of Russian at school.

As Wennerstrom (2001, 230–243) in her survey observes, some previous studies on L2 prosody have focused on describing L2 intonation, whereas others have taken a more evaluative approach, focusing on foreign accent, comprehensibility or fluency. In my view, most previous studies on L2 intonation have been based on the auto-segmental metrical approach (see, e.g., Pierrehumbert 1987) and used the ToBI transcription system (see p. 39) for the phonological transcription of intonation (e.g., Grosser 1993; 1997; Jun & Oh 2000; Chen & Mennen 2008; Chen 2009). For example, Mennen (1999) investigated the realisation of Greek YNQs by advanced Dutch learners and found, first, that it is hard to produce a native-like contour and, second, that it was easier to produce a pitch contour that does not correspond to any contour in L1 than a contour that is similar to a contour in L1.

Several studies based on the auto-segmental metrical theory have focused on peak alignment, i.e., the position of the peak. Mennen (1998b; 2004; 2007) found that correct positioning of the peak caused difficulties even for proficient L2 learners. This seems to be a particularly difficult aspect of L2 intonation as learners fail to achieve native-like values even after many years of experience of the L2 (from 12 to 35 years). In fact, Mennen (1998b) concludes that the influence of L1 continued to remain great for highly proficient L2 speakers. Interestingly, Mennen (2004) also found that learning L2 influences the timing in L1.

The auto-segmental metrical theory has also been used to investigate foreign accent. Jilka's (2000; 2007) studies focused on foreign accent, i.e., what pitch movements were perceived as foreign by American English speakers in the speech of German learners. Using ToBI, he identified the difficulties of German speakers in producing American English intonation (Jilka 2000, 93–96). In the perception experiment, both German and American speakers were best at recognizing their fellow L1 speaker, which shows listeners' great sensitivity to pitch contours, being able to distinguish between familiar ones (of their L1) and "foreign ones" (those of L2) (Jilka 2000, 160).

Holm's (2008) study had much in common with Jilka's, but in addition to intonational features it also focused on durational features in the perception of foreign accent. Interestingly, the study had seven different L1s and the stimuli were manipulated to investigate how intonation and duration affect the degree of foreign accent and intelligibility in Norwegian. It can be concluded from this study that the ways in which different L1s influence L2 prosody vary.

The degree of success in the production of L2 intonational meanings has also been investigated, e.g., by Hewings (1995; 1998). His aim was to determine what kind of tone choices non-native speakers make in L2 English both in spontaneous and in read-aloud speech[34]. His study operated more on the

[34] The L1s of the speakers were Korean, Greek and Bahasa Indonesia.

phonological level rather than acoustic level, but some interesting findings were presented: 1) L2 speakers had shorter intonation units than L1 speakers, 2) L2 speakers realized prominence later than L1 speakers, 3) L2 speakers used level tone more often than L1 speakers, 4) L1 and L2 speakers used similar contours for information structuring, but L2 speakers use different contours than L1 speakers for expressing lack of confirmation.

Studies have also been conducted on the perception of intonational meaning in L2. For example, Cruz-Ferreira (1987) compared L1 and L2 speakers' perception of intonational patterns in English and Portuguese. The study showed that L2 speakers had difficulties in assigning intonational meaning to sentences that used an intonational pattern that did not exist in their L1 or was used differently in their L1. These results cannot, however, be directly applied to production of L2 intonation, but they give an interesting insight into passive prosodic competence in L2.

Other studies on L2 intonation have focused on how intonation is learnt. For example, Ramírez Verdugo (2006) described the benefits of a computer-assisted learning environment in learning L2 intonation. This study found evidence for the computer-assisted learning environment (involving visual representation of the pitch contours, auditory comparison of the L2 speaker with an L1 speaker as well as the possibility of the learner to record speech) helping the L2 speaker to progress both in relation to the acoustic realisations of pitch contours and to native speakers' judgements. Ramírez Verdugo (2006) suggests that increasing intonation awareness would improve L2 intonation. More importantly, the students were also able to transfer the skills acquired in a laboratory setting into more spontaneous conversation. Similar results have been obtained by de Bot & Mailfert (1982) who found that training in the perception of L2 intonation also resulted in improvement in its production. Some L2 studies have examined the behavioural and cortical effects in perceiving L2 prosody. For example, Sereno & Wang (2007) compared perception to production and concluded that production of the L2 pitch contours improved without production training after the perceptual training. Hence, the results emphasise the role of learning to perceive phonological contrasts in learning to produce them.

There are few studies in which Finnish has been the L1 of the subjects. Hirvonen (1970, 80) describes the difficulties of Finnish speakers in learning the English rising intonation as follows: Finnish L2 speakers maintain or even increase the intensity level up to the end of the utterance, whereas L1 speakers usually have a fall in the intensity contour towards the end. This leads to the impression of L2 learners pronouncing the last syllable as stressed (Hirvonen 1967, 47). Hirvonen (1970, 76) showed that as Finnish does not use intonation to signal differences between YNQs, question-word questions and imperatives, like English, a Finn learning English is likely to face difficulties in making this distinction by intonational means. For learning the rising intonation of English, Hirvonen (1970, 80) proposes singing, as this would force the Finnish learner to increase the tension of the vocal cords to produce the rise.

The largest series of studies concerning Finnish students' prosodic production in L2 is that of Toivanen. The main findings of Toivanen (1999, 209–250) were that Finnish speakers' English is characterized by 1) a lower mean F0 than in L1 speakers' speech, 2) a narrower pitch range than L1 speakers, and 3) a smaller standard deviation in pitch than L1 speakers' speech. All these differences were statistically significant. When studying intonation Toivanen found that Finns used simple falls and level tones more than L1 speakers but fewer fall-rise and fall-plus-rise patterns. In a later study Toivanen (2003) obtained similar results. In addition, he found that in read-aloud English YNQs (polar questions) L1 speakers used a falling tone slightly more often than a rising one (53 % vs. 47 %) but L2 speakers clearly preferred the falling tone (65.3 % vs. 24 %). In statement-questions[35] L1 speakers used a rising contour slightly more often than L2 speakers (L1: 74 %, L2: 57.2 %), but in contrast to the findings or some earlier studies L2 speakers used it very frequently. Finns had difficulties in expressing continuation intonation as they used a fall-rise extensively in statements despite their pragmatic function. (Toivanen 2003.) Finnish speakers' spontaneous production in L2 English was characterized by a considerable number of final falls, but also some final rises in the utterance-final position (Toivanen 2004). Toivanen (2006a; 2007) also showed, e.g., that Finnish students of English used a fall-rise pitch contour incorrectly when in a conversation with other L2 speakers in an academic context. In addition, Toivanen (2006b) claims that there seems to be a connection between L2 proficiency and the ability to use rising tones in a spontaneous conversation.

Furthermore, Toivanen & Waaramaa (2005) reported that rising tones in general and especially falling-rising pitch patterns were rare in L2 English spoken by Finns. In addition to that, Finns tended to use creaky voice towards the end of the speaking turn. Toivanen's (2009) more recent study presents somewhat different results: there was less creaky voice (only 4 % of the data) and more rising (15 %) and falling-rising (16 %) patterns. The explanation given for the contradictory results is that Toivanen (2009) studied spontaneous data, i.e., an authentic conversation between an L2 speaker and an L1 speaker, whereas the previous studies had used read-aloud speech as data. Toivanen's study showed that when engaged in meaningful interaction with a L1 interlocutor, a Finnish L2 speaker of English was prosodically more adequate and lively than in previous studies where L2 speakers have been conversing with each other.

In studies of *Russian L2 prosody*, the following results have been presented. Keijsper (1983, 125) stated that producing Russian YNQs is difficult for L2 learners because the pitch fall was completed only during the postcentral part and not the intonation centre. Zinder (1980, 128–129) reported that L2 learners of Russian do not usually face great difficulties in learning to produce intonation in YNQs, explaining this by the fact that interrogative intonation contours are somewhat similar across languages. However, certain differences

---

35 The statement-question in Toivanen's terminology is an utterance which has the form of a statement but which functions as a question.

between L2 and L1 speakers were observed: L2 speakers did not always produce the pitch peak in the intonation centre and the pitch peak was not as high in L2 speech than in L1. L1 speakers received 95–97 % recognition rates when producing YNQs, but L2 beginner learners got a little over 50 %, while advanced learners could achieve up to 90 %.

Finnish students' intonation in Russian has been the subject of a few previous studies. Kuosmanen & de Silva (2003; 2007) studied Finnish university students' (n = 10) productions of read-aloud Russian. The students were studying Russian at university and most of them were recorded after their semester in Russia. They conducted both acoustic analysis of the pitch contours as well as a native speakers' (n = 6) perception experiment. The two most striking observations to emerge from their study was that a high rising pitch pattern and the interrogative intonation in general caused difficulties for Finnish students of Russian. The main focus of Kärkkäinen et al. (2006) and Kärkkäinen (2009) was division of intonation unit and realisation of prominence in L1 and L2 Russian. Kärkkäinen (2009) observed that L2 Russian speech of Finns differed from L1 speech in that intonation units were longer in seconds but shorter in words. Furthermore, the most common position of prominence was the beginning of the intonation unit in L2 speech, but the end in L1 speech.

Shcherbakova (2002) remains the only study so far to describe Finnish students' pitch contours in spontaneous Russian, but as her data is based on a sole speaker, the results can be considered as indicative only. The results showed that the Finnish speaker realized the pitch contour in utterances expressing continuity in a fairly appropriate manner but struggled in producing the contour of discontinuity; the intonation centre was not realized prominently enough and there was sometimes a final rise instead of a fall.

Hirvonen (1970, 77), Toivanen (1999, 409–410) and de Silva & Volskaya (2005) emphasise the importance of learning to use the appropriate pitch contour from a pragmatic point of view. Using the wrong pitch contour can create additional meanings that the learner does not wish to convey. This is consistent with the conclusions drawn by Ramirez Verdugo (2005). For example, according to de Silva & Volskaya (2005), the level tone, which is pronounced with a low tone of voice, is very common in Finnish. Hence, a L2 learner of Finnish, who does not know that, might interpret it as having a negative emotional meaning.

All in all, a number of conclusions emerge on the basis of the studies above. First, several authors have argued that L2 intonation is difficult to learn and that L1 influences the production of L2 intonational patterns. Second, many studies offer results that could be applied in L2 teaching and L2 learning, mainly with the help of multimedia. Third, they emphasise the need for more research in the field.

3 EXPERIMENTAL DESIGN, MATERIAL AND METHODS

This chapter presents the materials and methods of the study. The study consists of six longitudinal experiments (Studies I–IV). In order to investigate fluency and intonation development during SA, two research phases were implemented. In the first, students were recorded, and the recordings analysed acoustically for pausing and speech rate and Russian teachers living in Finland were asked to evaluate the fluency of the students' speech samples online. The results of the first phase are reported in the present author's Licentiate thesis (Ullakonoja 2009a). The second stage involved further acoustic analysis of the students' data from the point of view of intonation and a perceptual evaluation by L1 speakers, who rated the successfulness of the question mode in the students' YNQs. To give an overview of the data, Figure 3 below presents the different data used in this study and how they were analysed. The principal data consist of recorded speech. In addition, background information on the subjects was collected and listening evaluation tasks of teachers and L1 speakers were used.

The chapter begins with an introduction to the speech corpus (section 3.1), including information on the subjects and the dialogues that were recorded. I also describe the recording procedure and discuss the use of read-aloud speech as data. Section 3.2 deals with the acoustic methods used and the use of the Praat programme. The fluency evaluation task is introduced in section 3.3. Section 3.4 explains the task where L1 Russian speakers evaluated the interrogativity of the students' YNQs perceptually. The last section (3.5) briefly describes the statistical methods used in the study on a general level; but these are described in more detail in each paper (see Appendices 1-6).

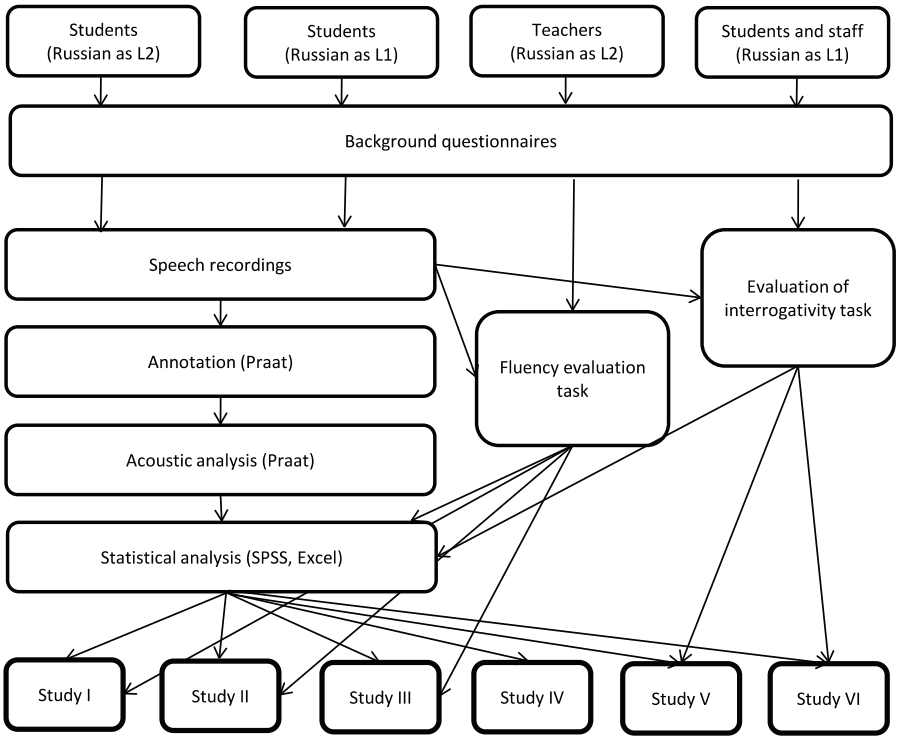


FIGURE 3 Data collection and analytical procedures.

3.1 Speech Corpus

Here, I present the data collection procedure for the phonetic analysis. First, I introduce the subjects and the questionnaires that were used for collecting the background information. I then describe the texts on which the phonetic data are based, and explain the recording procedure. The section ends with a discussion on the features of read-aloud data and the processes of reading aloud in L2.

The corpus was collected for two groups of students (who started their university studies in consecutive years) in different stages of their university studies. Table 7 (p. 66) summarises the recordings and participants of this speech corpus.

3.1.1 Finnish Speakers

As age and gender are known to affect pitch (see, e.g., Hollien & Paul 1969; Trollinger 2003), all the speakers in this study were the same sex and around the same age, viz. 19 to 24-year-old female undergraduate students of Russian at a Finnish university. The students were from two intakes: seven had started their university studies in 2004 and five in 2005.

The students were all native Finnish speakers who reported having no hearing or speaking disabilities. All the subjects participated in the same SA programme during the second year of their university studies, and studied at the same Russian university in Tver for 3.5 months. Prior to their stay in Russia, they had taken one course of Russian phonetics during their first year, where they had been taught the basic segmental and intonational features of the language. Some students resided with a Russian host family during their stay in Russia, whereas others stayed in the foreign-student dormitories. One student moved to a dormitory in the middle of her stay. During their 3.5-month-stay in Russia they had no formal instruction in phonetics, but participated in several Russian linguistics courses for L2 learners. In the Russian classroom, the students were often asked to read texts aloud. Interviews with their Russian teachers[36] and observations of their lessons revealed that teachers differed in their feedback to the students on pronunciation: some teachers corrected mispronunciations, especially word stress, whereas others paid hardly any attention to correcting pronunciation mistakes, but instead focused on correcting grammatical errors.

Before moving to central Finland for their university studies, the students had lived most of their lives in southern or south-eastern Finland (5 students), in central Finland (4 students), near or in Oulu (3 students) and in eastern or north-eastern Finland (3 students). On a subjective evaluation regional variation was not noticeable in their Finnish pronunciation. One student originating from south-western Finland had a strong dialectal pronunciation and was left out of the corpus. As dialectal variation in pitch has not been studied consistently in Finnish[37], it is unknown whether the place of origin of the students would affect the results of this study.

As the students did not participate in any language skills tests for the purposes of this study, their Russian competence is not discussed here nor taken into account in the analysis. However, according to my subjective evaluation, no noticeable differences were observed in pronunciation between students who had studied Russian for 7–10 years and students who had studied Russian for less than 7 years. Most of the students had not been exposed to the Russian language community before their 3.5-month stay in Russia during their second year, with the exception of a few short trips. The motivation towards learning Russian was considered to be equally strong for all the students. After

---

36 The interviews were conducted by the author in October 2005 and 2006.

37 Although some studies (Penttilä 1958; Wiik 1988; Pallonen& Yli-Luukko 1995; Yli-Luukko 2001; Ylitalo 2009; Aho 2010) have focused on the prosody of particular Finnish dialects.

all, most were students majoring in Russian and all had participated in a study in Russian programme.

In Studies I, II and III, twelve Finnish students majoring in Russian participated as speakers. They had studied Russian on average for 4.17 years (std. = 2.368) prior to their university studies. Only one subject had studied Russian as her first L2, starting in the 3rd grade, that is 10 years before she started university studies. One student had studied Russian as her second L2, starting in the 5th grade, whereas the rest had studied Russian only in high school, in one case only for a year. Most speakers (n = 10) had studied English as their first L2 and Swedish as their second L2[38]. Hence they had Russian as their third or fourth L2. The first three studies (Studies I–III) used the highest number of speakers (n = 12) to ensure a comparison of interspeaker differences in fluency. In the last three studies (Studies IV–VI), my aim was to study pitch features more closely, hence I chose to study fewer speakers but more data.

Pseudo-nym	Age	Academic Major	Previous Study of Russian	Position of Russian in their L2 studies
Fi1	25	Russian Language and Culture	4 years language courses, 1 year university	4th L2
Fi2	20	Russian Language and Culture	3 years high school, 1 year university	4th L2
Fi3	23	Russian Language and Culture	7 years school, 1 year university	2nd L2
Fi4	23	Other	2 years language courses, 1 year university	6th L2
Fi5	21	Russian Language and Culture	1 year language course, 1 year university	4th L2
Fi6	21	Russian Language and Culture	5 years school, 1 year university	3rd L2
Fi7	24	Other	6 years language courses, 1 year university	3rd L2
Fi8	20	Russian Language and Culture	3 years high school, 1 year university	4th L2
Fi9	22	Russian Language and Culture	3 years high school, 1 year university	4th L2

TABLE 4 Finnish participants' demographic information and academic profiles before study abroad (Study IV).

In Study IV nine students from the first intake year group were studied. The students whose speech was used were the same as those in the previous three studies, i.e., students who had started university in 2004 with addition of two students studying Russian as a minor subject. Only the students from the one of the intake year groups were used, as the recordings of the other intake year group were still under way when Study IV, chronologically the first study, was completed. Participants' demographic and academic profiles are given in Table 4. The Finnish participants of Study IV had studied Russian in different

[38] In Finland Swedish is called the *second national language* but most students regard it is as a foreign language.

institutions and for different lengths of time, and the age at which they started learning Russian also varied widely.

For Studies V and VI, only students with a similar language learning background were chosen from the corpus. This selection was made in order to obtain a group of students as homogeneous in their Russian L2 skills as possible. The participants were six female major students who had studied Russian for three years prior to university studies as their third or fourth foreign language. They were from 20 to 24 years old before SA.

The pseudonyms used for the Finnish students in the different studies and the corpus are summarised in Table 5 below. The pseudonyms used for the students were different in different studies for practical reasons. The data were used selectively so that only three students (Kati, Marjo and Liisa) participated in all six studies.

Pseudonym in the corpus	Pseudonym in Studies I-III	Pseudonym in Study IV	Pseudonym Study V-VI
Sanna	Fi1	-	Fi6
Kati	Fi2	Fi8	Fi2
Marjo	Fi3	-	Fi4
Ilona	Fi4	Fi9	Fi3
Ritva	Fi5	-	Fi5
Liisa	Fi6	Fi2	Fi1
Petra	Fi7	Fi3	-
Hanna	Fi8	Fi6	-
Elsi	Fi9	Fi5	-
Aamu	Fi10	-	-
Jonna	Fi11	-	-
Noora	Fi12	Fi1	-
Tiina	-	Fi4	-
Outi	-	Fi7	-

TABLE 5 Finnish participants' pseudonyms in the different studies.

There was no control group of Finnish students remaining at home during the time the others were studying in Russia because the aim was not to show statistically significant differences between students who went abroad and students who stayed at home but rather to examine the development in students' speech during SA. Also, finding a control group of Finnish students not participating in a study-in-Russia programme would have been difficult if not impossible, as at the time, studying in Russia was an integral (and compulsory) part of university studies in all Finnish universities for students majoring in Russian.

As it is, the sample can be considered to be fairly representative of the Russian students at the university in question in terms of their reading aloud and pronunciation skills. Most students spoke Russian with a Finnish accent, some having more difficulties than others in reading the texts. During the two-year period, all the Russian majors at the university were given the possibility to participate in the study. The subjects were recruited on a volunteer basis and

motivated by offering them the possibility to receive feedback on their pronunciation after all the recordings had been done.

3.1.2 Russian Speakers (Studies IV and VI)

In Studies IV and VI native Russian speakers were used as a control group of L1 speakers for comparison. I recorded a total of seven L1 Russian speakers, who read in pairs the same Russian dialogues as the Finnish students. All the speakers were used in Study IV, but one speaker's recording was left out of Study VI for technical reasons. The pseudonyms used for the Russian speakers in the two studies are shown in Table 6 below.

Pseudonym Study IV	Pseudonym Study VI	Age	Home city	Place of academic studies	Occupation
Ru1	-	23	St. Petersburg	Philological Faculty	Designer
Ru2	Ru2	28	St. Petersburg	Pedagogical Faculty	Student
Ru3	Ru3	24	St. Petersburg	Philological Faculty	Translator, interpreter
Ru4	Ru4	24	St. Petersburg	Technical University	Manager
Ru5	Ru5	19	Tver	Faculty of foreign languages	Student
Ru6	Ru6	20	Tver	Faculty of foreign languages	Student
Ru7	Ru1	23	St. Petersburg	Philological Faculty	Translator

TABLE 6 Russian participants' demographic information, academic profiles and pseudonyms in the different studies.

The Russian speakers were phonetically untrained native female speakers of Russian from St. Petersburg (five speakers) and Tver (two speakers) aged 19–28 (mean 23 years) at the time of the recordings. They had lived most of their lives in the region either of St. Petersburg or Tver. The dialectal differences in prosody between the speakers from the two cities are not great according to a subjective evaluation[39]. Most of the Russian subjects had a Bachelor of Arts degree from a philological faculty of a Russian university and were currently either students or working with L2s[40]. Table 6 shows the Russian participants demographic and academic profiles in more detail.

---

39 The phonology of Russian dialects has mostly been studied on the segmental level, and the main regional differences between St. Petersburg and Tver are observed in vowel reduction (see Panov 1968, 183–211, for a more detailed description). Prosodically, these dialects have been said to differ in speech rate and grouping speech into syntagmas (i.e., intonational units) (Shaul'skiy & Knyazev 2006, 4), and in the position or realisation of word-stress (Kolesov 1972, 86–90).

40 Note that there is a mistake in Study VI, where they are all said to be university students.

3.1.3 Background Questionnaires

All the recorded speakers were asked to fill in background questionnaires. For the Finnish students these helped to determine their speaking activity with L1 speakers and their fluency self-assessment. The questionnaires were filled out in Finnish in connection with each recording session (either on paper or online). The students were asked various questions (both open and multiple choice) about their language learning background and for self-assessment of their pronunciation skills and development. There were about 35 questions in total. Some of the information obtained through the questionnaires is not reported here. In most studies, the following background information was used: the students' age, the length of studying Russian prior to university studies, their previous visits to Russia, their L1 and their accommodation in Russia (host family vs. foreign-student dormitories) during SA. The questions concerning the students' self-assessment and language behaviour in Russia were addressed in Study III. These questions dealt with the students' perceptions regarding their language use in Russia, improvement in their pronunciation skills and their approach to learning pronunciation. Other background information on the Finnish subjects is presented in two conference papers (Ullakonoja 2007a; Ullakonoja 2008a) that are not included in the present study.

The Russian speakers filled in a short background questionnaire in Russian, in which they were asked, e.g., where they have lived during their lives, age, occupation and educational background. The background information obtained through the questionnaires was used only to be able to describe the Russian speakers' background briefly in the present chapter and to confirm that they were L1 speakers and had lived most of their lives in the St. Petersburg or Tver regions.

3.1.4 Texts for the Reading Task

In the recordings, the Finnish subjects were asked to read written dialogues in pairs (two in Russian and one in Finnish). The Russian dialogues were telephone conversations (dialogues 46 and 100) taken from Russian as L2 teaching material (Shilova & Usmanova 1990). In the dialogues two assumedly middle-aged women have telephone conversations about everyday life. The lexical stress was marked in the original texts as it usually is in Russian L2 materials. The dialogues were chosen, first, because they contained different utterance types (including YNQs) and second, because their lexical content was deemed suitable for the speakers' proficiency level.

The Finnish dialogue was written by the present author. The goal was to design a text that would be close to the students' everyday speech and would contain different clause types and take the form of a dialogue between two people where overlapping speech could be avoided. The style of the Finnish text differed from that of the Russian texts. The Finnish text was closer to a spoken dialogue between two young people. The texts were different, because I wanted texts that would be "easy to read". It was thought that the students

would be more comfortable reading a text in Finnish which was written "in the way young people speak" than one written in a more formal style[41]. In Russian, on the contrary, it might have been hard for the students to read a text written in the way young people would speak; hence, a text from teaching material was used, in the belief that it would better correspond to the style they were more used to reading.

All the dialogues were analysed entirely only in Study IV. In the other studies only parts of the dialogues were used. In Studies I, II and III one turn in one of the Russian dialogues was used and in Studies V and VI only the YNQs of the two Russian dialogues.

To keep the fluency evaluation task to a reasonable duration, for each student only one turn in one of the Russian dialogues was chosen for the analysis in Studies I, II and III. However, to be able to study fluency, both perceptually and acoustically, a turn as long as possible was needed. The longest continuous sequence of speech in the Russian dialogues consisted of six sentences. It was a response to the interlocutor's question *Слушай, а как мы раньше жили без телефона? (Slushay, a kak my ran'she zhili bez telefona?) Listen, how were we able to live without a telephone before?*

The turn chosen for the analysis in Studies I, II and III was:

> Не представля́ю себе. Ну ла́дно. А заче́м я, со́бственно, тебе́ звоню́? Ах да, насчёт Анны. Она́ уезжает не сего́дня ве́чером, а за́втра у́тром. Так что, если хо́чешь её проводи́ть, приходи́ к нам у́тром, часов в де́вять.
>
> Ne predstavlyayu sebe. Nu ladno. A zachem ya, sobstvenno, tebe zvonyu? Ah da, naschet Anny. Ona uezzhaet ne segodnya vecherom, a zavtra utrom. Tak chto, esli hochesh' ee provodit', prikhodi k nam utrom, chasov v devjat'.
>
> 'I can't imagine. Oh well. And why am I calling you in the first place? Oh yes, about Anna. She is not leaving tonight, but tomorrow morning. So if you want to see her off, come to our place in the morning at about nine o'clock.'

In Study II, an extract from the Finnish read-aloud dialogue was also analysed for comparison. As no single turn was as long as the Russian turn under scrutiny, in the Finnish texts two turns[42] by each speaker were chosen. These

---

41 Spoken Finnish differs from written standard Finnish a great deal.

42 The two Finnish turns analysed in Study II were:
Ai oli vai? Ei musta... Musta tuntu etten mä osannu mitää. Hyvä nyt kysyy jotai ihmeen zoologisia teorioita, joista mä en oo koskaa kuullukaa… 'You think so? I don't. I think I couldn't answer any question correctly. What's the point in asking about some zoological theories that I've never even heard of? (Previous turn of the interlocutor was: Ai nii, olinhan mä. Se oli kyllä tosi helppo! 'Oh yes, I was too. It was really easy!'.)
No en varmaa osta! Mulla menee se kokonaa elämiseen. Mä lähen nyt kotii. Nähääks huomenna? 'I certainly won't buy that. It'll all go on living costs. I'm going home now. Will we meet tomorrow?'(Previous turn of the interlocutor was: No mut kohtaha tulee taas opintotuki, ostasit vaikka sellasen DVD-soittimen, ku niillähä voi soittaa CDtä. 'Well, but you'll soon receive your monthly study allowance, why don't you buy a DVD player, 'cause they'll play CDs'.)

two turns were selected because they were fairly long and corresponded approximately, as measured in word length, to the Russian turn above.

In Studies V and VI, which investigated pitch patterns in YNQs, only the Russian dialogues were used. In fact, the studies employed only those YNQs of the data that depending on the pitch contour could also be interpreted as statements. As the analysed questions are listed in Study V (Study V: Table 1) and Study VI (Study VI: Table 1), along with their English translations, I have not rewritten them here.

Thus, the corpus was used selectively. However, the Russian texts were purposefully longer than needed in order to be able to analyse the corpus further in future works. Furthermore, they provided enough material in an authentic textbook context (closely resembling an authentic telephone conversation) for the studies presented here.

3.1.5 Recording Procedure

In this section I report on how the reading of the texts described above was recorded. As Table 7 (p. 66) shows, Finnish subjects were recorded either three or four times at different stages of their university studies:

1) beginning of their university studies (half of the group, those who started their studies in 2005)
2) prior to the stay in Russia (at the end of the first year of their studies),
3) after about one month's stay in Russia (during their second year) and
4) following the stay.

As the first recording was done only for half of the Finnish students, only the data on the Finnish dialogues were analysed from it. That is, Study II used the data from the first time each intake year group read the Finnish text (i.e., either the recording done in the beginning of their university studies or prior to the stay in Russia, depending on the group). For practical reasons the recordings done prior to the stay abroad took place four months before the students went to Russia and the recordings done following the stay took place about a month after their return to Finland. Although this is not the ideal setting, the students did not spend the intervening time in Russia or were not even in contact with Russians to any considerable extent.

During the recordings the students read the three dialogues described above (one in Finnish, about 3 minutes, and two in Russian, together about 7 minutes) in pairs. The pairs were self-selected in the first recording session, and if possible, the same pairs were recorded in the remaining sessions. The Russians also self-selected their partners: they could bring a friend with them to the recording sessions. For the purposes of the study it was considered best to record the dialogues separately for the Finnish and Russian speakers. Having a Russian interlocutor in the Russian dialogues would have given a pronunciation model for the L2 learner, and hence, it would have been

impossible to control if the learner was mimicking the native speaker or producing speech "on her own". Of course, the chosen setting means that the Finnish speakers might have adjusted their speech towards Finnish when speaking Russian, and possibly the results would have been different if a native Russian interlocutor had been used instead (see Grosjean (2001) for more discussion on possible accommodation depending on the language of the interlocutor).

Recording	Referred to in Studies I-VI	Time of the recordings	Place of the recordings	Participants	Task
1	Study II: Finnish recordings (half of the group)	Beginning of 1st year at university (September 2005)	Jyväskylä	some of the Finnish students Study II (n = 7)	Reading the Finnish and Russian dialogues with another Finnish student
2 & 6	Study I, II & III, IV: before the stay	End of the 1st year at university (April, May 2005-2006)	Jyväskylä	all the Finnish students Study I, II & III (n = 12)	Reading the Finnish and Russian dialogues with another Finnish student
3 & 7	Study I, II, & III: middle of stay Study V & VI: T1	In the middle of the stay in Russia during the 2nd year at university (October 2005-2006)	Tver	all Finnish students Study I, II & III (n = 12) Study V & VI (n = 6)	Reading the Finnish and Russian dialogues with another Finnish student
4	Study IV: native Russian, Study VI: L1	October 2005	St. Petersburg	native Russian speakers Study IV (n = 5) Study VI (n = 4)	Reading the Russian dialogues with another Russian speaker
8	Study IV: native Russian, Study VI: L1	October 2006	Tver	native Russian speakers Study IV & VI (n = 2)	Reading the Russian dialogues with another Russian speaker
5 & 9	Study I, II, III & IV: after the stay Study V & VI: T2	After the 3.5-month-stay in Russia (January 2006-2007)	Jyväskylä	all Finnish students Study I, II & III (n = 12) Study V & VI (n = 6)	Reading the Finnish and Russian dialogues with another Finnish student

TABLE 7 Summary of the recordings.

The same texts were read in all the recording sessions and the texts were given to the speakers one at a time. The subjects were not told that they were to read the same texts each time, nor were they given access to the texts during the intervening period. The Russian speakers read only the Russian dialogues. All

the recordings were done in the same way: the pair read the Finnish dialogue first (often without rehearsing), then the first Russian dialogue and, after that, the second Russian dialogue. The same Finnish dialogue was recorded each time for the Finns, to ensure the same recording context. The dialogues were read so that after the first reading, the subjects changed roles. Thus, both speakers read all the turns of both interlocutors in the dialogues. The subjects were instructed to concentrate on utterances and intonation rather than on single sounds, to repeat the whole turn/sentence in the event of mispronunciations or hesitation, and in general to speak as naturally as possible.

The subjects were given time to prepare for their performance. They could also practise reading the texts as many times as they wished, and ask the researcher for the meaning or the pronunciation of a single word. Also, the time they could spend on reading the material was not limited: they could read and reread the texts as many times as they wanted until they were satisfied with the result. However, many of them were satisfied with the first recording and did not wish to rerecord. According to Blum & Koskinen (1991) and Golman Eisler (1968, 15), rereading the text and familiarity with its content increases reading fluency and decreases the frequency and duration of pauses. Hence, it can be suggested that the present students' performance was about as good as they were capable of since they were able to familiarise themselves with the text before the actual recording took place.

In Finland, all the recordings were done in the same studio and each time with the same equipment (computer equipped with Adobe Audition 1.0 and 2.0, microphones AKG GN30). Each subject had her own microphone. The recordings done in Russia (for all the Russian speakers and in the middle of the stay for the Finns) were done under different circumstances. In 2005 in St. Petersburg and Tver the speakers were recorded with a Sony TCD-D3 DAT recorder and a Sony ECM-959A microphone. In 2006, in Tver, a Roland Edirol 24-bit Wave/MP3 digital recorder R-09 was used with a Sony ECM-959A microphone for the recordings, which were done in a quiet hotel room. The two channels were extracted from the stereo sound file with Adobe Audition 1.0 and Audacity 1.2.4. In the recordings the sample rate was set to 44100 Hz with a 16 bit resolution on the computer on the Roland Edirol digital recorder.

To recapitulate, Table 7 (p. 66) gives a summary of all the recordings done during the present study. The recordings took place altogether nine times during the three years, because the Finnish participants were from two consecutive intakes.

3.1.6 Read-aloud Speech as Data

The present speech corpus consists of recordings of read-aloud speech. However, today in phonetic research spontaneous data is often preferred to read-aloud speech. This seems to be the case, especially when investigating prosody and fluency. Read-aloud speech has traditionally been used in segmental studies, where it is useful for analysing the same sounds in the same context. I chose nevertheless to use read-aloud speech in this prosodic study.

First, this enables comparison of the pronunciation of a given text and fluency of reading rather than aspects of speech planning. Given that the previous fluency studies (Riggenbach 1991; Freed 1995) have shown that fluency ratings are also affected by the linguistic choices of the speaker, in read-aloud speech it is possible to control for linguistic content. Second, I wanted to investigate interspeaker variation with respect to the same utterances in different stages of L2 learning. Moreover, the choice of read-aloud material also facilitated the comparison of prosodic development.

Despite its many advantages, the limitations of read-aloud data should, however, be acknowledged. First, the naturalness of the speech produced is limited because the learners do not focus on what they are saying, but rather how they are saying it. Second, the speaking context itself is not natural because the speakers are not used to recording their speech in a studio. They might not pause or use pitch contours that are similar to those they would use in a real-life situation. In the present study the speakers were acquainted with the laboratory setting and they were also given time to practise reading. Also, the Finnish subjects were used to reading Russian texts aloud in the classroom. Further, spontaneous speech studies also frequently use a laboratory setting because of the quality requirements of the recordings.

A potential disadvantage of using read-aloud speech is that one can never be sure whether the results reflect fluency of reading or fluency of pronunciation. However, they do not reflect knowledge of syntax or lexis to the same extent as spontaneous data. Also, when speaking spontaneously, speakers pay attention to different aspects of their L2: some concentrate on grammar while others focus on finding the right words (Möhle 1984) and hence, their difficulties in speaking fluently may be due to different causes. When reading aloud, the speakers are not focusing on these aspects, although some might be worried about correct pronunciation.

When reading aloud a speaker does not structure the utterance, or plan the content, but merely decodes what is written and articulates it. In contrast, in spontaneous speech there are several stages before the actual articulation of speech sounds, such as, e.g., constructing an utterance on the basis of the speaker's communicative intention with the help of the semantic and syntactic information the speaker has (Levelt 1989). Although Levelt's (1989) speech production model illustrates spontaneous speech, he describes the reading task aptly when contrasting it to spontaneous speech (Levelt 1989, 259):

> "In reading, the speaker can rely heavily on the printed materials. Lexical retrieval and the building of syntactic constituents can be based largely on parsing of the visual input. Reading aloud is primarily a perceptual, phonological, and articulatory task." (Levelt 1989, 259.)

In other words, reading aloud consists of the processes of 1) perception of the text, 2) phonological encoding of it, and 3) articulation. If one is interested in studying articulation, using read-aloud speech as data would make it less complicated in the sense that there are, according to Levelt (1989, 259), fewer processes involved than in spontaneous interaction, where the speaker needs to

plan the lexical, semantic and grammatical content of his/her message before phonological encoding and articulation.

I now turn to L2 reading aloud and the factors that may have influenced the speech of my subjects in the recordings. One potential factor that influences the reading process is the alphabet, because the subjects are learning an L2 with a different alphabet than their L1 or other L2 they have studied previously. For example, Perfetti (2003) claims that reading involves the cooperation of two systems: a language, and the writing system that encodes it. The verbal processes present in a reading task are "general symbol activation and retrieval, letter recognition, word decoding and semantic access" (Perfetti 1985, 169). With respect to the reading task in the present research (L1 Finnish students reading a Russian text), it is possible to say that at least the first two of these processes are affected by the Cyrillic alphabet of L2 Russian that differs from the Roman alphabet of their L1 or other L2s. It is evident that this feature of the text has an effect on its reading and, hence, on pronunciation and perhaps also on speaking rate. It can be assumed that reading a L2 text with a Roman alphabet would be easier as a reading process for Finnish students than reading a Cyrillic text (see, e.g., Perfetti 1985, 88–90 for the importance of knowing the orthographic rules of the language being read).

Finnish is a language spelled with an orthography that has a highly regular correspondence between letters and sounds (VISK § 7). Russian, on the other hand, is more complex in this sense as it does not have as a high correspondence between orthography and pronunciation, and hence the Russian orthography can be called morpho-phonemic (see, e.g., Kasatkin 2003, 210–216). There are some phonotactic rules in Russian that are not visible in the orthography. For instance, Russian vowel articulation is determined by word stress[43] in that unstressed vowels are reduced in quality and quantity. Also, Russian consonant articulation is affected by the regressive voicing assimilation, which means that consonant becomes voiced if the following consonant is voiced (with the exception of resonants and voiced labio-dental fricatives) and vice versa: a consonant is devoiced if the following consonant is voiceless (see, e.g., Burton & Robblee 1997; Bondarko 1998, 119–122). Thus, the differences between the Finnish and Russian orthographies are likely to present a challenge for Finnish L2 learners of Russian.

As adult L1 speakers of different languages have been found to read at a very similar speed in their L1, the writing system itself does not seem to affect the reading rate once it is mastered (Gray 1956, cit. by Just & Carpenter 1987, 290). However, the number of characters in a writing system as well as similarities between characters can influence the beginner reader (Just & Carpenter 1987, 290). This can also influence L2 readers (see, e.g., Varyushenkova & Lyubimova 1986). In the present study, the Finnish students who were reading in Russian were, of course, familiar with the Cyrillic alphabet, but may have been distracted by similarities between their L1 and L2 alphabets.

---

43 In most textbooks and other L2 Russian for foreigners materials word stress is marked with ´, but in regular Russian texts it is not.

For example, the graphemes *p*, *y* and *c* are the same in both alphabets in most fonts, but correspond to different sounds in Russian and in Finnish. Some graphemes of the two alphabets are alike and represent similar sounds (for example *a, o* when they occur in the stressed position in Russian), other graphemes are alike but represent completely different sounds (for example *m, x*) and some graphemes in the two languages are completely different (for example *ä, ö* and *я, г*).

Finally, it is interesting to consider how SA might influence reading skills. I agree with Huebner (1995), who concluded that the L2 environment offers many possibilities to learn to read a foreign script, which again helps the learner to improve his/her reading skills. He adds that being in the SA context emphasises the need to be literate in L2, which promotes the reading skills of an SA group. The text then can become easier to read because the students' vocabulary expands and their language skills improve. It can be concluded that reading in L2 differs from reading in L1 in a number of ways and can be further complicated by different L1 and L2 alphabets.

3.2 Acoustic Methods

This section explains the principles of segmentation and the acoustic analysis of the speech data. As Studies I–VI themselves do not provide a detailed description of the acoustic methods used, I discuss this background here more thoroughly. The section commences with an introduction of the Praat software used for acoustic analysis. Then, the data annotation procedure in Praat is described. The last two subsections focus on ways of measuring pausing, speech and articulation rates and pitch.

3.2.1 Analysis Tool – Praat

For the segmentation and acoustic analysis, the computer software developed for phonetic analysis, Praat (Boersma & Weenik 2009, versions 4.3–5.1), was used. With the programme one "can analyze, synthesize, and manipulate speech, and create high-quality pictures" (Boersma 2007). Only a small part of Praat's features, namely annotation, measuring pitch and duration, were used for studies reported in this dissertation. One reason for choosing Praat as an analytical tool was that it is widely used in experimental phonetic research all around the world. Apart from its free availability, its other advantages are easy download and installing, and regular upgrades. It works very reliably and has been credited, e.g., with having an efficient F0 analysis algorithm (Boersma 1993) that uses autocorrelation (see, e.g., Ladefoged 1996, 148–151); this was used in Study VI. Praat works well both on Macintosh and Windows operating systems, which was also an important advantage. The authors of the programme as well as an Internet user group provide support for its users.

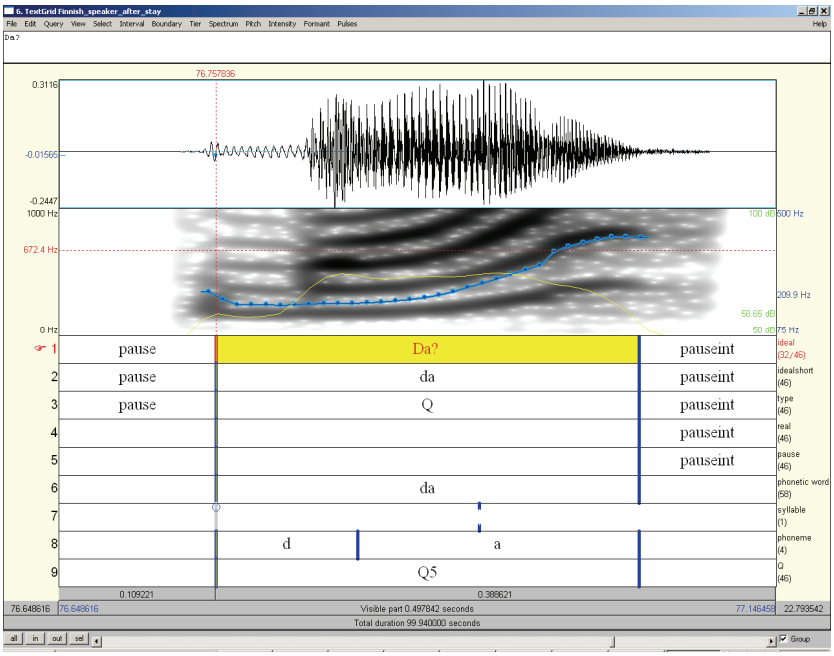


FIGURE 4 Window of the Praat programme representing sound and TextGrid files of the sentence *Да?* (*Da*?[44]) in Russian produced by a Finnish speaker.

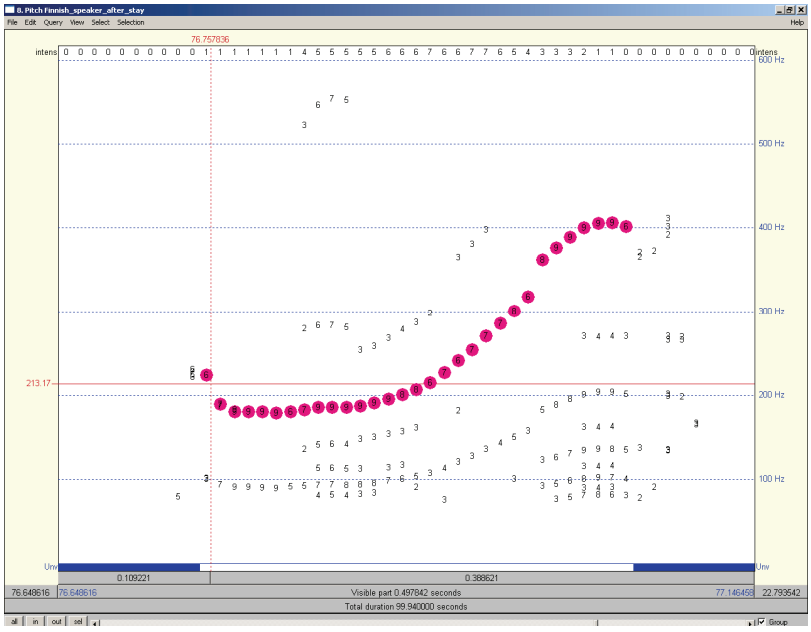


FIGURE 5 Window of the Praat programme representing the PitchObject of the sentence *Да?* (*Da*?[44]) in Russian produced by Finnish speaker.

44 In English: Yes?

Furthermore, the possibility to use simple text-based scripts with Praat was one reason for choosing it for this study. Together with the above-mentioned arguments in favour of Praat, the facts that I had used Praat when working on my Master's thesis and that I have been able to take a course on Praat scripting, were also reasons for preferring it over other programmes[45]. Figure 4 illustrates a Praat analysis window with both the sound and TextGrid files. The higher part of the screen represents the waveform, the spectrogram is seen in the middle and the lower part shows the annotation of the TextGrid (in this case on 9 tiers). On the top of the spectrogram (middle part) the pitch is represented with a blue line and intensity with a yellow line. The TextGrid is a text file containing the labelling and boundaries of each marked interval in the sound file. The purpose of the TextGrid is to help the user to find the correct places in the sound file after having marked boundaries and added labels relevant to his/her research. The TextGrid also contains a time scale, so that the time location of each boundary and the duration of the interval can be measured from the TextGrid alone.

In the pitch calculations, Praat was used to create a PitchObject for each of the sound files. It is possible manually to check and correct the pitch, if the PitchObject is open in one window and the sound file (together with TextGrid) in another. Figure 5 shows the PitchObject editor window (same sentence of the same speaker as in Figure 4) where the programme has marked all the possible candidates for pitch (different numbers), but estimated the most likely contour (in pink). The user can click on the numbers to change the estimated pitch contour. The blue line at the bottom of the screen shows the unvoiced parts. Furthermore, Praat was used in the present study to draw pictures of the pitch contours in the Praat picture window.

3.2.2 Data Annotation

Most of the speech data were segmented manually in Praat into the TextGrid files. A total of nine tiers[46] were annotated for each sound file in TextGrids: 1 – "sentence", 2 – "sentence short", 3 – "clause type", 4 – "real", 5 – "pause", 6 – "phonetic word", 7 – "syllable" 8 – "phoneme" and 9 – "Q" (short for question) (see the horizontal blue bars in Figure 3). Tiers 3, 5, 6, 7 and 9 are the most relevant for this study.

Some segmentation was done first automatically with scripts, and then checked manually. A *script* is a text file that gives commands to the programme and thus automates processes that the user would otherwise carry out manually step by step. The user can utilise already existing ready-to-use scripts that are widely available on the Internet (see, e.g., Lennes 2007), or modify them or write completely new scripts, for example, with the help of the history-

---

45 There are also other efficient programmes for speech analysis, e.g., Intelligent Speech Analyser™ (ISA) (Toivonen 2007), COLEA: A Matlab Software Tool for Speech Analysis (Loizou 2008), Speech Filing System (SFS) (UCL 2008), WaveSurfer (Sjölander & Beskow 2006) and EDSW (DSP Center 2007).

46 A tier is "a level or layer of segmentation".

command that saves everything the user does with the programme into text format. The annotation was done following the instructions of Lennes & Ahjoniemi (2005).

First, the script "mark_pauses.praat" (Lennes 2007) was used to mark intervals longer than 200 ms as pauses. The annotation was verified and corrected manually. The actual annotation process commenced on the basis of the written text. The original text was transcribed from the Cyrillic alphabet using the transcription conventions generally used in the Anglo-Saxon world (PCGN 2009), as these do not require the use of diacritics or other special symbols difficult to use in Praat. For some of tiers (i.e., levels of segmentation) it was possible to label the marked intervals automatically with a script "label_from_textfile.praat" (Lennes 2007). Of course, while it was not possible to label all the tiers automatically (e.g., pausing, phonetic word and syllables), the automatic labelling, which created similarly labelled intervals for all the subjects, helped, e.g., to find the parts needed for analysis.

The text was then divided into sentences (1st tier), and the clause types were roughly grouped into three categories: Q = question, D = declarative and E = exclamation (3rd tier). This utterance type classification has been used, e.g., by Makarova (2001). Also, to be able automatically (with a script) to extract utterances and name the resulting sound files, a shorter annotation for each utterance was needed (2nd tier). The abovementioned annotations were the same for each speaker and done on the basis of the original text, not the actual acoustic signal. An utterance was used as the unit of intonation research, instead of prosodic units (such as tone group/unit, intonation unit, etc.) to be able to compare intonation and pausing between speakers and not the way they structure speech (i.e., where they have boundaries).

The 4th tier (*real*) was annotated only for some speakers, and for some of the material for a different study (Kärkkäinen et al. 2007) on the basis of subjective perception. On the 5th tier (*pauses*) the turn-internal pauses were classified (following Riggenbach 1991) into disfluent or fluent and labelled accordingly. Pauses shorter than 200 ms were detected auditorily: everything that was subjectively perceived as a pause was annotated without using, e.g., a threshold value for pause duration. As more precisely defined in subsection 2.1.1 (p. 29), a fluent pause refers here to a pause that occurs at syntactic or phrasal boundaries whereas a disfluent pause is a pause existing elsewhere. Therefore, following Strangert (1990) a pause here is a "perceived pause" rather than an acoustically silent interval. After pause annotation each sound file was edited so that the other speaker's voice was removed from the file. In other words, pauses that were not turn-internal were set to zero with a script (set_pauses_to_zero.praat). This script was written from scratch by the author in collaboration with Hanna Anttila (Ullakonoja 2009b). The script checked the annotation tier, and if there was a 'pause' in the annotation, it set the sound to zero at that point, leaving a completely silent interval on the sound file. Turn-internal pauses were labelled 'pauseint' in order for them not to be removed.

On the 6th tier *phonetic words* were annotated. The phonetic words in the Russian material correspond to what is called *фонетическое слово* (*foneticheskoe slovo*) in the Russian research tradition (see, e.g., Avanesov 1956, 61). In the western tradition the term *prosodic word* usually refers to a similar unit. A phonetic word usually corresponds to a lexical word, but it may also refer to some two-word combinations where an unstressed particle or a preposition is pronounced together with the main word. For example, in the present material the preposition and pronoun *к нам*[47] (*k nam*) are treated as a phonetic word. A phonetic word has one word stress (or lexical stress). In the Finnish sample, it was decided that lexical words always correspond to phonetic words in the annotation. This might have affected the results, since sometimes the three-word sequence *mä en oo*[48] was pronounced more like [men:o:] and could perhaps also have been treated as one phonetic word. The choice of annotating the lexical words in Finnish was made in order to be more systematic: it would have been impossible to define exactly when the sequence above would be one, two or three phonetic words if the annotation of phonetic words had not been done on the basis of the lexical word principle. On this tier, possible creaky voice was marked in brackets after the annotation. Also, turn-internal pauses and their duration in seconds were annotated on this tier for the parts that were used in Studies V and VI.

On the 7th tier (*syllables*) (which was a point tier[49]) the syllable nucleus was marked with a point. That is, the exact syllable boundaries were not determined because it was possible to calculate the speech and articulation rates by comparing the number of syllables (i.e., syllable nuclei[50]) with the sample duration and pause duration.

The 8th tier (*phonemes*) was annotated only for some of the data for Study VI. This was done in cases where the position of the peak was not clear from the phonetic word annotation. They were annotated manually by listening to the sound file and examining the spectrogram. The Roman alphabet was used in the annotation. In addition to Study VI, some of the phonemes were annotated for the purposes of a different study (Kärkkäinen et al. 2007).

The 9th tier (*questions*) was added for the purposes of creating the perception experiment in Study V and running the script in Study VI. There, YNQs of the data were coded as Q1–Q7.

Thus, the data annotation was done selectively so that only the first three tiers were annotated for the entire corpus. The other tiers were annotated only when needed. The annotation will be continued in the future when more analyses are done on the basis of the corpus.

---

47 In English: 'to us'.
48 In English: 'I'm not'.
49 The other tiers were interval tiers, where the boundaries and the intervals between them were annotated. A point tier is a tier where the boundaries are not marked, but instead a certain place, a point in the acoustic signal, is labelled.
50 Counting syllable nuclei instead of syllables has been used by, e.g., Simões (1996).

3.2.3 Analysis of Pausing and Speech and Articulation Rates

Below I explain the computing of pause and syllable frequencies and speech and articulation rates. The number of syllable nuclei and the number of phonetic words, as well as the total duration of pauses and total duration of speaking time were obtained automatically by using Praat scripts. The script gives the result in the form of a text file containing the duration and label of each interval. The resulting text file can be exported to Excel or SPSS for further analysis. A script calculate_segment_durations.praat (Lennes 2007) was used in Studies I and II after it had been modified to better suit the purposes of this research.

The calculation of speech and articulation rates can be expressed with the help of the following equations (see, e.g., Grosjean & Deschamps 1975; Grosjean 1980b, 40–41; Towell et al. 1996):

Speech rate (syll/s) = Number of syllable nuclei/total speaking time
Speech rate (PW/s) = Number of phonetic words/total speaking time
Articulation rate (syll/s) = Number of syllable nuclei/(total speaking time - total pause time)
Articulation rate (PW/s) = Number of phonetic words/(total speaking time - total pause time)

The analysis of speech and articulation rates includes the comparison between L1 and L2 and comparison between the rates of the same student at different stages of stay abroad. This was computed by ranking the students from fastest to slowest and comparing the rankings within the group.

In Study II, I measured speech rate both in phonetic words per second (PW/s) and syllables per second (syll/s) as I wanted to avoid the influence of language-specific features on the results that have been encountered in similar comparative studies such as that of Lehtonen (1981). As Crystal (1969) has pointed out, in addition to measuring words per time unit, pausing and the timing of syllables should also be taken into account when defining speech rate. It was considered sufficient to restrict the analyses to syllables and phonetic words in this study. However, it has to be acknowledged that, among others, Pfitzinger (1999) has defined speech rate in terms of phonemes per second and syllables per second, claiming that combining these two would best correspond to the perceived speech rate, and thus would be a more suitable measure for comparing different languages than using phonemes or syllables alone (see also Pfitzinger & Tamashima 2006). Here calculating the number of phonemes would have been difficult in some cases. For example, when in Finnish, *mä en oo* [51] was pronounced more like [mäen:o:] or in L2 Russian, where mispronunciation, e.g., *торжественному* (*torzhestvennoti* [52]) pronounced

[51] In English: 'I'm not'.
[52] In English: 'festive'.

[tʌrʒevᵉnːoᵐ] or self-correction, e.g., *у тебя* (*u tebya*[53]) pronounced [utıp utıbja] caused sequences where the number of phonemes was unclear.

3.2.4 Pitch Analysis

Researchers have long been in search of pitch measurement techniques that would convey linguistically relevant pitch movements. So far, no single model or technique has been found confidently to determine the changes in pitch that would be linguistically significant. The fact is that acoustic pitch contours show a lot of information that is linguistically wholly irrelevant, whereas phonological transcriptions (such as, e.g., ToBI) simplify the picture and have not been widely applied to all language or L2 speech. Hence, there is a multitude of ways in which pitch contours can be analysed. In this subsection, I briefly explain and justify my choices in the pitch analysis.

I begin with the calculation of pitch. In Study IV and VI pitch was calculated for each sound file using Praat's autocorrelation[54] method (Boersma 1993). The same settings were used in calculating pitch for all the speakers (Time step automatic, pitch floor: 75 Hz, pitch ceiling: 600 Hz). That is to say, the settings were not adjusted to each speaker's pitch range. This, however, could be done in future analysis of the corpus, for example by using the two-pass method suggested by Hirst (2007)[55]. A floor value of 75 Hz was used, because Finnish women can have relatively low pitch. Other methods of pitch calculations, such as the Target approximation model developed by Prom-on, Xu & Thipakorn (2009) and Momel by Hirst (2007) are useful in modelling intonation, but Praat's method was considered sufficient for this study. Praat's pitch calculations were checked manually by observing the spectrogram together with the PitchObject file.

However, creaky phonation, which has both previously and in this study been found to be a typical feature of Finnish (see, e.g., Ogden 2003; 2004) (but not Russian) subjects' speech, influenced the calculations done by Praat. In Study VI, creaky voice was manually corrected to the "normal" and "continuous" pitch level of the utterance on the basis of the spectrogram, or in cases where this was not possible, it was unvoiced for the pitch measurements. In Study IV, I wanted to show the influence of creaky voice on the Finnish subjects' pitch, and therefore no corrections were made manually to the pitch objects Praat produced. The third possibility for dealing with creaky voice would have been to unvoice all the parts where it occurred. However, as Anttila's (2008, 100–117) study showed, there was no great difference in the F0 results of L1 Finnish utterances depending on a method chosen, even though it affects them to some extent.

Pitch normalisation vis-à-vis pitch was not conducted in this study. Pitch could have been normalised, e.g., in relation to mean pitch, median pitch,

---

53 In English: 'you' (in genitive preceeded by preposition 'u' meaning have)
54 For different ways of automatic measurement of pitch see, e.g., 't Hart (1990, 22–23).
55 This, however, was not available the time Study IV was published.

maximum pitch, minimum pitch or the pitch range of each speaker (see, e.g., Crystal 1971; Jassem & Kudela-Dobrogowska 1980). In the present interspeaker comparisong it was considered sufficient to use a logarithmic scale without normalisation. Hence, in this work pitch has mainly been measured using semitones[56] (ST), as this has proved to provide the best representation of auditory pitch perception (Nolan 2003). However, in Study IV, Hertz (Hz) values are reported when determining the typical pitch level of a speaker.

In Study IV, the script "draw_f0_curves_from_files.praat" (Lennes 2007) was slightly modified by the author for drawing the superimposed pitch contours (see Figures 1–4, Study IV). The script provided time normalisation of the pitch contour so that it constrained all the utterances under investigation to be of the same duration. Furthermore, it made it possible to draw different utterance types in different colours. For running the script, the utterances were first extracted automatically from the long sound file into shorter sound files using another script, "save_labeled_intervals_to_wav_sound_files.praat" (Lennes 2007), which saved each utterance of the sound file with the label of the 3rd tier (Q, D or E, see p. 73). In Study IV the pitch calculations were done for each utterance separately and not corrected manually.

For pitch range, different statistical values have been used in earlier investigations. Some scholars have used the mean (Bezooijen 1995) or median (Carlson et al. 2004; Lennes 2009), others (van Heuven & van Zanten 2005) 95 % of the pitch values around the mean, or the range between the lower 25 % and upper 75 % of the pitch points (Carlson et al. 2004). Mennen et al. (2007), on the other hand, have calculated the 80 % range (excluding the top and bottom 10 % from the analysis), interquartile range and std (+/- 2) around the mean in ST for determining pitch range.

Mennen et al. (2008) as well as Patterson (2000) and Patterson & Ladd (1999) have also developed methodologies for investigating pitch range. They conclude that the above-mentioned long-term distributional (LTD) measures (mean F0, F0 maximum and minimum, 90 % or 80 % range) do not best suit cross-language comparisons of pitch range, because they rely on the assumption that the F0 data is normally distributed. Instead, Mennen et al. (2008) and Patterson (2000) propose using different linguistic measures for between-speaker and cross-language comparisons. Patterson (2000, 89–90, 148, 153) suggests that the best measure of between-speaker comparisons of pitch range is the difference between two measuring points in pitch: non-sentence initial high and the post accent valley. Mennen et al. (2008, 529) propose the use of the "difference between the average of post-accent peaks (H) and post-accent valleys (L)" for comparison of German and English pitch range. The main argument put forward by the two studies is, however, the same: investigation and modelling of linguistically significant pitch turning points are more important than absolute F0 values.

---

[56] Other logarithmic scales, such as the ERB and Mel scales, also reflect relative perceived pitch (Lieberman & Blumstein 1988, 154).

There would have been a number of other ways to study pitch contours. For example, Iivonen (2001c; 2005, 123) has successfully visualised three utterances of the same speaker by measuring the F0 syllable by syllable. Thus, the established three pitch contours were drawn by having all the pitch contours of the speaker in the background. In this way, the position of the utterances within the speaker's pitch range could be clearly seen. This method, however, would have required syllable level segmentation, which was thought to be both methodologically difficult and very time-consuming, and hence was not done in the present study. My decision to compare pitch contours simply by drawing them and visually inspecting them originated from Aho & Toivola (2008), who compared native Finnish speakers' pitch contours to L2 speakers, whose L1 was Russian. Their interest was on the foreign accent of the L2 speakers, but the research setting was somewhat similar to mine: they compared L1 speakers' perception experiments to the acoustic realisation of L2 pitch contours.

However, in addition to drawing the pitch contours as suggested in Aho & Toivola (2008), I also wanted to measure pitch acoustically to be able more quantitatively to compare L2 with L1 speakers as well as L2 speakers with the ratings given to them. Hence, I followed Toivanen's (1999, 209–227) example in measuring mean F0, std. of F0 and F0 range for each utterance[57]. In addition to that, I measured the mean absolute slope of F0, as besides the above-mentioned measurements Anttila (2008, 41, 58) had found it useful in acoustic comparison of the pitch contours of different speakers.

For Study VI a script "measure_and_draw_pitch.praat" (Ullakonoja 2009b) was written to automatically measure mean pitch, pitch range, standard deviation of pitch and mean absolute slope[58] for each question separately. In addition to that, the script was used to draw the pitch contours together with the spectrogram and TextGrid for analysing the form of the pitch contours visually. Another script "pitch_max_placement.praat" (Ullakonoja 2009b) was written to detect the position of the peak in the questions. The position of the peak was determined both as the word that the peak is placed on as well as the timing of the peak within that word.

Other studies (Kuosmanen & de Silva 2003; Asu 2004; Kuosmanen & de Silva 2007) on intonation have used different measurements of the pitch contour. These include:

- beginning of the intonation unit
- beginning of the stressed vowel
- mid-point of the stressed vowel
- end of the stressed vowel
- peak (i.e., maximum)
- mid-points of the unstressed vowels in different parts of the unit.

---

57 Toivanen's (1999) subjects were Finnish students of English.
58 Mean absolute slope measures the mean variability in pitch (Boersma 2005).

In addition to measuring pitch at these points, the following pitch calculations can also be made:

- F0 slope between minimum and maximum in the nuclear syllable
- mean absolute F0 slope
- F0 difference between end of the nuclear syllable and end of the following syllable
- F0 slope, range and standard deviation in the nuclear vowel

Section 3.2 above focused on the acoustic methods that were used to analyse the speech corpus, and described how the programme Praat was used in analysing pausing, speech and articulation rates and pitch.

3.3 Fluency Evaluation Task (Studies I, II & III)

The aim of this section is to explain the procedure of the fluency evaluation task where 30 teachers of Russian evaluated perceptually the fluency of the Finnish students' speech. In Studies I and II the perceptual evaluations of fluency were compared to the acoustic analysis of the speech samples. In Study III the perceptual fluency evaluations were recalculated (using z-scores normalisation) and compared to the students' self-evaluation. First, the participants are characterised and then the procedure of the task is described.

3.3.1 Finnish Listeners – Teachers of Russian

The judges who took part in the fluency evaluation task were 30 foreign language teachers of Russian living in Finland. The listeners' L1s are given in Table 8 below. Most of the teachers (n = 25) were L1 female speakers of Finnish. When comparing the reliability of the ratings of the teachers with different L1s, no noticeable differences were found (Cronbach's alpha[59] for Finnish = 0.917, for Finnish and Swedish = 0.919, for all = 0.918). Because the L1 of the listeners did not seem to influence the fluency ratings, all the respondents, including the two L1 speakers of Russian, were included in the study.

The age of the listeners is given in Table 9 below. Listeners from all age groups were represented. The majority of the listeners (47 %) were between 36 and 49 years old, but some were also under 35 years or over 50 years old. Listeners over 50 years old were in the minority.

Finnish	Russian	Swedish	No response	Total
25	2	2	1	30

TABLE 8 Mother tongue of the listeners.

[59] See section 3.5 for the definition of Cronbach's alpha.

Under 35 yrs.	36-49 yrs.	over 50 yrs.	Total
9 (30 %)	14 (47 %)	7 (23 %)	30 (100 %)

TABLE 9 Age of the listeners.

Table 10 below presents the listeners' teaching experience. It ranged from 1 to over 30 years with an average of 13 years (std. = 9.2) in any L2. Hence, it can be said that they were on average fairly experienced teachers. The participants had slightly less experience in teaching L2 Russian than teaching L2s in general, with a mean of 11.5 years (std. = 9.6; range from 1 to over 30 years).

	Mean	Std.	n
Years of teaching any L2	13.1	9.2	30
Years of teaching Russian as L2	11.5	9.6	30

TABLE 10 Teaching experience of the listeners.

Using L2 teachers of Russian as judges was not a self-evident choice. It would also have been possible to use L1 Russian speakers or even L1 speakers of Finnish (not knowing Russian). The justification for choosing L2 teachers of Russian who lived in Finland was that I wanted the judges to be used to listening Russian spoken with a Finnish accent in order for them to focus their evaluation on fluency, and not other, e.g., segmental, features. Some studies (Koster & Koet 1993; Okamura 1995) have found language teachers to be more critical in their evaluations than nonteachers. Okamura (1995) found that nonteachers were less strict in their judgements but were able to provide a clearer distinction between advanced and average learners, especially when evaluating fluency. Koster & Koet (1993) report similar findings, in that non-native teachers were stricter in their judgements than L1 speakers of the language and they also spotted more errors on both the segmental and prosodic levels in L2 English. However, the choice of L2 teachers also stems from the fact that they are used to evaluating the fluency of their students and can do it without requiring a definition of fluency.

3.3.2 Procedure

On the basis of the recorded material (see 3.1.4), a fluency evaluation task was designed. Only one turn in a Russian dialogue per speaker (n = 12) was used from the recordings done prior to, during and following the stay. This resulted in a total of 36 stimuli. In the task, the listeners were simply asked to evaluate the fluency of the L2 speakers' speech samples.

The material for the fluency evaluation task was prepared in Praat such that the extracts (on average 19 s each) were set in a randomised order. The resulting sound file had the number of the stimulus in Russian first, then the stimulus, a 7-second pause, and then a sound marking the start of the next stimulus. Before the actual listening task, a test file, containing only reading

aloud of the digits *один* (*odin*[60]), *два* (*dva*[61]), *три* (*tri*[62]), was presented to the subjects so that they could adjust the volume of their headsets to a convenient level.

All the judges were asked to evaluate the fluency of the samples by responding to a questionnaire in Finnish where they were first asked to provide some background information on themselves. As Derwing et al. (2004) have argued, even untrained listeners seem to attend to similar features of speech when asked to rate the fluency of a speech sample. Hence, instead of giving the teachers a definition of fluency, they were asked to define what they understood by fluent speech and write down their definitions (see Study I & Ullakonoja 2009, 40–42, for a summary)[63].

The actual evaluation task was simply to rate the fluency of each stimulus on a 1 (not fluent) to 5 (very fluent) scale as L2 learner's speech. As many as 23 of the 30 judges used the full evaluation scale (1–5). Each judge was asked to rate each stimulus only once. However, still better reliability could have been attained by asking the judges to do the ratings twice (with some time in between) and then comparing the ratings of each judge. Also, three teachers did not rate all the stimuli (had missed out one or two stimuli). After the listening task the subjects were requested to write down the factors they thought disturbed the fluency of the stimuli.

The task was first piloted in a language lab for three Russian teachers. The actual evaluation task was administered online through a web publishing platform (Moniviestin, developed by the Virtual University Project at the University of Jyväskylä), which allowed for password protection of the sound file, while the questionnaire was available freely online. The questionnaire site was created with the SPSS Data Entry programme (for more information see http://www.spss.com/Data_Entry). The online questionnaire was tested by one teacher, and as she reported no problems, the request to participate in the task was sent out to a mailing list (with about 270 members) of teachers of Russian in Finland in August 2007. In addition, the questionnaire was also sent directly to 15 teachers of Russian the author knew. By the first deadline (in 3 weeks), 17 teachers had filled in the questionnaire. The questionnaire was re-sent through the same mailing list and to the mailing list (about 200 members, partly overlapping the other list) of teachers of Russian maintained by the Ministry of Education. By the new deadline (a month later), 9 more teachers had answered. Thus, I obtained the fluency ratings of a total of 30 teachers. An average of all the ratings for each sound sample was calculated and thus a mean fluency determined for each sample. This rating could be then compared to the acoustic analysis in Studies I and II.

It is impossible to calculate the response rate owing to overlapping membership of the two lists the request was sent to, but it is unlikely to be very

---

60 In English: 'one'.
61 In English: 'two'.
62 In English: 'three'.
63 The participating students (whose the speech was evaluated) were also asked what they think constitute fluent L2 speech (Ullakonoja 2008c).

high (perhaps around 10 %). Hence, the sample is not very representative. However, because the purpose of the present study was not to analyse the fluency evaluations as such, but instead to use them as a tool for measuring fluency, even a low response rate does not falsify the results. The online data collection enabled the judges to perform the task at the time and place most suitable for them. This also made it very unlikely that the judges could have conferred with each other about the ratings they were giving. The participants evaluated the speech samples individually and were not aware of the fact that for the same speakers there were multiple samples.

3.4 Evaluation of Interrogativity Task (Studies V & VI)

The second evaluation task of the present study focused on the perception of interrogativity in YNQs. The purpose was, as in the previous task, to compare the acoustic measurements to perceptual judgements. Specifically in this task, the aim was to determine which utterances produced by the L2 speakers were perceived as questions by the L1 speakers.

3.4.1 Russian Listeners

The listeners were 40 students and staff members at the philological faculty of St. Petersburg State University (36 women, 4 men). All were self-reported L1 speakers of Russian none of whom were paid for their contribution. The evaluation task was conducted in two parts: the recordings of the Finnish students in the middle of their stay in Russia and those following it were evaluated separately. The first group of listeners (n = 19) listened to the recordings made in the middle of the stay and the second group (n = 21) those made following the stay.

	Under 20 yrs.	20-29 yrs.	30-39 yrs.	40-49 yrs.	Total
Men	4 (10 %)	0 (0 %)	0 (0 %)	0 (0 %)	4 (10 %)
Women	28 (70 %)	5 (13 %)	2 (5 %)	1 (3 %)	36 (90 %)
Total	32 (80 %)	5 (13 %)	2 (5 %)	1 (3 %)	40 (100 %)

TABLE 11 Age and gender of the Russian listeners.

The clear majority of the participants in both groups were female students under 20 years of age (Tables 11–12).

Group 1: students	Group 1: other	Group 2: students	Group 2: other	Total
17 (43 %)	2 (5 %)	20 (50 %)	1 (3 %)	40 (100 %)

TABLE 12 Occupation of the Russian listeners.

As Table 13 shows, 49 % of them considered themselves as used to hearing foreign-accented Russian. However, when asked about their amount of exposure to L2 Russian, most of them (53 %) said that they were exposed to L2 Russian rarer than once a month (Table 14). Only three participants said that they heard foreign-accented Russian daily. There was also a question asking if the listeners had sometimes in their lives taught Russian to foreigners, and three participants reported that they had.

Group 1: used to hearing L2 Russian	Group 1: not used to hearing L2 Russian	Group 2: used to hearing L2 Russian	Group 2: not used to hearing L2 Russian	Total
12 (31 %)	6 (15 %)	7 (18 %)	14 (36 %)	39[64] (100 %)

TABLE 13 Listeners' familiarity with non-native Russian.

Everyday	Once a week	Once a month	Rarer	Total
3 (8 %)	10 (25 %)	6 (15 %)	21 (53 %)	40 (100 %)

TABLE 14 Listeners' frequency of exposure to non-native Russian.

3.4.2 Procedure

As described in Study V, the YNQs that could be interpreted either as questions or statements[65] depending on their pitch contour were chosen from the two read-aloud dialogues for the evaluation task. The 84 stimuli consisted of the reading aloud of the questions by each of the six Finnish students in the two separate recording sessions: in the middle of their stay in Russia and following it. The stimuli were concatenated into two sound files: one for each recording session (middle of stay and following stay). The duration of the sound files was 11 minutes each. Because of the large number of stimuli required for the task, the experiment was run in two separate parts. One group of listeners (n = 19) listened to the file containing the recordings done following the stay in Russia, whereas the other group (n = 21) rated the stimuli from the previous recording session (i.e., middle of the students' stay in Russia). The extracted utterances were presented to the listeners once in randomised order with an interstimulus interval of 11 ms, each preceded by a sound marking the start of the stimulus. The task was performed in a language lab in St. Petersburg (November 2008 and February 2009) over headphones to enable the listeners to adjust the volume and concentrate on listening without disturbances.

On the evaluation form the listeners were asked to answer four questions each time they heard a stimulus: 1) determine if they heard a falling, rising or level tone 2) choose the sentence mode (question or non-question) 3) determine the degree of goodness of the stimulus as a question (on 1–5 scale, where 1 was "not a question" and 5 was "a question") 4) determine the emotionality of the

64 One listener did not reply to this question.
65 The exact utterances can be found in Study V (Study V: Table 1) and Study VI (Study VI: Table 1).

sample (by marking a "+" if they considered it emotional). The listeners were advised to focus on questions 2 and 3 and answer the rest only if they had time. The form also included questions on the listeners' background to answered prior to listening. As it turned out, questions 1 and 4 were fairly difficult to answer, and with some listeners not answering them at all, and thus these were not analysed in Studies V and VI. Answers to questions 2 and 3 were entered into SPSS for further analysis. The average recognition rate (percentage of answers detecting a question to question 2) and average acceptability ratings (mean of answers to question 3) were then calculated.

The listeners were told before listening that all the stimuli were intended as questions by Finns, hence, they were not forced to recognise some of the stimuli as statements. Also the choice was to be made between question and non-question, not question and statement, the purpose being to induce the listeners to concentrate on interrogativity rather than sentence type. Although this may not be the most commonly used design for a perceptual evaluation task, the idea was to evaluate how successful the students' were in uttering interrogatives, i.e., "goodness of fit". According to Volskaya (2010), a similar design has been used by Abramova (1999).

Some, but not a considerable amount of, erroneous answers can be expected on the basis of an experiment by Chernigovskaya et al. (2000), who asked 50 native listeners to evaluate each native speakers' stimulus from five different aspects: 1) communicative type, 2) completeness/incompleteness, 3) intonation unit division 4) position of the intonational centre, and 5) emotionality. Across these different tasks 18 % of the L1 speakers answered incorrectly.

In sum, the task was indeed rather multifarious for the listeners. They were asked to evaluate mainly whether they heard a question or not and if they were led to think of the stimulus as an interrogative, how good a question it was (on 1 to 5 scale). Additionally, if they had time and if they felt they could, they were also asked to determine the pitch movement and emotionality of the stimulus.

3.5 Statistical Methods

The purpose of this last section of the chapter is to expand the methodological discussion to include the statistical methods used. These are in fact, the methods most often used in the analysis of phonetic data. For the statistical analysis of the data I used Microsoft Office Excel 2003–2008 and SPSS 14.0–16.0. If a script (see p. 72) was used, Praat gave the data in a text file which could be imported into an Excel or SPSS data sheet. Organising and categorising the data along with some calculations of frequencies and building of charts were done in Excel. SPSS was used mainly for testing the statistical differences and correlations between the variables, but also for descriptive statistics of the data. As a more detailed description of the statistical methods used is found in each

study (Studies I–VI), I explain here only the common principles followed in the analysis.

The statistical analysis always commenced by ascertaining if the data were normally distributed. Besides inspecting the histogram this was verified by the Kolmogorov-Smirnov test. If the test indicated that the distribution of the variables was not normal, non-parametric tests (e.g., Wilcoxon signed-ranks test) were used for testing the statistical significances between the means. Vice versa, if the Kolmogorov-Smirnov test indicated normally distributed variables, parametric tests, usually the paired samples t-test, were used. Paired samples tests were used when the same speaker's speech was compared at different times, whereas independent samples were used, e.g., when different L1 and L2 speakers were compared.

Correlation coefficients were used to define the relationship between two variables. Pearson's correlation was used when the data under investigation were normally distributed, and Spearman's correlation when the data were not, as it does not assume normal distribution of the data (e.g., Study I). The existence of a correlation was verified in scatter plot graphs. If there had been more speakers, it would have been possible to compare the host-family group and the dormitories group for statistically significant differences, but with only six students in each group this was not done (see, e.g., Heikkilä 2004).

In Study V, in the question categorisation task, the interrater reliability was determined by using Cohen's Kappa statistics that can be used for nominal measures (Kraemer 1983). Cohen's Kappa was computed with a SPSS macro by Dates (2006), available online. Values are usually between 0 and 1, where 0 indicates that the agreement of the raters is purely coincidental, and 1 a complete agreement between the raters (Sim & Wright 2005). Universally accepted principles about how to interpret the Kappa coefficient do not seem to exist. Landis & Koch (1977, 165) propose that values from 0.41 to 0.60 indicate moderate interrater agreement. On the other hand, Sim & Wright (2005) draw attention to the fact, that Kappa is influenced, e.g., by the number of categories used as well as the sample size. Leech et al. (2005) report that Kappa should be higher than 0.70, but point out, that because Kappa accounts for chance it tends to be lower than the other interrater reliability measures. In Study V, the Kappa yielded 0.56 for the stimuli recorded in the middle of the stay and 0.59 for the stimuli recorded after the stay. Hence, it can be concluded that the present interrater agreement does not exist purely by chance.

In Studies I and V, Cronbach's alpha was determined to evaluate the interrater reliability of ordered categorical variables (see, e.g., Bryman & Cramer 2001, 62). Cronbach's alpha is generally used to determine interrater consistency. It means that "it is not really necessary for two listeners to share a common meaning of the rating scale, so long as each listener is consistent in classifying the phenomenon according to his or her own definition of the scale" (Stemler 2004). This means that two listeners can judge a sample differently but if they were, e.g., to rank the samples, they would put them in a similar order. A satisfactory value of Cronbach's alpha is generally considered to be higher

than 0.8 (see, e.g., Nunnally & Bernstein 1994, 252; Bryman & Cramer 2001, 62). The value of Cronbach's alpha in Study I for the fluency evaluation task was 0.92. This is high enough to allow the conclusion that the raters were consistent in their judgements. In Study V, Cronbach's alpha for the acceptability ratings was 0.96 for the raters rating the samples recorded in the middle of the stay in Russia and 0.86 for the raters rating the samples recorded following the stay.

In previous L2 studies higher and lower Cronbach's alphas for listeners' ratings have been reported. According to Derwing et al. (2004), inter-listener agreement can be quite high (Cronbach's alpha = 0.95), also for untrained listeners. In contrast, Cucchiarini et al. (2002), who studied the oral fluency of L2 Dutch speech that was rated on a 1–10 scale, found that phoneticians' judgements were very reliable (Cronbach's alpha = 0.96), but less trained listeners did not achieve such high reliability ratings (Cronbach's alpha from 0.82 to 0.88). As pointed out by Derwing et al. (2004, 658), reliability comparison of the studies by Lennon (1990), Freed (1995) and Riggenbach (1991) with that by Cucchiriani et al. (2002) is difficult because they do not specify the value of Cronbach's alpha, and have estimated reliability in different ways.

Given that the ratings of the listeners for each sample were fairly similar, it can be concluded that the fluency evaluation task was quite an efficient and a reliable way of measuring the perceived fluency of the samples. The consistency of the raters in the evaluation of interrogativity task was verified by calculating the correlations between the question categorisation and acceptability tasks (Study V: Table 4). It showed that L1 speakers of Russian agree on what is a question and what is not, and are consistent in giving their ratings in both tasks.

It has to be pointed out that when comparing the mean fluency ratings of each student prior to and following the stay in Russia in Studies I and III, the significance level was set to $p < 0.005$ to keep the results reliable and unaffected by minor differences. However, if the significance level had been $p < 0.05$, the results of Study III would have been somewhat different; using z-scores, the mean fluency would have increased for 9/12 students and for 8/12 students.

In short, the statistical methods applied to the present study were mostly used for detecting statistical differences in means between two groups. In addition to this, they served as an important way to measure the reliability of the two perceptual evaluation tasks.

4 DISCUSSION AND CONCLUSIONS

In this chapter the main results of the six studies are discussed and summarised. First, I examine pausing and speech/articulation rate as acoustic correlates of reading aloud fluency (Studies I–III). This is followed by a synthesis of the studies addressing pitch range (Study IV) and pitch contours in YNQs (Studies V–VI). The overall design, findings and applications are discussed in section 4.4. Finally, the implications to future work are presented and conclusions drawn, and a synthesis of each of the research questions is given. In referring to individual subjects, I use the pseudonyms they had in the corpus together with those used in each separate study (see Table 5, p. 61).

4.1 Fluency

4.1.1 Summary of the Main Findings (Studies I, II and III)

In this subsection I summarise Studies I, II and III, which focused on fluency, and discuss their results in the light of previous research. Studies I, II and III investigated the L2 learners' improvement in fluency during SA in terms of pausing and speech/articulation rate. Study I consisted of two sub-studies: teachers' evaluations of students' fluency, and an acoustic analysis of pausing, and discussed the relationship between the two. Study II examined speech and articulation rates, which were measured both in phonetic words per second and syllables per second. Also studied was the relationship between speech and articulation rates and the fluency ratings obtained in Study I. Study II also addressed the question of whether speech and articulation rate were speaker-specific and/or language-specific. Study III recalculated the fluency ratings of Study I using normalisation, and investigated the students' self-assessment in relation to the fluency ratings.

As in, for example Lehtonen (1977; 1978; 1979; 1981), I also attempted to find acoustic correlates of fluency. My original aim was to develop L2 teaching

and learning by understanding what acoustic features contribute to the perception of read-aloud speech as fluent, not to develop automatic ways of evaluating spoken language tests. My data consisted of read-aloud dialogues, and thus fluency refers to read-aloud speech only. In the present study, I focused on the development of the Finnish students' Russian fluency when studying in Russia. I chose to study pause frequency and duration as well as speech and articulation rates because, according to Cucchiarini et al. (2000; 2002), a fast speech rate and low pause frequency are the most important factors for perceiving read-aloud speech as fluent. These features of speech often do not get enough attention in L2 classrooms, and L2 speech typically differs from native speech in these respects.

As might be expected, it was found in Studies I and III that the majority of the speakers improved their perceived reading fluency during their SA experience (Study I: Figure 1). However, this improvement was not as systematic as in the study by Freed et al. (1995), in which the weaker students improved their fluency more than the better ones. One somewhat unexpected finding in Study I was that two students (Sanna-Fi1 and Ilona-Fi4) showed a progressive decline in read-aloud fluency during SA (Study I: Figure 1). One possible reason for the decline in Sanna's skills is that she had possibly become more aware of her pronunciation and thus started monitoring it. As a result she made more self-corrections. Thus, her speech would have contained more repairs and more disfluent pauses (following the stay more students had repairs in their speech than prior to the stay). Hence, it can be argued that an intensive focus on correct pronunciation may also result in more disfluencies (such as self-repairs and hesitations), and thus in lower fluency ratings. The other student, Ilona, however, was evaluated as rather fluent in all the recordings (mean ratings of over 4), and hence, the decline in her speech may either represent normal variation in her skills or result from the earlier finding that highly proficient students do not benefit as much as from SA as less proficient ones.

Furthermore, in Study I it was found that speech containing a high number of pauses was perceived as less fluent than speech with few pauses. This was especially true for disfluent pauses (which often occurred together with repairs, repetitions and other hesitation phenomena). Interestingly five samples contained no disfluent pauses, but yet did not receive a very high fluency rating. Therefore, it cannot be said that speech with no disfluent pauses will always be perceived as fluent. This indicates that pause frequency is not the only feature contributing to the perception of speech as fluent. Furthermore, it was found that the more fluent the speaker was estimated to be, the shorter was her disfluent pause duration (both in absolute and relative values).

Consistently with Towell et al. (1996, 103), increased speech rate was found to be more significant than articulation rate in determining the L2 fluency of the speakers in Study II. It was concluded that faster L2 speech and articulation rates are evaluated as more fluent than slower ones. This seems to confirm earlier findings that L1 speakers generally react more positively to a

faster L2 speech rate (Munro & Derwing 1998; 2001, 464; Paananen-Porkka 2007, 340), which supports the idea that L2 speakers should aim to speak faster.

In Study III the results concerning the teachers' ratings of the students' fluency obtained in Study I were recalculated using z-score normalisation as this improved the comparability of the listeners' ratings. The recalculation confirmed the earlier findings: the majority of the students were evaluated as significantly more fluent readers following the stay than prior to it. However, unlike in Study I, only one student showed a significant ($p = 0.0001$) improvement in teacher-rated fluency following her stay than in the middle of it. Otherwise, the fluency results were consistent with those of Study I. Study III also examined the students' self-assessments, comparing them with their fluency ratings. The main finding was that the students who said that their pronunciation had improved and who showed interest in learning and practising were judged on average to be more fluent readers.

To sum up the results of Studies I and III on fluency, the interesting finding was that all the students evaluated themselves as more fluent after their stay than prior to it (Studies I & III). This implies that the students themselves saw SA as a way of improving their fluency. As the teachers' ratings also showed a significant improvement in fluency for the majority of the students, the studies corroborate the earlier findings that there indeed is a relationship between L2 fluency and SA.

4.1.2 Pausing

I now turn to the results obtained on pause frequency and duration. My pilot study (Ullakonoja 2007b) on pausing showed that there were fewer and shorter pauses in L1 Russian speakers' speech when compared with Finnish L2 learners of Russian. In the pilot study, pausing by three learners' in read-aloud Russian (two dialogues, about 4.5 min per speaker) was compared to that of a L1 speaker of Russian (1 min 15 s.). The main results were that following the stay in Russia, there were less chunking of disfluencies (all 3 speakers), fewer disfluent pauses (2/3 speakers) and shorter pause duration (2/3 speakers). The study also showed that even L1 speakers sometimes produce disfluent pauses, e.g., when hesitating.

However, in Study I pause frequency in L1 and L2 was not compared. Instead, the frequency distribution of fluent and disfluent pauses [66] was measured three times (prior to the stay in Russia, middle of stay and following it). First, Study I showed that in total fluent speakers produced fewer pauses in their speech than speakers who were evaluated as disfluent. The total mean pause frequency of the speakers as a group decreased as the length of experience increased. Speakers had on average 12.8 pauses prior to the stay, 11.7 in the middle of the stay, and 11.0 pauses following the stay. Disfluent pause frequency decreased by over 50 % more than fluent pause frequency

[66] See p. 22 for definitions of fluent and disfluent pauses.

during the stay. The majority of the speakers (9/12) had fewer disfluent pauses following the stay in Russia than before it (Study I: Table 1).

Figure 6 below illustrates pause frequency development in different stages of the stay on the individual level. The frequency of fluent pauses is shown in different shades of blue and green, whereas the frequency of disfluent pauses is in different shades of red and orange. There are no clear common patterns of development in pause frequency. For twelve students, eight different development patterns were indentified in the frequency of fluent pauses. Development in the frequency of disfluent pauses showed six different patterns of which the most common one (for three students: Noora-Fi12, Aamu-Fi10 and Kati-Fi2) is a gradual decrease in frequency.

The analysis of pause duration in Study I showed that on average the speakers used shorter pauses following the stay than before it. This can also be seen in Figure 7 below, which shows the development of mean pause duration at different stages of the stay. Most students had the shortest mean pause duration in the middle of their stay (T2). However, Ritva's (Fi5) pause duration was shortest prior to the stay and Liisa's (Fi6) and Elsi's (Fi9) following the stay.

When the results on pause duration were compared to those of previous studies on Russian L1 speakers, the students' pauses appeared to be longer than in Volskaya (2004) but shorter than in Riazantseva (2001). This may be due to the different durational thresholds used. In Volskaya (2004), mean pause duration was 173.5 ms (range 153–188 ms) in read-aloud speech of L1 Russian speakers when the duration of all perceived pauses less than 250 ms was measured acoustically. However, Riazantseva (2001) obtained quite different values for spontaneous Russian: 767 ms for a topic narrative task and 822 ms for a cartoon description task. The different values of Volskaya (2004) and Riazantseva (2001) can be explained by Riazantseva's choice of analysing monologue data and measuring all the silent intervals between 100 ms and 3000 ms as pauses.

In Paananen-Porkka (2007, 240) students' ratio of pausing time in spontaneous speech varied widely between speakers: from 26 % to 57 % in L2 English. In the studies by Temple (1992, 32; 2000), on the other hand, L2 speakers spent 38 % of their speaking time pausing (this included both silent and filled pauses). She also found that L2 speakers' frequency of filled pauses was significantly higher than that of L1 speakers. In comparison, Study I (Study I: Table 2) showed that my subjects spent a lot less time pausing (range from 6 % to 34 %). The results are not entirely comparable, however, as I distinguished between fluent and disfluent pauses and studied read-aloud speech.

In Study I, pause duration was also measured for fluent and disfluent pauses separately. When the two pause categories were compared for duration, a relationship between pause type and duration was found: when learners' fluent pauses were long, disfluent pauses also tended to be long and vice versa. The smallest relative pause duration was observed in the recordings done in the middle of the stay for the majority of the speakers. This could be explained by

the fact that they were used to reading in Russian when in Russia, and hence used a faster speech rate than in Finland.

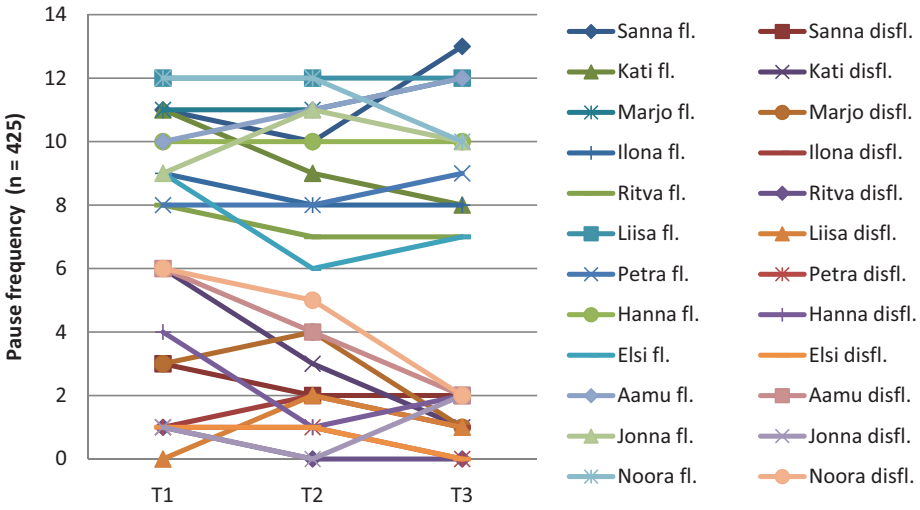


FIGURE 6 Pause frequency (fl. = fluent pauses, disfl. = disfluent pauses) at different stages of stay (T1 = prior to the stay, T2 = in the middle of the stay, T3 = following the stay).

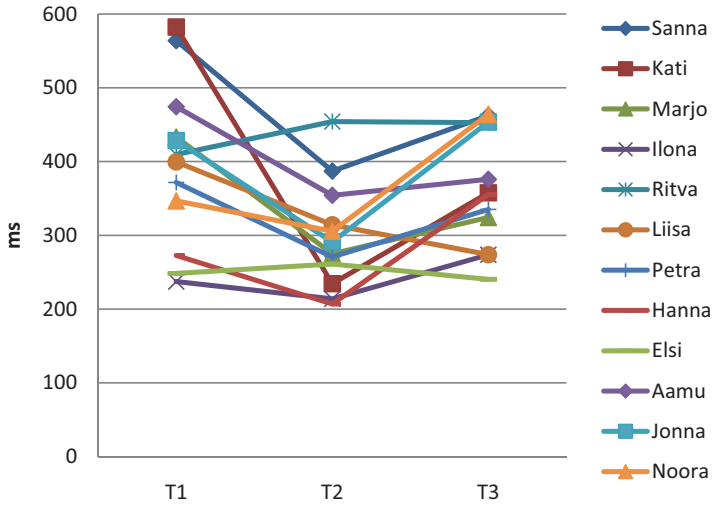


FIGURE 7 Mean pause duration at different stages of stay (T1 = prior to the stay, T2 = in the middle of the stay, T3 = following the stay) in ms.

4.1.3 Speech and Articulation Rate

Turning now to speech and articulation rates (Study II), the main finding was that the majority of the students significantly increased their L2 speech and articulation rates during the 3.5-month stay in Russia. At the same time their perceived fluency increased. This clearly shows that the students benefited from their stay in Russia in terms of faster and more fluent L2 reading aloud.

My results for speech and articulation rates (Study II) can be compared to those of other researchers. Table 15 presents the experimental data on L2 speech and articulation rates in the SA context obtained in different studies (Möhle 1984, 30; Lennon 1990, 404; Towell et al. 1996, 98; Freed et al. 2004; Segalowitz & Freed 2004, 195).

Study	Speech rate prior to the stay	Speech rate following the stay	Articulation rate prior to the stay	Articulation rate following the stay
Freed: L1 = English L2 = French (n = 8)	106.78 WPM = 1.78 words/s	113.33 WPM = 1.89 words/s	-	-
Lennon: L1 = German L2 = English (n = 4)	84 WPM = 1.4 words/s	97 WPM = 1.62 words/s	96 WPM = 1.6 words/s	110 WPM = 1.83 words/s
Möhle: L1 = French L2 = German (n = 3)	120.36 syll/min = 2.01 syll/s	120.18 syll/min = 2.0 syll/s	3.22 syll/s	3.78 syll/s
Möhle: L1 = German L2 = French (n = 3)	175.18 syll/min = 2.92 syll/s	201.26 syll/min = 3.35 syll/s	4.50 syll/s	4.85 syll/s
Segalowitz: L1 = English L2 = Spanish (SA group) (n = 22)	55.63 WPM = 0.93 words/s	80.63 WPM = 1.34 words/s	-	-
Segalowitz: L1 = English L2 = Spanish (at home group) (n = 18)	51.07 WPM = 0.85 words/s	52.51 WPM = 0.88 words/s	-	-
Towell et al.: L1 = English L2 = French (n = 12)	136.61 syll/min = 2.28 syll/s	156.88 syll/min = 2.61 syll/s	3.85 syll/s	4.17 syll/s
Ullakonoja: L1 = Finnish L2 = Russian (n = 12)	**3.46 syll/s 1.61 PW/s**	**3.93 syll/s 1.79 PW/s**	**4.58 syll/s 2.12 PW/s**	**5.00 syll/s 2.28 PW/s**

TABLE 15 Comparison of mean L2 speech and articulation rates in SA context of the studies by Freed (2004), Lennon[67] (1990), Möhle (1984), Segalowitz & Freed (2004) and Towell et al. (1996) with the results of Study II[68] in syll/s and PW/s.

67 Lennon's values have been calculated on the basis of the mean values he has have given for each speaker.

68 My results have been presented without distinguishing between the two groups discussed in Study II (host-family group and dormitories group).

Table 15 shows that in my Study II the speech and articulation rates in L2 were higher than in the other studies. The comparison needs to be interpreted with some caution, since the other studies are on spontaneous speech, which is likely to be slower. As Table 15 shows, in earlier studies either speech or articulation rate, or both, were faster following the stay than prior to it. In Study II both were on average faster after the stay. Also, on the individual level, the majority (8/12) of the students had a faster speech and articulation rate following the stay than prior to it.

Figure 8 below illustrates the individual development patterns in speech and articulation rate in different stages of stay measured in syllables/second. 6/12 students show a gradual increase in speech rate and 8/12 in articulation rate during the stay. Other rather common pattern is increase during the first half of the stay (T1–T2) and then decrease during the second half (T2–T3) so that the values following the stay are however higher than prior to the stay. Only two students (Jonna-Fi11 and Hanna-Fi8) have a different pattern either in speech rate or articulation rate.

Figure 9 below is not radically different from Figure 8 but it presents the measurements in phonetic words/second, which gives slightly different results. Speech rate develops in three patterns: 1) gradual increase during the stay (7/12 students), 2) increase during the first half of the stay (T1–T2) and decrease during the second half (T2–T3), so that the values in T3 remain higher than those in T1 (Sanna-Fi1 and Aamu-Fi10), and 3) increase during the first half (T1–T2) and decrease below the values of T1 during the second half (T2–T3) (Jonna-Fi11, Hanna-Fi8 and Ilona-Fi4). Five different patterns can be seen in the development of the articulation rate. The most common of these is a gradual increase during the stay (4/12 students); however quite a few students (Jonna-Fi11, Elsi-Fi9, Petra-Fi7, Hanna-Fi8 and Ilona-Fi4) diverge from this pattern.

As was suggested in section 2.1. (p. 29), people may read words that are unfamiliar to them more slowly than words that are familiar to them, and also, that L2 knowledge seems to predict L2 reading skills. Bearing this in mind, it may be possible that the speech rate of a learner over a longer stretch of speech also reflects the size of his/her lexicon. Would then the slower speakers of my study also be those whose Russian vocabulary is not as large as that of the faster speakers? If so, would that also imply that their Russian skills are poorer, if size of vocabulary is one measure of language skills in general? Assuming this is the case, is what I ended up measuring as fluency, nevertheless affected in some way by the students' language proficiency, which I did not measure here? Möhle (1984) argues, that an increase in students' speech rate during a semester abroad could be a result of the broadening of their lexical knowledge. This also offers a possible explanation for the fact that the speech rate of my students accelerated during their SA. If the students' vocabulary had expanded during the semester, they would have recognised more words in the texts of the reading task, and hence they would have read familiar words faster than unfamiliar ones.

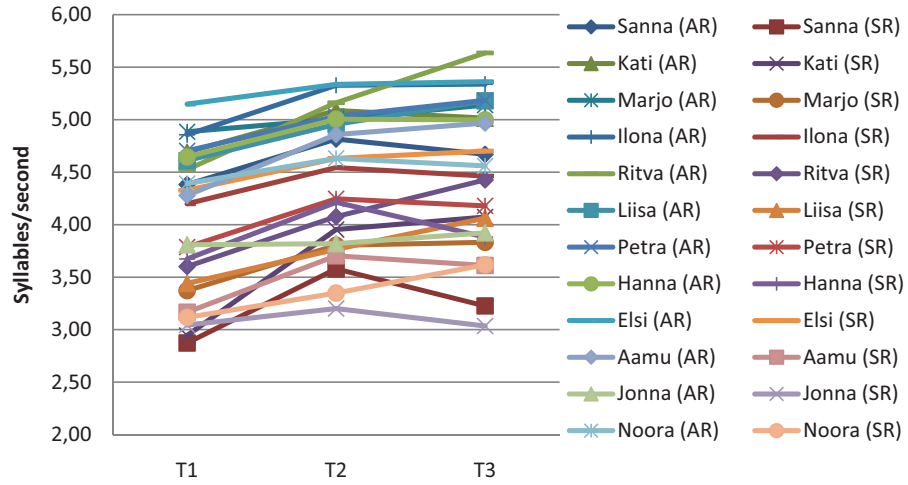


FIGURE 8 Mean articulation rate (AR) and mean speech rate (SR) at different stages of stay (T1 = prior to the stay, T2 = in the middle of the stay, T3 = following the stay) in syllables/second.

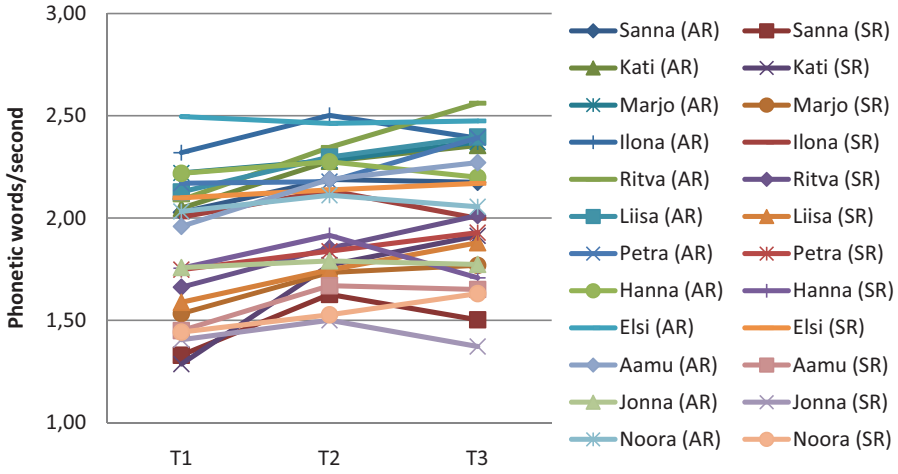


FIGURE 9 Mean articulation rate (AR) and mean speech rate (SR) at different stages of stay (T1 = prior to the stay, T2 = in the middle of the stay, T3 = following the stay) in PW/s.

If we compare my L2 results (in Table 15) with the results obtained from Russian L1 speakers, my students were, not surprisingly, a lot slower than native speakers. In native Russian speech, the speech rate was 6.5 syllables/s in read-aloud speech and 6.2 syllables/s in spontaneous speech (Volskaya 2009a, 137). An interesting comparison can also be made with results from the study by Toivola et al. (2009b). Their six Russian speakers had a mean articulation rate

of 3,73 syllables/s and speech rate of 3,05 syllables/s in L2 Finnish[69]. The speakers were not highly proficient in Finnish, yet, their speech rate is rather similar to that of my L2 learners of Russian.

Next, Table 16 below presents an overview of the results for speech and articulation rates in L1 Finnish. The results of these studies vary widely. This inconsistency may be due to rather different data analysed in the studies and, in some cases, to a small number of informants. In read-aloud speech the rates from previous studies are less than 1 syllable/s slower than in Study II. My study is the only one with a dialogue setting, which may have influenced the speech rate, as the speaker had time to pause and inhale while the other speaker was speaking. In a monologue read-aloud or spontaneous setting, a speaker needs to pause for physiological reasons more than in a dialogue. The articulation rates in Table 16 show less variation, suggesting that a typical articulation rate for Finnish is somewhat over 6 syllables/second. When the results from Study II were compared with Lehtonen's (1978) figures, it was found that the L1 Finnish reading rate was similar when measured in syllables/s, but faster in my study when measured in phonetic words/s.

Study	Speech rate	Articulation rate
Paananen-Porkka: L1 = Finnish, spontaneous speech (n = 6) monologue	5.11 syll/s	6.21 syll/s
Sallinen-Kuparinen: L1 = Finnish, vocational school students', read-aloud speech (n = 30[70]?) monologue	289 syll/min = 4.82 syll/s	5.9 syll/s
Sallinen-Kuparinen: L1 = Finnish, high school students' read-aloud speech (n = 30[70]?) monologue	319 syll/min = 5.32 syll/s	6.7 syll/s
Moore: L1 = Finnish, TV broadcasters' spontaneous speech (n = 1), monologue	3.64 syll/s	5.20 syll/s
Moore: L1 = Finnish, Radio announcer's spontaneous speech (n = 1), monologue	5.63 syll/s	6.48 syll/s
Lehtonen: L1 = Finnish, read-aloud speech (n = 5) monologue	330 syll/min = 5.5 syll/s	400 syll/min = 6.67 syll/s
Lehtonen: L1 = Finnish, spontaneous speech (n = 5) monologue	196 syll/min = 3.27 syll/s	317 syll/min = 5.28 syll/s
Toivola et al. L1 = Finnish, read-aloud speech (n = 6)	5.29 syll/s	6.28 syll/s
Ullakonoja: **L1 = Finnish, read-aloud speech (n = 12) dialogue**	**5.77 syll/s**	**6.63 syll/s**

TABLE 16 Comparison of mean speech and articulation rates in L1 Finnish in the studies by Lehtonen (1978), Moore (1990), Paananen-Porkka (2007), Sallinen-Kuparinen (1979) and Toivola et al. (2009b) with the results of Study II in syll/s.

69 These values have been recalculated by the present author on the basis of the data presented in the paper by Toivola et al. (2009b).

70 Sallinen-Kuparinen's total number of students was 60, but she does not state how many students were in each group.

Finally, Study II also addressed the question of whether speech and articulation rates are speaker-specific. Consistently with Towell et al. (1996, 96), strong evidence was found that the speech and articulation rates in L1 Finnish of each speaker were related to their speech and articulation rates in L2 Russian. Therefore, it can be concluded that speech and articulation rate were speaker-specific in this sample (n = 12). In other words, speakers who spoke slowly in L1 were also likely to speak slowly in L2. On the basis of earlier research (e.g., Paananen-Porkka 2007), an expected finding was that L1 was spoken faster than L2.

In the contrastive study conducted by Grosjean & Deschamps (1975) with L1 speakers of English and French, the speech rates and articulation rates of the two languages were found to be very similar in a spontaneous interview setting. However, the two languages differed in pausing: in English the pauses were shorter but more frequent than in French, resulting in similar total pausing time. The authors concluded that the reason for this was in the different syntactic and morphological structure of the two languages. This is precisely what renders the present cross-language comparison of L1 Finnish and L2 Russian difficult. Finnish and Russian being so different, is it possible to say what the differences found really reflect – differences between the languages, speakers or something else?

To sum up, it seems difficult to compare speech and articulation rates in different languages, in different speech styles and different texts. Furthermore, they seem somewhat to depend on personal and speaker-specific variables that cannot all be controlled for. However, it is clear that both speech and articulation rate play a crucial role in L2 fluency and that they often become faster as a result of SA.

4.1.4 Development in read-aloud fluency during SA in terms of acoustic measurements

The acoustic correlates of fluency in the present study include pausing (pause type, frequency and duration) and speech and articulation rates. Here, I summarise how Studies I, II and III contribute to the definition of fluency.

First, I examine individual fluency development during SA. Figure 10 shows the mean fluency rating of each speaker during different stages of the stay. The most common pattern (7/12 students) is a gradual increase in mean perceived fluency during SA (between T1 and T2 and T2 and T3). Sanna (Fi1) and Ilona (Fi4) show the opposite pattern: their fluency ratings decrease during SA. For Jonna (Fi11) there is a decrease between before the stay and middle of stay, but the values remain the same during the middle of stay and thereafter. Hanna (Fi8) and Petra (Fi7) represent the fourth possible pattern, which is an increase in fluency during the first half of the stay (T1–T2), but a slight decrease during the second half (T2–T3).

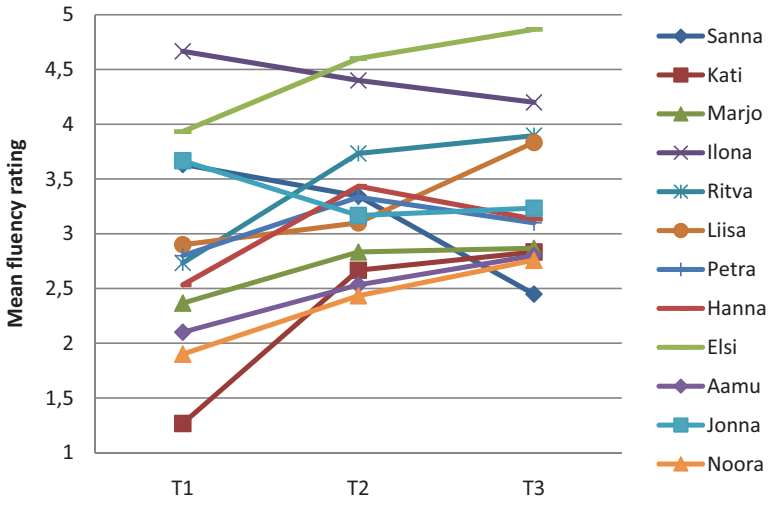


FIGURE 10 Mean fluency ratings at different stages of stay (T1 = prior to the stay, T2 = in the middle of the stay, T3 = following the stay).

Next, I discuss pause frequencies with respect to the two pause types (fluent and disfluent pauses). Figure 11 below illustrates the relationship between the number of fluent pauses and disfluent pauses produced by each speaker (n = 12) at each recording session compared with the mean fluency rating. Each mark represents one speech sample, in other words, there are three dots per speaker (one per recording session). The figure confirms the existence of the correlation earlier found between the frequency of both pause types and fluency ratings (the values of the correlation coefficients and statistical significances were reported in Study I). On the individual level (Study I: Figure 1 & Table 1), the least fluent speakers (Kati-Fi2, Aamu-Fi10 and Noora-Fi12) prior to the stay also had the most (6) disfluent pauses. Interestingly, however, the total absence of disfluent pauses did not result in markedly high fluency ratings. For example, Liisa (Fi6) had no disfluent pauses before the stay, but was still rated average in fluency (fluency rating = 2.9). Similarly, Jonna (Fi11) in the middle of the stay (fluency rating = 3.2) and Petra (Fi7) following the stay (fluency rating = 3.1) did not stand out has having very high fluency ratings. However, the other speakers with no disfluent pauses received a fairly high fluency rating.

Apart from pause frequencies, Study I also focused briefly on other disfluency features. First, I studied pause placement[71], and was able to indentify where disfluent pauses might typically occur. Interestingly, disfluent pauses occurred frequently in three places (Study I: Table 3). However, the reasons for pausing at these places remained unclear. These constructions did not, unlike in the previous studies (see, e.g., Lehtonen & Heikkinen 1981, 329–336; Guion et al. 2000, 209) contain rare, long lexical items that would have been difficult to

71 I use the term pause placement, by which I mean the positions of the pauses in the utterance (the term *pause distribution* has also been used in other studies).

articulate and thus viewed as possible reasons for disfluencies. There was no line feed either in the original texts at these places that could have explained the use of pauses. The last place (Study I: Table 3, C) is the most unexpected one because it is just before the end of the turn in a very commonly used construction *часов (pause) в девять 'chasov (pause) v devyat''*[72]. Perhaps students paused here to verify whether the text said *девять 'devyat''* or *десять 'desyat''*[73].

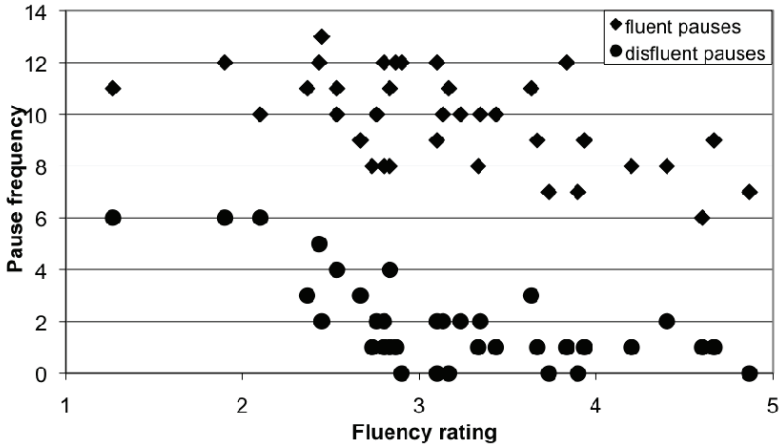


FIGURE 11 Relationship between pause frequency and mean fluency rating of the samples.

In addition, Study I measured the frequency of repairs and repetitions in order to identify possible *disfluency clusters* (as also shown by Riggenbach 2001) giving an impression of disfluency. This showed that in each recording session the two least fluent subjects, at least, had the most "disfluency clusters".

Next, I will compare the results of the acoustic analysis with the fluency ratings. Lennon (1990, 414) argues that interspeaker differences are mainly shown in differences in pausing rather than differences in the articulation rate. He proposes that for fluency acquisition it is pause placement, pause duration and the frequency of pauses that play a crucial role. The results of Study II can be seen as partly supportive of this finding, as it was shown that speech rate (inclusive of pauses) was more important than articulation rate (minus pauses) in determining the fluency of the speech samples. As Table 17 shows, it was, however, pause frequency that had the highest correlations with the fluency rating (Ullakonoja 2008b). The difference between the correlation of pause frequency with the fluency ratings and that of speech rate with the fluency ratings is nevertheless very small.

My findings strongly support the claim that pausing as well as speech and articulation rates are important in determining whether L2 learners' speech is perceived as fluent or not. The results contradict Lehtonen's (1978, 56) findings in an experiment where "a faster rate of speech or a smaller number of pauses

72 In English: 'nine o'clock'.
73 In English: 'ten'.

was not felt to be more fluent" and in which pausing in fluent speech (in L1 Finnish or L1 English) did not follow a specific pattern. Lehtonen (1978) argued that fluency is such a complex concept – involving also the linguistic content and communicative context – that it is not possible to define it merely by experimental phonetic means. However, as my results seem to suggest, phonetic factors are an important element of how fluency is perceived and evaluated. It is not only possible to measure prosodic factors, but they also seem important in defining fluency, even if other factors (lexical, grammatical, social etc.) are ignored. For example, as previously mentioned, Lehtonen (1981) implied that pause duration might be an indicator of fluency because it was shorter in fluent readers' read-aloud speech. My results confirm this finding (see Table 17). It is easy to agree with the conclusion drawn by Riazantseva (2001) that by teaching students the L2 pausing patterns they would be perceived as more native-like in their fluency.

		Mean perceived fluency rating		
		Pearson Correlation	**p**	**n**
Pause frequency		-0.742	0.001	36
Articulation rate	Phonetic words/s.	0.484	0.003	36
	Syllables/s.	0.416	0.012	36
Speech rate	Phonetic words/s.	0.722	<0.0001	36
	Syllables/s.	0.697	<0.0001	36

TABLE 17 Correlation of pause frequency, articulation rate and speech rate with perceived fluency rating (Ullakonoja 2008b).

It is clear that my results show that students should speak faster in order to sound more fluent. Other studies also support this claim. For example, it has been found that native speakers would like L2 users to speak with a speech rate about 10 % faster than they do (Munro & Derwing 1998; 2001, 464). They also evaluate a fast speech rate by an L2 speaker more positively than a slow one (Paananen-Porkka 2007, 340). Nevertheless, speaking faster should not be achieved at the expense of pronunciation accuracy. It is possible that SA increases students' speech rate as their language skills increase and they become more confident and faster readers. For example, Walsh (1994, 51–52) proposed that SA makes students speak faster but also with more errors in grammar, syntax, vocabulary and even pronunciation.

Study III addressed the self-evaluations of the students in comparison with the teachers' fluency ratings. I agree with Pellegrino (1998) that it is perhaps impossible to generalize students' self-perceptions as they reflect individual experiences. Self-evaluations should be interpreted with some caution as it has often been found that students are likely to evaluate their language development positively due to being content with the improvement of their language skills during their stay abroad (Huhta 1994; Pellegrino 1998). However, learners' self-perceptions can also be guided by the classroom-based idea of grammatical correctness (rather than, e.g., communicative competence).

Hence, they may perceive themselves as unsuccessful language users in the SA context. This in turn may lead to benefiting less and less from opportunities to use the L2. (Pellegrino 1998.) However, the individual differences may be summarised in accordance with Segalowitz et al. (2004, 14):

> "The more the adult learner is able to communicate in the target language the more he or she will do so. As a result, the very act of communicating will further enhance learning, leading to more communication, which should promote further learning".

The amount of L2 exposure is likely to be greater during SA than at home. In Russia the students have the possibility of receiving a wide variety of L1 speaker input in L2. However, most students might still be getting most of their L2 input from teachers (especially if living together with other Finns). Teachers are, of course, L1 models, but as has been suggested by Hatch (1983, 154–159), L1 speakers tend to speak to foreigners at a slower rate, and using more pauses, more intonational variation and greater intensity than they would when speaking to another L1 speaker. It is possible that this is also the case with teachers, who also often slow down their speech rate and make longer pauses when addressing L2 students.

To summarise, the overall findings (Studies I, II & III) showed that the SA context provides a learning context that is beneficial in many ways to adult L2 learners. When they are surrounded by the L2, I believe that they become more confident in using it (see also Segalowitz et al. 2004, 14), and hence also more fluent. In the studies reported here, the measurement of prosodic characteristics (namely pausing and speech rate) showed that fluency increased during the SA. It was also found that the students themselves felt that they increased their fluency during the SA. Moreover, the teachers' perceptions confirmed this finding when they evaluated samples of students' speech.

4.2 Pitch as an Intonational Feature

4.2.1 Summary of the Main Findings (Studies IV, V and VI)

First, I summarise the findings of the three studies (IV, V and VI) that concern pitch. I will begin with the results on pitch range. They showed, e.g., that most Finnish speakers had a narrower pitch range and a smaller standard deviation of pitch in Finnish than in Russian. The mean pitch of the Finnish speakers in the present study was 206 Hz in Finnish and 214 Hz in L2 Russian as compared to the 243 Hz of the L1 Russian speakers. When comparing the three utterance types (declaratives, questions and exclamations), Study IV showed, that in L1 Russian, declaratives differed significantly from questions and exclamations in mean pitch but not in pitch range. In contrast, in L1 Finnish, exclamations differed significantly from declaratives and questions in mean pitch and pitch range. In L2 Russian all three utterance types differed significantly in mean pitch, but not in pitch range.

I now turn to Russian YNQs and, more precisely, how they were produced by Finns and how L1 Russian speakers evaluated these productions. The main finding of Study V was that overall only 57 % of the utterances intended as questions by Finns were interpreted as such by L1 Russian listeners. In other words, 43 % of the Russian utterances produced by Finns were not perceived as questions by L1 speakers even though they were intended as such. Study V offers multiple explanations for this rather low overall value, but fails to give a single leading explanation. The acceptability ratings, i.e., when the native speakers were asked to rate how acceptable each stimulus would be as a question (on 1–5 scale), yielded similar results. The results from both tasks show great interspeaker and intraspeaker variation. In only one of the seven utterances was a question nearly always (in over 80 % of the cases) successfully produced (Q4: *Ty rada za menya?*).

Finally, Study VI investigated the L1 and L2 production of Russian YNQs, and also compared the latter to the evaluation of the interrogativity task presented above in Study V. Each question was studied separately and for different questions the cue for interrogativity varied: it could either be the sharpness, size or place of the peak in the pitch contour. In general, L2 YNQs were characterised by a flatter and lower peak that occurred earlier than in the L1 utterances. A peak-delay, i.e., the phenomenon where the maximum pitch is not on the stressed syllable (of the intonation centre) but on the unstressed syllable following it, was observed in the L1 but not L2 speech. The pitch measurements of Study VI confirmed the findings of Study IV of a lower mean pitch for L2 than L1 speakers in Russian. In addition, the measurements showed that maximum pitch and mean absolute slope are also significantly lower for L2 speakers. This in turn showed that the speed of pitch change was smaller for L2 speakers. By and large, there was great interspeaker variation as well as variation between the utterances in question.

4.2.2 Pitch Range

This subsection presents the main results for pitch range in Study IV. The different ways of defining pitch range in the earlier studies were presented in subsection 2.2.1. The method used in Study IV was calculation of the difference between maximum and minimum pitch in addition to measuring mean pitch[74]. Many researchers look at pitch range on the utterance, breath-group or tone-unit level (see, e.g., Crystal 1969, 143). I am more interested in the individual range or, more precisely, the range speakers choose to use when reading aloud in Finnish (L1) versus Russian (L2). Hence, the unit of investigation in Study IV was the whole speech sample of each speaker and individual utterances when studying pitch range in different utterance types. This subsection deals with pitch range, i.e., the difference between the maximum and minimum pitch as well as the mean pitch of individual speakers. Study IV discussed pitch range

[74] I also calculated the 95 % range, but as the comparison of the L1 Finnish and L1 and L2 Russian did not yield different results, it was left out of the final paper.

from two aspects: global pitch range and pitch range in different utterance types. For the sake of clarity the same structure is kept in this subsection. By pitch range I refer to mean pitch and pitch range.

First, I will characterize *global pitch range*, i.e., the speakers' pitch range across the entire recording. Table 18 below represents a comparison of the mean pitch values in L1 Finnish female speech reported in different studies. All the values are in Hz and have been obtained from read-aloud speech. As mentioned by Laukkanen & Leino (1999, 41, 101–103, 149), mean pitch is influenced by such individual differences between the speakers as age, educational and cultural background. They reported that Finnish female students' mean pitch varied from 151 Hz to 246 Hz and that there was great interspeaker variation, the typical female pitch being around 200 Hz.

The wide variation between the values obtained in different studies (Table 18) can also be explained by the fact that the studies differ in the number of subjects and in what they have measured. Whereas Syrjä (2007) and Lennes (2009) measured only vowels and de Silva et al. (2003) vowels and non-obstruents, Laukkanen et al. (1999), Rantala (2002), Valo (1994) and I measured all the voiced elements. For example, in Study IV, I measured all possible pitch points, including possible creaky voice. The mean pitch value found in Study IV is slightly higher than the values reported in most of the other studies, but the difference is not great. Hence, it is possible to conclude that the subjects of Study IV represent a fairly typical population of Finnish female speakers as far as mean pitch is concerned. The fact that their mean pitch is slightly higher than in previous studies could be due to their age: the speakers were all under 25 years old.

De Silva et al. (n = 3)	**Laukka-nen & Leino (n = 190)**	**Laukka-nen et al. (n = 46)**	**Lennes (n = 6)**	**Rantala et al. (n = 33)**	**Syrjä (n = 7)**	**Ullakon-oja Study IV (n = 9)**	**Valo (n = 4)**
179 Hz	194 Hz	190.8 Hz	8.92 ST = 167.5 Hz	236.5 Hz	189.5 Hz (1998 data) 190.7 Hz (1999 data)	**206 Hz**	184.3 Hz

TABLE 18 Female speakers' mean F0 in L1 Finnish in the studies by de Silva et al. (2003) Laukkanen et al. (1999), Laukkanen & Leino (1999, 149), Lennes (2009) Rantala et al. (2002), Syrjä (2007, 117), Valo (1994, 96–100) and Study IV.

De Silva et al. (2003) report a considerable difference between L1 Finnish and L1 Russian females in mean pitch in read-aloud speech (see Tables 18–19). Despite the fact that the data is from only two Russian and three Finnish speakers this finding raises an interesting question from the L2 point of view: do native Finnish speakers speak Russian with a higher mean pitch than Finnish and, if so, do they try to adapt to the L1 Russian mean pitch as their language proficiency increases during SA? This was in fact one of the research questions addressed in Study IV. The results of Study IV showed that the learners' mean pitch in

Russian (see Table 19) is considerably lower than that of the L1 speakers. It is, however, slightly higher than most of the mean pitches for Finnish female speakers summarised in Table 18 and, more importantly, higher than in their own Finnish (Study IV).

Table 19 presents a comparison of both native speakers' and L2 learners' mean pitch in Russian. It should be pointed out that average pitch values for female Russian speech have not been reported in many studies. Also, to my knowledge, Study IV is the only one to date where mean pitch in L2 Russian has been calculated. In Volskaya's (2009b) study the five female speakers' mean pitches varied widely from 195 Hz to 246 Hz in read-aloud speech, and equally as much in spontaneous speech. Speakers who were over 30 years old had lower pitches than the two younger ones, who both had a mean pitch of 240-250 Hz in read-aloud and spontaneous speech. Although Volskaya's (2009b) mean pitch values are much lower than those of other studies, comparison of Tables 18 and 19 clearly shows that the mean pitch for women tends to be higher in Russian than in Finnish. It should be stated that the existing studies on Russian are possibly based on only a small number of speakers, and hence, the results are not as reliable as the studies of Finnish mean pitch, most of which include a much greater number of speakers.

Study	**Mean pitch in L1 Russian**	**Mean pitch in L2 Russian**
De Silva et al. (n = 2)	245 Hz	
Kodzasov & Krivnova (n not specified)	260 Hz	
Martynov (n not specified)	240 Hz	
Mikheev (n not specified)	256 Hz	
Volskaya (n = 5) Read-aloud speech Spontaneous speech	 218 Hz 220 Hz	
Ullakonoja: L1 (n = 7), L2 (n = 9)	**243 Hz**	**214 Hz**[75]

TABLE 19 Comparison of female speakers' mean F0 in L1 and L2 in the studies by De Silva et al. (2003), Kodzasov & Krivnova (2001, 110), Martynov (1971, cited by Shtern 1988, 203), Mikheev (1970, cited by Shtern 1988, 203) and Volskaya (2009b) with the results of Study IV.

To obtain a more complete picture, the values of the individual speakers of Study IV were also examined. As Figure 12 shows, there is wide interspeaker variation in mean pitch (Hz), both in L1 Finnish as well as in L1 and L2 Russian. The mean pitch for each Finnish speaker is always lower in Finnish and higher in the two Russian recordings (prior to the stay and following it) except for one speaker (Noora-Fi1), who has the same mean pitch in Finnish and in Russian following the stay, but whose pitch was also higher in Russian prior to the stay. This suggests that all the Finnish learners of Russian in Study IV seem to distinguish between the two languages in mean pitch by using a higher mean pitch in Russian than in Finnish. This might be considered a step in the right

75 The value was the same both prior to and following the students' stay in Russia.

direction, since as discussed above, Finnish is generally spoken at a lower mean pitch than Russian. Ilona (Fi9), in particular, uses a much higher mean pitch in Russian than in Finnish.

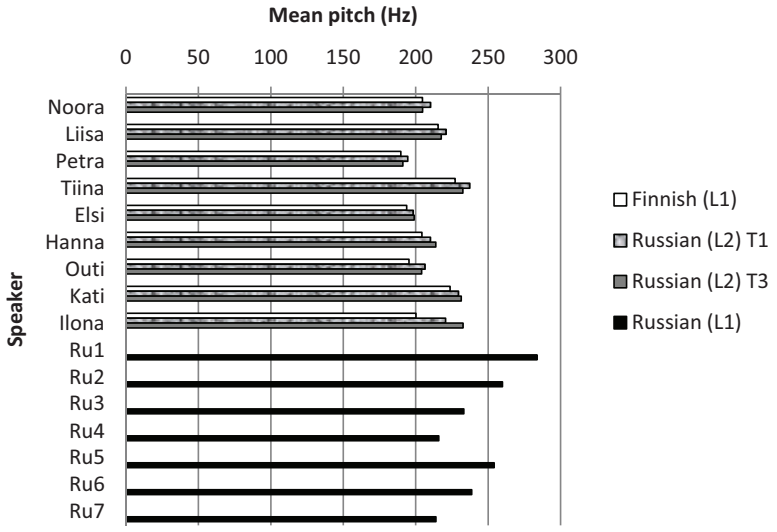


FIGURE 12 Mean pitch of the speakers in Study IV (T1 = prior to the stay in Russia, T3 = following the stay) in Hz.

Turning now to the differences between maximum and minimum pitch, Table 20 summarises the studies on pitch range in L1 Finnish and L1 and L2 Russian. Studies on pitch range in L1 Russian are scarce. In addition, investigation of pitch range in any L2 is relatively rare (see Mennen 2007, 55 for a short review). As it is hard to draw conclusions on the basis of such a small number of earlier studies, I will mostly focus on discussing my own results.

Study	Pitch range in Finnish	Pitch range in Russian (L1)	Pitch range in Russian (L2)
Laukkanen et al. (n = 46)	7 ST		
Syrjä (n = 7)	11 ST		
Martynov (n unknown)		7.0 ST	
Mikheev (n unknown)		19.6 ST	
Volskaya (n = 5)		Read-aloud speech: 7.8 ST Spontaneous speech: 7.1 ST	
Ullakonoja: Finnish (n = 9), Russian L1 (n = 7), Russian L2 (n = 9)	**209 Hz**	**249 Hz**	**215 Hz prior to the stay 221 Hz following the stay**

TABLE 20 Comparison of female speakers' pitch range in L1 and L2 in the studies by Laukkanen et al. (1999), Martynov (1971, cited by Shtern 1988, 203), Mikheev (1970, cited by Shtern 1988, 203) and Volskaya (2009b) with the results of Study IV.

Mennen (1998a), Aoyama et al. (2007) and Toivanen (1999) have all found that L2 learners of English do not use as wide a pitch range as L1 speakers. This finding might, however, be related to the pitch range in the learners' L1 and thus explained as transfer from L1, as argued by Mennen (2006). If languages do differ in pitch range, this would, of course, mean that the L2 learner would not always have a narrower pitch range in L2 than in L1. For example, L1 Russian speakers (with a relatively wide pitch range in L1) learning Finnish (a language with a relatively narrow pitch range) would need to learn to narrow instead of widen the pitch range they use in L1. Comparing the pitch ranges of L1 and L2 speakers is not, however, a straightforward task. As Toivanen (1999) observes, L1 and L2 subjects may have different expectations and skills with respect to the same reading task. Toivanen considers that L1 speakers are capable of acting out a dialogue (and acting usually involves the use of a wide pitch range and high pitch with a lot of variation) whereas L2 speakers may merely read the text.

The overall differences in pitch range between the different groups of speakers analysed in Study IV show that L1 Russian speakers used a wider pitch range than L2 speakers. Also, the students' mean pitch range was slightly narrower in Finnish than in Russian. Most students also were found to use a wider pitch range in Russian following their stay in Russia than prior to it (Study IV). As Figure 13 illustrates, there are also wide individual differences in pitch range both in L1 Finnish and Russian as well as in L2 Russian.

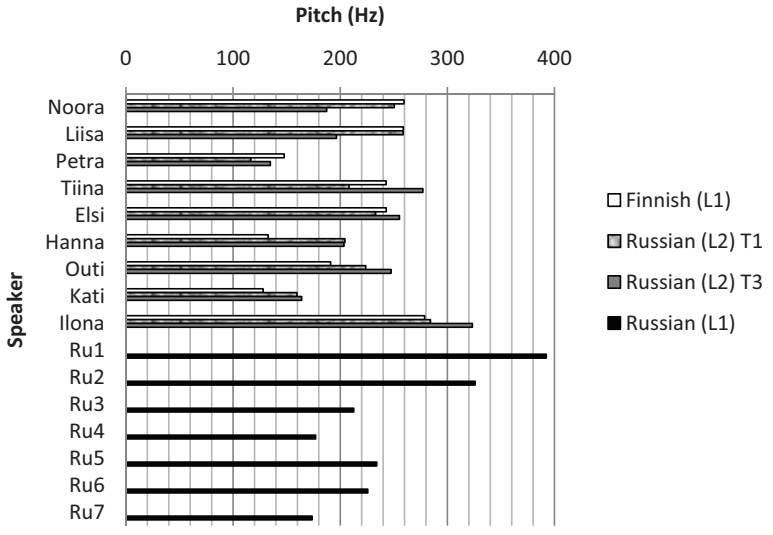


FIGURE 13 Pitch range of the speakers of Study IV (T1 = prior to the stay in Russia, T3 = following the stay) in Hz.

On average, in their L1s Russians had a wider pitch range than Finns (Figure 13). Four students (Hanna-Fi6, Outi-Fi7, Kati-Fi8 and Ilona-Fi9) had a narrower pitch range in Finnish than in either of the Russian L2 recordings. Two students

(Tiina-Fi4 and Elsi-Fi5) had a narrower pitch range in Finnish than in Russian following their stay[76]. Another student, Petra (Fi3) had a more native-like pitch range in Russian following her stay. Hence, it seems that six out of the nine students used a different pitch range for the two languages. Interestingly, however, two students (Noora-Fi1 and Liisa-Fi2) showed the opposite tendency. They both had almost the same pitch range in Finnish and Russian prior to the stay. However, their pitch range in Russian following the stay was much narrower. Perhaps the result would have been different, if pitch range in Finnish had also been measured following the stay instead of only prior to the stay. This could, in fact, be measured in the future as the recordings in Finnish made following the stay have been stored. Also, it would be interesting to complete the analysis with the L2 Russian recordings done in the middle of the stay (T2) as Study IV was only concerned with comparing the recordings done prior to and following the stay.

According to Laukkanen et al. (1999), the narrower pitch range of Finnish speakers in comparison with speakers of other languages is due to the relatively simple intonation system of Finnish that does not necessitate a wider pitch range. Volskaya's (2009b) results show a very narrow pitch range in L1 Russian. It is possible that these results have been calculated differently from the other studies in Table 20, i.e., in intonation units. Thus, if pitch range has been calculated for each intonation unit separately and then averaged, it will produce a different result from that obtained in the present study, where the pitch range is calculated over the entire speech sample. For example Odé (1989, 92) determined that a range of 10–23 ST is typical in Russian. Odé (1989, 115–125) found that pitch range was narrower for falling pitch contours (4–10 ST) than for rising pitch contours (10–17 ST).

Next, I review the results of Study IV in the light of previous studies on *pitch range in different utterance types*. In Study IV (on read-aloud speech) in L1 Finnish, the mean pitch was statistically significantly higher in exclamations than questions and statements[77]. The pitch range, on the other hand, was statistically significantly narrower in exclamations than questions and statements in L1 Finnish. This result is consistent with Anttila (2009) in that questions and statements did not differ in pitch range.

The comparison of pitch range and mean pitch in different utterance types in L1 Russian in Study IV showed that statements differ significantly from questions and exclamations in mean pitch but not in pitch range. Mean pitch was significantly lower in statements than in other utterance types. This finding contradicts that of Nikolayeva (1977, 84–85), who reported a wider pitch range in questions than statements in Russian.

In Study IV, in addition to calculating the pitch range of different utterance types, the pitch contours were also visually compared. Only three examples of superimposed pitch contours were given (Study IV: Figures 1–4).

---

[76] There is a typo in Study IV (p. 1702): "…for most speakers (**5**/9) it was narrower than in Russian". It should be for most speakers (**6**/9)…

[77] The term for a statement used in Study IV is 'declarative'.

When the superimposed contours were examined, first, some interspeaker variation was observed. Not all the Russian speakers had pitch contours as "lively" as the one in Study IV, Figure 1, but the contours for the other L1 Russian speakers resemble the one chosen for presentation. In Finnish, most speakers' superimposed pitch contours were similar as the one in Study IV, Figure 2. One exception was Kati (Fi8), but the difference may be due to problems in the pitch estimation algorithm, as the contours present a lot of jumps down to around 100 Hz. Of the L2 learners, only one Finnish speaker, Elsi (Fi9), had superimposed Russian contours close to those of the L1 speaker (Study IV: Figure 1). The contours for the remainder varied a great deal. Some were closer to Finnish (Study IV: Figure 2) whereas others were closer to Russian (Study IV: Figure 1). A comparison of the two recording sessions (prior to vs. following the stay in Russia), showed that some students improved in producing questions (namely pitch on the nucleus). To be able to draw more specific conclusions on the L2 pitch contours, a more detailed experimental study was performed in Study VI.

To conclude, together with the earlier studies Study IV showed that pitch range can be a language-dependent feature and that the L2 learners of Russian were trying to adapt to the wider and higher pitch range of Russian as their language skills improve.

4.2.3 Pitch Contours in YNQs

Turning now to studying pitch contours in Russian YNQs, more attention will be paid to pitch movements during a particular utterance. This is of interest because, as we saw earlier, in Russian questions may be distinguished from statements solely by prosody, mainly pitch, a fact clearly important for L2 Russian.

First, a pilot study (Ullakonoja et al. 2007) was conducted to compare the pitch contours in statements, exclamations, question-word questions and YNQs for two Finnish learners of Russian drawn from the corpus. It showed that out of all the utterance types the subjects had most difficulties with YNQs. Hence only YNQs were selected for further investigation in the present study. They were studied both from the recognition and production points of view. Production was studied by investigating the L2 speakers' productions acoustically, in Study VI. Recognition, on the other hand, meant the L1 speakers' evaluation of the L2 learners' production as questions, which was explored in Study V. The results of the acoustic study were compared to the evaluation ratings of L1 speakers obtained in Study V. The aim of these two studies was, through experimental research, to shed light on the possible difficulties L2 speakers face when producing YNQs in Russian.

4.2.3.1 Evaluation of Pitch Contours in YNQs by L1 speakers

A central issue in studying the pitch contours of L2 Russian speakers in YNQs was how these are perceived by L1 Russian speakers. The results of the

listening evaluation task reported in Study V are interesting in two ways: the sensitivity of the L1 Russian speakers to pitch changes and the L2 learners' capacity to produce acceptable interrogative intonation. As mentioned above, the main focus in Study V was the L2 Russian speakers' productions and their acceptability in the ears of native speakers as interrogatives. Investigation of Russian L1 speakers' perception of pitch contours per se would have needed another type of experiment involving pitch manipulation.

As described in section 3.4, in Study V, L1 speakers were asked to evaluate the students' productions in two tasks: categorisation (question or not) and acceptability rating (on a 1–5 scale, 1 = "not a question", 5 = "a question"). In Study V, I used the term recognition rate to refer to the percentage of "question" answers by the listeners in the categorisation task. This may be a slightly misleading choice, as the listeners were told before listening, that the speakers were trying to utter an interrogative, and hence it is not a question of recognising an interrogative, but rather deciding if it is good enough to be said to be one ('goodness of fit'). However, for the sake of clarity, I will also continue to use the term recognition rate in this subsection.

Both tasks yielded similar results. In the categorisation task, 57 % of the utterances intended as YNQs were perceived as questions by the L1 Russian speakers. In the acceptability ratings, the overall mean was 2.95, which is even slightly poorer than the result of the categorisation task. Both the recognition ratings and the acceptability ratings showed considerable interspeaker differences as well as differences between the questions.

One might have expected an overall higher recognition rate on the basis of, e.g., the studies by Kuosmanen & de Silva (2003; 2007). After all, the subjects were intermediate-advanced learners of Russian, who had studied the L2 intensively and successfully for over four years and who had also been extensively exposed to the L2 in a SA setting. Furthermore, they had had a course in phonetics (both practical exercises and theoretical lectures) during their first year at university, where they had explicitly been taught the different IKs of Russian. Study V clearly confirmed the earlier results of Kuosmanen & de Silva (2003; 2007) by underlining the great difficulty of the Finns in producing an acceptable interrogative in Russian.

Although the overall recognition rate was weak, it can also be compared to Shcherbakova's (2001) results on how native Russians (without prior knowledge of Finnish) recognised questions in Finnish. In that study, she found a recognition rate of 21 %. In comparison to the 57 % obtained in the present study for Russian YNQs uttered by Finns, it proves that Finnish speakers are definitely doing something different when uttering Russian YNQs than Finnish questions. The fact that the judges of my study were not exposed to L2 Russian very often in their every-day lives could be one explanation for the strictness of their ratings. That is, they might have rated the interrogatives as less acceptable than people who were used to hearing foreign-accented Russian. Also, as a study by Rietveld & Gussenhoven (1987) suggests, a particular intonational structure has a particular temporal structure, and hence the listeners in my

experiment may have also been disturbed, not only by the different pitch movements but also by a speech rate they were not used to associating with the particular pitch contour. The literature review (section 2.2.3) suggests that YNQs are spoken faster than statements and consequently a slower speech rate here might have increased the "non-question" ratings.

Furthermore, on the basis of the literature review, it can be proposed that, when speaking their L1 the subjects would be likely to use similar pitch contours (but perhaps for a different function) to those needed to utter an interrogative in the L2. In other words, one could argue that the Finnish learners are able to produce the pitch contours needed in Russian to express an acceptable YNQ, but as in their L1 the same pitch contour has a different function, they are unable to use that contour for this specific purpose. Also, they may be aware of the pitch contours that should be used in Russian YNQs, but cannot put this explicit knowledge into practice. Finally, it is also possible that they are not familiar with the pitch contour they should produce when the nucleus is in other than utterance-final position. This type of contour requires a very sharp peak on the nuclear syllable, which may be difficult for Finns. With respect to the recognition rate for each question, noticeable differences between the questions were observed. However, Study V did not give an explanation for success in some and failure in other questions. The explanations given in the study include syntax and the lexicon as potential influences along with the frequency of use of the constructions.

The learners' productions were also compared in the two recording sessions, the one in the middle of their stay in Russia (T1 in Study V, T2 here) and the one following it (T2 in Study V, T3 here). In Figure 14 the overall recognition rate is calculated for each speaker comparing the middle of the stay and following it. Study V focussed on the L1 Russian speakers' evaluation of the successfulness of the L2 learners' production. The change in the recognition rate during the second half of the stay can clearly be seen in the figure: both Ilona's (Fi3) and Marjo's (Fi4) T3 questions are recognised better than at T2. Sanna's (Fi6) recognition rates show a similar tendency but to a smaller extent. Liisa (Fi1), Kati (Fi2) and Ritva (Fi5), however, show the opposite tendency: the recognition rate of their questions decreases from T2 to T3.

The mean acceptability ratings, as illustrated by Figure 15, developed in three ways during the second half of SA. Here again, T2 refers to the middle of the stay and T3 to the recordings following the stay. For Ilona (Fi3), Marjo (Fi4) and Sanna (Fi6) an increase in the acceptability ratings can be seen. Ritva (Fi5), on the other hand, had almost the same acceptability rating in T2 and T3. The decline in the acceptability ratings was observed for two students (Liisa-Fi1 and Kati-Fi2).

In sum, the mean values of the categorisation task showed a significantly poorer recognition rate in T3 than T2, whereas the mean acceptability ratings showed a significantly better result in T3 than T2. Although the results are contradictory, the differences between T2 and T3 were not great in either task. Hence, we can speak about, if not the improved skills, then at least the retention

in learning of the pitch contours, as there was an interval of approximately one month between the students returning home and the making of the recordings. As Figures 14 and 15 showed, there were wide interspeaker differences in the development patterns.

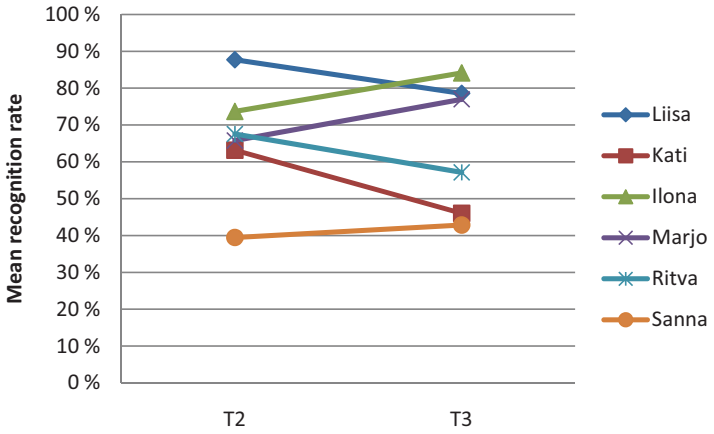


FIGURE 14 Recognition rate of the speakers of Study V (T2 = middle of the stay in Russia, T3 = following the stay).

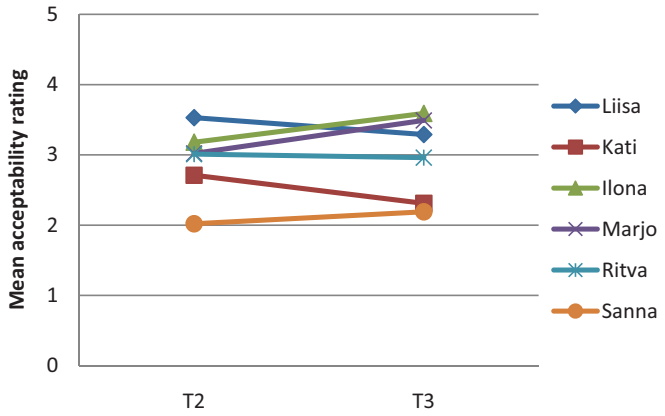


FIGURE 15 Mean acceptability ratings of the speakers of Study V (T2 = middle of the stay in Russia, T3 = following the stay).

4.2.3.2 Production of Pitch Contours in YNQs by L2 speakers

Hence, it became clear that L2 speakers were not highly successful in producing Russian YNQs according to native-speaker perceptual evaluation. The purpose of the last study (Study VI) was to trace the possible reasons for this through acoustic analysis of the students' productions. To a certain extent, Study VI was a continuation of Study V. In Study VI the acoustic realisation of L1 and L2

pitch contours was studied experimentally and compared with the recognition rate obtained in Study V. The acceptability ratings of Study V, however, were not compared with the students' acoustic realisations.

The existence of an "intonational grammar" or complete description of pitch contours in YNQs in Russian would have facilitated the study. Instead, the six L1 speakers of the corpus served as "a model" of the pitch contour for each utterance studied. However, it was found that the pitch contours produced by the native speakers were not always alike. It was hence sometimes difficult to decide what contour to use as the point of comparison. However, the listening experiment of Study V helped to determine the L2 contours that were perceived as interrogatives by native speakers and those that were not.

One goal of Study VI was to define what kind of pitch contours L2 speakers use in Russian YNQs. This was of particular interest as the perception results of Study V had showed that only 57 % of the YNQs produced by Finns were interpreted as such by L1 Russian speakers. The production results yielded great differences between different questions as well as between speakers. The comparison of the L1 and L2 pitch contours showed that in general L2 productions differed from L1 productions a great deal. In L2 utterances (YNQs) the peak was in general flatter and lower as well as positioned earlier than in the L1 utterances. The pitch measurements showed that the L2 speakers had a significantly lower mean pitch and narrower mean absolute slope than the L1 speakers. Also inspection of the seven questions (see Study VI: Table 1) separately showed that some of them had very rarely been identified as questions.

On the basis of Study VI, the factors that I consider may possibly account for the L2 speakers unsuccessful productions in Russian YNQs are: 1) height and sharpness of pitch peak, 2) peak position (including peak-delay), 3) speech rate, and 4) creaky-voice. Below, I will discuss each of these in more detail.

First, the analysis focused on the sharpness of the peak, its position and height in the pitch contours of YNQs. The study showed that in general the L2 questions that had low recognition rates had a flatter and lower peak than the L2 questions which received high recognition rates. The flatness of the peak was not measured in absolute values, but detected only by visual observation of the contours. Pitch height, which was given in Study VI as one explanation for the unsuccessful L2 productions, has been shown by Makarova (2007) to distinguish questions and exclamations from statements. Accordingly, in questions the pitch peak should be high. In some questions, L2 speakers were rather successful (and also received high recognition rates) whereas in other questions they were not successful at all. As reported in Study VI, the height of the pitch peak flagged a correlation with the recognition rate.

Second, another plausible explanation for the unsuccessful productions of the Finns may be the position of the peak, which may be placed on either a different word or on a different syllable of the appropriate word than in L1 speech. The relation between the position of the peak and the recognition rate was detected both visually from the pitch contours as well as from the

measurements (the word on which the peak is positioned was determined and the timing of the peak within that word was calculated). No correlations were calculated, but instead unsuccessful productions were compared against successful ones, resulting in the observation that peak position was inappropriate in the unsuccessful productions.

Problems in positioning the peak have been reported previously by Mennen (1998b), even for highly proficient speakers. In Study VI in the L2 speakers' unsuccessful productions, the incorrect word or several words were accentuated in the course of the utterance, whereas according to the traditional descriptions, only one word in the utterance should be accentuated. The problems with positioning the peak can also rise from peak-delay. As shown above, peak-delay[78] was typical in native Russian YNQs. Volskaya (2007), Meyer & Mleinek (2006) and Makarova (2007) also showed evidence that peak-delay exists in Russian and can be used in Russian to distinguish questions from declaratives and exclamations. In Study VI, peak-delay seemed very systematic in YNQs in L1 Russian whereas it was practically non-existent in L2 Russian. This may offer a further possible explanation for the rather low overall recognition rate of the utterances intended as questions by Finns.

Third, another argument that could explain the unsuccessful productions of the Finns concerns speech rate, which, as was mentioned above, is known to be faster in Russian YNQs than statements[79] (Svetozarova 1975; 1982). It is possible that, because L2 speakers in general speak slower than L1 speakers, it has an effect on the interpretation of their isolated production of questions. This is only speculative as the relationship between speech rate and recognition rate was not investigated in the study.

Fourth, creaky voice, which is typical at the end of Finnish questions (see, e.g., Anttila 2008, 54–56) was also observed in the Russian of the Finnish students. It can be speculated that creaky voice might have disturbed recognition of the utterance as a question in cases where the interrogative contour would have required a final rising contour at the utterance-final position.

The results of Study VI can be compared to Anttila's (2008, 77, 90) results on read-aloud Finnish as she used a part of the present Finnish corpus for her study. She reports the following pitch contours for the students' YNQs: 13 rise-fall, 5 final rise, 3 high initial and 12 high overall pitch contours. Some of the eight students used by Anttila are the same as those in Study V. The comparison shows that despite the fact that the students use final rise and rise-fall patterns in their L1, they fail to use them appropriately in Russian. This can be explained by the fact that their peaks in Finnish (L1) are not as sharp and high as is needed in Russian (L2).

---

78 See p. 42 for a definition.

79 Van Heuven & van Zanten (2005) have also found that questions are spoken faster than statements in Manado Malay (an Austronesian language), Orkney English and Dutch.

Finally, Kuosmanen & de Silva (2003; 2007) argue that longer Russian questions are harder than shorter ones for L1 speakers to recognise when they are produced by Finns. Study V did not support this finding, as the longest question in the data (Q7) was not the hardest to recognise and the shortest one (Q5) was not the easiest. However, turning to the results of the acoustic analysis in Study VI, the longest question was rather difficult for L2 learners to produce, as none of them produced a native-like or even near native-like contour. It also received rather poor recognition rates in Study V. Thus, L2 learners seemed to hesitate over peak placement and some ended up simply with having too many peaks on this particular question. Hence, in producing YNQs, the length of the utterance may possibly disturb production.

My results contradict Zinder's (1980, 128–129) claims that interrogative contours are rather similar across languages, and thus learning to produce Russian YNQs is not particularly challenging for L2 learners. However, Keijsper (1983, 125) claims that intonation in Russian YNQs is in fact difficult for L2 speakers, because in Russian YNQs the pitch fall extends to the postcentral syllable. My results are along similar lines: they seem to argue for the view that peak-delay (i.e., the phenomenon that the peak is postponed to the postcentral syllable) makes it difficult for L2 learners to produce an acceptable YNQ in Russian.

The Finns' difficulty of producing YNQ in Russian can also be explained in light of Haan's (2001) Functional Hypothesis, which argues that syntactical/lexical marking of a question predicts the presence of high pitch. In other words in YNQs with no syntactical/lexical marking high pitch is maximally present, whereas in question-word questions and statements less so. While Haan's (2001) work was based only on Dutch, it has been suggested that the phenomenon is typical of L2 prosody in general (Chen & Mennen 2008; Mennen et al. 2010). I will therefore briefly discuss my results in the light of this hypothesis. According to my interpretation, the hypothesis suggests that in Finnish YNQs (where interrogativity is marked by morphological means) high pitch is minimally present while in Russian YNQs (most often without syntactical/lexical/morphological interrogativity marking) high pitch would be maximally present. As, in addition to higher mean pitch, high pitch can also be manifested by wider pitch range, this interpretation also seems reasonable when the results of Study IV on mean pitch and pitch range in different utterance types are taken into account. That is, I reported in Study IV that, in Finnish, only exclamations differed statistically significantly from questions and declaratives in mean pitch and pitch range while, in Russian, questions and exclamations were uttered with a higher mean pitch and wider pitch range than declaratives. In other words, in Finnish, questions and declaratives did not differ from each other, as interrogativity is marked by morphological means, but in Russian they did, as they are less marked in this way.

All in all, producing pitch contours in Russian YNQs proved difficult for Finns. Although the subjects were intermediate-advanced speakers of Russian they were not highly successful in producing a pitch contour that would be

interpreted as interrogative by native speakers. This finding has been reported earlier by Kuosmanen & De Silva (2003; 2007). Study VI clearly showed a need for a thorough investigation of pitch contours in Russian (L1).

4.3 Summary of the individual development of fluency and intonation during SA

Above, I also discussed the individual development of prosody during SA in connection with each study. Here, the purpose is to identify the speakers who gained most during SA and who were the most successful by the measures of the present studies.

By most of the fluency measures (perceptually evaluated fluency, pause duration and speech and articulation rate) Elsi and Ilona were the most successful students. When disfluent pauses were measured, Ritva, Petra and Elsi were successful, having no or only one disfluent pause at T1, T2 or T3. When speech and articulation rate were measured, in addition to Elsi and Ilona other successful students were Marjo, Ritva, Petra, Hanna and Kati.

With respect to development during SA, Kati showed the most improvement in fluency during SA (T1–T3) across all the fluency measures. Her perceptual fluency ratings improved by 1.6. Liisa, Elsi and Noora also improved in perceptual fluency ratings (by 0.9). The frequency of disfluent pauses showed that Kati and Aamu improved the most during SA (as their disfluent pause frequency decreased from 6 to 1 and 6 to 2). Mean pause duration decreased most for Kati (224 ms), Liisa (126 ms), Marjo (109 ms) and Sanna (102 ms). In addition to Kati, Ritva also gained considerably among this group in speech and articulation rates.

The most successful students in all the pitch measurements were Ilona, Liisa, Tiina[80] and Kati. Tiina, Kati and Ilona had the highest mean pitch in their L2 Russian (Study IV). At T1 (prior to the stay) Ilona, Liisa and Noora had the widest pitch range, whereas at T3 (following the stay) Ilona, Tiina and Elsi had the widest pitch range in L2 Russian. With respect to pitch range in the different utterance types, Elsi was the most successful according to visual inspection of the contours. The recognition rates and acceptability ratings of YNQs in L2 Russian (at T2 and T3), showed that Liisa and Ilona were the most successful. Their overall recognition rates for all questions were rather high: 83 % for Liisa and 79 % for Ilona. The mean acceptability ratings were similarly above average for these students: 3.6 for Liisa and 3.2 for Ilona.

Last, I will look at the improvement during SA (T1–T3), i.e., which student seemed to gain most in pitch production according to the measures used in this study. Ilona clearly increased her mean pitch in Russian most during SA. With respect to pitch range, Tiina widened her pitch range most (by 69 Hz) during SA, followed by Ilona (by 39 Hz). When looking at the recognition rates and

[80] Tiina participated only in Study IV.

acceptability ratings in YNQs between T2 and T3, Marjo and Ilona improved the most: in recognition rates 11 % and 10 %, respectively, and in the acceptability ratings 0.47 and 0.41.

Thus, different students seem to improve different prosodic features during their SA. A few students stand out has having gained a lot during SA (e.g., Kati and Ilona) and others that were generally successful in the measures chosen (e.g., Elsi, Ilona, Tiina, Liisa). In sum, SA seems to influence students' prosody in different ways, and further investigation is needed to learn why these particular students gained a lot and the others did not.

4.4 Strengths, Limitations and Possibilities for Future Research

Finally, I will evaluate the present study. In addition to pointing out its strengths and limitations I will also discuss the possibilities of extending the scope of this research in the future.

The limitations of the present study are similar to those in experimental phonetic research in general: can the results obtained in a laboratory setting be applied in "real life" and do they reflect "real" phenomena present in "real" speech outside the laboratory? Speech performance in the laboratory can be affected by multiple factors, such as tension or unfamiliarity with the recording situation. It is also possible that, in particular, L2 speakers tend to monitor their speech in such a situation, which might make their speech too controlled and thus disfluent. (Lehtonen 1981, 331; Levelt 1989, 460–463.) The advantages of this research setting are, however, that the speech data collected was similar in lexical and syntactic content for all speakers and that they were not asked to formulate what they were saying. As L2 speakers, they were also all used to reading L2 texts aloud.

A potential limitation of the study is the small number of speakers investigated. Traditionally in phonetic research rather small corpora have been used. Today as the analysis (and partly the segmentation also) can be automatized, investigation of larger corpora has become possible. However, as the annotation was mostly done manually in the present study it was considered reasonable to use a smaller corpus. The size of the data is also a consequence of the decision to use students from a single Finnish university as subjects so as to be able to some degree to control for what they have been taught prior to SA. Also, they all went to the same university in Russia, where the curriculum was the same. In Studies I, II and III a considerable number of speakers (12) were used, but only a small extract from the data was analysed because of the requirements of the evaluation task. In Study IV, the entire dialogues for nine speakers were analysed. In Studies V and VI only six speakers were analysed, because I wanted to investigate a group of students similar in their Russian learning background. In those studies only YNQs were selected as the data. Furthermore, the amount of data in Studies V and VI had to be limited because it too was subjected to perceptual evaluation.

In both the perceptual evaluation tasks, the samples were presented to the listeners in the same randomized order. This was done to prevent the impact of the order of the stimuli on the ratings. However, it is possible that there was some learning or accustoming effect, which might have caused the listeners to rate the last speakers differently from the first ones. It is also possible that they did not give too good/too poor ratings at the beginning of the task because they could not know a priori the range of the speakers' skills, and perhaps were more likely to "save their best/worst rates" for later. Hence, it might have been possible to improve the reliability of the evaluations by giving the listeners some very poor and very good samples to listen to prior to listening to the samples to be evaluated, as, e.g., Cucchiarini et al. (2002) have done. This would have given the listeners an idea of the general proficiency level and the range of the speech samples. Furthermore, as the listeners were not given any definition of fluency in the fluency evaluation task, they may have applied different criteria for what is 'fluent'. On the other hand, as argued by Freed et al. (2003) listeners tend to use similar enough criteria in their evaluation. As the aim was to test whether speakers attend to pausing and speech and articulation rates when evaluating L2 fluency, giving a definition of fluency might in fact have jeopardized the research setting.

The design of the perceptual evaluation task of interrogativity was possibly not the most suitable one for investigating how successfully the students produced questions in Russian as the listeners were told a priori the speaker's intention (that all students were trying to utter a question on each occasion). However, I feel that it was successful for the purposes of this study, although it may not have been a wise choice to call the ratings obtained through this task "recognition ratings". In future studies, both declarative and interrogative versions of the same sentences could be included.

In addition, the present study did not examine all parts of the pitch contour in detail. As the literature review showed, there is no agreement among researchers on what parts of the contour other than the peak are significant in distinguishing utterance types. The findings of Volskaya (2001) and Kasatkin (2007) have shown that in addition to the nucleus, the prenuclear part, for example, is also important for question perception in Russian[81]. Svetozarova (1982) showed that native Russian speakers were able to perceive statements and YNQs correctly solely on the basis of the prenuclear part. Also Kuosmanen & de Silva (2003; 2007) have underlined the importance of investigating other parts of the contour than the peak alone. Hence, in future analysis other parts of the contour could also be analysed.

Furthermore, it would be interesting to use different measuring techniques than the ones used in this study. For pause duration, an analysis of pause duration distribution might have been more revealing than simply calculating the mean duration. For example, Hird & Kirsner (2010) have shown that pause distributions can be very skewed and that there are usually two

---

81 This has also been shown for Swedish and American English by Hadding-Koch & Studdert-Kennedy (1964).

pause duration distributions per speaker: short and long pauses. For pitch range, measuring, e.g., the median, 80% or 90% range or Patterson's (2000) model could have potentially been used. Using them could shed more light on the cross-language comparison of pitch range. Also, different methodologies for investigating intonation remain to be explored. For example, in the future it would be interesting to utilise the Momel or INTSINT (Hirst 2007) models or ToBI (or Odé's ToRI) transcription systems. Werner's (2001) idea of using self-organised maps in intonation research also deserves further applications in comparing L1 and L2 intonation.

In addition, more meticulous comparison of pitch contours following, e.g., Hermes (1998) would yield more detailed results. However, more detailed peak measurement would result in important methodological choices to be made: how to segment phonemes in L2 if it is not clear what the target phoneme is; or how to decide which syllable is stressed in L2, if the vowel reduction is realised in quality but not in quantity (realisations of the word *совесть*[82] '*sovest*'' such as [sɑ:vjest] instead of [ˈsovʲɪstʲ]). Furthermore, if only voiced segments are taken into account, mispronunciations and creaky voice would result in different voiced segments for different speakers, which would make the comparison more difficult.

This study investigated Russian intonation in YNQs only. Other sentence types could be studied in the future, as they are likely to be difficult for Finns as well. For example, Leed (1965) states that also the Russian declarative prosody is difficult for English speakers to learn. In the comparison of intonation in L1 and L2 speech, there are other possibilities yet to be explored. For example, intensity and duration were not examined in the present study, even though together with pitch they are well known acoustic correlates of intonation. Also, it was not studied how an utterance was situated in the speakers' voice profile in general, which could have been interesting from the point of view of different sentence types. For this, Iivonen's (2001b) Temporal Voice Range Profile (TVRP), for example, could be used. Furthermore, using measurements like this could show whether speakers utter interrogatives on a higher pitch in general than, e.g., statements.

In previous studies, as well as pausing and speech and articulation rates, intonational features have also been mentioned as important qualities of fluent speech (see, e.g., Anderson 1990; Wennerstrom 2000; Lauranto 2004). Further research comparing the fluency ratings obtained in Study I to the performance of the intonational contours of the same students (see, e.g., Ullakonoja 2010) is currently underway. In addition, using Strangert & Gustafson's (2008) argument that F0 measurements correlate highly with listener evaluations of a "good speaker", the measurements of, e.g., F0 range, min, max in my data (Study IV) could be compared with the fluency ratings of Study I.

The SA context itself involves various social and cultural factors (1998b; see, e.g., Wilkinson 1998a; Freed et al. 2004) that were left beyond the scope of the present study, as the emphasis was on phonetic analysis. All the students

---

82 In English: 'conscience'.

went to the same town in Russia at the same stage of their university studies, and about half of them stayed with a host family whereas the rest lived in foreign student dormitories. Obviously, it is impossible to claim that the learning environment or the amount of the L2 input in Russia would have been exactly the same for all participants. However, this study did not aim at covering the actual amount of spoken input of students or their behaviour while abroad. In future, it might be possible to look at the written input the students are exposed to (e.g., teaching materials, reading for pleasure) and how it might influence fluency and prosody development. Also, students' activities during SA could be mapped, e.g., by asking them how many hours per day they spend listening to spoken Russian on radio or TV, as, e.g., Derwing et al. (2004; 2006) have done, and compare that to their oral productions. Controlling students' behaviour and the input they are getting (in questionnaires, diaries, motivational questionnaires etc.) would enable further qualitative analysis of the learner profiles of successful and non-successful students.

Also, it is quite likely that the students' performance developed not only because of their SA, but also because of other factors such as, e.g., improvement in L2 proficiency, increased self-confidence, increased motivation or improvement in familiarity with Cyrillic text. These aspects were not studied here. There is also a possibility that the students' performance was affected by learning to read the texts in question: after all, they got more practice in reading the texts each time they were recorded. The recordings done prior to the stay and during it were separated by an interval of about five months, whereas the interval between the last two recordings was only three months. Herman's (1985) study could be seen as partly supportive of this claim as she suggests that repeated readings help students to become more fluent because their reading rate increases.

The research questions and theories that concern L2 learning/acquisition directly have deliberately been excluded from this study, although the results can be interpreted in the light of L2 development. This aspect could serve as an interesting starting point for future studies, especially since relatively little research has been done on multilingual L2 learners, for example on how previously learnt L2s possibly affect the pronunciation of the L2 being learnt (see Llama et al. 2008 and Major 2008 for a review). It would be interesting to further trace the individual learning paths of the students. The first three studies showed considerable interspeaker differences that might have been due to the differences in the language learning backgrounds of the subjects. For the last two studies (Study V and Study VI) the speakers were selected from the corpus on the basis of their similar language learning background (three years before university studies). Yet, the participants were not a homogeneous group in their realisation of Russian prosody according to the perception experiment and acoustic analysis. However, from the L2 point of view it would be, in a different type of study, interesting to find out why some speakers develop more than others in fluency and prosody during SA. If it is possible in the future to enlarge the scope of this research to encompass L2 learning and thus to apply

L2 learning theories to the results presented here, it would also be necessary to study how L2 speakers perceive the various intonational contrasts in Russian (in addition to how they produce them). For example Cruz-Ferreira (1987) has justified the importance of studying perception by the traditional claim that if learners cannot perceive the intonational contrasts of their L2, they are bound to fail in producing them.

As Adams (1979), Hieke (1984) and Aho (2010, 58) argue, the phenomenon of linking can also be problematic for L2 learners. Without doubt this applies to L2 learners of Russian. In my view linking is closely related to pausing in that when learners pause at inappropriate places they also fail to link together words that form a structural entity, such as a phonetic word, a noun phrase or a verb phrase. For example, my data show that *часов в девять* (*chasov v devyat'*[83]) was frequently produced with a pause between the words *часов* and *в девять*, despite the fact that the words are structurally and semantically related. Hence, in read-aloud speech, the question is how L2 speakers structure the text they are reading. Do they see it as consisting of separate words (in which case they are also failing to link the words in a way a native speaker would do) or of phrases or a combination of words? This would be another potential direction for future research.

Other possible directions for further research would be to study other prosodic features such as rhythm, word stress (which is closely related to vowel articulation in Russian), and voicing/unvoicing of consonants, and compare them with the fluency ratings or the question recognition rate (see, e.g., Meister & Meister 2007 for an example of error-analysis of Russian L2 learners of Estonian). Rhythm is one of the most intriguing possibilities for future studies, as the study by de Silva & Volskaya (2005), for example, showed that Finns used rhythmical characteristics of Finnish in their L2 Russian.

As Wennerstrom (2001, 247) points out very few studies have focused on the acquisition L2 prosody. Many studies simply state that L1 influences L2 prosody, but very few studies have, like the present one, investigated samples from the same speakers at different stages of their language development. Thus, this study contributes to a little researched topic.

The strengths of the present study are, first, that the study combines acoustic analyses with two perceptual evaluation tasks. This dissertation consists of a total of six published papers, five of which have been published in international anonymously reviewed journals/conference proceedings. These six studies focus on different aspects of the same research problem while each addresses a separate set of research questions (section 1.2). The dissertation expands the theoretical background of those studies as well as collates their results.

Second, methodologically, the study employed a number of methods. In addition to several acoustic and statistical methods, the two perceptual evaluation tasks were designed to evaluate fluency and interrogativity. Also,

---

83 In English: nine o'clock.

the students' self-evaluation of their fluency in comparison with the teacher's ratings was discussed (Study III).

Third, the speech corpus (section 3.1) collected for the study is large and provides a number of possibilities for future research. It consisted of recordings of Finns in the different stages of learning Russian as L2 (prior to, in the middle of SA and following SA) as well as L1 Russian speakers reading the same dialogue texts. The corpus was analysed and annotated selectively for the purposes of the present research.

Fourth, the number of judges in the evaluation tasks was high compared to many other studies (30 for the fluency evaluation task and 19+21 for the evaluation of the interrogativity task). This high number of judges was shown to give statistically reliable results. Also a novelty of the study is that in one of the perceptual evaluation tasks (fluency evaluation), the judges were L2 teachers, most of whom were not L1 speakers of Russian.

Fifth, the study focused on multilingual students' L2 prosody in the SA setting, where the L2 was not the first L2 learnt by the students. Furthermore, the L1 and L2 of the students were unrelated and neither the L1 or L2 was English, as in most previous studies. Also the language learning background of the students was rather similar, and they were studying Russian at the same universities in Finland and Russia.

In sum, as was shown in the literature review, both fluency and pitch have not been widely investigated in L2 production, and thus, the present study contributes to an infrequently researched domain. There are a number of potential directions for future work, including the use of more data and the analysis of other prosodic features (such as rhythm).

4.5 Implications

As mentioned above, the terms *learning* and *acquisition* have deliberately been left outside the scope of the present study. The term used here is *development* in comparing different stages of SA. However, it is clear that the present research also has implications for the field of L2 learning, and, consequently, in this section I will speculate about learning as well.

The implications of the present study for L2 learning are first, that students should be encouraged to spend some time in the country where the L2 is spoken. Secondly, it seems obvious, that in teaching more attention ought to be paid to prosody (especially pause placement and pitch contours in YNQs) in order to improve students' fluency, comprehensibility and pragmatic competence. When students are reading a text aloud they often seem to be focusing only on segmental pronunciation, whereas they could simultaneously be developing their prosodic skills. I believe that paying attention to prosody would help students to learn to structure the text better, and so to understand better what they are reading.

The research results will help to develop the teaching of Russian phonetics in Finland, and with more research could enable to the creation of computer-based learning programmes, where the student can develop the L2 prosodic contours through repetition and practice. It has already been shown, e.g., by several researchers (Weltens & Bot 1984; Straszer 2003; Ramirez Verdugo 2006; Rocca 2008) that a visual illustration of the intonation contour helps the L2 learner to acquire L2 intonation (see, e.g., Iivonen 1987, 70–71, for a summary). Furthermore, there have already been successful experiments in using technology[84] to teach L2 intonation (see, e.g., Hermes 1998; Levis & Pickering 2004; Martin 2010) as well as to evaluate L2 segmental production through automatic speech recognition techniques (Moustroufas & Digalakis 2007).

As Weber (1991) points out, in an academic context L2 reading is an important way to learn L2. Taillefer (2005, 521) also suggests that academic reading skills in L2 are essential in coping in the SA academic context. However, according to my own experience, the text books used in the Finnish classroom (whether at school or university) pay hardly any attention to teaching how to read in L2s, and seem to be based on the assumption that it is similar to reading in L1. Reading texts aloud, in my experience, is done in the L2 classroom context but fairly rarely outside it. L2 textbooks should therefore focus more on teaching how to read in L2, especially in an L2 with a different alphabet than their L1. Furthermore, in line with the views presented by Anderson (1994, 185), in teaching more attention should be paid to increasing the reading rate in the L2, not at the expense of reading comprehension or segmental production, but, perhaps, focusing occasionally on the reading rate rather than, e.g., reading accuracy (on efficient ways of teaching to improve the reading rate see, e.g., Nuttall 1982, 38–41; Jensen 1986; Mahon 1986). Also, students might also become more fluent when listening to someone read aloud fluently, as, e.g., Rasinski (2003, 38–40) suggests for L1 learners.

The results of this study can, up to a point, be applied to L2 learners of any language. However, one should remember that L1 speakers of Finnish (or a Finno-Ugric language) who are learning Russian (or any Indo-European language), face a different task than learners whose L1 is typologically related to the language they are learning (see, e.g., Ringbom 1987, 80; Koda 2007). As Ringbom (1987, 112–113) argues, previously learnt L2s can help to acquire a new L2. For example, most of my subjects have studied 2–3 L2s before Russian, and as their L1 is very different from their L2s, they can perhaps make more use of their other L2s when learning Russian than of their L1.

Other possible implications of this study include the possibility to develop the evaluation of oral skills in general and also in the high school examinations (matriculation examinations). In Finland oral proficiency is still not a part of the L2 matriculation examinations, but students can participate in a voluntary oral

---

84 See Chun et al. (2008) for a summary of the different software available for teaching prosody through technology.

skills test[85]. Unfortunately, this is reflected in the minor role that oral skills have in L2 teaching in secondary education in Finland. Studies such as the present one could be used in developing semiautomatic ways to evaluate oral skills in examinations. When measuring oral skills, fluency is an important variable. Also, the fact that in this study the teachers' ratings showed high consistency, indicates that teachers can reliably be used as evaluators of fluency. That is, as teachers are the evaluators of other language skills in the matriculation examination, they could equally be considered competent to evaluate oral skills.

This study also has implications for Russian intonation research in that it reveals that the peak-delay phenomenon seems to occur frequently in YNQs. More research is needed to determine if it is a norm or an artefact, and when exactly it occurs. Methodologically, this study has proposed a way of acoustically measuring read-aloud fluency. It has also underlined the need of a more thorough analysis of L1 Finnish, L1 Russian as well as L2 Russian intonation to be able to better understand the difficulties Finnish students face when learning the interrogative intonation in Russian.

4.6 Conclusions

This study contributes to literature on L2 prosodic characteristics. It may serve as the start of a series of further studies on L2 Russian of Finnish students. The six studies presented here have provided knowledge on reading aloud fluency and prosody that I now summarise with the help of the research questions presented in section 1.2 above.

Question 1. How do Finnish L2 speakers of Russian change in their read-aloud fluency during the SA period? (Studies I and III)

Most students were more fluent following the stay than prior to it. Fluency development was not always linear across the students; that is the students did not show equal improvement on the same fluency rating scale. The fluency ratings were different in the middle of the stay and following it. Most students were judged more fluent following the stay than in the middle of it, which seems to indicate that SA was beneficial to most students.

Question 2. Do the temporal/acoustic variables studied (speech and articulation rates and pausing) correspond to the fluency ratings? (Studies I and II)

Speech rate, articulation rate, pause frequency and pause duration correlated with the fluency ratings (Table 17, p. 99). Hence, they can all be regarded as correlates of read-aloud fluency.

---

85 From autumn 2010 onwards, a new national oral skills test has been taking place in some L2s in high school, but it is still a voluntary and not a part of matriculation examinations.

Question 3. Is there a relationship between a speaker's self-assessment and language behaviour in Russia and her fluency rating? (Study III)

Such a relationship was found that the students who paid more attention to pronunciation and tried to make contact with native speakers were judged by the teachers as more fluent. Furthermore, the students' self-evaluations of their language skills were realistic: for example, those who said that their pronunciation had improved were also judged as more fluent by the teachers.

Question 4. Are speakers perceived as more fluent in their L2 if they use fewer and shorter pauses and pause at syntactically appropriate locations? (Study I)

According to the teachers' evaluations of the students' fluency, the samples with a smaller pause frequency, with shorter pauses and with pauses at syntactical boundaries were more fluent than those with many pauses, long pauses and with pauses that were situated elsewhere than at syntactical boundaries. The results were similar for relative and absolute pause durations.

Question 5. Are speakers evaluated as more fluent in their L2 if their speech and/or articulation rate is faster? (Study II)

The students whose speech and articulation rates (as measured both in phonetic words/second and syllables/second) were faster were also rated as more fluent by the teachers than the students with slower rates.

Question 6. Are speech and/or articulation rate speaker-dependent? (Study II)

Most speakers were ranked similarly among the group at all three stages of recordings, i.e. a speaker that was a slow speaker in her L1 was a slow speaker in L2 also and vice versa. Hence, in this study speech and articulation rates were speaker-dependent.

Question 7. Are mean pitch and pitch range different in L1 Finnish and L1 and L2 Russian for female speakers? (Study IV)

The analysis showed that mean pitch was lower in L1 Finnish than in L1 Russian and the pitch range was narrower in L1 Finnish. In L2 Russian the mean pitch was slightly higher and pitch range somewhat wider than in the L1 Finnish of the same speakers, but not as high or wide as in L1 Russian.

Question 8. If the answer to the question 7 is positive, do Finnish L2 learners of Russian develop a more native-like mean pitch and pitch range in Russian during their stay in Russia? (Study IV)

Some individual students developed a more native-like mean pitch and pitch range in Russian during SA. However, as a group the L2 speakers' mean pitch was not different in the recordings conducted prior to and following the stay. However, they had a slightly wider pitch range following their stay in

Russia than prior to it, which would indicate the development of a somewhat more native-like pitch range.

Question 9. How do native Russian speakers perceive and judge Finnish L2 learners' utterances intended as YNQs? (Study V)

Only 57 % of the YNQs produced by Finns were perceived as such by the Russian listeners. Also the mean acceptability rating (from 1 to 5) of the utterances as a question was rather low (2.95). Thus, Finnish learners' YNQs were evaluated as rather poorly acceptable in general, but there were wide interspeaker differences as well as differences between the items.

Question 10. Are L2 learners more successful in producing YNQs perceived as questions following their stay in Russia than during it? (Study V)

The learners were statistically significantly less successful in producing YNQs following the stay than during it according to the question recognition rate (the categorisation task). The difference was small. In contrast, the acceptability ratings yielded the opposite result. These contradictory results can be interpreted to indicate that overall the learners had "forgotten" very little of what they had learnt in Russia even after a month at home and not using the L2. Again, great interspeaker differences were observed.

Question 11. Do the acoustic pitch measurements of L1 and L2 Russian speakers differ? (Study VI)

L2 speakers of Russian had statistically significantly lower mean pitch and smaller mean absolute slope (indicating that the pitch contour had a shallower slope) than L1 speakers of Russian.

Question 12. What characterises the relationship between recognition rate and intonation in L2 questions? (Study VI)

A statistically significant association was found between recognition rate and pitch standard deviation. Visual comparison of the pitch contours, however, showed clearly that the L2 contours that were very dissimilar to L1 speaker contours were not recognised as questions.

To sum up the results of the six studies, the results were in line with expectations and mostly confirmed the results of the earlier studies. Nevertheless, more research on L2 prosodic production is needed, particularly research combining acoustic methods and the SA context.

In Finland, it is very common nowadays for L2 learners to take advantage of the possibilities of SA at some stage of their studies. In fact, from the year 2000 onwards, 20 % more Finnish students have enrolled in SA programmes each year, making a total of over 8,000 students per year (Korkala 2008, 6, 8). In 2007 only 249 of these students chose to study in Russia, whereas the most popular countries were in Central Europe. Because of its popularity, it is necessary for both L2 learners and their teachers to better understand the

processes involved in SA so that students can profit from the opportunity as much as possible. On the basis of the research presented here SA can be strongly recommended to L2 students.

Although the teaching phonetics is important, I believe that by staying abroad, a Finnish university student, who has already had training on the phonetic contrasts of Russian, can benefit from a Russian-speaking environment in that s/he will also learn "the natural way", as children do, by listening and repeating. In Russia, the learner is completely surrounded by the L2, immersed in it and – what is crucial – has a real need for L2 skills to "survive" in everyday life. From the point of view of prosody, the hearing of L2 – hearing enough of it – and thus, being able to better perceive its prosodic features, is strongly enhanced by the SA experience.

YHTEENVETO

Da. Eto vopros! Suomalaisten opiskelijoiden venäjän lukupuhunnan prosodinen kehittyminen vaihto-opiskelujakson aikana Venäjällä

Tausta ja tavoitteet

Da. Eto vopros! voisi olla relevantti ilmaus suomalaisen ja venäläisen välisessä viestintätilanteessa, jossa suomalainen mielestään tuottaa kysymyksen *Da?* (Niinkö?). Koska lause kuulostaa venäläisestä pikemminkin väitelauseelta siinä käytetyn intonaation takia, joutuu suomalainen vielä tarkentamaan *Eto vopros!* (Tämä on kysymys). Nämä suomalaisten kokemat haasteet kysymyslauseiden intonaation tuottamisessa ovat yhtenä lähtökohtana tälle tutkimukselle, joka käsittelee prosodian, tarkemmin sanottuna intonaation, ja sujuvuuden kehittymistä vaihto-opiskelujakson aikana Venäjällä.

Väitöskirjassa tarkastellaan vieraskielisen lukupuhunnan sujuvuutta ja intonaatiota sekä akustisesta että perseptuaalisesta näkökulmasta katsottuna. Tavoitteena on selvittää vieraan kielen sujuvuuteen vaikuttavia prosodisia piirteitä sekä intonaation tuottamista. Teoriatausta koostuu soveltavan kielitieteen ja fonetiikan alaan sijoittuvista tutkimuksista, joissa on tutkittu sujuvuutta, intonaatiota tai vieraan kielen prosodiaa. Tutkimuksessa ei oteta kantaa toisen tai vieraan kielen oppimisteorioihin sinänsä eikä sovelleta mitään yksittäistä oppimista koskevaa teoreettista viitekehystä.

Aiemmissa tutkimuksissa (esim. Freed 1995; 1998; Towell ym. 1996; Collentine & Freed 2004; Lafford 2004; Trofimovich & Baker 2006) on todettu, että oppijoiden suullinen kielitaito tai sujuvuus suullisessa ilmaisussa paranee vaihto-opiskelun aikana. On myös osoitettu, että vaihto-opiskelu on perinteistä kotimaan luokkahuonetta parempi vieraan kielen oppimisympäristö (Walsh 1994; Harley & Hart 2002; Segalowitz & Freed 2004). Sujuvuutta on aiemmissa tutkimuksissa mitattu yleensä joko kuulonvaraisella arvioinnilla tai akustisesti mittaamalla. Tässä väitöstutkimuksessa käytössä ovat nämä molemmat menetelmät: Suomessa asuvat venäjän kielen opettajat arvioivat kuuntelukokeessa näytteiden sujuvuutta, ja akustisen analyysin keinoin mitataan tauotusta sekä puhe- ja artikulaationopeutta. Tauotus ja puhe- ja artikulaationopeus on valittu sujuvuuden akustisiksi mittareiksi aiempien sujuvuustutkimusten perusteella (esim. Riggenbach 1991; Cucchiarini ym. 2002).

Vieraan kielen oppijoiden puheen tuottamista on tutkittu eniten äännetasolla, vaikka monet nykytutkijat ovatkin sitä mieltä, että yksittäisiä äänteitä tärkeämpää on oppia tuottamaan kohdekielen prosodiset piirteet oikein. Venäjän ja suomen intonaatiorakenteiden eroista on toistaiseksi vasta vähän tutkimustietoa. Aiemmat tutkimukset suomalaisista venäjä vieraana kielenä -oppijoista ovat osoittaneet, että venäjän vaihtoehtokysymysten (englanniksi: *yes/no questions*) tuottaminen on heille vaikeaa (esim. Kuosmanen & de Silva 2003; 2007; Ullakonoja ym. 2007). Niinpä tässäkin tutkimuksessa keskitytään

juuri vaihtoehtokysymyksiin. Suomessa vaihtoehtokysymys tuotetaan morfologisin keinoin, kun taas venäjässä yleisimmin kysyvän intonaation avulla. Tässä väitöstutkimuksessa intonaatiota tutkitaan sekä syntyperäisten kielenpuhujien kuulonvaraisella arvioinnilla että akustisesti mittaamalla perustaajuutta.

Tutkimus sijoittuu venäjän kielen tutkimuksen, soveltavan kielitieteen, kokeellisen fonetiikan ja vieraan tai toisen kielen oppimisen rajapinnoille. Tutkimuksen tavoitteena on selvittää, 1) kuinka puheen tauotus sekä puhe- ja artikulaationopeus toimivat sujuvuuden mittareina vieraskielisessä lukupuhunnassa, 2) kuinka sujuvuus mahdollisesti kehittyy vaihto-opiskelujakson aikana akustisten mittareiden ja opettajien sujuvuusarvioiden perusteella ja 3) miten oppijat käyttävät perustaajuutta (pitch) venäjässä yleensä sekä tarkemmin vaihtoehtokysymyksissä verrattuna syntyperäisiin kielenpuhujiin. Tarkemmat tutkimuskysymykset on muotoiltu sivulla 18 englanniksi. Vieraan kielen prosodian tuottamista vaihto-opiskelujakson aikana ei ole aiemmin tutkittu pitkittäistutkimuksena kaikkien näiden akustisten parametrien (tauotus, puhe- ja artikulaationopeus, perustaajuus) osalta.

Aineisto ja menetelmät

Tutkimuksen koehenkilöt ovat suomea äidinkielenään puhuvia naisopiskelijoita (iältään 19–24 vuotta), jotka opiskelevat venäjää vieraana kielenään[86] yhden suomalaisen yliopiston ainelaitoksella ja viettävät yhden lukukauden (3,5 kk) Venäjällä 2. opintovuoden aikana. Ensimmäisenä yliopistovuotena koehenkilöt olivat opiskelleet yhden fonetiikan kurssin, joka sisälsi ääntämisharjoituksia ja fonetiikan teoriaa. Venäjällä oleskelun aikana he opiskelivat venäjää vieraana kielenä pienryhmissä heille räätälöidyn ohjelman mukaan, johon ei kuulunut ääntämisen tai fonetiikan kursseja. Osa opiskelijoista asui vaihto-opiskelujakson aikana ulkomaalaisten opiskelijoiden asuntolassa ja osa venäläisissä isäntäperheissä. Väitöskirjan eri osatutkimuksissa on mukana 6–12 suomalaista opiskelijaa. Lisäksi joissain osatutkimuksissa on käytetty vertailukohtana 6–7 syntyperäistä venäläistä naista (iältään 19–28 vuotta).

Tutkimuksen aineistona on käytetty sekä nauhoitettua lukupuhuntaa että kahta havaintokoetta. Lukupuhuntakorpusta varten opiskelijat lukivat pareittain yhden suomenkielisen ja kaksi venäjänkielistä dialogia eri vaiheissa yliopisto-opintojaan: ennen 3,5 kk:n vaihto-opiskelujaksoa Venäjällä, sen puolivälissä ja sen jälkeen. Dialogit ovat samat kaikilla nauhoituskerroilla. Venäjänkieliset dialogit ovat puhelinkeskusteluja venäjä vieraana kielenä –oppimateriaalista (Shilova & Usmanova 1990). Suomenkielisen dialogin tutkimuksen tekijä on laatinut itse. Myös syntyperäisiltä venäläisiltä puhujilta nauhoitettiin pareittain samat venäjänkieliset dialogit yhteen kertaan. Suomessa nauhoitukset tehtiin studio-olosuhteissa suoraan tietokoneelle Adobe Audition 1.0–2.0 -ohjelmalla kahdella AKG GN30 -mikrofonilla. Venäjällä nauhoitukset tehtiin käytännön syistä kahdella eri tavalla: Sony TCD-D3 DAT-nauhurilla ja Roland

---

86 Vieraana kielenä tässä ymmärretään mikä tahansa äidinkielen jälkeen opittava uusi kieli.

Edirol -digitaalitallentimella. Molempien kanssa käytettiin yhtä Sony ECM-959A -stereomikrofonia.

Menetelminä tässä kokeellis-foneettisessa tutkimuksessa ovat kuulonvarainen arviointi, akustinen ja tilastollinen analyysi. Akustisessa analyysissä, aineiston annotoinnissa ja havaintokoeärsykkeiden laatimisessa käytetään Praat-ohjelmaa (Boersma & Weenik 2009). Tilastollista analyysiä käytettään mm. eri nauhoituskertojen vertailuun ja tulosten yleistettävyyden määrittämiseen ja se on tehty pääosin SPSS-ohjelmalla.

Olennaisena osana tutkimusta on myös kahden havaintokokeen tulosten vertaaminen akustisiin mittauksiin (tauoista, puhe- ja artikulaationopeudesta, perustaajuudesta ja sen vaihtelusta). Ensimmäisessä havaintokokeessa 30 suomalaista venäjän opettajaa arvioi yhden dialogipuheenvuoron jokaiselta opiskelijalta kultakin nauhoituskerralta. Opettajia pyydettiin arvioimaan näytteiden sujuvuutta asteikolla 1–5 (1 = ei sujuva, 5 = erittäin sujuva). Toinen havaintokoe käsittelee vaihtoehtokysymysten arviointia. Siinä 40 syntyperäistä venäläistä määritti opiskelijoiden tuottamien kysymyksistä 1) arvioivatko he ilmauksen kysymykseksi vai eivät ja 2) jos ilmaus olisi kysymys, kuinka hyvä kysymys se heidän mielestään olisi asteikolla 1–5. Tässä havaintokokeessa on mukana vain vaihto-opiskelujakson puolivälissä ja jakson jälkeen tehdyt nauhoitukset. Lisäksi aineistona käytetään opiskelijoiden tekemiä itsearviointeja.

Osatutkimusten tulokset

Väitöskirja koostuu kaikkiaan kuudesta osatutkimuksesta sekä niiden yhteenvedosta (159 sivua). Kaikki osatutkimukset on raportoitu englannin kielisissä artikkeleissa (liitteet 1–6), jotka on julkaistu 2007–2010. Osatutkimukset on koottu päätuloksineen taulukkoon 1. Ensimmäiset kolme osatutkimusta käsittelevät sujuvuutta ja olivat mukana tekijän julkaisemattomassa lisensiaatintutkimuksessa (Ullakonoja 2009). Seuraavat osatutkimukset pureutuvat perustaajuuteen ja ovat edellisten lisäksi osa tätä väitöskirjaa.

Ensimmäinen artikkeli *Pausing as an indicator of fluency in the Russian of Finnish learners* käsittelee opiskelijoiden sujuvuutta 30 suomalaisen venäjänopettajan arvioimana sekä näiden sujuvuusarvioiden mahdollista yhteyttä tauotukseen (taukojen määrä, kesto ja sijainti). Toisen artikkelin *Speech rate as an indicator of fluency in the Russian of Finnish learners* aiheena on puhe- ja artikulaationopeus ja sen yhteys ensimmäisessä artikkelissa tehtyyn sujuvuusarviointiin. Kahdessa ensimmäisessä artikkelissa analysoidaan myös sujuvuuden kehittymistä kieliharjoittelun aikana. Kolmannessa artikkelissa *Perception of L2 fluency in study abroad context* perehdytään sujuvuuteen peilaten opiskelijoiden itsearviointeja opettajien sujuvuusarvioihin.

Ensimmäisissä kolmessa artikkelissa raportoidut tulokset osoittavat mm. että sekä opettajien arvioiden että tuotosten prosodisten piirteiden (tauotus ja

TAULUKKO 1 Väitöstutkimuksen artikkelien viitetiedot ja päätulokset.

Osatutkimus	Artikkelin bibliografiset tiedot	Päätulokset
Study I	R.U. 2008. Pausing as an indicator of fluency in the Russian of Finnish learners. Teoksessa Barbosa, Plinio A. & Madureira, Sandra & Reis, César (toim.) *Proceedings of the Speech Prosody 2008 Conference. Campinas, Brazil.* São Paolo: Editora RG/CNPq. 339–342.	- Suurimmalla osalla opiskelijoista sujuvuus kasvoi ja taukojen määrä väheni vaihto-opiskeluskelujakson aikana.
Study II	R.U. 2009. Speech rate as an indicator of fluency in the Russian of Finnish learners. Teoksessa O'Dell, Michael & Nieminen, Tommi (toim.) *Fonetiikan päivät 2008 – The Phonetics Symposium 2008.* Tampere Studies in Language, Translation and Culture, Series B. 97–109.	- Suurimmalla osalla opiskelijoista puhenopeus kasvoi vaihto-opiskelujakson aikana.
Study III	R.U. & H. Dufva. 2008. Perception of L2 fluency in study abroad context. *Academic Exchange Quarterly,* Fall 2008 (12) 3. 62–66.	- Opiskelijoiden itsearviointi vastasi tietyssä määrin opettajien sujuvuusarviointeja.
Study IV	R.U. 2007. Comparison of Pitch Range in Finnish (L1) and Russian (L2). Teoksessa Trouvain, Jürgen & Barry, William J. (toim.) *Proceedings of the 16th International Congress of Phonetic Sciences, 6–10 August 2007, Saarbrücken, Germany*. Saarbrücken: Universität des Saarlandes. 1701–1704.	- Suurinosa opiskelijoista luki venäjää korkeammalla äänenkorkeudella ja suuremmalla vaihteluvälillä eli enemmän syntyperäisen kielenpuhujan kaltaisesti vaihto-opiskelujakson jälkeen.
Study V	R.U. 2010. How do Native Speakers of Russian Evaluate Yes/no questions Produced by Finnish L2 Learners? *Rice Working Papers in Linguistics*, Vol. 2. 92–105.	- Syntyperäiset kuulijat luokittelivat kysymyksiksi vain 57 % opiskelijoiden ääneen lukemista vaihtoehtokysymyksistä.
Study VI	R.U. 2010. Pitch Contours in Russian Yes/no Questions by Finns. Teoksessa Hasegawa-Johnson, Mark, Bradlow, Ann, Cole Jennifer, Livescu, Karen, Pierrehumbert, Janet & Shih, Chilin (toim.) *Proceedings of the Speech Prosody 2010 conference, Chicago*.	- Havaittiin paljon puhujien välistä vaihtelua, mutta yleisesti ottaen opiskelijoille oli haasteellista tuottaa tarpeeksi terävä perustaajuushuippu ja oikeassa paikassa.

puhenopeus) mukaan suurimmalla osalla opiskelijoista luetun venäjän sujuvuus kehittyy kieliharjoittelun aikana. Lisäksi opiskelijoiden itsearviointi vastaa jonkin verran opettajien antamia sujuvuusarvioita. Toisaalta sujuvuuteen liittyvät osatutkimukset kertovat myös tauotuksen ja puhe- ja artikulaationopeuden haasteellisuudesta. Epäsujuvat opiskelijat tauottavat puhettaan niin, että se todennäköisesti häiritsee kuulijoita. Lisäksi heidän puhe- ja artikulaationopeutensa on hidasta.

Kolme jälkimmäistä artikkelia käsittelee perustaajuutta (F0, pitch). Niistä ensimmäisessä artikkelissa *Comparison of pitch range in Finnish (L1) and Russian (L2)* analysoi opiskelijoiden keskimääräistä äänenkorkeutta sekä sen vaihteluväliä heidän äidinkielessään suomessa ja vieraassa kielessä venäjässä. Opiskelijoiden puheesta mitattuja arvoja verrataan myös syntyperäisten kielenpuhujien puheeseen. Viidennessä artikkelissa *How do native speakers of Russian evaluate Yes/no questions produced by Finns?* syvennytään toisen havaintokokeen tuloksiin. Siitä käy ilmi, miten syntyperäiset venäläiset kuulijat arvioivat suomalaisopiskelijoiden tuottamien vaihtoehtokysymysten hyväksyttävyyttä ja hyvyyttä kysymyksinä. Kuudennessa eli viimeisessä artikkelissa *Pitch patterns in Russian Yes/no questions by Finns* analysoidaan opiskelijoiden vaihtoehtokysymysten perustaajuuskontuureja ja verrataan niitä sekä syntyperäisten venäläisten tuotoksiin että viidennessä artikkelissa raportoituihin kysymysarvioihin.

Kahdessa viimeisessä artikkelissa verrataan siis vain kahden viimeisen nauhoituskerran aineistoja, koska tarkoituksena on selvittää kuinka pysyvää opiskelijoiden intonaation tuottaminen on. Vaihto-opiskelujakson jälkeinen nauhoitus tehtiin käytännön syistä vasta, kun opiskelijat olivat jo olleet Suomessa noin kuukauden (joululomalla), jona aikana he eivät olleet juurikaan käyttäneet venäjää. Niinpä voitiin olettaa, että heidän tuotoksensa perustui siihen mitä he olivat pysyvästi oppineet venäjän ääneenlukemisesta.

Kolme viimeistä osatutkimusta osoittavat, että opiskelijoilla on vaihto-opiskelujakson jälkeenkin vaikeuksia tuottaa vaihtoehtokysymyksiä niin, että venäläiset tulkitsisivat ne kysymyksiksi. Venäläiset kuulijat arvioivat kysymyksiksi 57 % opiskelijoiden tuottamista vaihtoehtokysymyksistä. Akustisesti analysoiduissa vaihtoehtokysymyksissä haasteellista näyttäisi olevan erityisesti perustaajuushuipun korkeus ja paikka. Toisaalta monen opiskelijan tuotokset myös paranivat merkitsevästi vaihto-opiskelujakson loppupuolella. Opiskelijoiden keskimääräinen äänenkorkeus oli venäjässä selvästi korkeampi kuin suomessa ja perustaajuuden vaihteluväli venäjässä suurempi kuin suomessa, mikä viittaa syntyperäisen puhujan äänenkorkeuden ja vaihteluvälin tavoitteluun.

Tutkimustulokset osoittavat siis, että vaihto-opiskelujakson aikana koehenkilöiden lukupuhunta sujuvoitui ja vaihtoehtokysymysten intonaatio kehittyi tunnistettavampaan suuntaan. Lisäksi suurimmalla osalla keskimääräinen äänenkorkeus ja perustaajuuden vaihteluväli lähenivät syntyperäisten venäläisten vastaavia mittausarvoja.

Lopuksi

Tutkimuksessa selvitettiin siis sujuvuutta ja intonaatiota yhden opiskelijaryhmän venäjänkielisessä lukupuhunnassa. Jatkossa tutkimusta voisi laajentaa moneenkin suuntaan. Esimerkiksi intonaatiota tulisi tutkia muissakin lausetyypeissä kuin vaihtoehtokysymyksissä ja tässä olisi syytä käyttää laajempaa korpusta sekä useampia analyysimenetelmiä. Tähän mennessä saatuja tutkimustuloksia voidaan kuitenkin jo soveltaa suoraan venäjä vieraana kielenä opetuksessa: tauotus liittyy kiinteästi puheen jaksotteluun ja sitä kautta rytmiin ja sujuvuuteen, intonaatio puolestaan liittyy puheen ymmärrettävyyteen ja viestinnän onnistumiseen. Oman haasteensa venäjän prosodian oppimisen jatkotutkimuksille asettaa se, että suomen kielen prosodian tutkimus on vielä hyvin vähäistä eikä olemassa olevia tutkimuksia ole aina mahdollista luotettavasti verrata keskenään. Venäjän prosodiaa on sen sijaan tutkittu enemmän, mutta aineistot ovat joskus melko suppeita ja erityisesti kokeellinen intonaation tutkimus on vähäistä.

Tutkijat ovat erimielisiä siitä, millaista koeasetelmaa esim. havainto-kokeissa olisi kulloinkin tarkoituksenmukaisinta käyttää (ks. esim. Toivola 2010). Tässä tutkimuksessa haluttiin elisitoida mahdollisimman luonnollista puhetta kuitenkin niin, että kaikki koehenkilöt tuottivat samat lauseet. Aineiston nauhoituksen pohjana käytettiin kirjoitettuja dialogeja, jotka otettiin suoraan venäläisestä oppimateriaalista, jossa esimerkiksi leksikaalisen aineksen arvioitiin olevan opiskelijoiden tasolle sopivaa.

Tämän väitöskirjatutkimuksen perusteella voidaan aiempien tutkimusten tavoin sanoa, että kysyvän intonaation tuottaminen vaihtoehtokysymyksissä on vaikeaa suomalaisille venäjän oppijoille. Tutkimuksen pohjalta voidaan suositella, että ääntämisen opetuksessa kiinnitettäisiin välillä huomiota äännetason piirteiden asemasta prosodiaan ja harjoiteltaisiin, paitsi intonaatiota, myös puheen oikeaa tauotusta ja jaksottelua.

Tutkimuksen perusteella voidaan selvästi osoittaa, että vaihto-opiskelujaksolla on suurimmalle osalle opiskelijoista merkitystä ääneen lukemisen sujuvuuden kehittymisessä. Lisäksi jotkut opiskelijat myös tuottavat vaihtoehtokysymyksiä paremmin vaihto-opiskelujakson jälkeen kun sen puolivälissä. Tämä tutkimus on siis osoittanut, että vaihto-opiskelulla on merkitystä ääntämiselle, ainakin sujuvuuden ja prosodian kannalta. Todennäköisesti vaihto-opiskelu kehittää opiskelijan kielitaitoa ja kulttuurikompetenssia paljon laajemminkin, mutta tutkimustietoa toisen tai vieraankielen oppimisesta vaihto-opiskelujakson aikana on toistaiseksi niukasti. Tietoa olisi tärkeä paitsi tuottaa lisää, myös jakaa vaihto-opiskelijoille, jotta he osaisivat hyödyntää ulkomaan kokemuksensa mahdollisimman tehokkaasti.

BIBLIOGRAPHY

Abercrombie, D. 1967. Elements of general phonetics. Edinburgh: Edinburgh University Press.

Abramova, I. E. 1999. Faktory, vliyayuschie na formirovanie angliyskogo proiznosheniya u russkih shkol'nikov (eksperimental'no-foneticheskoe issledovanie na materiale angliyskogo konsonantizma). Avtoreferat kandidatskoy dissertacii. Sankt-Peterburgskiy gosudarstvennyy universitet.

Adams, C. 1979. English speech rhythm and the foreign learner. The Hague: Mouton.

Aho, E. 2010. Spontaanin puheen prosodinen jaksottelu. Nykykielten laitos/Kielitiede: Helsingin yliopisto. Doctoral dissertation. Available at: https://oa.doria.fi/bitstream/handle/10024/63264/spontaan.pdf?sequence=1 Last retrieved: 9th October 2010.

Aho, E. & Toivola, M. 2008. Venäläisten maahanmuuttajien suomen prosodiasta. Virittäjä 1/2008, 3–23.

Aho, E. & Yli-Luukko, E. 2005. Intonaatiojaksoista. Virittäjä 2/2005, 201–220.

Akamatsu, N. 2005. Effects of second language reading proficiency and first language orthography on second language word recognition. In V. Cook (ed.) Second language writing systems. Multilingual matters, 238–259.

Altenberg, E. P. 2005. The perception of word boundaries in a second language. Second Language Research 21(4), 325–358.

Altenberg, E. P. & Ferrand, C. T. 2006. Fundamental frequency in monolingual English, bilingual English/Russian, and bilingual English/Cantonese young adult women. Journal of Voice 20(1), 89–96.

Anderson, L. 1990. Intonation, turn-taking and dysfluency: Non-natives conversing. In M. Hewings (ed.) Papers in discourse intonation. Birmingham: University of Birmingham, 102–113.

Anderson, N. J. 1994. Developing active readers: A pedagogical framework for the second language reading class. System 22(2), 177–194.

Anttila, H. 2008. The effect of interrogative function on intonation in spontaneous and read Finnish. Department of Speech Sciences, University of Helsinki. Master's thesis. Available at: http://urn.fi/URN:NBN:fi-fe200807301740. Last retrieved: 7th August 2009.

Anttila, H. 2009. Interrogative intonation in spontaneous Finnish. 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, 167–176.

Aoyama, K. & Guion, S. G. 2007. Prosody in second language acquisition: Acoustic analyses of duration and F0 range. In O. Bohn & M. J. Munro (eds.) Language experience in second language speech learning: In honor of James Emil Flege. Amsterdam: Benjamins, 281–297.

Aoyama, K., Guion, S. G., Flege, J. E., Yamada, T. & Akahane-Yamada, R. 2008. The first years in an L2-speaking environment: A comparison of Japanese children and adults learning American English. IRAL 46(1), 61–90.

Asu, E. L. 2004. The phonetics and phonology of Estonian intonation. Department of Linguistics, University of Cambridge. Doctoral Dissertation.

Avanesov, R. I. 1956. Fonetika sovremennogo russkogo literaturnogo yazyka. Moskva: Izdatel'stvo Moskovskogo Universiteta.

Baker, W. & Trofimovich, P. 2001. Does perception precede production? Evidence from Korean-English bilinguals. LACUS Forum 27, 273–284.

Ballmer, T. T. 1980. The role of pauses and suprasegmentals in a grammar. In H. W. Dechert & M. Raupach (eds.) Temporal variables in speech. The Hague: Mouton, 211–220.

Baron, J. & Strawson, C. 1976. Use of orthographic and word-specific knowledge in reading words aloud. Journal of Experimental Psychology: Human Perception and Performance 2(3), 386–393.

Bent, T., Bradlow, A. R. & Smith, B. L. 2007. Segmental errors in different word positions and their effects on intelligibility of non-native speech. All's well that begins well. In O. Bohn (ed.) Language experience in second language speech learning: In honor of James Emil Flege. Amsterdam: John Benjamins, 331–347.

Bezooijen, R. V. 1995. Sociocultural aspects of pitch differences between Japanese and Dutch women. Language and Speech 38, 253–265.

Blum, I. H. & Koskinen, P. S. 1991. Repeated reading: A strategy for enhancing fluency and fostering expertise. Theory into Practice 30(3). 195–200.

Boersma, P. 1993. Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. IFA Proceedings. Institute of Phonetic Sciences, University of Amsterdam, 97–110. Available at: http://www.fon.hum.uva.nl/paul/papers/Proceedings_1993.pdf. Last retrieved: 16th June 2008.

Boersma, P. 2005. Calculating slope. Available at: http://uk.groups.yahoo.com/group/praat–users/message/1717 Last retrieved: 12th November 2009.

Boersma, P. 2007. Intro. Available at http://www.fon.hum.uva.nl/praat/manual/Intro.html. Last retrieved: 30th March 2009.

Boersma, P. & Weenik, D. 2009. Praat: Doing phonetics by computer. Available at: http://www.fon.hum.uva.nl/praat/ Last retrieved: 30th March 2009.

Bogomazov, G. M. 2001. Sovremennyy russkiy yazyk: Fonetika. Moskva: Vlados.

Bohn, O. & Flege, J. E. 1990. Perception and production of a new vowel category by adult second language learners. In J. Leather & A. James (eds.) New sounds 90, 38–56.

Bolinger, D. 1978. Intonation across languages. In J. H. Greenberg, C. A. Ferguson & E. A. Moravcsik (eds.) Universals of human language. Volume 2: Phonology. Stanford: Stanford University Press, 471–524.

Bolinger, D. 1986. Intonation and its parts: Melody in spoken English. London: Arnold.
Bondarko, L. V. 1998. Fonetika sovremennogo russkogo yazyka. Sankt-Peterburg: 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.
Botinis, A., Granström, B. & Mobius, B. 2001. Developments and paradigms in intonation research. Speech Communication 33(4), 263–296.
Brazil, D. 1997. The communicative value of intonation in English. Cambridge: Cambridge University Press.
Brecht, R. D., Davidson, D. E. & Ginsberg, R. B. 1995. Predictors of foreign language gain during study abroad. In B. F. Freed (ed.) Second language acquisition in a study abroad context. Amsterdam: John Benjamins, 37–66.
Brecht, R. D. & Robinson, J. L. 1995. On the value of formal instruction in study abroad: Student reactions in context. In B. F. Freed (ed.) Second language acquisition in a study abroad context. Amsterdam: John Benjamins, 317–334.
Brosnahan, L. F. & Malmberg, B. 1970. Introduction to phonetics. Cambridge: Cambridge University Press.
Bryman, A. & Cramer, D. 2001. Quantitative data analysis with SPSS release 10 for Windows: A guide for social scientists. Great Britain: Routledge.
Bryzgunova, E. A. 1972. Zvuki i intonaciya russkoy rechi. Moskva.
Bryzgunova, E. A. 1975. The declarative-interrogative opposition in Russian. Slavic and East European Journal 19, 155–161.
Bryzgunova, E. A. 1977. Zvuki i intonaciya russkoy rechi. Moskva: Russkiy yazyk.
Bryzgunova, E. A. 1982. Intonatsiya. In N. Y. Shvedova (ed.) Russkaya grammatika I. Moskva: Nauka, 96–122.
Burton, M. W. & Robblee, K. E. 1997. A phonetic analysis of voicing assimilation in Russian. Journal of Phonetics 25(2), 97–114.
Butcher, A. 1980. Pause and syntactic structure. In H. W. Dechert & M. Raupach (eds.) Temporal variables in speech. The Hague: Mouton, 85–90.
Carlson, R., Elenius, K. & Swerts, M. 2004. Perceptual judgments of pitch range. In B. Bel & I. Marlin (eds.) Proceedings of Speech Prosody 2004, Nara, Japan, 689–692.
Cenoz, J. 2000. Pauses and hesitation phenomena in second language production. ITL: Review of Applied Linguistics 127–128, 53–69.
Chen, A. 2009. Perception of paralinguistic intonational meaning in a second language. Language Learning 59(2), 367–409.
Chen, A. & Mennen, I. 2008. Encoding interrogativity intonationally in a second language. In P. A. Barbosa, S. Madureira & C. Reis (eds.) Proceedings of the Speech Prosody 2008, Campinas, Brazil: Editora RG/CNPq, 513–516.

Chen, S. & Fon, J. 2007. The effects of phonetic distance, learning context and learner proficiency on L2 perception of English liquids. In J. Trouvain & W. J. Barry (eds.) Proceedings of the 16th ICPhS, 6–10 August 2007, Saarbrücken. Saarbrücken: Universität des Saarlandes, 1721–1724.

Chen, S. & Fon, J. 2008. The peak alignment of prenuclear and nuclear accents among advanced L2 English learners. In P. A. Barbosa, S. Madureira & C. Reis (eds.) Proceedings of the Speech Prosody 2008, Campinas, Brazil: Editora RG/CNPq, 643–646.

Cheremisina-Enikolopova, N. V. 1999. Zakony i pravila russkoy intonacii. Moskva: Flinta, Nauka.

Chernigovskaya, T. V., Svetozarova, N. D., Tokareva, T. I., Tret'yakov, D. A., Ozerskii, P. V. & Strel'nikov, K. N. 2000. Specialization of cerebral hemispheres in the perception of Russian intonations. Human Physiology 26(2), 142–147.

Chun, D. M., Hardison, D. M. & Pennington, M. C. 2008. Technologies for prosody in context – past and future of L2 research and practice. In J. G. Hansen Edwards & M. L. Zampini (eds.) Phonology and second language acquisition. Amsterdam: John Benjamins, 63–94.

Cohen, A. D., Collier, R. & 't Hart, J. 1982. Declination: Construct or intrinsic feature of speech pitch? Phonetica 39(4), 254–273.

Coleman, J. A. 1998. Language learning and study abroad: The European perspective. Frontiers: The Interdisciplinary Journal of Study Abroad 4, 167–203.

Collentine, J. & Freed, B. F. 2004. Learning context and its effects on second language acquisition. Studies in Second Language Acquisition 26, 153–356.

Cooper, W. E. 1981. Fundamental frequency in sentence production. New York: Springer.

Council for Cultural Co-operation. Education Committee, Modern Languages Division, Strasbourg & Council of Europe. 2001. Common European framework of reference for languages: learning, teaching, assessment. Cambridge: Cambridge University Press.

Crown, C. L. & Feldstein, S. 1985. Psychological correlates of silence and sound in conversational interaction. In D. Tannen & M. Saville-Troike (eds.) Perspectives on silence. Norwood, NJ: Ablex, 31–54.

Cruttenden, A. 1986. Intonation. Cambridge: Cambridge University Press.

Cruz-Ferreira, M. M. F. 1989. Non-native comprehension of intonation patterns in Portuguese and in English. United States. Doctoral Dissertation.

Cruz-Ferreira, M. 1987. Non-native interpretive strategies for intonational meaning: An experimental study. In A. James & J. Leather (eds.) Sound patterns in second language acquisition. Dordrecht: Foris, 103–120.

Crystal, D. 1969. Prosodic systems and intonation in English. London.

Crystal, D. 1971. Relative and absolute in intonation analysis. Journal of the International Phonetic Association 1(1), 17–28.

Crystal, D. 1975. The English tone of voice: essays in intonation, prosody and paralanguage. London: Arnold.

Crystal, T. H. & House, A. S. 1990. Articulation rate and the duration of syllables and stress groups in connected speech. Journal of the Acoustical Society of America 88(1), 101–112.

Cucchiarini, C., Strik, H. & Boves, L. 2000. 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.

Cucchiarini, C., Strik, H. & Boves, L. 2002. Quantitative assessment of second language learners' fluency: Comparisons between read and spontaneous speech. Journal of the Acoustical Society of America 111(6), 2862–2873.

Cutler, A. 1983. Speakers' conceptions of the function of prosody. In A. Cutler & D. R. Ladd (eds.) Prosody: Models and measurements. Berlin: Springer, 79–91.

Daneman, M. 1991. Individual differences in reading skill. In R. Barr, M. L. Kamil, P. B. Mosenthal & D. Pearson (eds.) Handbook of reading research, volume II. New York: Longman, 512–538.

Dates, B. G. 2006. Cohen's kappa. Available at:
http://www.spsstools.net/Syntax/Matrix/CohensKappa.txt.
Last retrieved: 1st July 2009.

Davidson, D. E. 2007. Study abroad and outcomes measurements: The case of Russian. Modern Language Journal 91(2), 276–280.

Davidson, D. E. 2010. Study abroad: When, how long, and with what results? New data from the Russian front. Foreign Language Annals 43(1), 6–26.

de Bot, K. & Mailfert, K. 1982. The teaching of intonation: Fundamental research and classroom applications. TESOL Quarterly 16(1), 71–77.

de Silva, V. 1999. Quantity and quality as universal and specific features of sound systems: Experimental phonetic research on interaction of Russian and Finnish sound systems. Studia philologica Jyväskyläensia, Jyväskylä: University of Jyväskylä. Doctoral Dissertation.

de Silva, V., Iivonen, A., Bondarko, L. V. & Pols, L. C. W. 2003. Common and language dependent phonetic differences between read and spontaneous speech in Russian, Finnish and Dutch. In M. J. Solé, D. Recasens & J. Romero (eds.) Proceedings of the 15th ICPhS, Barcelona, 3–9 August 2003, 2977–2980.

de Silva, V. & Shcherbakova, L. P. 1998. Aktsentno-ritmiceskie struktury finskih i russkih slov (v usloviyah obuceniya nerodnomu yazyku), Vestnik SPbGU Ser, 2 23(4), 52–61.

de Silva, V. & Ullakonoja, R. 2009. Introduction: Russian and Finnish in contact. 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, 15–20.

de Silva, V. & Volskaya, N. B. 2005. Mesto fonetiki v standartah RKI i v praktike vladenya yazykom. In N. P. Isayev, A. Mustajoki & E. Protasova (eds.) Russkiy yazyk kak inostrannyy: Problemy izucheniya, prepodavaniya,

ocenki v svete obschsheevropeyskih kompetenciy vladeniya inostrannym yazykom. Moskva: MAKS Press, 61–73.

de Silva, V., Mäkilä, K., Kärkkäinen, H. & Ullakonoja, R. 2010. Uusi venäjän ääntämisopas. Helsinki: Finn Lectura.

Deese, J. 1980. Pauses, prosody, and the demands of production in language. In H. W. Dechert & M. Raupach (eds.) Temporal variables in speech. The Hague: Mouton, 69–84.

Derwing, T. M., Rossiter, M. J., Munro, M. J. & Thomson, R. I. 2004. Second language fluency: Judgments on different tasks. Language Learning 54(4), 655–679.

Derwing, T. M., Thomson, R. I. & Munro, M. J. 2006. English pronunciation and fluency development in Mandarin and Slavic speakers. System 34(2), 183–193.

Drommel, R. H. 1980. Towards a subcategorization of speech pauses. In H. W. Dechert & M. Raupach (eds.) Temporal variables in speech. The Hague: Mouton, 227–238.

DSP Center 2007 = Centr cifrovoy obrabotki signalov Sankt–Peterburgskiy gosudarstvennyy universitet telekommunikaciy im. prof. M.A. Bonch-Bruevicha. EDSW version 1.043 Available at: http://www.dsp.sut.ru/rus/products/edsw/index.html. Last retrieved: 16th June 2008.

Fant, G., Kruckenberg, A. & Barbosa-Ferreira, J. 2003. Individual variations in pausing, a study of read speech. In M. Heldner (ed.) Proceedings of the Swedish phonetics meeting in Umeå, PHONUM 2003. Umeå: Umeå University, Department of Philosophy and Linguistics, 193–196. Available at: http://www.ling.umu.se/fonetik2003. Last retrieved: 30th March 2009.

Favreau, M. & Segalowitz, N. S. 1982. Second language reading in fluent bilinguals. Applied Psycholinguistics 3(4), 329–341.

Flege, J. E. 1993. Production and perception of a novel, second-language phonetic contrast. Journal of the Acoustical Society of America 93(3), 1589–1608.

Flege, J. E., Birdsong, D., Bialystok, E., Mack, M., Sung, H. & Tsukada, K. 2006. Degree of foreign accent in English sentences produced by Korean children and adults. Journal of Phonetics 34(2), 153–175.

Flege, J. E. & Hillenbrand, J. 1986. Differential use of temporal cues to the /s/-/z/ contrast by native and non-native speakers of English. Journal of the Acoustical Society of America 79(2), 508–517.

Flege, J. E., MacKay, I. R. A. & Meador, D. 1999. Native Italian speakers' perception and production of English vowels. Journal of the Acoustical Society of America 106(5), 2973–2987.

Flege, J. E., Munro, M. J. & MacKay, I. R. A. 1995. Factors affecting strength of perceived foreign accent in a second language. Journal of the Acoustical Society of America 97(5), 3125–3134.

Fougeron, I. 1999. Intonacia i edinstvo teksta. In R. F. Kasatkina (ed.) Problemy fonetiki III. Moskva: Nauka, 216–227.

Fougeron-Benenson, I. 1971. De l'intonation dans les phrases interrogatives russes. Paris: Université de Paris III. Doctoral Dissertation.

Freed, B. F. 1995. What makes us think that students who study abroad become fluent? In B. F. Freed (ed.) Second language acquisition in a study abroad context. Amsterdam: John Benjamins, 123–148.

Freed, B. F. 1998. An overview of issues and research in language learning in a study abroad setting. Frontiers: The Interdisciplinary Journal of Study Abroad 4. 31–60.

Freed, B. F. 2000. Is fluency, like beauty, in the eyes (and ears) of the beholder? In H. Riggenbach & R. Schmidt (eds.) Perspectives on fluency. Ann Arbor: University of Michigan Press, 243–265.

Freed, B. F., Dewey, D. P., Segalowitz, N. & Halter, R. 2004. The language contact profile. Studies in Second Language Acquisition 26(2), 349–356.

Freed, B. F. & Ferguson, C. A. 1995. Second language acquisition in a study abroad context. Amsterdam: John Benjamins.

Freed, B. F., Segalowitz, N. & Dewey, D. P. 2004. Context of learning and second language fluency in French: Comparing regular classroom, study abroad, and intensive domestic immersion programmes. Studies in Second Language Acquisition 26, 275–301.

Freed, B. F., So, S. & Lazar, N. A. 2003. Language learning abroad: How do gains in written fluency compare with gains in oral fluency in French as a second language? ADFL Bulletin 34(3), 34–40.

Frieda, E. M. & Nozawa, T. 2007. You are what you eat phonetically. The effect of linguistic experience on the perception of foreign vowels. In O. Bohn & M. J. Munro (eds.) Language experience in second language speech learning: In honor of James Emil Flege. Amsterdam: John Benjamins, 79–96.

Glushko, R. J. 1981. Principles for pronouncing print: The psychology of phonography. In A. M. Lesgold & C. A. Perfetti (eds.) Interactive processes in reading. Hillsdale, New Jersey: Lawrence Erlbaum, 61–84.

Goldman Eisler, F. 1968. Psycholinguistics: experiments in spontaneous speech. London: Academic Press.

Gorsuch, G. J. & Taguchi, E. 2008. Repeated reading for developing reading fluency and reading comprehension: The case of EFL learners in Vietnam. System 36(2), 253–278.

Grabe, W. 1991. Current developments in second language reading research. TESOL Quarterly 25(3), 375–406.

Grabe, W. 2009. Reading in a second language. Moving from theory to practice. New York: Cambridge University Press.

Graesser, A. C., Hoffman, N. L. & Clark, L. F. 1980. Structural components of reading time. Journal of Verbal Learning and Verbal Behaviour 19, 135–151.

Grosjean, F. 1980a. Linguistic structure and performance structures: Studies in pause distribution. In H. W. Dechert & M. Raupach (eds.) Temporal variables in speech. The Hague: Mouton, 91–106.

Grosjean, F. 1980b. Temporal variables within and between languages. In H. W. Dechert & M. Raupach (eds.) Towards a cross-linguistic assessment of speech production. Frankfurt: Peter Lang, 39–53.

Grosjean, F. 2001. The bilingual's language modes. In J. Nicol (ed.) One mind, two languages, bilingual language processing. Oxford: Blackwell, 1–22.

Grosjean, F. & Deschamps, A. 1975. Analyse contrastive des variables temporelles de l'anglais et du français: Vitesse de parole et variables composantes, phénomènes d'hésitation. Phonetica 31, 144–184.

Grosjean, F. & Lane, H. 1974. Effects of two temporal variables on the listeners' perception of reading rate. Journal of Experimental Psychology 102(5), 893–896.

Grosser, W. 1993. Aspects of intonational L2 acquisition. In B. Kettemann & W. Wieden (eds.) Current issues in European second language acquisition research. Tübingen: Narr, 81–93.

Grosser, W. 1997. On the acquisition of tonal and accentual features of English by Austrian learners. In A. James & J. Leather (eds.) Second-language speech: Structure and process. Berlin: Mouton de Gruyter, 211–228.

Guion, S. G., Flege, J. E., Liu, S. H. & Yeni-Komshian, G. H. 2000. Age of learning effects on the duration of sentences produced in a second language. Applied Psycholinguistics 21(2), 205–228.

Gut, U., Trouvain, J. & Barry, W. J. 2007. Bridging research on phonetic descriptions with knowledge from teaching practice – the case of prosody in non-native speech. In J. Trouvain & U. Gut (eds.) Non-native prosody. Phonetic description and teaching practice. Berlin: Mouton de Gruyter, 3–21.

Haan, J. 2001. Speaking of questions: an exploration of Dutch question intonation. Utrecht: LOT.

Hadding-Koch, K. & Studdert-Kennedy, M. 1964. An experimental study of some intonation contours. Phonetica 11, 175–185.

Hammerly, H. 1991. Fluency and accuracy: toward balance in language teaching and learning. Clevedon: Multilingual Matters.

Harjula, H. & Manninen, S. 1994. A study year abroad: Exchange students' assessments of their English language proficiency. FINLANCE: A Finnish Journal of Applied Linguistics 13, 134–149.

Harley, B. & Hart, D. 2002. Age, aptitude, and second language learning on a bilingual exchange. In P. Robinson (ed.) Individual differences and instructed language learning. Amsterdam: Benjamins, 301–330.

Hatch, E. M. 1983. Psycholinguistics: a second language perspective. Rowley: Newbury House.

Hawkins, P. R. 1971. The syntactic location of hesitation pauses. Language and Speech 14(3), 277–288.

Hedge, T. 1993. Key concepts in ELT. ELT Journal: English Language Teachers Journal 47(3), 275–277.

Heikkilä, T. 2004. Tilastollinen tutkimus. Helsinki: Edita.

Herman, P. A. 1985. The effect of repeated readings on reading rate, speech pauses, and word recognition accuracy. Reading Research Quarterly 20(5), 553–565.

Hermes, D. J. 1998. Measuring the perceptual similarity of pitch contours. Journal of Speech, Language, and Hearing Research 41(1), 73–82.

Hewings, M. 1995. The English intonation of native speakers and Indonesian learners: A comparative study. RELC Journal 26(1), 27–46.

Hewings, M. 1998. Intonation choices in the English of non-native speakers: An exploratory study. In A. Sanchez-Macarro & R. Carter (eds.) Linguistic choice across genres: Variation in spoken and written English. Amsterdam: John Benjamins, 317–333.

Hieke, A. E. 1981. Audio-lectal practice and fluency acquisition. Foreign Language Annals 14(3), 189–194.

Hieke, A. E. 1984. Linking as a marker of fluent speech. Language and Speech 27. 343–354.

Hieke, A. E. 1985. A componential approach to oral fluency evaluation. Modern Language Journal 69, 135–142.

Hieke, A. E. 1987. Absorption and fluency in native and non-native casual English speech. In A. James & J. Leather (eds.) Sound patterns in second language acquisition. Dordrecht: Foris, 41–58.

Hird, K. & Kirsner, K. 2010. Objective measurement of fluency in natural language production: A dynamic systems approach. Journal of Neurolinguistics 23, 518–530.

Hirst, D. 2007. A praat plugin for momel and INSTINT with improved algorithms for modelling and coding intonation. In J. Trouvain & W. J. Barry (eds.) Proceedings of the 16th ICPhS, 6–10 August 2007, Saarbrücken. Saarbrücken: Universität des Saarlandes, 1233–1236.

Hirst, D. & Di Cristo, A. 1998. A survey of intonation systems. In D. Hirst & A. Di Cristo (eds.) Intonation systems: a survey of twenty languages. England: Cambridge University Press, 1–44.

Hirvonen, P. 1967. On the problems met by Finnish students in learning the rising interrogative intonation of English. Publications of the Phonetics Department of the University of Turku 2. Licentiate Thesis.

Hirvonen, P. 1970. Finnish and English communicative intonation. Publications of the Phonetics Department of the University of Turku 8. Doctoral Dissertation.

Hollien, H. & Paul, P. 1969. A second evaluation of the speaking fundamental frequency characteristics of post-adolescent girls. Language & Speech 12(2), 119–124.

Holm, S. 2008. Intonational and durational contributions to the perception of foreign-accented Norwegian. An experimental phonetic investigation. Department of Language and Communication Studies, Norwegian University of Science and Technology. Doctoral Dissertation.

Huebner, T. 1995. The effects of overseas language programmes. Report on a case study of an intensive Japanese course. In B. F. Freed (ed.) Second

language acquisition in a study abroad context. Amsterdam: John Benjamins, 171–193.
Huhta, A. 1994. Finnish exchange students' self-assessed language proficiency. FINLANCE: A Finnish Journal of Applied Linguistics 13, 41–68.
Humalajoki, R., Peltola, M. S. & Savela, J. 2006. Vieraan kielen vokaalikategorioiden havaitseminen ja oppiminen. The Phonetics Symposium 2006. Helsinki: University of Helsinki, 44–47.
Igarashi, Y. 2006. Intonational patterns in Russian interrogatives – phonetic analyses and phonological interpretations. In Y. Kawaguchi, I. Fónagy & T. Moriguchi (eds.) Prosody and syntax: Cross-linguistic perspectives, 175–196.
Iivonen, A. 1979. Is there interrogative intonation in Finnish? In E. Gårding, G. Bruce & R. Bannert (eds.) Nordic Prosody. Lund: Department of Linguistics, Lund University, 43–53.
Iivonen, A. 1983. On explaining the sentence initial pitch height in Finnish. Helsinki: University of Helsinki.
Iivonen, A. 1987. Puheen intonaatio. Helsinki: Gaudeamus.
Iivonen, A. 1998. Intonation in Finnish. In D. Hirst & A. Di Cristo (eds.) Intonation systems a survey of twenty languages. Cambridge: Cambridge University Press, 311–327.
Iivonen, A. 2001a. Intonation of Finnish questions. In W. A. van Dommelen & T. Fretheim (eds.) Nordic prosody, Proceedings of the VIIth Conference, Trondheim 2000. Frankfurt am Main: Peter Lang, 137–151.
Iivonen, A. 2001b. Kommunikatiivisten aktien ja puhetyylien prosodiikasta, In S. Ojala & J. Tuomainen (eds.) The 21st Phonetics Symposium, Turku, Publications of the Department of Finnish and General Linguistics of the University of Turku, 1–18.
Iivonen, A. 2001c. Speaker's temporal voice range profile as a background for intonation analysis. In L. V. Bondarko (ed.) 100 let eksperimental'noy fonetike v Rossii. Materialy mezhdunarodnoy konferencii 1–4 fevralya 2001. Sankt-Peterburg: Filologicheskiy fakul'tet SPbGU, 86–89.
Iivonen, A. 2005. Intonaation käsitteen täsmennystä. In A. Iivonen (ed.) Puheen salaisuudet. Helsinki: Gaudeamus, 93–128.
Iivonen, A. 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.
Iivonen, A. 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.
Iivonen, A., Niemi, T. & Paananen, M. M. 1995. Comparison of prosodic characteristics in English, Finnish and German radio and TV newscasts. In K. Elenius & P. Branderud (eds.) Proceedings of the 13th ICPhS, Stockholm, Sweden, 13–19 August 1995, 382–385.

Isabelli-García, C. L. 2003. Development of oral communication skills abroad. Frontiers: The Interdisciplinary Journal of Study Abroad 9, 149–173.

Itkonen, T. 1972. Kuoreveden ja keuruun murretta. Tekstejä ja sandhiseikkojen tarkastelua. Suomi 117(1).

Ivanova-Lyukyanova, G. N. 2003. Kul'tura ustnoy rechi, intonaciya, pauzirovanie, logicheskoe udarenie, temp, ritm. Moskva: Flinta, Nauka.

Jassem, W. & Kudela-Dobrogowska, K. 1980. Speaker-independent intonation curves. In L. R. Waugh & C. H. V. Schooneveld (eds.) The melody of language. Baltimore: Univ. Park Press, 135–148.

Jensen, L. 1986. Advanced reading skills in a comprehensive course. In F. Dubin, D. E. Eskey & W. Grabe (eds.) Teaching second language reading for academic purposes. Reading, Massachusetts: Addison-Wesley, 103–124.

Jilka, M. 2000. The contribution of intonation to the perception of foreign accent. Arbeiten des Instituts für Maschinelle Sprachverarbeitung (AIMS) Doctoral Dissertation. Available at: http://ifla.uni–stuttgart.de/institut/mitarbeiter/jilka/papers/diss.pdf. Last retrieved: 27th June 2009.

Jilka, M. 2007. Different manifestations and perceptions of foreign accent in intonation. In J. Trouvain, U. Gut & W. J. Barry (eds.) Non-native prosody: Phonetic description and teaching practice. Berlin: Mouton de Gruyter, 77–96.

Jun, S. & Oh, M. 2000. Acquisition of second language intonation. Proceedings of the 6th ICSLP 2000, Beijing, China, October 16–20. 2000. 73–76. Available at: http://www.humnet.ucla.edu/humnet/linguistics/people/jun/ICSLP–L2intonation.pdf. Last retrieved: 27th June 2009.

Just, M. A. & Carpenter, P. A. 1987. The psychology of reading and language comprehension. Boston: Allyn and Bacon.

Kasatkin, L. L. 2003. Fonetika sovremennogo russkogo literaturnogo yazyka. Moskva: Izdatel'stvo MGU.

Kasatkin, L. L. 2007. Russkaya intonacia: Tonalnye kontury. In R. F. Kasatkina (ed.) Problemy fonetiki V. Moskva: Nauka, 250–266.

Keijsper, C. E. 1983. Comparing Dutch and Russian pitch contours. Russian Linguistics: International Journal for the Study of the Russian Language 7(2), 101–154.

Kendall, T. S. 2009. Speech rate, pause, and linguistic variation: An examination through the sociolinguistic archive and analysis project. Department of English, Duke University. Doctoral Dissertation.

Kenny, K. D. 1996. Language loss and the crisis of cognition: between socio- and psycholinguistics. Berlin: Mouton de Gruyter.

Koda, K. 2007. Reading and language learning: Crosslinguistic constraints on second language reading development. Language Learning 57, 1–44.

Kodzasov, S. V. & Krivnova, O. F. 2001. Obshchaya fonetika. Moskva: Rossiiskiy gumanitarnyy universitet.

Kolesov, V. V. 1972. Fonetika. In N. A. Mescherskiy (ed.) Russkaya dialektologia. Moskva: Vysshaya shkola, 49–121.

Koponen, M. 1992. Glottal boundary markers and aspects of disfluency in Finnish English and epilogue (from 1990 to 1992), University of Jyväskylä, Department of English. Licenciate Thesis.

Korkala, S. 2008. Kansainvälinen liikkuvuus yliopistoissa ja ammattikorkeakouluissa 2007. Cimo publications 1/2008. Available at: http://www.cimo.fi/dman/Document.phx/%7Epublic/Julkaisut+ja+tilastot/Raportit+ja+selvitykset/cimo_publications_1_2008.pdf. Last retrieved: 28th February 2009.

Kormos, J. 2006. Speech production and second language acquisition. Mahwah: Lawrence Erlbaum.

Korpijaakko-Huuhka, A. 1996. Sujuvasti kertomalla – metodisia vihjeitä kuntoutumisen arviointiin. In K. Hyttinen, T. Jääskeläinen, L. Korjus-Julkunen, A. Timonen & P. Toivanen (eds.) Puheterapian uudet suunnat – logopedinen tutkimus ja kuntoutus tänään. Helsinki: Puheterapeuttien Kustannus, 126–133.

Koster, C. J. & Koet, T. 1993. The evaluation of accent in the English of Dutchmen. Language Learning 43(1), 69–92.

Kraemer, H. C. 1983. Kappa coefficient. In S. Kotz & N. L. Johnson (eds.) Encyclopedia of statistical sciences, Volume 4. New York: John Wiley and Sons, 352–354.

Kuosmanen, A. & de Silva, V. 2003. Why don't Russians answer my questions? Finnish students' problems in producing Russian interrogative intonation. In M. J. Solé, D. Recasens & J. Romero (eds.) Proceedings of the 15th ICPhS, Barcelona, 3–9 August 2003, 523–526.

Kuosmanen, A. & de Silva, V. 2007. Problemy finskih studentov pri ovladenii russkoy voprositel'noy intonaciey. In R. F. Kasatkina (ed.) Problemy fonetiki V. Moskva: Nauka, 297–307.

Kärkkäinen, H. 2009. Intonation units in Russian spontaneous speech spoken by Finnish Russian language learners. In M. Vainio, R. Aulanko & O. Aaltonen (eds.) Nordic prosody. Proceedings of the Xth Conference, Helsinki 2008. Frankfurt am Main: Peter Lang, 127–138.

Kärkkäinen, H., de Silva, V. & Ullakonoja, R. 2006. Havaintoja perustaajuudesta spontaanin puheen jaksottelussa suomessa ja venäjässä. In R. Aulanko, L. Wahlberg & M. Vainio. The Phonetics Symposium 2006. Publications of the Department of Speech Sciences, University of Helsinki 53, 84–92.

Kärkkäinen, H., Ullakonoja, R. & de Silva, V. 2007. Acoustic correlates of word stress and accent in Russian – analysis of Finnish FL learners of Russian. In R. Potapova (ed.) SPECOM 2007. Moscow: Moscow State Linguistic University, 233–238.

Ladd, D. R. 1996. Intonational phonology. Cambridge: Cambridge University Press.

Ladefoged, P. 1996. Elements of acoustic phonetics. Chigado: The University of Chigaco Press.

Lado, R. 1961. Language testing. The construction and use of foreign language tests: a teacher's book. London: Longmans.

Lado, R. 1964. Language teaching. A scientific approach. New York: McGraw-Hill.

Lafford, B. A. 2004. The effect of the context of learning on the use of communication strategies by learners of Spanish as a second language. Studies in Second Language Acquisition 26(2), 201–225.

Landis, J. R. & Koch, G. G. 1977. The measurement of observer agreement for categorical data. Biometrics 33(1), 159–174.

Lane, H. & Grosjean, F. 1973. Perception of reading rate by speakers and listeners. Journal of Experimental Psychology 97(2), 141–147.

Language Testing International 2004. Oral Proficiency Interviews (OPI), Available at: http://www.languagetesting.com/acad_opi.htm#opi. Last retrieved: 16th November 2008.

Lapkin, S., Hart, D. & Swain, M. 1995. A Canadian interprovincial exchange: Evaluating the linguistic impact of a three-month stay in Quebec. In B. F. Freed (ed.) Second language acquisition in a study abroad context. Amsterdam: John Benjamins, 67–94.

Laukkanen, A. & Leino, T. 1999. Ihmeellinen ihmisääni: äänenkäytön ja puhetekniikan perusteet, arviointi, mittaaminen ja kehittäminen. Helsinki: Gaudeamus.

Laukkanen, A., Mäki, E., Pukander, J. & Anttila, I. 1999. Vertical laryngeal size and the lowest tone in the evaluation of the average fundamental frequency (F0) of Finnish speakers. Logopedics Phoniatrics Vocology 24(4), 170–177.

Lauranto, Y. 2004. Puheen jaksotus, intonaatio ja välikielen pragmatiikan jäljitelty puhe. In B. Straszer & A. Brown (eds.) Suomen kielen prosodian opettamisen ja oppimisen kysymyksiä. Helsinki: Helsingin yliopiston suomen kielen laitos, 47–116.

Lauranto, Y. 2005. Sujuvuuden mittoja. In L. Kuure, E. Kärkkäinen & M. Saarenkunnas (eds.) Language and Social Action, AFinLA Yearbook 2005, 127–147.

Laver, J. 1994. Principles of phonetics. Cambridge: Cambridge University Press.

Leech, N. L., Barrett, K. C. & Morgan, G. A. 2005. SPSS for intermediate statistics: use and interpretation. Mahwah: Lawrence Erlbaum.

Leed, R. L. 1965. A contrastive analysis of Russian and English intonation contours. The Slavic and East European Journal 9(1), 62–75.

Lehtonen, J. 1978. On the problems of measuring fluency. In M. Leiwo & A. Räsänen (eds.) AFinLA Yearbook 1978. Jyväskylä: University of Jyväskylä, Department of English, 53–68.

Lehtonen, J. 1979. Speech rate and pauses in the English of Finns, Swedish-speaking Finns, and Swedes. Turku: Åbo Akademi.

Lehtonen, J. 1981. Problems of measuring fluency and normal rate of speech. In J. Savard & L. Laforge (eds.) Proceedings of the 5th Congress of

l'association internationale de linguistique appliquée. Montreal: Presses de l'Université Laval, 322–332.

Lehtonen, J. 1987. Communicative approaches to speech, language and the study of grammar. In H. W. Dechert & M. Raupach (eds.) Psycholinguistic models of production. Norwood, NJ: Ablex, 35–42.

Lehtonen, J. & Heikkinen, H. 1981. Anglismit ja tekstin luettavuus. Virittäjä 25(4), 327–339.

Lehtonen, J., Sajavaara, K. & May, A. 1977. Spoken English: the perception and production of English on a Finnish-English contrastive basis. Jyväskylä: Gummerus.

Lennes M. 2007. Mietta's Praat scripts. Available at http://www.helsinki.fi/~lennes/praat–scripts/. Last retrieved: 19th Dec 2007.

Lennes, M. 2009. Segmental features in spontaneous and read-aloud Finnish. 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, 145–166.

Lennes, M. & Ahjoniemi, S. 2005. Puheaineiston annotaatio eli nimikointi. Available at: http://www.helsinki.fi/~lennes/annotation_guide/. Last retrieved: 16th October 2009.

Lennon, P. 1984. Retelling a story in English as a second language. In H. W. Dechert, D. Möhle & M. Raupach (eds.) Second language productions. Tübingen: Narr, 50–68.

Lennon, P. 1990. Investigating fluency in EFL: A quantitative approach. Language Learning: A Journal of Applied Linguistics 40(3), 387–417.

Lennon, P. 2000. The lexical element in spoken second language fluency. In H. Riggenbach & R. Schmidt (eds.) Perspectives on fluency. Ann Arbor: University of Michigan Press, 25–42.

Levelt, W. J. M. 1989. Speaking: from intention to articulation. Cambridge: MIT Press.

Levis, J. & Pickering, L. 2004. Teaching intonation in discourse using speech visualization technology. System 32(4), 505–524.

Lieberman, P. 1963. Some acoustic measures of the fundamental periodicity of normal and pathologic larynges. Journal of the Acoustical Society of America 35(3), 344–353.

Lieberman, P. & Blumstein, S. E. 1988. Speech physiology, speech perception, and acoustic phonetics. Cambridge: Cambridge University Press.

Llama, R., Walcir, C. & Collins, L. 2008. The roles of typology and L2 status in the acquisition of L3 phonology: The influence of previously learnt languages on L3 speech production. In A. S. Rauber, M. A. Watkins & B. O. Baptista (eds.) New Sounds 2007. Florianópolis: Federal University of Santa Catarina, 313–323.

Llanes, À. & Muñoz, C. 2009. A short stay abroad: Does it make a difference? System 37(3), 353–365.

Loizou, P. 2008. COLEA: A Matlab Software Tool for Speech Analysis. Available at: http://www.utdallas.edu/~loizou/speech/colea.htm. Last retrieved: 16th June 2008.

Lyubimova, N. A. 1988. Foneticheskyy aspekt obshchenya na nerodnom yazyke. Leningrad: Izdatel'stvo LGU.

MacKay, I. R. A., Flege, J. E., Piske, T. & Schirru, C. 2001. Category restructuring during second-language speech acquisition. Journal of the Acoustical Society of America 110(1), 516–528.

Magen, H. S. 1998. The perception of foreign-accented speech. Journal of Phonetics 26(4), 381–400.

Mahon, D. 1986. Intermediate skills: Focusing on reading rate development. In F. Dubin, D. E. Eskey & W. Grabe (eds.) Teaching second language reading for academic purposes. Reading: Addison-Wesley, 77–102.

Major, R. C. 2008. Transfer in second language phonology. In J. G. Hansen Edwards & M. L. Zampini (eds.) Phonology and second language acquisition. Amsterdam: John Benjamins, 63–94.

Makarova, V. 1999. Pitch peak alignment in Russian declaratives, interrogatives, and exclamations. In J. J. Ohala, Y. Hasewaga, M. Ohaja, D. Granville & A. C. Bailey (eds.) Proceedings of the 14th ICPhS, San Francisco, 1–7 August 1999. Berkeley: University of California, 1173–1176.

Makarova, V. 2001. Perceptual correlates of sentence-type intonation in Russian and Japanese. Journal of Phonetics 29(2), 137–154.

Makarova, V. 2007. The effect of pitch peak alignment on sentence type identification in Russian. Language and Speech 50(3), 385–422.

Marriott, H. E. 1995. The acquisition of politeness patterns by exchange students in Japan. In B. F. Freed (ed.) Second language acquisition in a study abroad context. Amsterdam: John Benjamins, 197–224.

Martin, P. 2010. Learning the prosodic structure of a foreign language with a pitch visualizer. In M. Hasegawa-Johnson, A. Bradlow, J. Cole, K. Livescu, J. Pierrehumbert & C. Shih (eds.) Proceedings of Speech Prosody 2010, Chigaco. Available at: http://speechprosody2010.illinois.edu/papers/100980.pdf. Last retrieved: 11th June 2010.

Martynov, V. C. 1971 (cited by Shtern 1988). Raspredeleniye dlitel'nostey tonal'nyh i shumovyh uchastkov rechi. VII Vsesoyuznaya Acousticheskaya Konferenciya: Tezisy Dokladov.

McAllister, R., Flege, J. E. & Piske, T. 1999. The acquisition of Swedish long vs. short vowel contrasts by native speakers of English, Spanish and Estonian. In J. J. Ohala, Y. Hasewaga, M. Ohaja, D. Granville & A. C. Bailey (eds.) Proceedings of the 14th ICPhS, San Francisco, 1–7 August 1999. Berkeley: University of California, 751–754.

Meister, L. & Meister, E. 2007. Perceptual assessment of Russian-accented Estonian. In J. Trouvain & W. J. Barry (eds.) Proceedings of the 16th ICPhS, 6–10 August 2007, Saarbrücken. Saarbrücken: Universität des Saarlandes, 1717–1720.

Mennen, I. 1998a. Can language learners ever acquire the intonation of a second language? Proceedings of the ESCA Workshop on Speech Technology in Language Learning, 17–20.

Mennen, I. 1998b. Second language acquisition of intonation: The case of peak alignment. Papers from the Regional Meetings, Chicago Linguistic Society, 34(2), 327–341.

Mennen, I. 1999. The realisation of nucleus placement in second language intonation. In J. J. Ohala, Y. Hasewaga, M. Ohaja, D. Granville & A. C. Bailey (eds.) Proceedings of the 14th ICPhS, San Francisco, 1–7 August 1999. Berkeley: University of California, 555–558.

Mennen, I. 2004. Bi-directional interference in the intonation of Dutch speakers of greek. Journal of Phonetics(32), 543–563.

Mennen, I. 2006. Phonetic and phological influences in non-native intonation: An overview for language teachers. QMUC Speech Science Research Centre Working Paper WP9.

Mennen, I. 2007. Phonological and phonetic influences in non-native intonation. In J. Trouvain, U. Gut & W. J. Barry (eds.) Non-native prosody: Phonetic description and teaching practice. Berlin: Mouton de Gruyter, 53–76.

Mennen, I., Chen, A. & Karlsson, F. 2010. Characterising the internal structure of learner intonation and its development over time. In K. Dziubalska-Kołaczyk, M. Wrembel & M. Kul (eds.) Proceedings of New Sounds 2010: The sixth international symposium on the acquisition of second language speech. University of Poznan. Available at: http://www.bilingualism.bangor.ac.uk/documents/Mennen%20et%20al%20NewSounds.pdf Last retrieved: 28th November 2010, 319–324.

Mennen, I., Schaeffler, F. & Docherty, G. 2007. Pitching it differently: A comparison of the pitch ranges of German and English speakers. In J. Trouvain & W. J. Barry (eds.) Proceedings of the 16th ICPhS, 6–10 August 2007, Saarbrücken. Saarbrücken: Universität des Saarlandes, 1769–1772.

Mennen, I., Schaeffler, F. & Docherty, G. 2008. A methodological study into the linguistic dimensions of pitch range differences between German and English. In P. A. Barbosa, S. Madureira & C. Reis (eds.) Proceedings of the Speech Prosody 2008, Campinas, Brazil: Editora RG/CNPq, 527–530.

Meyer, R. & Mleinek, I. 2006. How prosody signals force and focus – A study of pitch accents in Russian yes-no questions. Journal of Pragmatics 38(10), 1615–1635.

Mikheev, Y. V. 1970 (cited by Shtern 1988). Statisticheskyy zakon raspredeleniya periodov osnovnogo tona russkoy rechi. Akusticheskyy Zhurnal 16(4), 558–562.

Mixdorff, H., Vainio, M., Werner, S. & Järvikivi, J. 2002. The manifestation of linguistic information in prosodic features of Finnish. In B. Bel & I. Marlien (eds.) Proceedings of the Speech Prosody 2002, Aix-en-Provence, 11–13 April 2002, Université de Provence, 515–518.

Moore, K. 1990. On prosodic elements in television and radio sports narrations. In K. Wiik & I. Raimo (eds.) Nordic Prosody V. Turku: University of Turku, 219–235.

Moore, K. & Korpijaakko-Huuhka, A. 1996. The clinical assessment of Finnish fluency. In M. J. Ball & M. Duckworth (eds.) Advances in clinical phonetics. Amsterdam: John Benjamins, 171–196.

Moustroufas, N. & Digalakis, V. 2007. Automatic pronunciation evaluation of foreign speakers using unknown text. Computer Speech & Language 21(1), 219–230.

Munro, M. J. 1995. Nonsegmental factors in foreign accent: Ratings of filtered speech. Studies in Second Language Acquisition 17(1), 17–34.

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. & Derwing, T. M. 1995. Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. Language and Speech 38(3), 289–306.

Munro, M. J. & Derwing, T. M. 1998. The effects of speaking rate on listener evaluations of native and foreign-accented speech. Language Learning 48(2), 159–182.

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(4), 451–468.

Möhle, D. 1984. A comparison of the second language speech production of different native speakers. In H. W. Dechert, D. Möhle & M. Raupach (eds.) Second language productions. Tübingen: Narr, 26–49.

Nenonen, S. 2001a. Venäläiset suomenoppijat suomen kielen pitkien painottomien vokaalien havaitsijoina. In M. Charles & P. Hiidenmaa (eds.) Tietotyön yhteiskunta – kielen valtakunta. AFinLA Yearbook 2001, 11–31.

Nenonen, S. 2001b. Äännekestot venäläisen suomenoppijan spontaanissa suomenkielisessä puheessa. The 21st Phonetics Symposium, Turku, Publications of the Department of Finnish and General Linguistics of the University of Turku, 30–39.

Nenonen, S., Shestakova, A., Huotilainen, M. & Näätänen, R. 2003. Linguistic relevance of duration within the native language determines the accuracy of speech-sound duration processing. Cognitive Brain Research 16, 492–495.

Nenonen, S., Shestakova, A., Huotilainen, M. & Näätänen, R. 2005. Speech-sound duration processing in a second language is specific to phonetic categories. Brain and Language 92(1), 26–32.

Neufeld, G. & Schneiderman, E. 1980. Prosodic and articulatory features in adult language learning. 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: Newbury House, 105–109.

Nevalainen, T. 1990. Some aspects of convergence in Finnish and English prosody. In K. Wiik & I. Raimo (eds.) Nordic Prosody V. Turku: University of Turku, 236–255.

Nikolayeva, T. M. 1977. Frazovaya intonatsiya slavyanskih yazykov. Moskva: Nauka.

Nolan, F. 2003. Intonational equivalence: An experimental evaluation of pitch scales. In M. J. Solé, D. Recasens & J. Romero (eds.) Proceedings of the 15th ICPhS, Barcelona, 3–9 August 2003, 771–774.

Nunnally, J. C. & Bernstein, I. H. 1994. Psychometric theory. New York: McGraw-Hill.

Nuttall, C. 1982. Teaching reading skills in a foreign language. London: Heinemann Educational Books.

Odé, C. 1989. Russian intonation: a perceptual description. Amsterdam: Rodopi.

Odé, C. 2003. Toward ToRI, a manual transcription system of Russian intonation. IFA Proceedings 25, 1–14. Available at: http://213.207.98.211/ode/pdf/Proceedings25–ToRI.pdf. Last retrieved: 9th March 2007.

Odé, C. 2005. Neutralization or truncation? The perception of two Russian pitch accents on utterance-final syllables. Speech Communication 47(1–2), 71–79.

Ogden, R. 2003. Voice quality as resource for the management of turn–taking in Finnish talk-in-interaction. In M. J. Solé, D. Recasens & J. Romero (eds.) Proceedings of the 15th ICPhS, Barcelona, 3–9 August 2003,123–126.

Ogden, R. 2004. Non-modal voice quality and turn-taking in Finnish. In E. Couper-Kuhlen & C. E. Ford (eds.) Sound patterns in interaction: Cross-linguistic studies from conversation. Amsterdam: John Benjamins, 29–62.

Ogden, R. & Routarinne, S. 2005. The communicative functions of final rises in Finnish intonation. Phonetica 62, 160–175.

Ohala, J. J. 1983. Cross-language use of pitch: An ethological view. Phonetica 40, 1–18.

Ohala, J. J. 1984. An ethological perspective on common cross-language utilization of F0 of voice. Phonetica 41, 1–16.

Ohara, Y. 2001. Finding one's voice in Japanese: A study of the pitch levels of L2 users. In A. Pavlenko (ed.) Multilingualism, second language learning and gender. Berlin: Mouton de Gruyter, 231–254.

Okamura, A. 1995. Teachers' and nonteachers' perception of elementary learners' spoken Japanese. Modern Language Journal 79(1), 29–40.

Olynyk, M., Anglejan, A. D. & Sankoff, D. 1987. A quantitative and qualitative analysis of speech markers in the native and second language speech of bilinguals. Applied Psycholinguistics 8(2), 121–136.

Paananen, M. M. 1998. Pausing in non-native spontaneous speech. In S. Werner (ed.) Nordic Prosody: Proceedings of the VIIth Conference, Joensuu 1996. Frankfurt am Main: Peter Lang, 237–245.

Paananen–Porkka, M. M. 2007. Speech rhythm in an interlanguage perspective: Finnish adolescents speaking English. PIC monographs. Helsinki: University of Helsinki. Doctoral Dissertation.

Pallonen, J. & Yli-Luukko, E. 1995. Hietamäen murretta. Suomen kielen näytteitä 42. Helsinki: Painatuskeskus.

Panov M. V. 1968. Fonetika sovremennogo russkogo literaturnogo yazyka. Moskva: Nauka.

Patterson, D. 2000. A linguistic approach to pitch range modelling. University of Edinburg. Doctoral Dissertation.

Patterson, D. & Ladd, D. R. 1999. Pitch range modelling: Linguistic dimensions of variation. In J. J. Ohala, Y. Hasewaga, M. Ohaja, D. Granville & A. C. Bailey (eds.) Proceedings of the 14th ICPhS, San Francisco, 1–7 August 1999. Berkeley: University of California, 1169–1172.

PCGN = The permanent committee on geographical names for British official use 2009. Romanization system for Russian. Available at: http://www.pcgn.org.uk/ROMANIZATION_RUSSIAN.pdf. Last retrieved: 12th August 2009.

Pellegrino, V. A. 1998. Student perspectives on language learning in a study abroad context. Frontiers: The Interdisciplinary Journal of Study Abroad 4, 91–120.

Penttilä, A. 1958. Intonaatiotutkimuksia. Virittäjä 62, 1–18.

Perales, J. & Cenoz, J. 1996. Silence, communicative competence and communication strategies in second language acquisition. In G. M. Grabher & U. Jessner (eds.) Semantics of silences in linguistics and literature. Heidelberg: Winter, 67–87.

Perfetti, C. A. 1985. Reading ability. New York: Oxford University Press.

Perfetti, C. A. 2003. The universal grammar of reading. Scientific Studies of Reading 7(1), 3–24.

Perfetti, C. A. & Roth, S. F. 1981. Some of the interactive processes in reading and their role in reading skill. In A. M. Lesgold & C. A. Perfetti (eds.) Interactive processes in reading. Hillsdale, New Jersey: Lawrence Erlbaum, 269–297.

Pfitzinger, H. R. 1999. Local speech rate perception in German speech. In J. J. Ohala, Y. Hasewaga, M. Ohaja, D. Granville & A. C. Bailey (eds.) Proceedings of the 14th ICPhS, San Francisco, 1–7 August 1999. Berkeley: University of California, 893–896.

Pfitzinger, H. R. & Tamashima, M. 2006. Comparing perceptual local speech rate of German and Japanese Speech. In R. Hoffmann & H. Mixdorff (eds.) Proceedings of the Speech Prosody 2006 Conference, 2–5 May 2006, Dresden, Germany, 105–108.

Pierrehumbert, J. B. 1987. The phonology and phonetics of English intonation. Bloomington: Indiana University Linguistics Club.

Pilon, R. 1981. Segmentation of speech in a foreign language. Journal of Psycholinguistic Research 10(2), 113–122.

Piske, T. & MacKay, I. R. A. 1999. Age and L1 use effects on degree of foreign accent in English. In J. J. Ohala, Y. Hasewaga, M. Ohaja, D. Granville & A. C. Bailey (eds.) Proceedings of the 14th ICPhS, San Francisco, 1–7 August 1999. Berkeley: University of California, 1433–1436.

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

Prom-on, S., Xu, Y. & Thipakorn, B. 2009. Modeling tone and intonation in Mandarin and English as a process of target approximation. Journal of the Acoustical Society of America 125(1), 405–424.

Quené, H. & van Delft, L. E, 2010. Non-native durational patterns decrease speech intelligibility. Speech Communication 52, 911–918.

Ramirez Verdugo, D. 2005. The nature and patterning of native and non-native intonation in the expression of certainty and uncertainty: Pragmatic effects. Journal of Pragmatics 37(12), 2086–2115.

Ramirez Verdugo, D. 2006. A study of intonation awareness and learning in non-native speakers of English. Language Awareness 15(3), 141–159.

Rantala, L., Vilkman, E. & Bloigu, R. 2002. Voice changes during work: Subjective complaints and objective measurements for female primary and secondary schoolteachers. Journal of Voice 16(3), 344–355.

Rasinski, T. V. 2003. The Fluent Reader. Oral reading strategies for building word recognition, fluency and comprehension. New York: Scholastic Professional Books.

Raupach, M. 1980. Temporal variables in first and second language speech production. In H. W. Dechert & M. Raupach (eds.) Temporal variables in speech. The Hague: Mouton, 263–270.

Reagan, V. 1995. The acquisition of sociolinguistic native speech norms: Effects of a year abroad on second language learners of French. In B. F. Freed (ed.) Second language acquisition in a study abroad context. Amsterdam: John Benjamins, 245–267.

Riazantseva, A. 2001. Second language proficiency and pausing: A study of Russian speakers of English. Studies in Second Language Acquisition 23(4), 497–526.

Rietveld, A. C. M. & Gussenhoven, C. 1987. Perceived speech rate and intonation. Journal of Phonetics 15(3), 273–285.

Riggenbach, H. 1991. Toward an understanding of fluency: A microanalysis of nonnative speaker conversations. Discourse Processes 14(4), 423–441.

Riggenbach, H. 2001. Hesitation phenomena in second-language fluency. In A. Wennerstrom (ed.) Music of everyday speech: Prosody and discourse analysis. Cary: Oxford University Press, 252–257.

Ringbom, H. 1987. The role of the first language in foreign language learning. Clevedon: Multilingual Matters.

Rocca, P. A. D. 2008. New trends on the teaching of intonation of foreign languages. In A. S. Rauber, M. A. Watkins & B. O. Baptista (eds.) New sounds 2007, 420–428. Available at: http://www.nupffale.ufsc.br/newsounds/Papers/37.Rocca_Paulina.pdf. Last retrieved: 27th June 2009.

Rossi, M. 1981. L'intonation, de l'acoustique à la sémantique. Paris: Klincksieck.

Routarinne, S. 2003. Tytöt äänessä: Parenteesit ja nouseva sävelkulku kertojien vuorovaikutuskeinoina. Suomalaisen kirjallisuuden seuran toimituksia. Helsinki: University of Helsinki. Doctoral Dissertation.

Routarinne, S. 2008. Miksi intonaatio nousee? In S. Routarinne & T. Uusi-Hallila (eds.) Nuoret kielikuvassa. Kouluikäisten kieli 2000-luvulla. Helsinki: Suomalaisen kirjallisuuden seura, 125–145.

Sajavaara, K. & Lehtonen, J. 1980. The analysis of cross-language communication: Prolegomena to the theory and methodology. In H. W. Dechert & M. Raupach (eds.) Towards a cross-linguistic assessment of speech production. Frankfurt: Peter Lang, 55–76.

Sallinen-Kuparinen, A. 1979. Hur talar en finne? In E. Gårding, G. Bruce & R. Bannert (eds.) Nordic Prosody. Lund: Department of Linguistics, Lund University, 255–261.

Scharff-Rethfeldt, W., Miller, N. & Mennen, I. 2008. Unterschiede in der mittleren Sprechtonhöhe bei Deutsch/Englisch bilingualen Sprechern – Speaking fundamental frequency differences in highly proficient bilinguals of German/English. Sprache, Stimme, Gehör 32, 123-128.

Segalowitz, N. 1986. Skilled reading in the second language. In J. Vaid (ed.) Language processing in bilinguals: Psycholinguistic and neuropsychological perspectives. Hillsdale: Lawrence Erlbaum, 3–19.

Segalowitz, N. 2000. Automaticity and attentional skill in fluent performance. In H. Riggenbach (ed.) Perspectives on fluency. Ann Arbor: University of Michigan Press, 200–210.

Segalowitz, N. 2007. Access fluidity, attention control, and the acquisition of fluency in a second language. TESOL Quarterly 41(1), 181–185.

Segalowitz, N. 2010. Cognitive bases of second language fluency. New York: Routledge.

Segalowitz, N. & Freed, B. F. 2004. Context, contact, and cognition in oral fluency acquisition: Learning Spanish in at home and study abroad contexts. Studies in Second Language Acquisition 26(2), 173–199.

Segalowitz, N., Freed, B. F., Collentine, J., Lafford, B. A., Lazar, N. A. & Diaz-Campos, M. 2004. A comparison of Spanish second language acquisition in two different learning contexts: Study abroad and the domestic classroom. Frontiers: The Interdisciplinary Journal of Study Abroad 10(4), 1–18.

Segalowitz, N., Poulsen, C. & Komeda, M. 1991. Lower level components or reading skill in higher level bilinguals: Implications for reading instruction. AILA Review 8, 15–30.

Sereno, J. A. & Wang, Y. 2007. Behavioral and cortical effects of learning a second language. In O. Bohn (ed.) Language experience in second language speech learning: In honor of James Emil Flege. Amsterdam: John Benjamins, 239–258.

Serrano, R., Tragant, E. & Llanes À. 2010. The effect of 'length of stay' on the development of L2 oral performance in the case of Erasmus students. A

paper read at the Eurosla 2010 conference, 3rd September, 2010, Reggio Emilia, Italy.

Shaul'skiy E. V. & Knyazev S. V. 2006. Russkaya dialektologiya, fonetika. Moskva. Available at: http://home.medewerker.uva.nl/c.ode/bestanden/dialektologi_2utgave.pdf Last retrieved: 26th November 2010.

Shcherba, L. V. 1955. Fonetika frantsuzkogo yazyka. Moskva: Izdatel'stvo literatury na inostrannyh yazykah.

Shcherbakova, L. P. 2001. Vospriyatie intonacionnyh tipov nerodnogo yazyka. In L. V. Bondarko (ed.) 100 let eksperimental'noy fonetike v Rossii. Materialy mezhdunarodnoy konferencii 1–4 fevralya 2001. Sankt-Peterburg: Filologicheskiy fakul'tet SPbGU, 173–175.

Shcherbakova, L. P. 2002. Intonacionnoe oformlenie russkoy rechi finskih studentov (na primere spontannoy rechi), Problemy i metody eksperimental'no-foneticheskih issledovaniy. K 70-letiyu professora kafedry fonetiki i metodiki prepodavaniya inostrannyh yazykov L. V. Bondarko, 270–275.

Shilova, K. A. & Usmanova, E. E. 1990. 100 dialogov po telefonu. Moskva: Russkiy yazyk.

Shtern, A. S. 1988. Nekotorye statisticheskye harakteristiki russkoy spontannoy rechi. In N. D. Svetozarova (ed.) Fonetika spontannoy rechi. Leningrad: Izdatel'stvo Leningradskogo universiteta, 196–224.

Silverman, K., Beckman, M. E., Pitrelli, J., Ostendorf, M., Wightman, C., Price, P., Pierrehumbert, J. & Hirschberg, J. 1992. TOBI: A standard for labeling English prosody. Second international conference on spoken language processing (ICSLP'92), Banff, Canada, October 13–16, 1992, 867–870.

Sim, J. & Wright, C. C. 2005. The kappa statistic in reliability studies: Use, interpretation, and sample size requirements. Physical Therapy 85(3), 257–268.

Simões, A. R. M. 1996. Phonetics in second language acquisition: An acoustic study of fluency in adult learners of Spanish. Hispania 79(1), 87–95.

Sjölander, K. & Beskow, J. 2006. WaveSurfer. Available at: http://www.speech.kth.se/wavesurfer/. Last retrieved: 16th June 2008.

Skrelin, P. & Volskaya, N. B. 2004. Russian read and spontaneous speech: Prosodic data analysis. In G. Bruce & M. Horne (eds.) Nordic Prosody, Proceedings of the IXth Conference, Lund 2004. Frankfurt am Main: Peter Lang, 235–244.

Stahl, S. A. & Heubach, K. 2006. Fluency-oriented reading instruction. In M. C. McKenna & K. A. D. Stahl (eds.) Reading research at work: Foundations of effective practice. New York: Guilford Publications, 177–204.

Stemler, S. E. 2004. A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. Practical Assessment, Research & Evaluation 9(4). Available at: http://PAREonline.net/getvn.asp?v=9&n=4. Last retrieved: 29th June 2009.

Strangert, E. 1990. Pauses, syntax and prosody. In K. Wiik & I. Raimo (eds.) Nordic Prosody V. Turku: University of Turku, 294–305.

Strangert, E. 1991. Pausing in texts read aloud. In B. Bel & I. Marlien (eds.) Proceedings of the Speech Prosody 2002, Aix-en-Provence, 11–13 April 2002, Université de Provence, 238–241.

Strangert, E. & Gustafson, J. 2008. Acoustic-prosodic features behind the impression of speaker skills. Paper read at the 10th Nordic Prosody Conference, Helsinki 6th August 2008.

Straszer, B. 2003. Havaintoja suomen kielen prosodisten piirteiden oppimisesta tietokoneavusteisella alkeiskurssilla. Virittäjä 107(4), 588–596.

Suomi, K. 2007. Accentual tonal targets and speaking rate in northern Finnish. TMH-QPSR 50(1), 109–112.

Svetozarova, N. D. 1975. The inner structure of intonation contours in Russian. In G. Fant & M. A. A. Tatham (eds.) Auditory analysis and perception of speech. London: Academic Press, 499–510.

Svetozarova, N. D. 1982. Intonacionnaya sistema russkogo yazyka. Leningrad: Izdatel'stvo Leningradskogo universiteta.

Svetozarova, N. D. 1998. Intonation in Russian. In D. Hirst & A. Di Cristo (eds.) Intonation systems: a survey of twenty languages. England: Cambridge University Press, 261–274.

Syrjä, T. 2007. Vieras kieli suussa: Vieraalla kielellä näyttelemisen ulottuvuuksia näyttelijäopiskelijan äänessä, puheessa ja kehossa. Tampere: Tampereen yliopisto. Doctoral Dissertation. Available at: http://acta.uta.fi/pdf/978-951-44-7116-2.pdf. Last retrieved: 13th July 2009.

't Hart, J., Collier, R. & Cohen, A. 1990. A perceptual study of intonation: an experimental-phonetic approach to speech melody. Cambridge: Cambridge University Press.

Taguchi, E. & Gorsuch, G. J. 2002. Transfer effects of repeated EFL reading on reading new passages: A preliminary investigation. Reading in a Foreign Language 14(1), 43–64. Available at:
http://nflrc.hawaii.edu/rfl/April2002/taguchi/taguchi.pdf.
Last retrieved: 23rd October 2007.

Taguchi, E., Takayasu-Maass, M. & Gorsuch, G. J. 2004. Developing reading fluency in EFL: How assisted repeated reading and extensive reading affect fluency development. Reading in a Foreign Language 16(2), 70–96. Available at:
http://www.nflrc.hawaii.edu/RFL/October2004/taguchi/taguchi.html.
Last retrieved: 30th September 2008.

Taillefer, G. F. 2005. Foreign language reading and study abroad: Cross-cultural and cross-linguistic questions. The Modern Language Journal 89(4), 503–528.

Tannen, D. 1985. Silence: Anything but. In D. Tannen & M. Saville-Troike (eds.) Perspectives on silence. Norwood: Ablex, 93–111.

Temple, L. 1992. Disfluencies in learner speech. Australian Review of Applied Linguistics 15(2), 29–44.

Temple, L. 2000. Second language learner speech production. Studia Linguistica 54(2), 288–297.

Toivanen, J. 1998. Suomalaisten puhuman englannin sävelkulkuilmiöistä: Millaista on välikielen prosodiikka? In A. Iivonen & T. Nevalainen (eds.) Vieraan kielen fonetiikan opetuksen näkökohtia. Helsinki: Hakapaino Oy, 127–141.

Toivanen, J. 1999. Perspectives on intonation: English, Finnish and English spoken by Finns. Siilinjärvi: Oulun yliopisto. Doctoral Dissertation.

Toivanen, J. 2003. Tone choice in the English intonation of proficient non-native speakers. Phonum 9, 165–168.

Toivanen, J. 2004. Pitch dynamism of English produced by proficient non-native speakers: Preliminary results of a corpus-based analysis of second language speech. Proceedings Fonetik 2004. Available at: http://www2.ling.su.se/fon/fonetik_2004/toivanen_fonetik2004.pdf. Last retrieved: 26th November 2009.

Toivanen, J. 2006a. The fall-rise intonation in the English of highly proficient non-native speakers. The Phonetics Symposium 2006. University of Helsinki, 161–168.

Toivanen, J. 2006b. Prosodic variation in Finnish-English speech: A corpus-based investigation of two sub-varieties of an L2 speech form. The Phonetics Symposium 2006. University of Helsinki, 169–179.

Toivanen, J. 2007. Fall-rise intonation usage in Finnish English second language discourse. Proceedings of Fonetik 2007 50(1), 85–88.

Toivanen, J. 2009. Context affects the prosody of second language speech: An exploratory study of Finnish English. In M. Vainio, R. Aulanko & O. Aaltonen (eds.) Nordic Prosody. Proceedings of the Xth Conference, Helsinki 2008. Frankfurt am Main: Peter Lang, 227–236.

Toivanen, J. & Waaramaa, T. 2005. Tone choice and voice quality of dispreferred turns in the English of finns. Logopedics Phoniatrics Vocology 30(3–4), 181–184.

Toivola, M., Lennes, M. & Aho, E. 2009a. Onko maahanmuuttajan vieraalla aksentilla väliä? In M. O'Dell & T. Nieminen (eds.) The Phonetics Symposium 2008. Tampere Studies in Language, Translation and Culture, Series B. Tampere: Tampere University Press. 73–79.

Toivola, M., Lennes, M. & Aho, E. 2009b. Speech rate and pauses in non-native Finnish. Proceedings of the 10th Interspeech 2009 Conference, Brighton, 1707–1710.

Toivonen, R. O. 2007. Intelligent Speech Analyser™ (ISA) Software. Available at: http://www.sci.fi/~pitchsys/. Last retrieved: 16th June 2008.

Tortel, A. 2009. Evaluation qualitative de la prosodie d'apprenants français: apport de paramétrisations prosodiques. University of Provence. Doctoral Dissertation.

Tortel, A. & Hirst, D. 2010. Rhythm metrics and the production of English L1/L2. In M. Hasegawa-Johnson, A. Bradlow, J. Cole, K. Livescu, J.

Pierrehumbert, Janet & C. Shih (eds.) Proceedings of the Speech Prosody 2010 conference, Chicago. Available at: http://aune.lpl.univ-aix.fr/~sprosig/sp2010/papers/100959.pdf Last retrieved: 27th November 2010.

Towell, R., Hawkins, R. & Bazergui, N. 1996. The development of fluency in advanced learners of French. Applied Linguistics 17(1), 84–119.

Trofimovich, P. & Baker, W. 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.

Trofimovich, P., Lightbown, P. M., Halter, R. H. & Song, H. Comprehension-based practice. The development of L2 pronunciation in a listening and reading program. Studies in Second Language Acquisition 31(4), 609–639.

Trollinger, V. L. 2003. Relationships between pitch-matching accuracy, speech fundamental frequency, speech range, age, and gender in American English-speaking preschool children. Journal of Research in Music Education 51(1), 78–94.

Trouvain, J. 2004. Tempo variation in speech production. Implications for speech synthesis. Saarbrücken: Institute of Phonetics, Saarland University Doctoral Dissertation. Available at http://www.coli.uni-saarland.de/groups/WB/Phonetics/structure.php?page=Research/PHONUS_research_reports/phonus.php Last retrieved: 30th March 2009.

Tsao, Y., Weismer, G. & Iqbal, K. 2006. Interspeaker variation in habitual speaking rate: Additional evidence. Journal of Speech, Language & Hearing Research 49(5), 1156–1164.

UCL 2008. Speech filing system. Available at: http://www.phon.ucl.ac.uk/resource/sfs/. Last retrieved: 16th June 2008.

Ullakonoja, R. 2007a. Obuchenie russkoy rechevoy kul'ture. Aktual'nye problemy prepodavaniya russkogo yazyka kak inostrannogo. An invited paper read at Novoe vremja – novoe videnie seminar, 23rd November 2008, Tver, Russia.

Ullakonoja, R. 2007b. Tauot sujuvuuden määrittäjinä suomalaisten venäjänoppijoiden puheessa. A paper read at ViKiPeda 2007 Conference, 22th May 2007, Helsinki.

Ullakonoja, R. 2008a. Puheen sujuvuuden havaitseminen ja itsearviointi. A paper read at AFinLA autumn symposium 2008, 15th November 2008, Helsinki.

Ullakonoja, R. 2008b. Read-aloud fluency in Russian as a foreign language: Acoustic features. A poster presented at the X Nordic Prosody, 5th August 2008, Helsinki.

Ullakonoja, R. 2008c. Vaihto-opiskelu ja vieraan kielen sujuvuuden kehittyminen. A paper read at XXXV Kielitieteen päivät, 24th May 2008, Vaasa.

Ullakonoja, R. 2009a. Fluency development in L2 during study abroad: Finnish students of Russian. University of Jyväskylä. Licentiate Thesis. Available

at: http://urn.fi/URN:NBN:fi:jyu-200905221613. Last retrieved: 27th December 2010.

Ullakonoja, R. 2009b. Praat scripts by Riikka. Available at: https://www.jyu.fi/Members/riikkau/praat. Last retrieved: 30th March 2009.

Ullakonoja, R. 2010, Prosodic correlates of reading-aloud fluency in L2 Russian. A poster presented at Eurosla 2010 Conference, 3rd September 2010, Reggio Emilia, Italy.

Ullakonoja, R., Kärkkäinen, H. & de Silva, V. 2007. Havaintoja venäjän kielen lukupuhunnan sävelkulun ja spontaanin puheen jaksottelun oppimisesta. In O. Salo, T. Nikula & P. Kalaja (eds.) Language in Learning – AFinLA Yearbook 2007. Jyväskylä: Suomen soveltavan kielitieteen yhdistys AFinLA, 215–231.

Ultan, R. 1978. Some general characteristics of interrogative systems. In J. H. Greenberg, C. A. Ferguson & E. A. Moravcsik (eds.) Universals of human language. Volume 4: Syntax. Stanford: Stanford University Press, 211–248.

Vaissière, J. 1983. Language-independent prosodic features. In A. Cutler & D. R. Ladd (eds.) Prosody: Models and measurements. Berlin: Springer, 53–66.

Valo, M. 1994. Käsitykset ja vaikutelmat äänestä: Kuuntelijoiden arviointia radiopuheen äänellisistä ominaisuuksista. Jyväskylä: University of Jyväskylä. Doctoral Dissertation.

Valls-Ferrer, M. 2008. Language acquisition during a stay abroad period following formal instruction. Temporal effects on oral fluency development. Department of translation and language sciences. Pompeu Fabra University. Doctoral Dissertation. Available at: http://www.recercat.net/bitstream/2072/43320/1/Margalida+Valls+Ferrer_TA.pdf Last retrieved: 27th November 2010.

van Heuven, V. J. & van Zanten, E. 2005. Speech rate as a secondary prosodic characteristic of polarity questions in three languages. Speech Communication 47(1–2), 87–99.

Varyushenkova, N. N. & Lyubimova, N. A. 1986. Fonetiko-graficheskaya interferentsiya i ee proyavlenie v rechi na nerodnom yazyke. Russkiy yazyk za rubezhom 1(99), 88–90.

VISK = Hakulinen, A., Vilkuna, M., Korhonen, R., Koivisto, V. & Alho, I. 2004: Iso suomen kielioppi. Helsinki: Suomalaisen Kirjallisuuden Seura. Available at: http://scripta.kotus.fi/visk. Last retrieved: 17th October 2008.

Volskaya, N. B. 2001. Osobennosti intonacii i sintagmaticheskogo chlenenya voprositel'nyh vyskazyvaniy raznoy dliny. In L. V. Bondarko (ed.) 100 let eksperimental'noy fonetike v Rossii. Materialy mezhdunarodnoy konferencii 1–4 fevralya 2001. Sankt-Peterburg: Filologicheskiy fakul'tet SPbGU, 54–57.

Volskaya, N. B. 2002. O pauzah virtualnyh i realnyh. In N. B. Volskaya & N. D. Svetozarova (eds.) Problemy i metody eksperimental'no-foneticheskih issledovaniy. K 70-letiyu professora kafedry fonetiki i metodiki

prepodavaniya inostrannyh yazykov L. V. Bondarko. Sankt-Peterburg: Filologicheskiy fakul'tet Sankt-Peterburgskogo gosudarstvennogo universiteta, 165–170.

Volskaya, N. B. 2003. Virtual and real pauses at clause and sentence boundaries. In M. J. Solé, D. Recasens & J. Romero (eds.) Proceedings of the 15th ICPhS, Barcelona, 3–9 August 2003, 499–502.

Volskaya, N. B. 2004. O pauze i ne tol'ko o nej. In L. V. Bondarko (ed.) Foneticheskie chteniya v chest' 100-letiya so dnya rozhdeniya L. R. Zindera. Sankt-Peterburg: Filologicheskiy fakul'tet SPbGU, 129–136.

Volskaya, N. B. 2007. Variativnost' intonacionnykh modeley v spontannoy rechi i chtenii: Intonacionnyy center. Materialy XXXVI mezhdunarodnoy filologicheskoy konferencii, 12–27 marta. Vypusk 17. Fonetika, 16–25.

Volskaya, N. B. 2009a. Aspects of Russian intonation. In V. de Silva & R. Ullakonoja (eds.) Phonetics of Russian and Finnish, general introduction, spontaneous and read-aloud speech. Frankfurt am Main: Peter Lang, 37–46.

Volskaya, N. B. 2009b. Prosodic features of Russian spontaneous and read-aloud speech. In V. de Silva & R. Ullakonoja (eds.) Phonetics of Russian and Finnish, general introduction, spontaneous and read-aloud speech. Frankfurt am Main: Peter Lang, 133–143.

Volskaya, N. B. 2009c. Personal communication. 9th October 2009.

Volskaya, N. B. 2010. Personal communication. 17th November 2010.

Volskaya, N. B. & Skrelin, P. 2001. The utterance type and the size of the intonation unit. Text, speech and dialogue: Lecture notes in computer science. Berlin: Springer, 285–290. Available at: http://www.springerlink.com/content/vqr0ymck090ug95x/. Last retrieved: 5th June 2010.

Välimaa–Blum, R. 1993. Intonation: A distinctive parameter in grammatical constructions. Phonetica 50, 124–137.

Walsh, R. 1994. The year abroad – A linguistic challenge. Teanga: The Irish Yearbook of Applied Linguistics 14, 48–57.

Weber, R. 1991. Linguistic diversity and reading in American society. In R. Barr, M. L. Kamil, P. B. Mosenthal & D. Pearson (eds.) Handbook of reading research, Volume II. New York: Longman, 97–119.

Wells, J. 2006. English intonation: an introduction. United Kingdom: Cambridge University Press.

Weltens, B. & Bot, K. D. 1984. The visualisation of pitch contours: Some aspects of its effectiveness in teaching foreign intonation. Speech Communication 3(2), 157–163.

Wenk, B. J. 1985. Speech rhythms in second language acquisition. Language and Speech 28(2), 157–175.

Wennerstrom, A. 2000. The role of intonation in second language fluency. In H. Riggenbach & R. Schmidt (eds.) Perspectives on fluency. Ann Arbor: University of Michigan Press, 102–127.

Wennerstrom, A. 2001. Music of everyday speech: Prosody and discourse analysis. Cary: Oxford University Press.

Werner, S. 2001. Itseorganisoituvat kartat intonaatiotutkimuksen apuvälineenä. Puhe ja kieli 21(4), 152–162.

Wilkinson, S. 1998a. On the nature of immersion during study abroad: Some participant perspectives. Frontiers: The Interdisciplinary Journal of Study Abroad 4, 121–138.

Wilkinson, S. 1998b. Study abroad from the participants' perspective: A challenge to common beliefs. Foreign Language Annals 31(1), 23–39.

Wiik, K. 1988. F0:n huipun sijainti suomessa. In M. Karjalainen & U. K. Laine (eds.) Fonetiikan päivät – Espoo. Papers from the 15th meeting of Finnish phoneticians. Otaniemi: Teknillinen korkeakoulu, 215–229.

Yanko, T. E. 2008. Intonacionnye strategii russkoy rechi v sopostavitel'nom aspekte. Moskva: Yazyki slavyanskih kul'tur.

Yli-Luukko, E, 2001. Yläsatakuntalainen intonaatiokuvio. Virittäjä 105, 2–21.

Yuen, I. 2007. Declination and tone perception in Cantonese. In T. Riad & C. Gussenhoven (eds.) Tones and Tunes: Experimental studies in word and sentence prosody. Berlin: Mouton de Gruyter, 63–77.

Ylinen, S., Huotilainen, M. & Näätänen, R. 2005a. Phoneme quality and quantity are processed independently in the human brain. Neuroreport: Lippincott Williams & Wilkins 16(16), 1857–1860.

Ylinen, S., Shestakova, A., Alku, P. & Huotilainen, M. 2005b. The perception of phonological quantity based on durational cues by native speakers, second-language users and nonspeakers of Finnish. Language and Speech 48(3), 313–338.

Ylinen, S., Shestakova, A., Huotilainen, M., Alku, P. & Näätänen, R. 2006. Mismatch negativity (MMN) elicited by changes in phoneme length: A cross-linguistic study. Brain Research 1072(1), 175–185.

Ylitalo, R. 2009. The realisation of prominence in three varieties of standard spoken Finnish. Acta Universitatis Ouluensis B Humaniora 88. Oulu: University of Oulu. Doctoral dissertation. Available at: http://herkules.oulu.fi/isbn9789514291142/. Last retrieved: 7th November 2010.

Yokoyama, O. 2001. A reinterpretation of the IK system in Russian. Die Welt Der Slaven XLVI, 1–26.

Zampini, M. L. 2008. L2 speech production research – findings, issues, and advances. In J. G. Hansen Edwards & M. L. Zampini (eds.) Phonology and second language acquisition. Amsterdam: John Benjamins, 219–249.

Zinder, L. R. 1979. Obshaya fonetika. Moskva: Vysshaya Shkola.

Zinder, L. R. & Bondarko, L. V. 1980. Problemy i metody eksperimental'no-foneticheskogo analiza rechi. Leningrad: Izdatel'stvo Leningradskogo universiteta.

Zlatoustova, L. V., Potapova, R. K. & Trunin-Donskoy, V. N. 1986. Obshchaya i prikladnaya fonetika. Moskva: Izdatel'stvo Moskovskogo Universiteta.

ORIGINAL PAPERS

I

PAUSING AS AN INDICATOR OF FLUENCY IN THE RUSSIAN OF FINNISH LEARNERS

by

Riikka Ullakonoja 2008

Proceedings of the Speech Prosody 2008 Conference, Campinas, Brazil. Barbosa, Plínio, Madureira, Sandra & Reis, César (eds.). São Paolo: Editora RG/CNPq, 339–342.

Pausing as an Indicator of Fluency in the Russian of Finnish Learners

Riikka Ullakonoja

Department of Languages
University of Jyväskylä, Finland
riikka.ullakonoja@campus.jyu.fi

Abstract

Previous research shows that pausing and disfluencies are common in non-native speech. The aim of this study was to investigate the relationship between fluency and pausing in Russian read-aloud speech of 12 Finnish university students and examine their fluency development during a 3.5-month study-period in Russia. To assess fluency, 30 Russian teachers rated the students' speech samples (on a 1–5 scale). The samples were then analysed perceptually and acoustically for pause frequency, duration and placement. Results show that pausing can be an indicator of foreign language fluency and that most students develop considerably in their Russian read-aloud fluency during their stay in Russia. Hence, when teaching students to read aloud in a foreign language, pausing should be emphasized as a way to become a fluent reader.

1. Introduction

Fluency is often mentioned as an aim of foreign language (FL) teaching. It has also been shown in few studies [5], [13] that when FL learners spend some time in the country where the target language is spoken, their speech becomes more fluent. Fluency has been defined in many ways e.g. by the number of pauses, their place and duration; speech rate, rhythm and hesitation [3], [11], [14]. The features of speech that make it fluent are situation and text dependent, and hence, speech with few pauses is not necessarily always perceived as fluent [7], [8]. In this study, fluency is used to refer to the fast, smooth reading aloud. As pause frequency and speech rate have been found to be the most important temporal correlates for read-aloud speech fluency perception [3], pausing is investigated here and speech rate will be discussed in a parallel study [15]. This is a follow-up study that concentrates on learner's speech production, which is not a very common approach in the field of FL prosody.

As shown by a number of previous studies (see e.g. [10], [2], [9]) extensive pausing is typical for non-native speech. Pauses occur together with hesitation, repetition or repair. According to Riggenbach [10] the "chunking together" of disfluencies (several disfluencies in a three word sequence) can be an important indicator of fluency. Pause duration is affected e.g. by the sentence length and pause placement [4].

The purpose of the study was to find out whether speakers are thought to be more fluent in their FL if they have a more native-like pause duration and placement. This article concentrates on the place, duration and frequency of pauses in the learner's speech. The main hypotheses were 1) learners' fluency improves during study abroad experience 2) learners with less pauses and/or shorter pauses are rated to be more fluent in Russian.

2. Material

The 12 subjects were 19–24 year-old female undergraduate major students of Russian. They were native Finnish speakers who reported having no hearing or speaking disabilities. Most of them had studied Russian as their 3rd or 4th FL (in Finland it is common to study 3–4 FLs). Half of the students stayed with a Russian host family during their stay in Russia (all of them participated in the same study abroad program) whereas the rest resided in foreign-student dormitories. A student moved from the host family to the dormitories in the middle of her stay. Each student was recorded three times reading the same dialogue with another student: before, during and after the 3.5-month-stay in Russia. Only the longest (and a difficult) turn of the dialogue (6 sentences) was chosen for the analysis. The total duration of analysed read-aloud speech was c. 12 mins. Students' speaking activity with native Russians and fluency self-evaluation was determined with the help of questionnaires.

3. Methods

The pauses were segmented in Praat [1] according to the auditory analysis. The perceived pauses were labelled as fluent (juncture) or disfluent (non-juncture) pauses [10], [6]. Pauses occurring at the sentence or phrasal boundary were fluent, whereas others were often disfluent sounding. The traditional classification of silent and filled pauses was not respected here because the latter were scarce in the material and because it was not considered useful in measuring fluency. The common minimum pause duration of 200 ms. was not used either. The pause duration was automatically measured in textgrids with a script. The quantitative analysis and graphical representation of the results was conducted in Excel and the statistical analysis in SPSS. Students' speech was compared with each others in different recording sessions and with the fluency rating each sample received in the fluency evaluation task.

Expert judges, 30 Finnish teachers of Russian as a FL, rated the fluency of the students' speech samples by perception. Teachers were from different age groups and had different amounts of experience in teaching Russian as a FL. They heard the stimuli (n = 36, each student in each recording session) once in a randomized order and rated the fluency of each sample on a 1–5 scale (1 = very disfluent, 5 = very fluent). Most teachers participated in the experiment by filling out a web-based questionnaire and listening to the sound file on their PC. The rest did the evaluation in a language lab. Teachers were also asked to give a definition of fluency and, after listening, determine the factors hindering fluency.

Thus, each sample received an average fluency measure, which was later compared to the acoustic analysis. The

interjudge reliability was evaluated by determining the reliability coefficient (the value of Cronbach's alpha) which yielded 0.92. Hence, the reliability of the ratings was high and most judges had a similar idea of what is fluent speech. The average rating for all the judges and all the speakers was 3.17 (std = 1.05) which indicates that the judges used more or less the whole scale in their fluency evaluations.

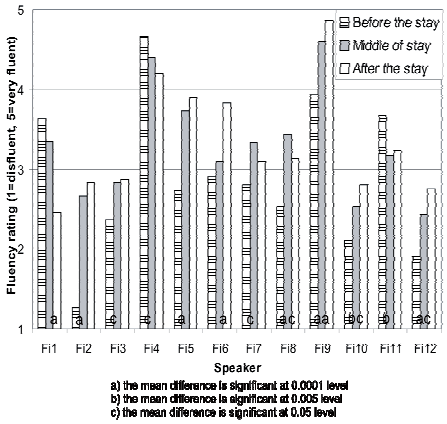
4. Results

4.1. Fluency perception

In the open questions prior to listening, the teachers defined fluent reading in a FL as speech that has a native-like pronunciation of segments, intonation, word stress and short pauses at correct places (over 10 mentions each). In addition, after listening they mentioned that monotonous speech and faltering made the samples sound disfluent.

As Figure 1 shows, the learners' fluency develops during their stay in Russia. 9/12 learners received a lower fluency rating before their stay in Russia than in the middle of it and 7 of them even improved their rating at the recording after their stay. 9/12 learners had a better fluency rating following their stay in Russia than prior to it.

Figure 1: *Fluency of the speakers at different stages of learning*



Most (16/24) mean differences were statistically significant at least at the 0.05 level (Figure 1). This means that 8/12 learners improved their fluency significantly by the middle of their stay and 3 of them even improved their fluency significantly after that. When comparing only the fluency ratings before the stay and after it, it was found that the majority (8/12) of the learners received a statistically significantly better fluency rating after their stay than before it ($p < 0.005$ for all).

4.2. Students' self-evaluation and exposure to Russian

When asking the subjects following their stay in Russia whether they could speak and read Russian more fluently now than before their stay, all responded affirmatively. Half of them (6/12) said that their pronunciation had developed noticeably. Some (5/12) said that they still had trouble producing the intonation in the way they wished. The students had different amounts of contact with Russians during their stay. Half of the students stayed with a host family where naturally they had possibilities to practice oral skills. The majority (11/12) of the students also spoke at least a little with their teachers outside the classroom. Four students said that they did not know any Russians they could talk to in the town in which they were staying. Only 4 students said that they tried actively to get in contact with native speakers. The students who lived with a host family did not get significantly better fluency ratings than those residing in the dormitories. In fact, students living in the dormitories were more fluent in each recording session and they improved as much as those living with a host family.

4.3. Pausing

4.3.1. Pause frequency

Firstly, the frequency distribution of the two pause types (fluent and disfluent pauses) was studied. The total number of pauses varied, because sometimes the speakers did not pause e.g. at the phrase boundary (as might traditionally be expected) but indicated the boundary by other prosodic means. Individual differences in pause frequency were found, but on average, the frequency of the fluent pauses remained the same and the frequency of the disfluent pauses decreased as the amount of experience increased (Table 1). 7/12 speakers had less disfluent pauses in the middle of their stay than before it. 8/12 speakers had less disfluent pauses after their stay than in the middle of it. The majority of the learners (9/12) had less disfluent pauses following the stay than prior to it. The distribution of fluent and disfluent pauses in different stages of stay did not differ statistically significantly between the speakers (Pearson's Chi-Square for fluent pauses χ^2 (22) = 2.358, p = 1.00, for disfluent pauses χ^2 (22) = 13.901, p = 0.905).

Table 1: *Frequency of different pause types (fl. = fluent pauses, disfl. = disfluent pauses).*

Speaker	Before the stay		Middle of the stay		After the stay	
	fl.	disfl.	fl.	disfl.	fl.	disfl.
Fi1	11	3	10	2	13	2
Fi2	11	6	9	3	8	1
Fi3	11	3	11	4	12	1
Fi4	9	1	8	2	8	1
Fi5	8	1	7	0	7	0
Fi6	12	0	12	2	12	1
Fi7	8	1	8	1	9	0
Fi8	10	4	10	1	10	2
Fi9	9	1	6	1	7	0
Fi10	10	6	11	4	12	2
Fi11	9	1	11	0	10	2
Fi12	12	6	12	5	10	2
Mean	10.0	2.8	9.6	2.1	9.8	1.2

There was a relationship between the frequency of different pause types and fluency ratings. A statistically significant negative linear correlation was found between the

mean fluency rating and frequency of fluent pauses (Pearson's Correlation = -0.586, $p < 0.001$). The correlation existed also between the mean fluency rating and the frequency of disfluent pauses (Spearman's Correlation = -0.657, $p < 0.001$) and between the mean fluency rating and the total frequency of pauses (Pearson's Correlation = -0.742, $p < 0.001$).

4.3.2. Pause duration

Secondly, pause duration (absolute and relative durations) was measured in the two pause types and compared to fluency ratings. Absolute durations of disfluent pauses were in average shorter than fluent pauses (Table 2). A correlation was found between the mean absolute durations of different pause types (Pearson's correlation = 0.426, $p < 0.05$).

The relative durations were calculated by proportioning the duration of each pause with the total duration of the sample. Thus, the number indicates the percentage of pausing in total utterance duration and allows the interspeaker comparison (Table 2). The majority of the speakers have the smallest relative duration of fluent pauses (9/12 speakers) and disfluent pauses (8/12 speakers) in the middle of the stay.

Table 2: *Mean duration of different pause types (fl. = fluent pauses, disfl. = disfluent pauses): absolute duration in ms. and relative duration in % of the utterance duration.*

Speaker	Before the stay		Middle of the stay		After the stay	
	fl.	disfl.	fl.	disfl.	fl.	disfl.
Fi1 (ms.)	619	364	442	110	467	432
Fi1 (%)	27	16	24	6	21	19
Fi2 (ms.)	416	888	273	118	374	225
Fi2 (%)	16	33	16	7	23	14
Fi3 (ms.)	457	347	318	156	335	192
Fi3 (%)	23	18	18	9	20	11
Fi4 (ms.)	240	211	209	234	296	95
Fi4 (%)	15	13	14	16	20	6
Fi5 (ms.)	420	323	454		453	
Fi5 (%)	23	18	28		30	
Fi6 (ms.)	400		343	138	283	165
Fi6 (%)	21		19	8	17	10
Fi7 (ms.)	346	577	290	113	335	
Fi7 (%)	20	34	18	7	22	
Fi8 (ms.)	333	123	211	161	313	566
Fi8 (%)	18	7	13	10	16	29
Fi9 (ms.)	261	133	292	73	240	
Fi9 (%)	17	9	21	5	17	
Fi10 (ms.)	443	527	355	351	408	183
Fi10 (%)	17	20	19	18	21	9
Fi11 (ms.)	426	455	291		409	677
Fi11 (%)	20	21	15		17	28
Fi12 (ms.)	405	231	318	352	342	266
Fi12 (%)	19	11	16	17	17	14
Mean(ms.)	402	430	317	218	358	352
Mean (%)	20	18	18	12	20	17

When the mean pause durations were compared to the fluency ratings, it was found that the most fluent speakers (Fi4 and Fi9) had a fairly short mean relative disfluent pause duration. Mean absolute durations of both fluent and disfluent pauses indicated significant negative correlations with the fluency rating (for fluent pauses Pearson's Correlation = -0.393, $p < 0.05$; for disfluent pauses Pearson's Correlation = -0.478, $p < 0.01$). Mean relative durations of disfluent pauses showed a similar relationship (Pearson's Correlation = -0.372, $p < 0.05$), but the fluent pauses did not (Pearson's Correlation = 0.072).

4.3.3. Pause placement

Fluent pauses occurred at phrasal and sentence boundaries whereas disfluent pauses were situated in the middle of the word (when there was hesitation, repetition or repair), in the middle of the noun phrase, or between the verb and its complement.

Table 3: *Frequency of disfluent pauses at most common places of the utterance (for all speakers).*

Phrase	Before the stay	Middle of the stay	After the stay	Total
A. ona uyezzhaet (pause) ni segodnya ...	4	6	0	10
B. yesli khochesh (pause) eyë provodit…	3	1	3	7
C. chasov (pause) v devyat	3	4	1	8

Disfluent pause placement was very much speaker dependent, however there were three places that were common (more than three occurrences) for disfluent pauses (Table 3). It is interesting that in phrases A and C there was considerably fewer disfluent pauses after the stay than before or middle of it. Perhaps this indicates that students had (either through experience or repetition of the same text) learnt not to pause in the middle of these constructions. Overall, there were repairs or repetitions in the speech of 3–4 subjects before the stay and in the middle of it. After the stay however, 7/12 students used repairs. It was found that before the stay it was the 3 least fluent subjects (Fi2, Fi12 and Fi10), in the middle of the stay the two least fluent (Fi12 and Fi10) and after the stay the three least fluent (Fi1, Fi12 and Fi10) that had "disfluency clusters" (several disfluencies in a three word sequence).

5. Discussion and Conclusions

As previous studies [5], [13] have shown and as it was hypothesized in this study, the learners' fluency improves during their study abroad experience. As the amount of experience increases, the fluency also improves. There was no systematic development in the way, as Freed [5] has found that weaker students would develop in their fluency more significantly than better ones. Certainly students who were already quite fluent prior to their stay in Russia (Fi4 and Fi9) could not improve as much as the weaker students on this scale, which evaluated all students' fluency. The student who improved her fluency the most was a student (Fi2) who received a very low rating before her stay. Some students (Fi4 and Fi11) achieved lower fluency ratings following their stay than prior to it. The explanations for this decline can be that they have become more conscious of their pronunciation, and hence, are trying to self-correct more, which causes more

repairs and disfluent pauses (after the stay more students used repairs in their speech than before the stay). The other explanation for fluency decline can also be the limited duration of the speech samples. The students may have spoken more fluently in general, but by chance had more disfluencies in this particular sample. The finding that students residing with a host family did not improve their fluency more/were not more fluent than the group living in the dormitories, is consistent with another study [12].

The other hypothesis was that FL speakers using less and/or shorter pauses are rated to be more fluent in Russian. The study showed that the speakers' fluency developed during their study abroad experience, hence they used less disfluent pauses after their stay. Speakers' pause frequency distributions were in fact rather similar, which could have been predicted due to the fact that the subjects were reading the same text. Therefore, speech with multiple pauses was perceived as less fluent than speech with few pauses. Particularly the high number of disfluent pauses (that often occurred together with repairs, repetitions and other hesitation phenomena) created a less fluent impression. Interestingly though, there were 5 samples with no disfluent pauses that did not, however, receive a very high fluency rating (2.9–3.9). Therefore, it cannot be said that speech with no disfluent pauses would always be perceived as very fluent. This indicates that the pause frequency is not the only feature contributing to the perception of speech as fluent.

There was individual variation in pausing (see also e.g. [4]). When comparing the duration results to native speakers, whose mean pause duration was in Volskaya's study [16] 173.5 ms. (range 153–188 ms.), we can see that students' pauses are longer, perhaps because of their slower speech rate. If learners' fluent pauses are short, disfluent pauses tend to be short also and vice versa. It should be noted that even very short disfluent pauses were easily detected in the auditory analysis because they caused interruption of the speech flow (e.g. in the middle of the sentence) whereas very short fluent pauses may go unnoticed. The majority of the speakers had the smallest relative pause duration in the middle of the stay. This may be due to e.g. a faster speech rate, which they have become used to using in Russia. Furthermore, it was found that the more fluent the speaker, the shorter her disfluent pause duration is (both in absolute and relative values).

For pause placement, it can be concluded that it is indeed the "disfluency clusters" (as also Riggenbach [10] has shown) that give an impression of disfluency. This was proven because in each recording session at least the two least fluent subjects had the most "disfluency clusters".

The study can be criticised for only having the author (a non-native speaker) to conduct the perceptual pause detection. The perceptual analysis however, was verified acoustically. The perceptual pause detection and acoustic analysis were completed prior to the fluency ratings and therefore could not affect the perceptual pause classification.

The implications of this study to FL learning are that firstly, we should encourage our students to spend some time in the country where the target language is spoken. Secondly, in teaching more attention ought to be paid to pause placement in order to improve fluency. When students are reading a text aloud they are often focusing on pronunciation and could simultaneously be developing their pausing skills.

In conclusion, this study has shown that fluency improves during the study abroad experience and that pausing is an indicator of fluency. Further research should consider other prosodic factors, e.g. speech rate and intonation, which potentially influence the fluency evaluations.

6. References

[1] Boersma, P., Weenik, D. 2007. Praat 4.6. [computer software] Downloadable at www.praat.org.

[2] Cenoz, J. 2000. Pauses and hesitation phenomena in second language production. *ITL: Review of Applied Linguistics* (127–128), 53–69.

[3] Cucchiarini, C., Strik, H., Boves, L. 2002. Quantitative assessment of second language learners' fluency: comparisons between read and spontaneous speech. *The Journal of the Acoustical Society of America* 6 (111), 2862–2873.

[4] Fant G., Kruckenberg A., Barbosa-Ferreira J. 2003.Individual variations in pausing, a study of read speech. *Proc. of the Swedish Phonetics meeting in Umeå, Phonum 2003*, 193–196.

[5] Freed, B.F. 1995. What makes us think that students who study abroad become fluent? In *Second Language Acquisition in a Study Abroad Context,* B.F. Freed & C.A. Ferguson (ed.). Amsterdam: Benjamins, 123–148.

[6] Kenny, K.D. 1996. *Language loss and the crisis of cognition: Between socio- and psycholinguistics.* Berlin: Mouton de Gruyter.

[7] Lehtonen, J. 1978. On the problems of measuring fluency. In *AFinLA:n vuosikirja 1978,* M. Leiwo & A. Räsänen (ed.). Jyväskylä: Jyväskylän yliopisto. Englannin laitos, 53–68.

[8] Lennon, P. 2000. The lexical element in spoken second language fluency. In *Perspectives on Fluency,* H. Riggenbach & R. Schmidt (ed.). Michigan: Ann Arbor, 25–42.

[9] Paananen-Porkka, M.M. 2007. *Speech rhythm in an interlanguage perspective: Finnish adolescents speaking English.* Helsinki: University of Helsinki.

[10] Riggenbach, H. 1991. Toward an Understanding of Fluency: A Microanalysis of Nonnative Speaker Conversations. *Discourse Processes* 4 (14), 423–441.

[11] Riggenbach, H. 2001. Hesitation phenomena in second-language fluency. In *Music of everyday speech: prosody and discourse analysis,* A. Wennerstrom (ed.).Cary: Oxford University Press, 252–257.

[12] Segalowitz, N., Freed, B.F. 2004. Context, contact, and cognition in oral fluency acquisition: learning Spanish in at home and study abroad contexts. *Studies in Second Language Acquisition* 2 (26), 173–199.

[13] Towell, R., Hawkins, R., Bazergui, N. 1996. The development of fluency in advanced learners of French. *Applied Linguistics* 1 (17), 84–119.

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

[15] Ullakonoja R. (forthcoming). Speech rate as an indicator of fluency in the Russian of Finnish learners. *Paper read at the Finnish Phonetics Symposium 2008, 12th Jan 08.*

[16] Volskaya, N.B. 2004. O pauze i ne tol'ko o ney. In *Foneticheskye chitenya v chest' 100-letya so dnya rozhdenya L.R. Zindera,* L.V. Bondarko (ed.). St. Petersburg: Filologicheskiy fakul'tet SPbGU, 129–136.

II

SPEECH RATE AS AN INDICATOR OF FLUENCY IN THE RUSSIAN OF FINNISH LEARNERS

by

Riikka Ullakonoja 2009

Fonetiikan päivät 2008 – The Phonetics Symposium 2008. Tampere Studies in Language, Translation and Culture, Series B. O'Dell, Michael & Nieminen, Tommi (eds.), 97–109.

Michael O'Dell & Tommi Nieminen (eds.), *Fonetiikan päivät 2008*.
Tampere Studies in Language, Translation and Culture, Series B 3.
Tampere: Tampere University Press, 97–109.

Speech rate as an indicator of fluency in the Russian of Finnish learners

Riikka Ullakonoja
University of Jyväskylä

Abstract

This study focuses on the speech rate development of 12 Finnish university students of Russian during their 3.5-month-study abroad experience. Speech and articulation rates are measured in phonetic words per second and syllables per second in the Russian read-aloud speech of the subjects. This is done at three recordings: prior to, during and following their stay in Russia. The results are compared to their read-aloud Finnish speech. The students are also compared depending on the residence (host-family vs. dormitories) in Russia. The study shows that speech and articulation rates correlate with the evaluated fluency of the speech samples. It was found that speech rate is a better indicator of fluency than articulation rate in non-native read-aloud speech. The results also show that articulation rate in mother tongue (Finnish) and foreign language (Russian) correlate with each other more than speech rate.

Keywords: speech rate, fluency, Finnish (L1), Russian (L2)

1 Introduction

When asking foreign language learners what aspects they consider important in learning the new language, their answers might include a desire to become fluent in that language. Also in the words of their teacher, in the syllabus and in also the Common European framework of reference for languages (Council for Cultural Co-operation. Education Committee, Modern Languages Division, Strasbourg and Council of Europe 2001) the term fluency and its derivations occur frequently. However, when teaching oral skills, it is perhaps not the fluent features of speech that are in the focus of attention, but instead the grammatical and lexical features or the pronunciation of segments. The purpose of the study is to follow the fluency development of 12 Finnish students of Russian during their 3.5-month-stay in Russia by studying their speech and articulation rates and comparing them to fluency evaluations of teachers.

Fluency can be defined in a number of ways, e.g. by studying pausing (pause frequency, duration and placement), hesitations or tempo (se e.g. Cucchiarini *et al.* 2002,

Lauranto 2005, for a review). In this study speech rate is regarded as an important factor of fluency. Cucchiarini *et al.* (2002) have shown that speech rate and pause frequency are the most important factors in read-aloud speech fluency perception. Also Riggenbach (1991) concluded that the central elements of foreign language (L2) fluency are pausing, speech rate and repairs. Moreover, several researchers (Riggenbach 1991, Freed 1995, Towell *et al.* 1996) have found that as L2 fluency increases, the speech rate increases also. My previous study (Ullakonoja 2008) focused on pausing and its relationship to foreign language fluency. In this paper, the same data is studied, but speech and articulation rates are regarded as acoustic correlates of fluency.

The speech rate (tempo) indicates the total time of a speaker uttering his speech, including pauses whereas the term articulation rate is commonly used to refer to the speech rate without pauses. In this study speech rate refers to reading rate. There are multiple factors affecting the habitual speech rate of individual speakers, and speakers can also vary their speaking rate in different situations (see Trouvain 2003 for a review). In this study the speaking context and content are the same for all speakers at all recording sessions. The speech and articulation rates of a L2 learner are often shown to be slower than these of a native speaker (e.g. Riggenbach 1991, Cenoz 2000, Paananen-Porkka 2007). In addition, learners possibly transfer the prosodic characteristics (e.g. stress) of their mother tongue to the language they are learning:

> When the Finn transfers the habit of pronouncing all of the syllables of each word unreduced and manifesting word boundaries with phonetical juncture segments (instead of linking) the rate of his speech is inevitably slower (Lehtonen 1981, p. 331).

A foreign language learner often has the impression that native speakers of the language speak very fast (Abercrombie 1967, p. 96). Also, when native speakers are listening to L2 speech, they would often prefer about 10 % faster speech rate than what the learner is producing (Munro & Derwing 2001, p. 464).

It has been found in several studies (Simoes 1996, Freed *et al.* 2004, Lafford 2004, Trofimovich & Baker 2006) that a good way to improve fluency in L2 is to spend some time in the country where L2 is spoken. For example Segalowitz & Freed (2004) established that the students who studied abroad improved their fluency more (on several measures including speech rate) than the students who stayed at home. Trofimovich & Baker (2006) found that L2 learners could not achieve a native speech rate no matter how long they stayed in the country of the L2 language. On the contrary, a study by Freed *et al.* (2004) suggests that the study abroad did not result in better fluency than an "intensive domestic immersion" context. In their study it was in fact the immersion context that turned out to be the most effective in fluency learning. To summarize, all the studies show the positive influence of L2 context to the fluency development.

There have been a few studies (e.g. Lehtonen 1979, Iivonen *et al.* 1995, Moore & Korpijaakko-Huuhka 1996, Suomi 2007) about speech rate in native Finnish speech. In Russian, pausing and its influence on prosodic phrasing and speech rate have been researched also in spontaneous speech (e.g. Shtern 1988, Volskaya forthcoming). To my knowledge the current paper is the first study investigating non-native speech rate in Russian and comparing it to the speakers' native language, Finnish, and contrasting different stages of learning. The aim of this study was to find out, firstly, whether speakers who are considered fluent speak/read aloud faster than disfluent speakers (both in terms of speech and articulation rates). In other words, speakers with faster speech or/and articulation rates are evaluated more fluent than slower speakers. Secondly, the speech and articulation rates in Finnish (mother tongue, L1) were compared to speech and articulation rate in Russian (L2) to find any similarities between the two.

2 Material

12 native Finnish students of Russian read two Russian and one Finnish dialogue in pairs. The reading was recorded in different stages of their university studies: prior to, in the middle of and following their stay in Russia. Only the longest turn of the Russian dialogues and two turns of the Finnish dialogue were analyzed of each student. The Russian material, hence, includes the reading of the same text three times (c. 11 minutes in total), whereas the Finnish material is from the first recording session (c. 3 minutes in total). The students are undergraduate major students of Russian who have studied Russian for 1–10 years prior to university studies. At the beginning of their 2nd year of university studies they participated in a 3.5-month-study-abroad-program. Half of the students (subjects Fi3, Fi4, Fi5, Fi7, Fi9 and Fi10) resided in the dormitories for foreign students during their stay in Russia with the remaining (subjects Fi1, Fi2, Fi6, Fi8, Fi11 and Fi12) living with a host family. The two groups were compared for speech and articulation rates development where applicable.

3 Methods

For evaluating the perceptual fluency of the speech samples, 30 Russian as a foreign language teachers in Finland were asked to determine the fluency of each sample on 1–5 scale (1 = not fluent, 5 = very fluent). Teachers listened to the samples in a random order without knowing that multiple samples of the same speaker were included. The reliability of the fluency ratings was good (Cronbach's alpha = 0.92). The procedure of the fluency evaluation task is more thoroughly reported in a parallel study (Ullakonoja 2008).

Segmentation and acoustic analysis of the samples were completed in Praat (Boersma & Weenink 2008). The segmentation consisted of annotation of phonetic words and syllables. The term 'phonetic word' comes from the Russian research tradition (e.g. Avanesov 1956, p. 61), and usually corresponds to a lexical word, but also to some two word combinations, where e.g. a preposition is pronounced together with the main word and where there is only one lexical stress. For example, in this data the preposition and pronoun *k nam* [knɑm] ('to us') are treated as a phonetic word. The term prosodic word has sometimes been used to describe the same phenomena in Finnish (see e.g. Aho & Yli-Luukko 2005). In Finnish, I decided that lexical words always correspond phonetic words in the annotation. The syllables were determined according to auditory analysis, hence the syllable means a realized syllable. Syllable nuclei were determined and proportioned with time (counting syllable nuclei instead of syllables has been used e.g. by Simoes 1996). In Russian the number of syllables corresponded the number of vowels in the utterance. In Finnish, single vowels were treated similarly as in Russian, as a syllable nucleus. Vowels in the vowel combinations in Finnish were mostly pronounced very closely together and consequently, they were also regarded as one syllable. Sometimes the syllabification in Finnish did not respect the traditional (or textual) syllabification, if e.g. the word *teorioita* ('theories (partitive case)') was pronounced [teorioɪɑ], it was considered trisyllabic: teo-rio-ta (speaker Fi7). Similarly also the phrase *mä en oo* ('I'm not') was pronounced mostly as [mæeno], [mæeo] or [menːoː] and in all cases it only had two syllables. Syllable omission was quite frequent in Finnish, e.g. *no en* [non] ('well no', Fi7), *huomenna* [huomen] ('tomorrow', Fi7).

The duration of phonetic words was measured with a script in Praat. Phonetic words per second and syllables per second were used for measuring speech and articulation rates (i.e. speech rate without pause time). Both measures were used in order to find out the differences, if any, between them and to make the language comparison as thorough as possible. Based on earlier results of a comparative study of English and Finnish speech rate (Lehtonen 1981), it was expected that the comparison of syllable-timed Finnish and stress-timed Russian would yield different results depending on the measure chosen. Syllables per second would show the influence of hesitation better, since hesitation is often not only one or two syllables but one phonetic word. Also syllables per second as a measure would show mispronunciations (e.g. omission of a syllable, see examples above) better than phonetic words per second. For example, following her stay in Russia speaker Fi12 has much hesitation in her speech and the segmentation gives quite different results depending on the measure chosen (Figure 1). The sentence has 6 phonetic words and 18 syllable nuclei, when the original text only had 5 phonetic words and 13 syllable nuclei.

Microsoft Excel was used for calculating speech rate and articulation rate as well as for the graphical representation of the results. SPSS was used to determine the correlations in the data and their statistical significances. The existence of linear

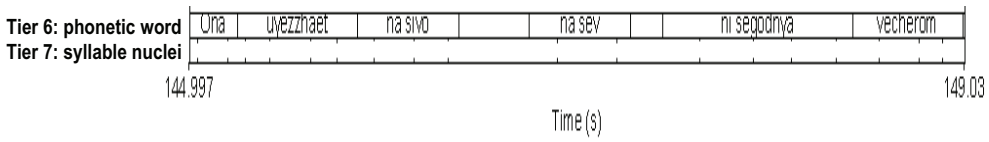


Figure 1: An example of the segmentation of the corpus *Ona uyezzhaet ne segodnya vecherom* 'She will leave not today at night' into phonetic words.

correlation was verified in scatterplot graphs. Paired samples t-test was used to find out the differences between different stages of learning. Speech and articulation rates of each sample were compared to its average fluency rating in order to determine the connection between speech and/or articulation rates and fluency. When comparing Finnish (L1) with Russian (L2) the individual variations in speech and articulation rates were minimized by comparing the within group ranking of each student in both languages (i.e. seeing whether the 2nd fastest student in Russian was also the 2nd fastest in Finnish etc.).

4 Results

In a previous study (Ullakonoja 2008), it was found that the majority of the speakers ($9/12$) developed in terms of their read-aloud fluency during the first half of their stay in Russia, and slightly over a half of them ($7/12$) further increased their perceived fluency during the rest of their stay. Furthermore, the study showed that pausing was closely related to read-aloud fluency in a foreign language.

4.1 Speech and articulation rates development during study abroad

In all subjects' speech the *speech rate* increased during the first half of their 3.5-month stay in Russia (0.2 phonetic words per second or 0.5 syllables per second on average) (Figures 2, 3; SR). Also, the majority of the subjects had a faster speech rate following their stay than before it (0.2 phonetic words per second or 0.5 syllables per second on average). Hence, the speech rate increases as the amount of experience increases. The development in speech rate is statistically significant ($p < 0.05$) when comparing before the stay results with middle of stay and before the stay results with after the stay in both phonetic words and syllables per second. However, the speech rate of some students ($4/12$ students when measuring phonetic words per second, $6/12$ students when measuring syllables per second) decreased slightly between the recordings done in the middle and after their stay. This decline is possibly due to the

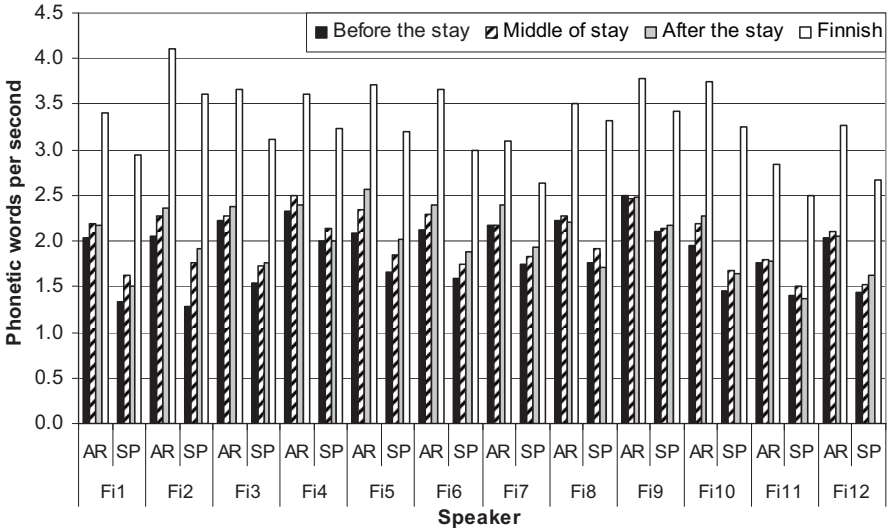


Figure 2: Articulation rate (AR) and speech rate (SR) in phonetic words per second in Finnish (L1) and in Russian (L2) at different stages of learning.

fact that their Russian reading was more "activated" while in the Russian speaking context than in the recording done following their stay.[1]

The measurement of *articulation rate* indicated a tendency similar to speech rate (Figures 2, 3; AR). Articulation rate also increased (0.1 phonetic words per second or 0.3 syllables per second on average) during the first half of the stay in the speech of most students ($9/12$). Between the 2nd and 3rd recordings, the articulation rate further increased for the majority ($7/12$) of the students (0.1 phonetic words per second on average), but also decreased or remained the same for some subjects. When comparing only the recordings done prior to and following the stay in Russia, it can be seen that the majority ($9/12$) of the students had a faster articulation rate after their stay than before it (0.2 phonetic words per second on average). The increase in articulation rate was statistically significant ($p < 0.05$) between before the stay and middle of stay results and between before the stay and after the stay results in both phonetic words and syllables per second.

The students were also divided into two groups according to their *residence* in Russia (host family vs. dormitories). The groups were neither balanced nor equal in their speech rate before their stay in Russia. When measuring phonetic words, students residing with a host family did not increase their speech rate on average

[1]The last recording was completed approximately one month after the students returned to Finland from Russia. It is possible that they had somewhat "forgotten" their Russian during that month, because some students had not used Russian at all after returning to Finland.

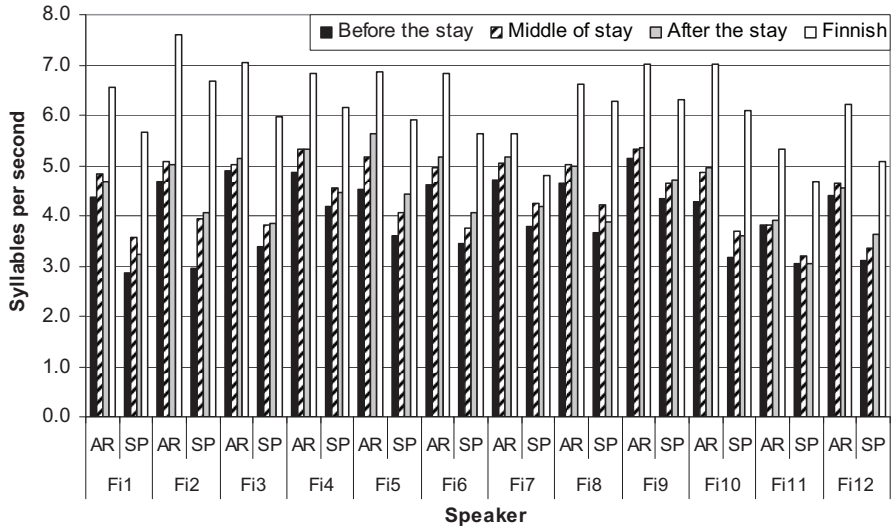


Figure 3: Articulation rate (AR) and speech rate (SR) in syllables per second in Finnish (L1) and in Russian (L2) at different stages of learning.

more than students living in the dormitories (Table 1). Contrary to what might have been expected, in syllables per second the dormitories group increased their speech rate more than the host-family group both during the first half and the whole length of their stay. In fact, the students residing with a host family had on average a slower speech rate at all recording sessions but as they also had a slower rate in Finnish, it seems that this is a random result. Similarly as in speech rate, the results of the articulation rate do not indicate that residence in the host family would make students speak faster during their stay in Russia. As a matter of fact, students residing in the dormitories increased their articulation rate more during the second half of their stay and during their entire stay in Russia (Table 1). The dormitories group might have had a better Russian competence and motivation already before the stay, which might have also been reflected in their speech rate.

4.2 Speech and articulation rates and fluency

What then is the relationship between speech or articulation rates and L2 fluency? The comparison of speech and articulation rates with perceived mean fluency rating flagged significant correlations (Table 2). The correlation was stronger between the speech rate and fluency rather than articulation rate and fluency. This indicates that pausing (hesitations and total pause time) also affects the fluency perception. The samples were also studied at the individual level where it was also noted that speech

Table 1: Mean speech and articulation rate of the students living with a host family and in the dormitories.

Residence	Before the stay	Middle of stay	After the stay	Finnish
	Speech rate: Phonetic words per second			
Host-family	1.47	1.68	1.67	3.01
Dormitories	1.75	1.89	1.92	3.14
	Speech rate: Syllables per second			
Host-family	3.18	3.68	3.65	5.66
Dormitories	3.74	4.17	4.20	5.87
	Articulation rate: Phonetic words per second			
Host-family	2.04	2.16	2.16	3.47
Dormitories	2.21	2.33	2.41	3.60
	Articulation rate: Syllables per second			
Host-family	4.42	4.72	4.72	6.52
Dormitories	4.73	5.12	5.27	6.73

rate correlates more reliably with the perceived fluency rating. For example, it was found that the least fluent (evaluated fluency = 1.3) sample was the speaker Fi2 prior to the stay. She was also the slowest of all speakers when measuring speech rate in phonetic words (Figure 2) and the second slowest when measuring speech rate in syllables (Figure 3). However, her articulation rate was not the slowest; in fact it was just below the average (Figures 2, 3). Correspondingly, the speaker who was evaluated the most fluent was Fi9 following their stay in Russia, who was also found to be the fastest of all speakers in speech rate and among the two fastest in articulation rate (Figures 2, 3).

4.3 Speech and articulation rates in Russian (L2) and Finnish (L1)

Next, speech and articulation rates in Finnish (L1) and Russian (L2) were compared. It was found that speech rate in Finnish correlates with the speech rate in Russian (Table 3). The correlation is however stronger between the articulation rate than speech rate in L1 and L2. This suggests that it is the amount of pause time that differs in L1 and L2, because the articulation rate indicates the speed of "uttering sounds," whereas speech rate includes pauses. As mentioned above, when comparing the in-

Table 2: Pearson correlations (R) between mean perceived fluency rating and speech and articulation rate.

	N cases	Correlation (R)	Significance (p)
Mean perceived fluency rating and articulation rate:			
Phonetic words/s	36	0.484	0.003
Syllables/s	36	0.416	0.012
Mean perceived fluency rating and speech rate:			
Phonetic words/s	36	0.722	< 0.001
Syllables/s	36	0.697	< 0.001

terspeaker performance, the speakers were ranked by speech rate and articulation rate from slowest to fastest in Finnish and at each recording session in Russian in order to be able to normalize the effect of differences in the structure of the two languages.

In Finnish (L1) the differences were small between syllables per second and phonetic words per second in articulation rate and speech rate. An individual speaker almost always received the same ranking position among the speakers in L1. In speech rate, $6/12$ speakers received a similar (maximum difference between ratings being 2) rating on average in Russian and in Finnish. In articulation rate $8/12$ speakers (when measuring phonetic words) and $7/12$ speakers (when measuring syllables) were ranked similarly in Finnish and Russian. This also indicates, that articulation and speech rates in L1 and L2 are related. Hence, speech rate seems to be a speaker-specific rather than a language-specific phenomenon.

5 Discussion and Conclusions

Overall, the majority of the students increased their L2 speech and articulation rates during their 3.5-month-stay in Russia statistically significantly as their perceived fluency increased also. This clearly shows that students seem to benefit from their stay in Russia so that they become faster and more fluent in Russian. Consistently with Towell *et al.* (1996, p. 103) the increased speech rate was found to be more significant than articulation rate in determining the L2 fluency of the speakers. When comparing the results with Lehtonen's (1978) study, it was found that the L1 Finnish reading rate was faster in this study when measuring phonetic words, but speech rates in syllables were similar in both studies.

The comparison of the students who stayed with a host family and students who resided in the dormitories was not very yielding as it turned out that the dormitories

Table 3: Pearson correlations for articulation rate (AR) and speech rate (SR) in phonetic words/s (pw) and syllables/s (syll) in Russian (L2) and Finnish (L1).

	Russian				Finnish		
	AR pw	AR syll	SR pw	SR syll	AR pw	AR syll	SR pw
Russian							
AR pw	1	0.966**	0.868**	0.861**	0.579**	0.556**	0.577**
AR syll	0.966**	1	0.811**	0.848**	0.586**	0.557**	0.574**
SR pw	0.868**	0.811**	1	0.985**	0.333*	0.282	0.424**
SR syll	0.861**	0.848**	0.985**	1	0.335*	0.279	0.423*
Finnish							
AR pw	0.579**	0.586**	0.333*	0.335*	1	0.985**	0.931**
AR syll	0.556**	0.557**	0.282	0.279	0.985**	1	0.913**
SR pw	0.577**	0.574**	0.424**	0.423*	0.931**	0.913**	1
SR syll	0.559**	0.552**	0.381*	0.376*	0.922**	0.929**	0.989**
N	36	36	36	36	36	36	36

** $p < 0.001$, * $p < 0.05$

group was already faster prior to the stay. Still, the results showed that in fact the students in the dormitories increased their speech and articulation rates more than the students living with host families. It can also be concluded that the speech and articulation rates in L1 are related to the speech and articulation rates in L2, consistently with Towell *et al.*'s study (1996, p. 96), where a strong correlation in L1 and L2 speech rate was established. Not surprisingly, the results also show that L1 is spoken faster than L2 (see e.g. Paananen-Porkka 2007).

The rhythmical features of speech were not taken into the account in this study. However, it is possible that the speech rate varies across the speech sample in the way as e.g. Deese (1980, pp. 74–76) has found that the majority of the faster sequences of speech occur either at sentence initial or terminal position. This study included recordings in Finnish only at the beginning and it was assumed that speech and articulation rates do not change significantly over time in one's L1 in the same reading task.

It has to be acknowledged that, naturally, there are other factors influencing speech and articulation rates and perceived fluency than the study abroad. Firstly, there is much individual variation in reading rate (even in L1). Also, in a reading task the subject might read very fast without comprehending everything being read (Lehtonen 1981, pp. 328–329; Perfetti 1985, p. 10) The student's motivation and interest are essential in L2 learning, therefore in this study also e.g. the motivation of the student towards Russian oral skills in general might have increased during the

stay in Russia. Furthermore, the findings concern only read-aloud speech in a laboratory setting and the analysis of spontaneous speech in a real communicative situation might have yielded different results.

It can be concluded that faster L2 speech (either in measures of speech or articulation rate) is perceived more fluent than slower L2 speech and that speech and articulation rates come closer to L1 speech and articulation rates as experience with L2 increases. Because native speakers of a language have been found to evaluate fast speech rate in non-native speech more positively than a slower speech rate (Munro & Derwing 1998; 2001, Paananen-Porkka 2007, p. 340), L2 teaching should pay more attention to practising appropriate speech rate in order to improve the communicative competence of the learners.

References

ABERCROMBIE, David 1967: *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.

AHO, Eija & YLI-LUUKKO, Eeva 2005: Intonaatiojaksoista. – *Virittäjä*, **109**(2): 201–220. (abstract: Intonation units).

AVANESOV, Ruben I. 1956: *Fonetika sovremennogo russkogo literaturnogo jazyka*. Moscow: Izdatel'stvo Moskovskogo Universiteta.

BOERSMA, Paul & WEENINK, David 2008: Praat—doing phonetics by computer. <http://www.fon.hum.uva.nl/praat/>.

CENOZ, Jasone 2000: Pauses and hesitation phenomena in second language production. – *ITL: Review of Applied Linguistics*, **127–128**: 53–69.

CUCCHIARINI, Catia, STRIK, Helmer & BOVES, Lou 2002: Quantitative assessment of second language learners' fluency: Comparisons between read and spontaneous speech. – *The Journal of the Acoustical Society of America*, **111**: 2862–2873.

COUNCIL FOR CULTURAL CO-OPERATION. EDUCATION COMMITTEE, MODERN LANGUAGES DIVISION, STRASBOURG AND COUNCIL OF EUROPE 2001: *Common European framework of reference for languages: learning, teaching, assessment.* Cambridge: Cambridge University Press.

DEESE, James 1980: Pauses, prosody, and the demands of production in language. – Hans W. Dechert & Manfred Raupach (eds.), *Temporal Variables in Speech*. The Hague: Mouton. 69–84.

FREED, Barbara F. 1995: What makes us think that students who study abroad become fluent? – Barbara F. Freed & Charles A. Ferguson (ed.), *Second Language Acquisition in a Study Abroad Context*. Amsterdam: Benjamins. 123–148.

FREED, Barbara F., SEGALOWITZ, Norman & DEWEY, Dan P. 2004: Context of learning and second language fluency in French: Comparing regular classroom, study abroad, and intensive domestic immersion programs. – *Studies in Second Language Acquisition*, **26**: 275–301.

IIVONEN, Antti, NIEMI, Tuija & PAANANEN, Minna 1995: Comparison of prosodic characteristics in English, Finnish and German radio and TV newscasts. – K. Elenius & P. Branderud (eds.), *Proceedings of the XIII International Congress of Phonetic Sciences, Stockholm, 13–19 Aug. 1995*, vol. 2. 382–385.

LAFFORD, Barbara A. 2004: The effect of the context of learning on the use of communication strategies by learners of Spanish as a second language. – *Studies in Second Language Acquisition*, **26**: 201–225.

LAURANTO, Yrjö 2005: Sujuvuuden mittoja. – Leena Kuure, Elise Kärkkäinen & Maarit Saarenkunnas (eds.), *AFinLA Yearbook 2005*. Jyväskylä: Jyväskylän yliopisto. 127–147.

LEHTONEN, Jaakko 1978: On the problems of measuring fluency. – Matti Leiwo & Anne Räsänen (eds.), *AFinLA Yearbook 1978*. Jyväskylä: Jyväskylän yliopisto. 53–68.

LEHTONEN, Jaakko 1979: Speech rate and pauses in the English of Finns, Swedish-speaking Finns and Swedes. – Rolf Palmberg (ed.), *Perception and Production of English: Papers on Interlanguage*. Åbo Akademi. 3–19.

LEHTONEN, Jaakko 1981: Problems of measuring fluency and normal rate of speech. – Jean-Guy Savard & Lorne Laforge (eds.), *Actes du 5e Congrès de l'Association internationale de linguistique appliquée*. Montreal: Presses de l'Université Laval. 322–332.

MOORE, Kate & KORPIJAAKKO-HUUHKA, Anna-Maija 1996: The clinical assessment of Finnish fluency. – Martin J. Ball & Martin Duckworth (eds.), *Advances in Clinical Phonetics*. Amsterdam: John Benjamins. 171–196.

MUNRO, Murray J. & DERWING, Tracey M. 1998: The effects of speaking rate on listener evaluations of native and foreign-accented speech. – *Language Learning*, **48**: 159–182.

MUNRO, Murray J. & DERWING, Tracey M. 2001: Modelling perceptions of the accentedness and comprehensibility of L2 speech: The role of speaking rate. – *Studies in Second Language Acquisition*, **23**: 451–468.

PAANANEN-PORKKA, Minna M. 2007: *Speech Rhythm in an Interlanguage Perspective: Finnish Adolescents Speaking English*. University of Helsinki.

PERFETTI, Charles A. 1985: *Reading Ability*. New York: Oxford University Press.

RIGGENBACH, Heidi 1991: Toward an understanding of fluency: A microanalysis of nonnative speaker conversations. – *Discourse Processes*, **14**: 423–441.

SEGALOWITZ, Norman & FREED, Barbara F. 2004: Context, contact, and cognition in oral fluency acquisition: Learning Spanish in at home and study abroad contexts. – *Studies in Second Language Acquisition*, **26**: 173–199.

SHTERN, A. S. 1988: Nekotorye statisticheskie harakteristiki russkoj spontannoj rechi. – Natalia D. Svetozarova (ed.), *Fonetika spontannoj rechi*. Leningrad: Izdatel'stvo Leningradskogo universiteta. 196–224.

SIMOES, Antonio R. M. 1996: Phonetics in second language acquisition: An acoustic study of fluency in adult learners of Spanish. – *Hispania*, **79**: 87–95.

SUOMI, Kari 2007: Accentual tonal targets and speaking rate in Nortern Finnish. – *Proceedings of Fonetik, TMH-QPSR, 50(1)*. 109–112.

TOWELL, Richard, HAWKINS, Roger & BAZERGUI, Nives 1996: The development of fluency in advanced learners of French. – *Applied Linguistics*, **17**: 84–119.

TROFIMOVICH, Pavel & BAKER, Wendy 2006: Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. – *Studies in Second Language Acquisition*, **28**: 1–30.

TROUVAIN, Jürgen 2003: *Tempo Variation in Speech Production: Implications for Speech Synthesis*. Institute of Phonetics, Saarland University.

ULLAKONOJA, Riikka 2008: Pausing as an indicator of fluency in the Russian of Finnish learners. – Plínio A. Barbosa, Sandra Madureira & César Reis (eds.), *Speech Prosody 2008: Fourth Conference on Speech Prosody, Campinas, Brazil*. 339–342.

VOLSKAYA, Nina forthcoming: Prosodic features of Russian spontaneous and read-aloud speech. – Viola de Silva & Riikka Ullakonoja (eds.), *Phonetics of Russian and Finnish, General Introduction: Spontaneous and Read-Aloud Speech*. Bern: Peter Lang.

III

PERCEPTION OF L2 FLUENCY IN STUDY ABROAD CONTEXT

by

Riikka Ullakonoja & Hannele Dufva 2008

Academic Exchange Quarterly, Fall 2008 (12) 3, 62–66.

Perception of L2 Fluency in Study Abroad Context

Riikka Ullakonoja, University of Jyväskylä, Finland
Hannele Dufva, University of Jyväskylä, Finland

Ullakonoja, M.A., is acting Lecturer in Russian, and, Dufva, Ph.D., is Professor of Language Education, at the Department of Languages, University of Jyväskylä, Finland.

Abstract
This article discusses the results of an on-going study investigating twelve Finnish university students' fluency in three Russian L2 reading aloud tasks administered before, during and after their period of study abroad. Here, we discuss the students' self-evaluated fluency and their responses to questions concerning the development of their pronunciation skills, comparing these with the fluency ratings given to them by Finnish teachers of Russian.

Introduction
This article deals with perceptions of Finnish university students' fluency in Russian - their foreign language (L2) - and the potential influence on their fluency of a period of study abroad. Aside from being a topic of research, L2 fluency is an explicit goal in foreign language education (e.g. CEF; Common European Framework of Reference for Languages, Council of Europe 2001). Teachers frequently use the term fluency to refer to their students' spoken or read-aloud production. It is also often mentioned in everyday life contexts and non-native speakers are commonly evaluated on the basis their spoken performance, e.g. the degree of foreign accent. The aim of this study is to investigate the possible relationship between evaluations of students' read-aloud fluency by teachers and students' own evaluations. An experimental phonetic analysis of fluency in students' read-aloud speech was previously conducted by the first author (Ullakonoja 2008a).

Here, L2 fluency is defined as perceived ease and fluidity of reading aloud. This would include such factors as an 'appropriate' speech rate, 'smooth' performance and the ability to pause at syntactically correct locations (Riggenbach 1991; Lennon 2000). We discuss perceived fluency both in terms of students' self-evaluations and teachers' ratings. Factors such as the students' language proficiency, linguistic accuracy, production of segmental or prosodic features or their overall oral performance are not discussed. Research seems to support the claim that study abroad increases learners' L2 fluency, sometimes significantly (e.g. Freed 1995; Towell et al. 1996; Freed et al. 2004; Lafford 2004; Trofimovich & Baker 2006). However, there are also studies suggesting that not all learners may equally benefit from it. As Simoes (1996) and Segalowitz & Freed (2004) argue, differences between speakers can be considerable. Further, Freed et al. (2004) showed that students in intensive domestic immersion gained most in terms of fluency when three contexts (at home, immersion and study abroad) were compared.

To sum up, the existing research seems to indicate that although study abroad may not be a sufficient condition for the development of L2 fluency, in most cases it is beneficial, particularly in that it offers the kinds of learning opportunities that may not exist at home or in the conventional domestic classroom setting. Abroad, learners are exposed to L2 in their everyday activities in various institutional, casual and media settings, and are bound to engage in interactions with native speakers. However, the

extensive exposure and the scope of potential communicative situations may not be beneficial to all. Some students may simply feel "overwhelmed by the amount, delivery rate, and complexity of the language that surrounds them" (Segalowitz & Freed 2004: 174).

Material and methods

Twelve Finnish female university students majoring in Russian who participated in the same 3.5 months' programme of study abroad during their second year of university studies participated in the study. Before university they had studied Russian as a foreign language from one to ten years. During their stay in Russia, they studied Russian for foreigners at Tver State University. Half of the students resided with a host family while the rest were accommodated in dormitories for foreign students. Their oral performance in Russian was recorded three times: prior to, during and following their stay in Russia. The task was to read aloud two dialogues in pairs. The longest continuous turn, consisting of six sentences from each student, from each recording session was chosen as a sample to be rated for fluency by 30 Finnish teachers of Russian. The teachers heard the stimuli (n=36) once in a randomized order, without knowing they contained multiple samples from the same speaker, and rated each sample. The sample size was limited, both because of the size of the student group participating in the exchange programme and because of the length of the listening task, which was around 20 min. This was considered optimal for the participants' concentration on the task and the reliability of their ratings.

On the basis of the teachers' evaluations, a mean fluency rating was obtained for each student for each recording session. In analysing the material, Ullakonoja (2008a) found that the students' L2 fluency, as evaluated by the teachers, increased during the 3.5 months' study abroad. For 8 out of 12 students, the increase was statistically significant (the difference between the ratings of each student prior to and following the stay were compared by the Wilcoxon Signed Ranks test, $p<0.005$). In the teachers' evaluation form used by Ullakonoja (2008a), the scale ran from one ('not fluent') to five ('very fluent') with no verbal descriptions given for the values between one and five. To see whether the potential variation in the use of the scale influenced the results and to normalize the ratings of each teacher, we have here performed a recalculation using z-scores where mean and standard deviation values are calculated across all the ratings given by teacher t (z equals mean subtracted from x multiplied by standard deviation, where x is a rating by teacher t).

Further, in this report we explore the students' self-evaluations of their L2 fluency development, using the responses to a questionnaire that they completed during the recording sessions. On a three-point scale of yes/no/don't know, they were asked to respond to statements concerning their views on the development pronunciation and the role of teaching and native speaker contacts therein. To analyse the responses, the students were divided into two groups according to the development of their fluency (as estimated by teachers). Group 1 (G1) consisted of students (n=8) whose teacher ratings of their L2 fluency were statistically significantly better after the stay in Russia than prior to it. Group 2 (G2) comprised the rest (n=4). For comparison of the mean fluency ratings of the students and their responses, the 'no' and 'don't know' answers were treated as one category. The reasoning behind this was that an affirmative answer indicated that the student had been attentive to the development of her oral skills and

aware of the possible changes while this was not necessarily implied by 'no' or 'don't know'. Thus we took the distinction to reflect - to a degree - self-awareness on the part of the student. Below, we compare the students' self-evaluations with the teachers' fluency ratings to determine whether a relationship exists between the two. Each subject's response to the statements (Q1-Q8) in the questionnaire is compared with their average teacher-rated fluency after the stay (both the mean and z-score mean) as well as with the difference between the fluency ratings prior to the stay and following it.

Results

First, we recalculated the results obtained by Ullakonoja (2008a) concerning the teachers' ratings of the students' fluency. When the statistical reanalysis (difference between means of each student in Paired Samples t-test) using the z-scores was performed, the results confirmed the earlier findings: 8/12 students had a higher mean fluency rating in the middle of the stay than prior to it and 9/12 students received a higher rating following their stay than prior to it. Also, for 7/12 students the teachers' rating was significantly higher following the stay than prior to it ($p<0.005$). Thus, most students were judged as more fluent after than before their stay abroad. Also, 8/12 were significantly more fluent readers in the middle of the stay than prior to it ($p\leq0.005$); however unlike in the earlier study, only one student showed a significant ($p=0.0001$) improvement in teacher-rated fluency after her stay. These results are consistent with those of the earlier study (Ibid.).

Second, we analysed the students' responses to the questionnaire and compared these to the teachers' evaluations of their performance. In Q1 (I have paid attention to practising my intonation and/or pronunciation outside the classroom), seven out of twelve students said they had practiced by themselves. No difference in the responses to Q1 was observed between G1 and G2. No association was found between the answers to Q1 and mean teacher-rated fluency after the stay. In answers to Q2 (I have noticed that my pronunciation improved noticeably during my stay in Russia), half of the students reported improvement. Again, groups G1 and G2 did not differ. However, mean teacher-rated fluency was higher (mean=3.5) for those who had noticed an improvement than for those who had not (mean=3.2) (Q2). All the students felt that their fluency had improved during their stay (Q3, I can now read and speak Russian more fluently than before my stay).

In Q4 (I actively tried to get into contact with Russians during my stay), 4/12 students reported they had attempted getting into contact with the local people whereas 4/12 had not and 4/12 answered 'don't know'. Only one student in G2 reported being active in seeking contact compared to three in G1. Interestingly, those who did not report seeking contacts with native speakers of Russian, had a higher mean fluency rating (mean=3.4) than those who did (mean=3.2). When asked about the role of their teachers (Q5, I would have wished that my Russian teachers would have paid more attention to teaching pronunciation and correcting my mispronunciations), most students (8/12) stated they would have wished more support from their teachers. The students who responded positively had a higher mean fluency rating (3.5) than those who responded negatively (3.1). Again, more students (75 percent) in G1 than in G2 (50 percent) answered affirmatively. In Q6 (I still have problems with producing intonation in the way I wish), some students (5/12), most from G2, felt that they still had not achieved their target intonation. Mean teacher-rated fluency was higher (mean=3.4) for those who did not report problems than for those who did (mean=3.2).

When asked about whether they had attempted to make their pronunciation better on their own (Q7, I tried independently to improve my pronunciation), most students (7/12) said that they had not. Here, half of G2 responded affirmatively as against only 25 percent of G1. However, mean teacher-rated fluency was higher (mean=3.5) for those who answered affirmatively than for those who did not (mean=3.3). In the final question, Q8 (I paid attention to practising correct pronunciation), only two students, both from G1, reported that they had tried to practise. They also had a clearly higher mean fluency rating (mean=3.9) than the rest (mean=3.2). To sum up, there were some differences between those who significantly improved their read-aloud fluency and those who did not. However, when all the questionnaire answers were compared with the student's teacher-rated fluency after her stay, no statistically significant relationships were found (Independent Samples t-test). When the two groups (G1 and G2) were compared using Pearson's Chi Square test for each question, no statistically significant differences were found. However, this may also be due to a sample size not large enough to establish statistical significance.

The students who had significantly improved their teacher-rated fluency in the reading aloud tasks (in G1) responded more or less as expected in four questions. Positive responses to questions Q5 and Q8 perhaps reflect students' awareness of and interest in developing their pronunciation skills whereas positive responses to Q4 and negative to Q6 may reveal more about their "self-esteem" and confidence as language users. Not surprisingly, the results suggest that improvement may be connected both with an interest in learning oral skills and in seeking native-speaker contacts. However, half of the students in G2 claimed they had tried to develop their pronunciation independently. Nevertheless, their ratings did not improve significantly. Although the possible reasons for this cannot be examined in detail here, it may be suggested that an intensive focus on correct pronunciation may also result in disfluencies such as self-repairs and thus in lower fluency ratings.

The comparison between the teacher-rated mean fluency of the students and their answers to each question showed that the self-evaluations did correspond - to a degree - to the fluency evaluations of the teachers: those who said that their pronunciation had improved and who claimed to have no problems in producing intonation also received higher mean fluency ratings (Q2 & Q4). Further, those who showed interest in learning and practising - and wanted their teachers to correct them - were judged on average as more fluent readers (Q5, Q7 & Q8).

Conclusion

The fact that all the students saw themselves as more fluent readers after their stay suggests that staying abroad is experienced as an important factor in improving fluency. Also in the teachers' ratings the majority of the students significantly improved their fluency. These results further support the findings of previous studies that suggest a relationship between studying abroad and L2 fluency. Also, a relationship between the students' self-evaluations and teachers' evaluations was found: the students who were judged to have improved their fluency were those who reported that they had paid attention to their pronunciation skills and aimed at improving them; these students were possibly also more capable of self-assessment.

However, it needs to be said that, first, the current study deals with read-aloud speech only, not spoken interaction. Second, we understand the development of fluency as a complex process, where features of articulation (such as e.g. pausing or patterns of intonation) intertwine with other (e.g. grammatical, lexical, pragmatic) characteristics. All these may be used in perceptions and evaluations of what is 'fluent'. The present study is part of a series of studies (Ullakonoja 2008a, 2008b) focusing exclusively on phonetic features. Third, the study abroad context itself involves

various social and cultural factors that could not be explored here (Freed et al. 2004, Wilkinson 1998). Fourth, more subjects should be investigated to be able to generalise the results.

Finally, many factors that are typically involved in the stay abroad context (for a review on the effects of study abroad on L2 learning, see Collentine & Freed 2004) are also considered important research foci in current studies on second language acquisition and foreign language education. In staying abroad, but in other contexts as well, it is not only the quantity of L2 input that is important but also its quality and meaningfulness. The overall role of the social and cultural environment, the significance of participating in the social networks of the target culture and the experience of authenticity and 'ownership' in one's learning process should not be forgotten. While recognizing the importance of these factors, we nevertheless feel that the study abroad context provides an excellent opportunity for a learner to become more fluent in L2. This opportunity could also be more systematically taken into consideration in university language education, for example in coaching students for their exchange visits and thus in helping them to gain optimal benefit from the target language environment.

References

Collentine, Joseph and Freed, Barbara F. "Learning context and its effects on second language acquisition," Studies in second language acquisition, Vol. 26, 153–171, 2004.

Council of Europe, Ed. Common European framework of reference for languages: learning, teaching, assessment. Cambridge: Cambridge university press, 2001.

Freed, Barbara F. "What makes us think that students who study abroad become fluent?" In Barbara F. Freed, and Charles A. Ferguson, Eds. Second language acquisition in a study abroad context. Amsterdam: Benjamins, 123–148, 1995.

Freed, Barbara F. and Segalowitz, Norman and Dewey, Dan P. "Context of learning and second language fluency in French: Comparing regular classroom, study abroad, and intensive domestic immersion programs," Studies in second language acquisition, Vol. 26, 275–301, 2004.

Lafford, Barbara A. "The effect of the context of learning on the use of communication strategies by learners of Spanish as a second language," Studies in second language acquisition, Vol. 26, 201–225, 2004.

Lennon, Paul. "The lexical element in spoken second language fluency." In Heidi Riggenbach and Richard Schmidt, Eds. Perspectives on fluency. Michigan: Ann Arbor, 25–42, 2000.

Riggenbach, Heidi. "Toward an understanding of fluency: A microanalysis of nonnative speaker conversations," Discourse processes, Vol. 14, 423–441, 1991.

Segalowitz, Norman and Freed, Barbara F. "Context, contact, and cognition in oral fluency acquisition: Learning Spanish in at home and study abroad contexts," Studies in second language acquisition, Vol. 26, 173-99, 2004.

Simoes, Antonio R.M. "Phonetics in second language acquisition: an acoustic study of fluency in adult learners of Spanish," Hispania, Vol. 79, No. 1, 87–95, 1996.

Towell, Richard, Hawkins, Roger and Bazergui, Nives. "The development of fluency in advanced learners of French," Applied linguistics, Vol. 17, 84–119, 1996.

Trofimovich, Pavel and Baker, Wendy. "Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech," Studies in second language acquisition, Vol. 28, 1–30, 2006.

Ullakonoja, Riikka. "Pausing as an indicator of fluency in the Russian of Finnish learners." In Plinio A. Barbosa, Sandra Madureira, and César Reis, Eds. Proceedings of the speech prosody 2008 conference. Campinas, Brazil. São Paolo: Editora RG/CNPq, 339–342, 2008a.

Ullakonoja, Riikka. "Speech rate as an indicator of fluency in the Russian of Finnish learners." In Michael O'Dell, and Tommi Nieminen, Eds. Fonetiikan päivät 2008 - The Phonetics symposium 2008. Tampere studies in language, translation and culture, series B, 2008b.

Wilkinson, Sharon. "Study abroad from the participants' perspective: A challenge to common beliefs," Foreign language annals, Vol. 31, No. 1, 23–39, 1998.

IV

COMPARISON OF PITCH RANGE IN FINNISH (L1) AND RUSSIAN (L2)

by

Riikka Ullakonoja 2007

Proceedings of the 16th International Congress of Phonetic Sciences, 6–10 August 2007, Saarbrücken, Germany, Trouvain, Jürgen & Barry, William J. (eds.). Saarbrücken: Universität des Saarlandes, 1701–1704.

Reprinted with kind permission by Jürgen Trouvain.

Comparison of Pitch Range in Finnish (L1) and Russian (L2)

Riikka Ullakonoja

Department of Languages, University of Jyväskylä, Finland
riikka.ullakonoja@campus.jyu.fi

ABSTRACT

The aim of the present study is to investigate whether the pitch range of a speaker can vary according to the language he speaks. The hypothesis is tested on Finnish university students studying Russian as a foreign language before and after their stay in Russia. First, the global pitch range (max – min) is determined. Second, the pitch range in different types of utterances (declarative, question, exclamation) is examined by superimposing pitch contours. The results indicate that the learners have a narrower pitch range and a less variable pitch than native speakers both in L1 and L2. However, the results also suggest that L2 experience seems to help most students to produce a more native-like pitch range, especially in questions.

Keywords: pitch range, foreign language acquisition, L2 experience, Finnish, Russian

1. INTRODUCTION

Research indicates that languages may differ in how pitch range is manifested [13]. For example, Russians are often referred to as using a wider pitch range and speaking on a higher tone than Finns whose speech is often characterized as monotonous. This is interesting considering foreign language (L2) learning: do Finnish L2 learners of Russian have to adjust their speech according to the pitch properties of the target language?

Pitch range can be defined in various ways, the simplest of which is the difference between F0max-F0min. However, this figure does not give full information about the distribution of the F0 values. Other ways to define pitch range is to include 95 % of the different pitch values around the mean or to study F0 differences between the overall level and the range of frequencies used (span) [3, 6]. Patterson [10] suggests measuring pitch range as "the difference between average non-initial accent peak and average post-accent valley", but that measurement was not adopted in this study. Also, pitch range can be investigated by graphically superimposing all pitch contours of a speaker with time normalization [4].

It has been discovered that a speaker's pitch range can vary according to e.g. emotions, speaking context and language [11, 8]. However, pitch range has hardly been studied in L2 context. In L2 speech e.g. a narrower pitch range for learners than native speakers has been observed [7, 1].

The aim of the research is to find out 1) whether the pitch range is different in Finnish and Russian and 2) if so, do learners acquire a more native-like pitch range of Russian during their stay in Russia?

2. ANALYSIS

2.1. Participants and material

The corpus consisted of 9 Finnish female university students' read-aloud speech, one dialogue in Finnish (c. 3 min/speaker with pauses) and two dialogues in Russian (c. 4 min/speaker with pauses). The Finnish dialogue consisted of 51 and the Russian dialogues of 50 utterances. The same texts were recorded before and after students' 4-month-stay in Russia. 7 Russian women read the same Russian texts.

Most students had studied Russian as their 4th foreign language. As they were not exposed to Russian in their everyday lives outside the university studies, the exchange period in Russia was essential for the acquisition of communicative competence and a key factor in learning Russian pronunciation, prosody in particular.

2.2. Procedure

The L2 material was recorded digitally on a computer (program Adobe Audition 1.0, sample rate 44100 Hz, 16 bit resolution) with AKG GN30 microphones in a recording studio. The students read the dialogues in pairs. The instruction was not to concentrate on pronunciation of single sounds but on presenting the dialogues naturally. The native speakers were recorded on a DAT-recorder with a Sony ECM-959A microphone.

2.3. Methods

The utterances and utterance types were segmented and annotated. The pitch was calculated with the autocorrelation method in Praat [2] and the pitch contours were graphically superimposed with time normalization. Furthermore, the results were statistically analyzed and tested (ANOVA) in SPSS.

3. RESULTS

The pitch range was investigated in the students' Finnish speech (L1), their Russian speech before and after their 4-month-stay in Russia, and in native Russian speech. Although the measuring unit is Hertz (Hz), the speakers' results are comparable since the variations of pitch are studied instead of absolute values.

3.1. Global pitch range

First, the global pitch range was measured by subtracting the minimum pitch value from the maximum pitch across the entire recording of each speaker. The pitch range of speakers varied according to the language and the amount of experience (Table 1).

Table 1: Pitch range (max – min) in Hz and [mean pitch] of all speakers in different recordings.

Finnish speaker	Finnish	Before the stay	After the stay	Russian speaker	Russian
Fi1	260[205]	250[210]	187[205]	Ru1	392[284]
Fi2	259[215]	259[221]	196[218]	Ru2	326[260]
Fi3	148[190]	116[194]	135[191]	Ru3	213[233]
Fi4	243[227]	208[237]	277[233]	Ru4	177[216]
Fi5	243[194]	233[198]	255[199]	Ru5	234[254]
Fi6	133[204]	204[210]	203[214]	Ru6	226[239]
Fi7	191[195]	224[206]	247[204]	Ru7	174[214]
Fi8	127[224]	160[229]	164[231]		
Fi9	279[200]	284[221]	323[233]		

Most Finnish speakers (8/9) used a different pitch range in Finnish than in Russian (either before their stay or after it): for most speakers (5/9) it was narrower than in Russian. In Russian, the majority (6/9) of the learners had a wider pitch range after their stay than before it. The native speakers' pitch range varied much, but its average (249 Hz) was wider than that of the learners (Finnish 209 Hz, Russian before the stay 215 Hz and after the stay 221 Hz).

All learners had a slightly lower mean pitch in Finnish than in Russian, with an average of 206 Hz. The mean pitch for all learners did not differ before and after their stay in Russia. The native Russians' mean pitch was 243 Hz. Some studies suggest that Russian women's mean pitch is usually a little higher, 260 Hz [5]. To summarize, the mean pitch of the Russians is much higher than that of the learners. However, on average, the learners use a higher mean pitch in Russian than in Finnish which seems to indicate that they have learnt to use a higher, more native-like pitch in Russian.

3.2. Local pitch range (different utterance types)

3.2.1. Pitch range (max-min) and mean pitch

Secondly, the pitch range was studied in different utterance types (declarative, question, and exclamation) by measuring the pitch range for each utterance separately. First, the mean pitch of each speaker varies according to the utterance type. Most Finnish informants had the highest mean pitch in Finnish in exclamations whereas in Russian it was found in questions. No great differences were found between the recordings before and after their stay in Russia. The native Russians had the highest mean pitch in questions and exclamations. Thus, as also the learners produced the highest mean pitch in questions, they seem to have acquired this feature of Russian.

Furthermore, the narrowest pitch range varied according to the speaker (Table 2).

Table 2: Utterance type (D=declarative, Q=question, E=exclamation) having the narrowest pitch range (Hz).

Finnish speaker	Finnish	Before the stay	After the stay	Russian speaker	Russian
Fi1	E(114)	Q(119)	E(110)	Ru1	Q(162)
Fi2	Q(106)	D(126)	E*(80)	Ru2	D(173)
Fi3	Q(129)	E(112)	Q+(128)	Ru3	D(174)
Fi4	E(123)	E(142)	D(181)	Ru4	E(120)
Fi5	E(147)	E(141)	D,E(143,145)	Ru5	D(140)
Fi6	D(157)	Q(160)	D,Q(198,202)	Ru6	-
Fi7	D,Q(167)	E(171)	Q(179)	Ru7	Q(145)
Fi8	E(92)	E(151)	E(133)		
Fi9	Q,E(141)	D(204)	D,Q(198,200)		

*) the difference between this utterance type and all other utterance types is significant at 0.032 level.
+) the difference between this utterance type and all other utterance types is significant at 0.005 level.

In Finnish, the learners had the narrowest pitch range in questions and exclamations. Before their

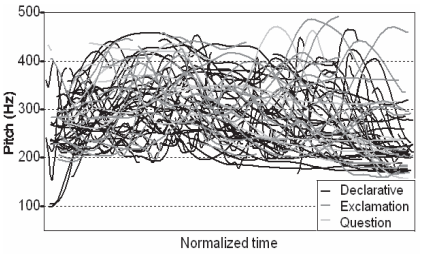
stay, most had the narrowest pitch range in Russian exclamations whereas after their stay there was not one utterance type that could be distinguished. Similarly, the native speakers did not have one single utterance type with the narrowest pitch range. Thus, the learners did not produce pitch the same way as they did before their stay (most of the students had the narrowest pitch range in a different utterance type before their stay than after it).

Considering the Finnish speakers as a group, it can be summarized that in Finnish the pitch range (Hz) and the mean pitch (Hz) depend on the type of the utterance so that exclamations differ statistically significantly from questions and declaratives in the mean pitch (p=0.000) and the pitch range (p=0.037). In their Russian, the utterance types differ significantly from each other (p=0.000) in the mean pitch both before and after the students' stay in Russia. In native Russian speech, the declaratives differ statistically significantly (p=0.000) from questions and exclamations.

3.2.2. Superimposed pitch contours

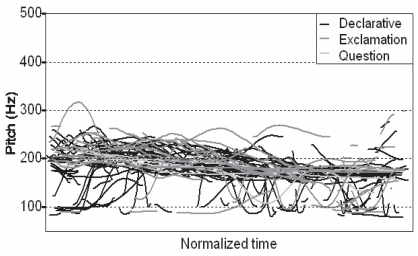
To get a better idea about the possible variations of the pitch in the utterance, all pitch contours of a speaker in each recording session were graphically compared by superimposing them with time normalization and using different colours for each utterance type (examples in Figures 1-4).

Figure 1: Pitch contours in native Russian utterances superimposed (speaker Ru1).



In general, the learners had less variation in their pitch contours both in Finnish and Russian. The native Russians typically used their whole pitch range more exhaustively, in a way that pitch values were spread more evenly around the whole range (Figure 1). In native Russian speech the high F0 peaks could occur anywhere in the utterance in all the utterance types, most often in questions.

Figure 2: Pitch contours in Finnish utterances superimposed (speaker Fi3).



In Finnish, there was a clear concentration of the pitch values around the mean (c. ± 30 Hz) with a declination of the pitch contour towards the end of the utterance and a tendency of high F0 peaks (if there was any) to be situated in the beginning or middle of the declaratives (Figure 2). It has to be pointed out that one speaker (Fi9) used a rather varied pitch range also in Finnish. In Finnish the declaratives and questions seemed to follow a similar (rather flat) pitch contour whereas in Russian there was a lot of fluctuation even in the declaratives.

Figure 3: Pitch contours in L2 Russian speech before the stay in Russia (speaker Fi3).

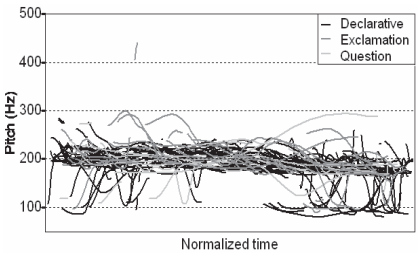
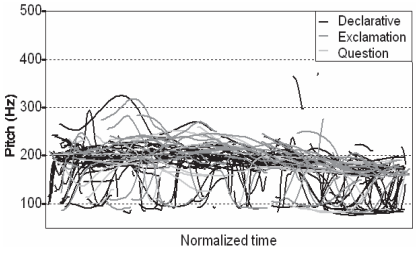


Figure 4: Pitch contours in L2 Russian speech after the stay in Russia (speaker Fi3).



In Russian, learners began all utterances each at almost the same pitch height and ended them slightly lower, whereas the native Russians had a tendency to alternate the pitch in the beginning and at the end of the utterances (Figures 3 and 4). The L2 speech could also be characterized by production of high F0 peaks in questions, but not so much in other utterance types. 5/9 students produced F0 peaks in questions before their stay in Russia and all of them after their stay. This suggests that the learners have learnt to vary their speech more, especially in questions, thanks to their L2 experience.

4. DISCUSSION AND CONCLUSIONS

To sum up, in Finnish (L1) the pitch range and mean pitch of exclamations differs statistically significantly from other utterance types. In Russian the mean pitch of the declaratives differ statistically significantly from exclamations and questions, but the pitch range does not seem to change according to the utterance type. Statistically significant mean pitch differences were also found in L2 speech in all utterance types.

Most L2 speakers of Russian have a narrower and less varied pitch range in Finnish than in Russian, and in both languages most of their pitch values are concentrated around the mean pitch. The native Russians have a wider pitch range than the learners, and their pitch values are more equally distributed around the whole pitch range. Furthermore, the L2 learners have a tendency to begin and end all their utterances at the same pitch level in L1 and L2 whereas the native Russians varied the utterance initial and final pitch height more.

The finding that the learners use a narrower pitch range in L2 than the native speakers is supported by earlier research [7, 1]. The results show that L2 experience seems to affect most students' L2 pitch range by making it more native-like (widening it and spreading the distribution of pitch values around the range). It also seems to encourage them to produce high F0 peaks in questions (in a native-like manner).

To conclude, further studies should consider the possibility that the Finnish informants' pitch could be affected by creaky phonation, which is a rather typical feature of Finnish [9] and L2 speech of Finns [12], but not of Russian. Currently, the research on voice quality is very scarce in both languages. It would also be interesting to investigate more speakers and possible subtypes of the utterances (e.g. different question types, responses). To determine the mean pitch and pitch range for Russian and Finnish women, larger speech corpora with speakers from different age groups should be used.

5. REFERENCES

[1] Aoyama, K., Guion, S.G. 2007. Prosody in Second Language Acquisition: An Acoustic Analysis on Duration and F0 Range. In: Bohn O.-S., Munro M. J. (eds.), *The Role of Language Experience in Second-Language Speech Learning - In Honor of James Emil Flege*. Amsterdam: John Benjamins, 281–297.

[2] Boersma, P., Weenik, D. 2006. *Praat: doing phonetics by computer* [computer software] Amsterdam: University of Amsterdam. http://www.praat.org visited 28-Feb-07.

[3] Haan, J., Van Heuven, V.J. 1999. Male vs. Female Pitch range in Dutch Questions. *Proc. 14th ICPhS* San Francisco, 1581–1584.

[4] Iivonen, A. 1999. F0 Contours of Utterances Superimposed on the Temporal Voice Range Profile of the Speaker. *Proc. 14th ICPhS* San Francisco, 953–956.

[5] Kodzasov, S.V., Krivnova, O.F. 2001. *Obschaya fonetika.* Moscow: Otkrytoye obshestvo.

[6] Ladd, D.R. 1996. *Intonational Phonology.* Cambridge: Cambridge University Press.

[7] Mennen, I., 1998. Can language learners ever acquire the intonation of a second language? *Proc. STILL,* Marholmen, Sweden, 17–20.

[8] Mennen, I. (to appear). Phonological and phonetic influences in non-native intonation. In: Trouvain, J., Gut, U. (eds), *Non-native Prosody: Phonetic Descriptions and Teaching Practice.* Mouton De Gruyter.

[9] Ogden, R. 2003. Voice Quality as a Resource for the Management of Turn-taking in Finnish Talk-in-interaction. *Proc. 15th ICPhS* Barcelona, 123–126.

[10] Patterson, D., Ladd, D.R. 1999. Pitch Range Modelling: Linguistic Dimensions of Variation. *Proc. 14th ICPhS* San Francisco, 1169–1172.

[11] 't Hart, J., Collier, R., Cohen, A. 1990. *A perceptual study of intonation.* Cambridge: Cambridge University Press.

[12] Toivanen, J., Waaramaa, T. 2005. Tone choice and voice quality of dispreferred turns in the English of Finns. *Logopedics Phoniatrics Vocology* 30, 181–184.

[13] Van Bezooijen, R. 1995. Sociocultural Aspects of Pitch Differences between Japanese and Dutch Women. *Language and Speech* 38 (3), 253–265.

Acknowledgements

The author wishes to thank MA Mietta Lennes for the help with the Praat scripting.

V

HOW DO NATIVE SPEAKERS OF RUSSIAN EVALUATE YES/NO QUESTIONS PRODUCED BY FINNISH L2 LEARNERS?

by

Riikka Ullakonoja 2010

Rice Working Papers in Linguistics vol. 2, 92–105.

How do native speakers of Russian evaluate yes/no questions produced by Finnish L2 learners?*

Riikka Ullakonoja
University of Jyväskylä

Abstract

This study analyzes native Russian speakers' evaluation of seven Russian yes/no-questions each produced by Finnish speakers in two sets of recordings (during a stay in Russia and after it). The Finnish speakers were six female university students of Russian. This research question is interesting because the two typologically unrelated languages differ in the prosody of yes/no-questions. In Russian a yes/no-question is created from a lexically and syntactically corresponding statement by means of intonation, whereas in Finnish the cue for questioning is an interrogative particle –ko/-kö instead of prosody. Hence, native Finnish speakers are likely to have difficulties in pronouncing Russian yes/no-questions. The aim was to find out how native Russian speakers recognise the intended questions produced by Finnish learners. First, the recognition rate of the different yes/no-questions was studied, and then the acceptability rating of questions was computed. The results show that in general students did not perform very well in producing a yes/no-question, but there was great variation depending on the question and learner. According to the successfulness of production two groups of utterances were established: successful and non-successful ones. The statistically significant difference between the two was explained by their syntactical and lexical content. The conclusions made are supportive of earlier findings, where Russian question intonation has been found difficult for Finns to learn.

Keywords: phonetics, speech perception, prosody, intonation, second language

1 Introduction

Perception of intonation in different languages has been the subject of extensive research and has been found to be a complex issue (Vaissière 2005). Yet a comprehensive theory of how intonation is perceived does not hitherto exist. The present study focuses on the perception of yes/no questions in Russian. The interest lies in the non-native Russian speech of advanced learners, whose mother tongue (L1) is Finnish. In this experiment native speakers of Russian were asked to evaluate a total

* I would like to thank Docent Nina B. Volskaya from St. Petersburg State University for the help with recruiting the subjects and conducting the listening experiment.

of 84 utterances meant as questions by Finns. The stimuli were extracted from Finnish students' read-aloud dialogues. The study is a part of a set of papers, where students' performance prior to, during and following their stay in Russia are compared. In this paper, the students were recorded during and following a stay in Russia to allow for comparisons between the two sets of recordings. The purpose of the study was to determine whether the students' pronunciation of yes/no questions remained consistent between the two recordings (during the stay in Russia and following it) coupled with native Russians' evaluations of their questions. In this paper, the recognition rate of the utterances intended as questions will be studied and comparisons between the two recordings will be made. Secondly, the ratings of the acceptability of the stimuli as questions will be analyzed and compared between the speakers and the two recordings.

The motivation for comparing the two recordings comes from the question of retention in learning to produce a question. An earlier study (Ullakonoja 2008) showed that most Finnish students displayed development in their fluency between the recordings done prior to the stay in Russia and following it. However, it also indicated that there was a slight decrease in fluency for some students between recordings done during the stay in Russia and after it. Furthermore, Ullakonoja (2009) showed that speech rate of some students decreased between the two recordings. Hence, it may be suggested that the decline in the skills of some students can be due to the fact that they "lose" some of the skills that they had learned in Russia after returning home. The present study aims to compare the recordings to be able to determine degree of retention in learning to produce acceptable intonation in Russian yes/no questions.

In Finnish yes/no questions, prosody does not play a distinctive role. Instead, yes/no questions are produced by morphological means (the interrogative particle *-ko/-kö* is attached to the sentence initial finite verb in neutral word order). Finnish interrogative intonation is typically characterized by a high-initial pitch and an intensity contour that roughly follows the pitch contour (Hirvonen 1970). An intonational grammar does not exist as such in Finnish, and a typical pitch pattern for Finnish yes/no questions has not been authoritatively defined. A recent empirical study can, however, shed some more light on this issue. Anttila (2008:64) claims that in Finnish, the question type affects mostly the pitch distribution, not the shape of the contour. In Anttila's (2008:76-77) study the most typical pitch contours for yes/no questions in Finnish read-aloud speech were a fall or a rise-fall. According to Hirvonen (1970) yes/no questions can be characterized by a relatively high pitch before the nucleus. When Anttila (2008:79,82-83) used Hirvonen's categories to regroup her data she found that the most frequent pitch contour for yes/no questions was indeed this longer high pitch followed by a fall for both men and women for both read-aloud and spontaneous speech.

In Russian, a yes/no question differs from a corresponding declarative only by prosody. In written texts and when using a formal register – less common in everyday speech – it is possible to distinguish yes-no/questions also grammatically with an interrogative particle *li*. The Russian intonation research rests on Bryzgunova's (1977) description of Russian intonational constructions (IKs). According to this theory, there are seven different IKs, each of which has distinctive functions and uses. Each IK also has a typical or range of typical intonation patterns. As Bryzgunova's theory was essentially meant for teaching purposes and as it was mostly based on the auditory observations of the author it has recently evoked some criticism for its lack of empirical data (see e.g. Yokoyama 2001). However, it is still the most widely used theory on Russian intonation as only a few empirical studies exist to date. According to Bryzgunova (1977) yes/no questions are usually pronounced with IK-3. This

intonation pattern is characterized by a sharp rise on the nuclear syllable or, if the nucleus is in the phrase-final position, a final rising pitch contour (see also Svetozarova 1998; Volskaya 2009). Also a delayed peak seems to be characteristic to Russian yes/no questions (Igarashi 2006; Meyer & Mleinek 2006). As for Finnish, the traditional view is that final rising contours do not occur in interrogatives (Hirvonen 1970). However, some recent findings (Iivonen 1998; 2001; Ogden & Routarinne 2005; Anttila 2008:70-76) suggest that final rising pitch contours can also exist, even though their role is not only to signal an interrogative. They can also be used to indicate address, emotionality or continuation (Iivonen 1979; 2001; Mixdorff et al. 2002). Despite this, as Mixdorff et al. (2002) observed, the final rise in questions was perceived as fairly unnatural by native Finnish listeners.

So far, large-scale contrastive studies on Russian and Finnish intonation have not been conducted and studies on non-native Russian intonation are also scarce. The main difference seems to be that a Finnish declarative rarely differs from a question only by its pitch contour (see e.g. Iivonen 1979), whereas in Russian intonation has a distinctive function (Bryzgunova 1977, Svetozarova 1982). Because both the acoustic features and linguistic functions of Russian intonation are different from Finnish (de Silva & Ullakonoja 2009), Finnish students face a challenge in learning Russian intonation (de Silva & Volskaya 2005). Kuosmanen & de Silva (2003; 2007) found that Russian yes/no questions are difficult for Finnish learners to pronounce due to the differing pitch contours. They found that only 63% of the eight interrogatives produced by ten Finnish students were recognized as questions by native speakers of Russian. In contrast, Toivanen's (2001) study of Finnish university students of English showed that they were able to produce English short questions (for example *Agree?, OK?*) rather well with a final rising pitch contour, which seems to show that the difficulty is not purely related to L1 influence but also to other factors. The present paper partly replicates the studies by Kuosmanen & de Silva (2003; 2007). Here, however, more judges were used and one of the goals was to compare whether the students productions during versus following stay are different.

Native speakers have been found to ignore some acoustic markers in speech perception because they rely also on lexical information when listening to the intonation of the stimuli (Lieberman 1965). However, in Russian, there are cases in which native listeners have to disambiguate sentence types on the basis of prosodic cues only. Interestingly, Finnish native listeners (not knowing Russian) were shown to perceive most Russian interrogatives as emotional speech whereas Russian native listeners (not knowing Finnish) perceived Finnish interrogatives as declaratives (Shserbakova 2001).

2 Methods

2.1 Speakers

In collecting the data for the listening experiment, six female Finnish (L1) undergraduate students (aged 19-25) were recorded. They were majoring in Russian and had studied Russian as their third or fourth foreign language (L2) for three years prior to attending university as well as one year at university. During their second year at university they spent one semester (3.5 months) in Russia. Half of the students resided with a Russian host family and the rest in the dormitories for foreign students. When asked, half of the students reported they had practiced pronunciation

independently. All students had some contact with and thus the possibility of communication with native speakers during their stay in Russia.

2.2 Speech data

Students were asked to read aloud two Russian dialogues which were originally short texts designed as telephone conversations (dialogues 46 and 100). The texts were taken from Russian as a foreign language teaching materials (Shilova & Usmanova 1990). Seven yes/no questions (see Table 1) from the dialogues were used as stimuli in the listening experiment. These particular questions were selected from the dialogues, because they can be understood either as declaratives or questions depending on the pitch contour used by the speaker. According to Bryzgunova's (1977) classification of Russian intonational constructions (IKs), these questions would normally be produced with the same intonational construction (IK-3) by native speakers. Because the students saw the texts in their written dialogue form, it was not possible for them to interpret the questions as declaratives.

	Russian	Russian (in Roman alphabet)	English translation
Q1	У тебя совесть есть?	U tebya sovest' yest'? *prep you conscience have*	Have you no conscience?
Q2	Соня?	Sonya? *sonya (proper name)*	Is that Sonya?
Q3	Чайный или столовый?	Chainyy ili stolovyy? *tea or dinner*	Tea or dinner service?
Q4	Ты рада за меня?	Ty rada za menya? *you happy for me*	Are you happy for me?
Q5	Да?	Da? *yes*	Yes?
Q6	Ты заболела?	Ty zabolela? *you be-sick*	Were you taken ill?
Q7	Судя по торжественному тону, ты хочешь сообщить мне нечто важное?	Sudya po torzhestvennomu tonu, *judging prep festive tone* ty khochesh' soobshchit' mne *you want inform me* nechto vazhnoye? *something important*	Judging by your festive tone of voice, you want to tell me something important?

Table 1: The utterances used in the experiment.

2.3 Recordings

The speech data were recorded at two recording sessions: 1) during the students' stay in Russia (referred to as T1) and 2) following their return to Finland (T2). The interval between the two recordings was approximately three months. The T1 recordings were conducted in the middle of the students' stay in Russia, whilst the T2 recordings were done at the beginning of the spring semester about a month after their return to Finland. Most of the students had not spoken Russian since returning home. Different recording equipment was used for practical reasons. In T1, the recordings were done

with a Sony TCD-D3 DAT-recorder or a Roland Edirol 24-bit Wave/MP3 R-09 digital recorder with a Sony ECM-959A microphone in a hotel room. In T2 the students were recorded using a computer equipped with the program Adobe Audition 1.0 and 2.0, and AKG GN30 microphones. The students were recorded in pairs because the texts used were dialogues. The dialogues were chosen in order to give a more realistic setting for the study.

2.4 Listening experiment

The listening experiment was conducted in two parts: T1 and T2 recordings separately. Hence there were two groups of listeners, one listening to the samples of the T1 recordings and the other listening to the T2 recordings. This was done to avoid too lengthy a task, as there were a total of 84 stimuli to be rated. The stimuli were presented in the same randomized order with approximately an 11-second interval between the stimuli. Prior to each stimulus listeners heard a sound indicating the start of the next stimulus as well as its number in Russian. The total duration of the task was approximately 11 minutes for each group. The listening experiment was prepared in the computer program Praat (Boersma & Weenik 2009) from the recorded sound files.

The judges were asked to define: (1) whether they perceived the stimulus a question or not (categorization, 1=non question, 2=question) and (2) whether the stimulus was successful as a question (acceptability rating, 1-5 scale: 1=not a question, 5=a question). The judges were told that the speakers intended all the utterances as questions. This was done to avoid forcing them to rate some stimuli as declaratives. If a more traditional setting for the listening task had been used (not informing the judges of the speaker's intent), the stimuli should also have included utterances intended as statements by the speakers. This however, would have resulted in a listening task perhaps too lengthy for the listener's concentration span.

2.5 Judges

Judges were either students or staff members at the philological faculty of St. Petersburg State University. There were a total of 40 listeners (19 in the first group, 21 in the second), all native speakers of Russian. Most of them were under 20 years old (80%), female (90%), students (92%). About half of them (48%) reported that they were used to hearing foreigners speak Russian. The amount of their exposure to non-native Russian varied from everyday (7.5%) to once a week (25%), once a month (15%) and rarer than once a month (52.5%). Only three judges had ever taught Russian to foreigners.

Not all listeners rated all the stimuli in all aspects: hence, there were some missing values that were excluded from the analysis. Cohen's Kappa for the categorization task was 0.563 in T1 and 0.588 in T2 and Cronbach's alpha for the acceptability ratings yielded 0.960 in T1 and 0.858 in T2. For the categorization task, following the principles of Landis & Koch (1977), it can be concluded that the values of Cohen's Kappa indicate moderate interjudge agreement. Cronbach's alpha, on the other hand, shows the interjudge consistency for the acceptability ratings. As for T1 and T2 in this study Cronbach's alpha yielded over 0.8, it can be considered good (see e.g. Bryman & Cramer 2001:62).

3 Results

3.1 Question categorization

First, I will discuss the results from the listening experiment from the point of view of question categorization. For this, the recognition rate (i.e. the proportion of the positive "yes" ratings) was calculated. When computing the recognition rates, missing values were excluded from the analysis. When looking at all the questions produced by all the speakers without distinguishing between the two recordings, it was found that the overall recognition rate was 57%. In other words, just over a half of the utterances intended as questions by Finns were recognized as such by native speakers of Russian.

Figure 1 depicts the mean recognition rate of the individual questions. As mentioned above, every utterance was pronounced twice by each of the six students. The utterance that was recognized the best as a question was Q4 (*Ty rada za menya?)*. The overall recognition rate was not very high: questions Q1, Q2 and Q7 have a recognition rate of below 50%. Hence, the utterances can be grouped into two categories according to their recognition rate: (1) those mostly understood as questions with a recognition rate between 56% and 99%, i.e. Q3, Q4, Q5 and Q6 (henceforth S=successful questions) and (2) those mostly recognized as non-questions with a recognition rate between 29% and 38%, i.e. Q1, Q2 and Q7 (henceforth NS=non-successful questions).

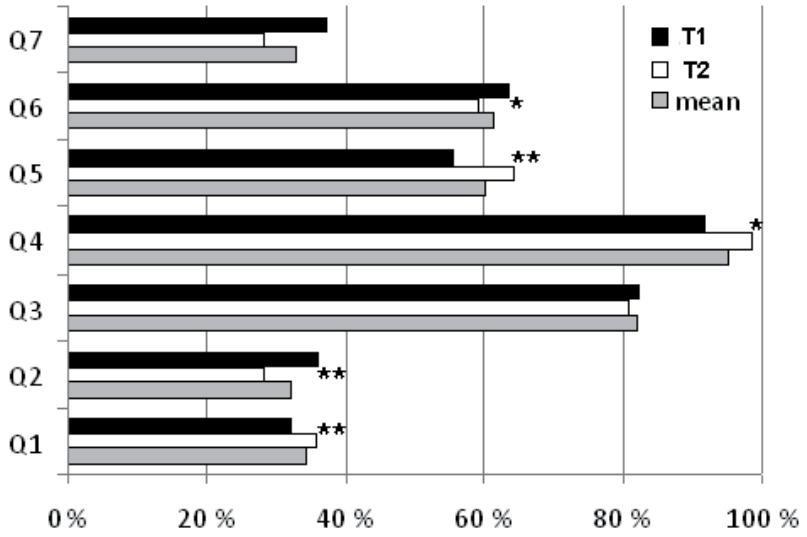


Figure 1: Native speakers' recognition rate of questions (n=84) produced by Finns during the stay (T1) and following it (T2) *) p<0.05, **) p>0.0001.

Moreover, Figure 1 compares the recognition rates in T1 and T2. The general recognition rate was slightly lower in T2 (56.4%) than in T1 (57.2%). The difference was statistically significant ($\chi^2(1)=153.566$; $p<0.0001$). When looking at the individual questions, three out of seven questions were recognized better in T2 than T1. The difference was the greatest in the shortest question Q5 (*Da?*), which had 6% better recognition rate in T2 than T1. The statistical significance of the differences between

T1 and T2 was tested in the Chi-Square test, which indicated that the majority of the differences between T1 and T2 were significant. In three questions (Q1, Q4 and Q5) there was significant improvement, whereas for Q2 and Q6 there was decline. While most questions showed a statistically significant difference between students' production in T1 and T2 as perceived by native speakers, there is no general tendency whether the difference is positive or negative between T1 and T2.

In Table 2 the recognition rate of the questions is considered from the point of view of the individual learners. Great interspeaker variation was observed. It is useful to examine the data by separating questions that were mostly rated successful questions (S) and those that were mostly rated unsuccessful (NS). Table 2 indicates that despite the fact that NS questions were generally recognized rather weakly, there were individual students (like Fi1 and Fi3 in Q1 and Q7 and Fi4 in Q2) who received rather high recognition rates. Both Q3 (S) and Q4 (S) had a high recognition rate for all speakers in both T1 and T2. Greater variation can be seen in Q5 (S) and Q6 (S).

Speaker	Question and category						
	Q1	Q2	Q3	Q4	Q5	Q6	Q7
	(NS)	(NS)	(S)	(S)	(S)	(S)	(NS)
Fi1 (T1)	90%	21%	68%	**84%**	95%	100%	**68%**
Fi1 (T2)	50%	0%	68%	**100%**	95%	86%	**86%**
Fi2 (T1)	**16%**	11%	84%	100%	11%	**78%**	89%
Fi2 (T2)	**24%**	11%	55%	100%	0%	**95%**	0%
Fi3 (T1)	**74%**	**21%**	**90%**	**83%**	**33%**	95%	**53%**
Fi3 (T2)	**86%**	**43%**	**91%**	**90%**	**90%**	55%	**62%**
Fi4 (T1)	5%	**95%**	**90%**	**90%**	**95%**	**16%**	5%
Fi4 (T2)	5%	**100%**	**100%**	**100%**	**100%**	**57%**	5%
Fi5 (T1)	**5%**	44%	**84%**	100%	95%	79%	**0%**
Fi5 (T2)	**10%**	14%	**90%**	100%	75%	57%	**10%**
Fi6 (T1)	**5%**	26%	**79%**	**95%**	**5%**	16%	11%
Fi6 (T2)	**43%**	5%	**81%**	**100%**	**25%**	5%	5%

Table 2. The recognition rate of questions of the individual speakers during the stay in Russia (T1) and following it (T2). The figures in bold indicate a higher recognition rate in T2 than T1. (NS=non successful questions, S=successful questions as perceived by native speakers).

When the development of individual speakers is compared, the student who scored a higher rate in most of the questions between T1 and T2 is Fi3. There is only one question in which she received a lower recognition rate in T2. There is also less variation in her utterances in T2 compared to T1, where the range was 43%-91%. Another successful speaker in T2 is Fi4, because four out of seven of her questions in T2 were recognized as an interrogative by all of the judges. There is, however, great variation in the speech of this subject: two of her intended questions were recognized as an interrogative by less than 6% of the judges. In fact, variation is typical for all speakers' productions: no single speaker achieved a high recognition rate in all questions, nor one did always have a low recognition rate. Hence, the results suggest that the target intonational construction IK-3 is difficult to learn to produce in all contexts. As mentioned above, according to the literature it is realized in at least two different kinds of pitch contours depending on the place of the nucleus. The students may struggle with the nucleus placement which would then lead to a contour choice, not perceived as interrogative by the listeners.

As Table 2 shows, only four out of 42 of the utterances intended as questions were *not* recognized as interrogatives by any of the judges, and 14 of the 42 utterances fall within the lowest 5% of recognition. It is also interesting to examine the utterances receiving 100% (or near) recognition rates. 11 out of 42 of the stimuli were interpreted as questions by all judges. Seven of these concern Q4 (*Ty rada za menya?)* which, as was shown in Figure 1, garnered the highest recognition rate for all speakers. It is also worthwhile to point out that over a half of the stimuli (27 out of 42) were recognized as questions by at least 90% of the listeners.

In Table 2 the comparison between the two recordings indicates that in about half (22 out of 42) of the questions the recognition rate is higher in T2 than T1, while in 16% (seven out of 42) it is the same and in 30% (13 out of 42) it is lower. The results would seem to suggest interspeaker differences. Some speakers have demonstrated improvement in their ability to pronounce questions during the last part of their stay in Russia, whereas for other speakers there was not such a significant change. As mentioned above speakers Fi3 and Fi4 seem to have benefited most from the study abroad: six out of seven stimuli of Fi3 have a better recognition rate in T2 than T1 and five out of seven stimuli of Fi4. The proportion of more successfully conveyed questions in T2 than T1 for Fi6 is about a half (four out of seven) but for Fi5 three out of seven and Fi1 and Fi2 only two out of seven. However, as in only 30% of the cases the recognition rate is lower in T2 than in T1, the students' performances show improvement and retention in learning. Hence, Table 2 clearly shows the complexity of the data, whereas Figure 1 simplifies it somewhat.

This raises the question as to why, then, are some questions (S) produced significantly more successfully than others (NS)? There are various factors that affect the success of the productions. For example, the possible explanations can be found in the syntax and the frequency of use of these constructions. Also fluency can offer an explanation; for instance, if a speaker struggles reading aloud a sentence (i.e. hesitates and pauses), it is difficult to produce acceptable sentence prosody.

The unsuccessfully produced questions in NS are in fact rather different from each other. For example, Q1 (*U tebya sovest' yest'*) is an idiomatic expression and resembles a rhetorical question. Another possible word order for a *yest'*-question would be *u tebya yest' sovest'*. This is more neutral and therefore might be used more often (Lobanova & Gorbachik 1976:6). It is possible students pronounced a pitch contour according to this word order which brought forward the nucleus. Perhaps this was then interpreted by the native speakers as a non-question, because the nucleus was on *sovest'* instead of *yest'*. To clarify, in order to be recognized as a question the nucleus should be on *yest'* despite the word order. The reasons for unsuccessful production of Q2 remain vague. The context *Allo! Sonya?* (Hello, is that Sonya?) clearly indicates an interrogative. However, as the line was in the very beginning of the second dialogue, the students were perhaps a bit out of touch or were merely using a L1 pitch contour. Using a L1 pitch contour in L2 could of course be an explanation for all the unsuccessful productions by the L2 speakers. Further analysis of the pitch contours will shed more light into this issue (Ullakonoja 2010). The unsuccessful production of Q7, on the other hand, can be also explained by its structure and lexicon: it is a very long question with some words that may be unfamiliar which may lead to the students focusing on the words rather than sentence prosody.

As mentioned above, the S questions were generally recognized well. Questions Q3 (*Chainyy ili stolovyy?)* and Q4 (*Ty rada za menya?*) are most likely to be interpreted as questions in any context perhaps due to their lexical content. Question Q5 (*Da?*) on the other hand is very frequent in everyday Russian, as a result students would have heard it regularly while in Russia. Utterance Q6 (*Ty zabolela?*) is a short

question with a rather simple lexical content and syntax, which perhaps facilitated its production.

3 Results

3.1 Acceptability ratings

Next, I will focus on the second goal of the listening experiment, i.e. determining how good the judges thought each stimulus was as a question. The acceptability rating of the question was investigated by examining the ratings of each question and each student at the two recordings on a scale of one to five (1=not a question, 5=a question). First, Figure 2 shows that in general the mean acceptability ratings were not very high. The overall mean of all questions was 2.95.

It is somewhat unexpected, when the results of Figure 2 are compared with Figure 1 that only question Q4 (*Ty rada za menya?)* reaches a mean acceptability of over four, while others are on average either between three and four or around two, which could be verbalized as "not very acceptable as a question". From the results presented in Figure 1, one could anticipate that Q4, which had the highest recognition rate, would have received an acceptability rating of near five.

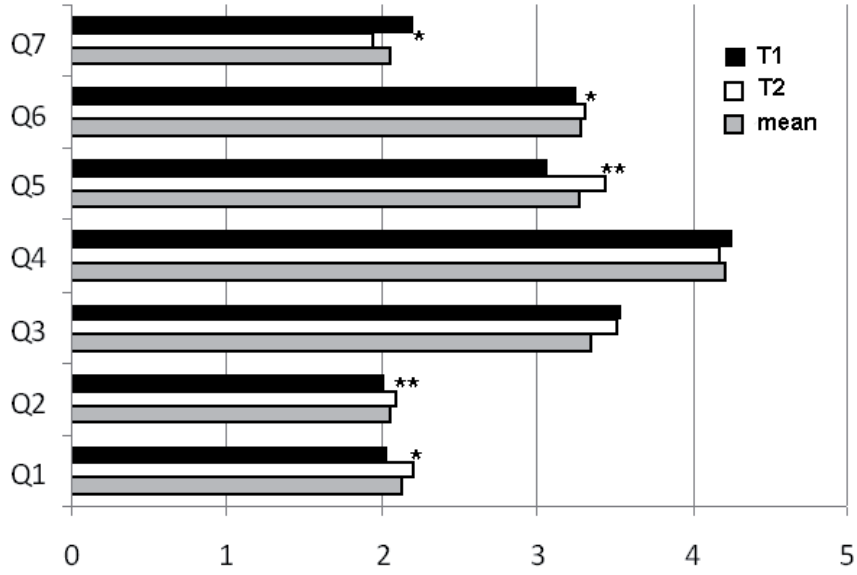


Figure 2: The mean acceptability rating of questions (n=84) during the stay in Russia (T1) and after it (T2) (0=not a question, 5=a question), *) $p<0.05$, **) $p>0.0001$.

The mean acceptability ratings (Figure 2) show a very similar pattern as the results from recognition rate measurement (Figure 1). The grouping of utterances into S and NS seems justifiable also in the acceptability ratings. In Figure 2 there are differences between T1 and T2 in acceptability ratings. The overall mean was slightly though reliably higher in T2 (2.99) than T1 (2.92), $\chi^2(16)=204.970$; $p<0.0001$. When the comparison of the acceptability ratings given by the judges for each question in the two recordings was further analyzed in Pearson's Chi Square test, it was found that for Q1, Q2, Q5 and Q6 the acceptability rating was reliably better in T2 than in T1.

Conversely, for Q7 it was significantly lower. Hence, the acceptability ratings show that the students were evaluated in general as producing more acceptable yes/no questions following their stay in Russia.

Table 3 presents the acceptability ratings for each question by individual speakers. Comparison of the two recordings yielded an unexpected result. The mean acceptability rating was lower for about half (23 out of 42) of the utterances in T2 than T1. However, for slightly less than a half (19 out of 42) it was higher. The comparison between the ratings of different questions shows, not surprisingly, that there is a tendency for some questions e.g. Q4 (*Ty rada za menya?*) to be rated as a fairly successful (3.6-4.8) question for all speakers. In other questions, e.g. Q2 (*Sonya?*) (1.1-4.3) and Q6 (*Ty zabolela?*) (1.3-4.7), there is more interspeaker variation, or variation between the two recordings. Thus, it cannot be concluded whether utterances other than Q4 would always be rated consistently.

Speaker	Question and category						
	Q1	Q2	Q3	Q4	Q5	Q6	Q7
	(NS)	(NS)	(S)	(S)	(S)	(S)	(NS)
Fi1 (T1)	3.8	1.7	3.2	**3.7**	4.6	4.5	3.2
Fi1 (T2)	2.5	1.1	2.7	**4.1**	4.5	4.3	3.4
Fi2 (T1)	**1.5**	**1.1**	3.8	4.5	1.1	**3.6**	3.6
Fi2 (T2)	**1.6**	**1.4**	2.4	4.1	1.1	**3.9**	1.1
Fi3 (T1)	**2.9**	**1.8**	**3.8**	4.3	**2.1**	4.7	2.5
Fi3 (T2)	**3.7**	**2.5**	**4.1**	4.0	**4.7**	3.7	2.4
Fi4 (T1)	1.6	**3.8**	**3.6**	**4.4**	4.7	**1.4**	**1.4**
Fi4 (T2)	1.5	**4.3**	**4.4**	**4.7**	4.5	**3.1**	**1.5**
Fi5 (T1)	**1.1**	2.6	**3.8**	**4.2**	4.3	3.8	**1.2**
Fi5 (T2)	**1.2**	1.8	**4.3**	**4.8**	3.6	3.2	**1.4**
Fi6 (T1)	**1.3**	1.1	3.2	4.5	**1.4**	1.3	**1.2**
Fi6 (T2)	**2.4**	1.1	3.1	3.6	**1.7**	1.3	**1.5**

Table 3. The mean acceptability rating of questions of the individual speakers during the stay in Russia (T1) and following it (T2). The figures in bold indicate a higher recognition rate in T2 than T1. (NS=non successful questions, S=successful questions as perceived by native speakers.)

As Table 3 shows, utterances that were, in general, recognized poorly as questions (NS) were still produced successfully by some speakers. For example, Q1 (*U tebya sovest' yest'*) is rated fairly good as a question for Fi1 (T1) and Fi3 (T2) and Q2 (*Sonya?*) for Fi4 (T1 and T2). When comparing different speakers, there is one student who received overall good ratings (Fi3, T2) and one who received consistently poor ratings (Fi6). As all learners succeeded in producing some of the questions well, this seems to show that they in principle knew how to produce the pitch contour of a yes/no question in Russian, but did not always succeed in doing so.

In Table 3 the improvement between T1 and T2 is indicated in bold. Improvement is observed in about half of the (21 out of 42) cases. In three out of 42 cases there is no difference between T1 and T2. The comparison of Table 2 and Table 3 indicates that the acceptability ratings sometimes show a different result than recognition rate. If examining the improvement, there are nine out of 42 cases where there is a difference in acceptability ratings compared to the recognition rate.

3.2 Reliability of the ratings

The results seem to suggest that there is variation in the ratings between the judges. In fact, the judges were rarely completely unanimous in their judgments. Despite this, as the values of Cohen's Kappa and Cronbach's alpha showed, the interjudge consistency was rather good. To further verify the ratings made by the judges, the categorization of the question and its acceptability rating was subjected to comparison. Table 4 presents the results obtained from the comparison of question categorization and acceptability rating of the interrogative without taking into the account the time of recordings (T1, T2). From this table it becomes apparent that the relationship between the two ratings is statistically significant. This indicates that the judges gave similar acceptability ratings to the questions whilst giving different ratings to the non-questions. Hence, it can be concluded that the ratings by the judges are reliable.

Speaker	Pearson's Chi Square correlation	df	p
Fi1	93.820	4	<0.0001
Fi2	87.136	4	<0.0001
Fi3	106.192	4	<0.0001
Fi4	108.742	4	<0.0001
Fi5	99.215	4	<0.0001
Fi6	70.813	4	<0.0001

Table 4: Pearson's Chi Square correlation between the question categorization and acceptability rating of the interrogative.

4 Discussion and conclusions

It has been shown that native speakers can read aloud the same written text with different pitch contours (Brazil 1984). It would therefore be incorrect to presume that even native speakers of Russian would always pronounce the sentences in exactly the same manner. They are however all prone to using an interrogative pitch contour, whereas Finnish learners have been shown to struggle with this, possibly because of the fact that their L1 lacks such contours.

The results of this study show that a great number of utterances intended as questions by Finns were not perceived as such by native speakers. This leads to a number of conclusions. As Hirvonen (1967:42) suggests, one explanation for this might be in the different approaches the individual judges took in accepting intonation that differs from native production. Some judges are perhaps more ready to accept non-native production whereas others are not. It should be remembered that in this study the majority of the judges were not acustomed to hearing foreign-accented Russian, which may have resulted in strict ratings. Furthermore, as the students were also speaking slower than native speakers (see Ullakonoja 2009) the judges could have been disturbed by the inappropriate temporal structure of the pitch contour. As Russian yes/no questions are spoken faster than declaratives (Svetozarova 1982:111-112), the judges could have favored the "non-question" rating in some cases where the students were speaking slowly. The variation between T1 and T2 could also reflect the difference in strictness of the two groups of judges, not only differences in the learners' productions.

Kuosmanen & de Silva (2003; 2007) and de Silva & Volskaya (2005) have shown, that Russian interrogatives are difficult for Finnish speakers. These findings are

further substantiated by this study: only half of the utterances intended as questions were recognized by 90% of the judges, and the acceptability ratings remained at an average level. The overall recognition rate of all recordings was 57%, slightly lower than in Kuosmanen & de Silva (2003; 2007). Hence, it can be concluded that even reasonably proficient Finns often fail to produce acceptable pitch contours in questions in Russian.

The results from the overall recognition rate (Figure 1) of individual questions, however, are not consistent with the earlier studies by Kuosmanen & de Silva (2003; 2007), who found that longer Russian questions produced by Finns were harder to recognize by native speakers than shorter questions. In this study, the long question (Q7, *Sudya po torzhestvennomu tonu, ty khochesh' soobshchit' mne nechto vazhnoye?*) was not the hardest to recognize, nor was the shortest question (Q5, *Da?*) the easiest. The contradictory results can partly be explained by the fact that in the present study the panel of judges consisted of a greater number of participants.

The general recognition rate of the questions was only slightly lower for items recorded in T2 than T1. The difference was small, but statistically significant. Therefore, it seems that in general the learners are not as good at producing yes-no/questions following their stay in Russia compared to during it. However, even after a month in Finland (T2), with hardly using or hearing any Russian, they are capable of reading the yes/no questions in the text not as well as but almost as successfully as during their stay in Russia (T1). Some explanations for the unsuccessful productions were offered on the basis of syntax and lexicon; however, a further acoustical analysis of the pitch contours is conducted for more evidence (see Ullakonoja 2010). The acceptability ratings yielded similar results as those obtained from the recognition rate analysis. There was great interspeaker and intraspeaker variation. The contradictory result is that for the acceptability ratings, a statistically significant improvement was observed from T1 to T2. This could suggest that there is in fact no "loss" but retention in learning.

To conclude, the findings of this study show that only one of the utterances (Q4, *Ty rada za menya?*) was consistently judged as a question. The other items displayed no general tendency either in the categorization task or the acceptability ratings. Great interspeaker differences were also found. The almost unanimous judgments of Q4 as a question can partly be explained by its lexical content as the utterance would likely be used more often as a question than a declarative in daily conversation.

To conclude, I would like to mention an interesting finding by Kuosmanen & de Silva (2003). They found that in the Russian yes/no question *Mozhno?*, Finnish students who used an incorrect final-rise instead of a correct rising-falling contour in the nuclear syllable were more likely to be recognized as pronouncing a question than those who used the correct contour. Thus, in the future, it is also important to study the pitch contours of the students' questions experimentally in order to determine what changes in F0 (and where) function as important cues to perceiving a question. It also remains to be explored how the differences between the realization of pitch contours in yes/no questions in Finnish and Russian affect learners' production, i.e. are the learners, for example, relying on L1 (or other L2s) when learning Russian L2 prosody.

References

Anttila, H. 2008. The effect of interrogative function on intonation in spontaneous and read Finnish. Master's thesis. Department of speech sciences, University of Helsinki. Available at: http://urn.fi/URN:NBN:fi-fe200807301740

Boersma, P. & Weenik D. 2009. *Praat: doing phonetics by computer.* (Versions 5.0-5.1) [computer program] Available at: www.praat.org

Brazil, D. 1984. *The intonation of sentences read aloud.* In: D. Gibbon & H. Richter (eds.). Intonation, accent and rhythm. Berlin: Walter de Gruyter. 46-66.

Bryman, A. & Cramer, D. 2001. *Quantitative Data Analysis with SPSS Release 10 for Windows: A Guide for Social Scientists.* Great Britain: Routledge.

Bryzgunova, E. A. 1977. *Zvuki i intonatsiya russkoy rechi.* Moscow: Russkiy yazyk.

de Silva, V. & Ullakonoja, R. 2009. *Introduction: Russian and Finnish in contact.* 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. 15-20.

de Silva, V. & Volskaya, N. B. 2005. *Mesto fonetiki v standartakh RKI i v praktike vladenya yazykom.* In: N. P. Isayev, A. Mustajoki & E. Protasova (eds.). Russkiy yazyk kak inostrannyy: Problemy izucheniya, prepodavaniya, otsenki v svete obshcheyevropeyskikh kompetentsiy vladeniya inostrannym yazykom. Moscow: MAKS Press. 61-73.

Hirvonen, P. 1967. *On the problems met by Finnish students in learning the rising interrogative intonation of English.* Publications of the Phonetics department of the University of Turku 2.

Hirvonen, P. 1970. *Finnish and English communicative intonation.* Publications of the Phonetics department of the University of Turku 8.

Igarashi, Y. 2006. *Intonational patterns in Russian interrogatives - phonetic analyses and phonological interpretations.* In: Y. Kawaguchi, I. Fónagy & T. Moriguchi (eds.). Prosody and syntax: Cross-linguistic perspectives. 175-196. Amsterdam: John Benjamins.

Iivonen, A. 1979. *Is there interrogative intonation in Finnish?* In: E. Gårding, G. Bruce & R. Bannert (eds.). Nordic prosody, papers from a symposium. Lund: Department of Linguistics, Lund University. 43-53.

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

Iivonen, A. 2001. *Intonation of Finnish questions.* W. A. van Dommelen & T. Fretheim (eds.) Nordic prosody: Proceedings of the VIIth conference Trondheim 2000. Frankfurt am Main: Peter Lang. 137-151.

Kuosmanen, A. & de Silva, V. 2003. *Why don't Russians answer my questions? Finnish students' problems in producing Russian interrogative intonation.* In: M. J. Solé, D. Recasens & J. Romero (eds.). Proceedings of the 15th International Congress of Phonetic Sciences: Barcelona 3-9 August 2003. 523-526.

Kuosmanen, A. & de Silva, V. 2007. *Problemy finskih studentov pri ovladenii russkoy voprositel'noj intonaciey.* In: R. F. Kasatkina (ed.). Problemy fonetiki V. Moscow: Nauka. 297-307.

Landis, J. R. & Koch, G. G. 1977. The measurement of observer agreement for categorical data. *Biometrics* 33(1). 159-174.

Lieberman, P. 1965. On the acoustic basis of the perception of intonation by linguists. *Word* 21. 40-54.

Lobanova, N. A. & Gorbachik A. L. 1976. Poryadok slov v russkom yazyke. Moscow: Izdatel'stvo moskovskogo universiteta.

Meyer, R. & Mleinek, I. 2006. How prosody signals force and focus - A study of pitch accents in Russian yes-no questions. *Journal of Pragmatics* 38(10). 1615-1635.

Mixdorff, H., Vainio, M., Werner, S. & Järvikivi, J. 2002. *The manifestation of linguistic information in prosodic features of Finnish.* In: B. Bel & I. Marlien (eds.). Proceedings of the Speech Prosody 2002, Aix-en-Provence 11-13 April 2002. 515-518.

Ogden, R. & Routarinne, S. 2005. The communicative functions of final rises in Finnish intonation. *Phonetica* 62. 160-175.

Shilova, K. A. & Usmanova, E. E. 1990. *100 dialogov po telefonu.* Moscow: Russkiy yazyk.

Shcherbakova, L. P. 2001. *Vospriyatie intonatsionnykh tipov nerodnogo yazyka.* In: L. V. Bondarko (ed.). 100 let eksperimental'noy fonetike v Rossii. Materialy mezhdunarodnoy konferentsii 1-4 fevralya 2001 g. St. Petersburg: Filologicheskiy fakul'tet SPbGU. 173-175.

Svetozarova, N. D. 1982. *Intonatsionnaya sistema russkogo yazyka.* Leningrad: Izdatel'stvo Leningradskogo universiteta.

Svetozarova, N. D. 1998. *Intonation in Russian*. In: D. Hirst & A. Di Cristo (eds.). Intonation systems – a survey of twenty languages. Cambridge: Cambridge University Press. 261-274.
Toivanen, J. H. 2001. *Perspectives on intonation: English, Finnish and English spoken by Finns*. Frankfurt am Main: Peter Lang.
Ullakonoja, R. 2008. Pausing as an indicator of fluency in the Russian of Finnish learners. In: P. A. Barbosa, S. Madureira & C. Reis (eds.). *Proceedings of the Speech Prosody 2008 conference*. Campinas, Brazil: Editora RG/CNPq. 339-342.
Ullakonoja, R. 2009. Speech rate as an indicator of fluency in the Russian of Finnish learners. In: M. O'Dell & T. Nieminen (eds.). *Fonetiikan päivät 2008 – the Phonetics symposium 2008*. Tampere studies in language, translation and culture, series B. 97-109. Available at: http://tampub.uta.fi/tup/978-951-44-7580-1.pdf.
Ullakonoja, R. 2010. Pitch contours in Russian yes/no questions by Finns. In: M. Hasegawa-Johnson, A. Bradlow, J. Cole, K. Livescu, J. Pierrehumbert & C. Shih (eds.) *Proceedings of the Speech Prosody 2010 conference*. Available at: http://speechprosody2010.illinois.edu/papers/100072.pdf
Vaissière, J. 2005. *Perception of intonation*. In: D. B. Pisoni & R. E. Remez (eds.). The handbook of speech perception. Malden, MA; Oxford, Carlton: Blackwell publishing. 236-262.
Volskaya, N. B. 2009. *Aspects of Russian intonation*. In: V. de Silva & R. Ullakonoja (eds.). Phonetics of Russian and Finnish. General introduction. Spontaneous and read-aloud speech. Peter Lang. 133-143.
Yokoyama, O. 2001. A reinterpretation of the IK system in Russian. *Die Welt Der Slave*n XLVI. 1-26.

VI

PITCH CONTOURS IN RUSSIAN YES/NO QUESTIONS BY FINNS

by

Riikka Ullakonoja 2010

Proceedings of the Speech Prosody 2010 Conference, Chicago, USA. Hasegawa-Johnson, Mark, Bradlow, Ann, Cole, Jennifer, Livescu, Karen, Pierrehumbert, Janet & Shih, Chilin (eds.).

Pitch contours in Russian yes/no questions by Finns

Riikka Ullakonoja

Department of Languages, University of Jyväskylä, Finland
riikka.ullakonoja@jyu.fi

Abstract

The aim of this paper is to determine the pitch contours Finns use when uttering yes/no questions in Russian. In addition, the pitch contours will be compared to native speech as well as subjected to native speaker evaluation. So far, there has been very little research on the prosody of Russian as a second language. L1 Finnish students are an interesting group to study because intonation in Finnish is not distinctive whereas in Russian it is.

Index Terms: Intonation, Russian, second language

1. Introduction

One of the challenges of L2 (second language) intonation research is identifying what constitutes acceptable variation of pitch contours for non-native speakers. In this study acoustic measurements were made of the pitch contours of Finnish learners which were then subjected to perceptual evaluation by native speakers. The focus is on yes/no questions in Russian that may differ from statements solely by prosodic means. Comparison of Finnish speakers' (L2 learners of Russian) Russian with that of L1 Russian speakers is interesting because in Finnish yes/no questions are marked with the interrogative particle, -ko/-kö, rather than prosody.

Unfortunately, no extensive Finnish-Russian comparative study on native speakers' pitch contours exists. Finnish interrogative intonation is not distinctive. A fall or rise-fall have been said to be common contours for yes/no questions in Finnish read-aloud speech [1] but a typical pitch pattern for such questions has not been defined. In Russian, on the other hand, yes/no questions can take the same lexical and syntactic form as declaratives, but the difference is realized by prosody. Yes/no questions commonly have the so called IK-3 pattern of Bryzgunova's [2] theory. The IK-3 is a pitch pattern that has a sharp rise on the nuclear syllable or, if the nucleus is in the phrase final position, a final rising pitch contour. In addition to that the prenuclear part is typically higher than the postnuclear part [3,4].

Finnish speakers' intonation in Russian has been a subject of a few previous studies. They have concluded that Finnish learners encounter difficulties in pronouncing Russian yes/no questions [5,6]. The present paper is a partial replication of those studies. In this study, however, Finnish students' Russian read-aloud speech will be analyzed acoustically in order to identify the pitch contours they use in Russian yes/no questions. The students' productions will be compared to each other as well as to native Russian speech. The students' speech will be rated by native speakers of Russian.

Table 1. *The yes/no questions (in bold) of the data in context* [7].

	Russian	Russian (in Roman alphabet)	English translation
Q7	А. Со́фья Па́вловна, я вас приве́тствую! В. Почему́ так до́лго не звони́ла? **Су́дя по торже́ственному то́ну, ты хо́чешь сообщи́ть мне не́что ва́жное?**	A. Sof'ya Pavlovna, ya vas privetstvuyu! B. Pochemu tak dolgo ne zvonila? **Sudya po torzhestvennomu tonu, ty khochesh' soobshchit' mne nechto vazhnoye?**	A. Sof'ya Pavlovna, I greet you! B. Why haven't you called me in such a long time? **Judging by your festive tone of voice, you want to tell me something important?**
Q4	А. Уга́дала! Мы сего́дня купи́ли ме́бельный гарниту́р . В. Ну, с тобо́й не соску́чишься! А. **Ты ра́да за меня́?**	A. Ugadala! My segodnya kupili mebel'nyy garnitur. B. Nu, s toboy ne soskuchish'sya. A. **Ty rada za menya?**	A. That's right! We bought a new suite of furniture today. B. Well, you are no bore. A. **Are you happy for me?**
Q5	В. Коне́чно, ра́да. Кста́ти, у меня́ то́же в кварти́ре есть ко́е-что́ но́венькое. А. **Да́?** Что́ же?	B. Konechno, rada. Kstati, u menya tozhe v kvartire yest' koye-chto noven'koye. A. **Da?** Chto-zhe?	B. Of course I'm happy. By the way, I also have something new in my apartment. A. **Yes?** What is it?
Q3	В. Мы купи́ли серви́з. А. **Ча́йный или столо́вый?** (1st dialogue continues…)	B. My kupili serviz. A.**Chainyy ili stolovyy?**	B. We bought a service. A. **Tea or dinner?** (1st dialogue continues…)
Q2	А. Алло́! **Со́ня?**	A. Allo! **Sonya?**	A. Hello! **Is that Sonya?**
Q1	В. Ната́ша, ты? Наконе́ц-то! Что́ случи́лось? Я жду це́лый день. Ты́ же обеща́ла позвони́ть у́тром. Я не отхожу́ от телефо́на. **У тебя́ со́весть е́сть?**	B. Natasha, ty? Nakonets-to! Chto sluchilos'? Ya zhdu tselyy den'. Ty zhe obeshchala pozvonit' utrom. Ya ne otkhozhu ot telefona. **U tebya sovest' yest'?**	B. Natasha, is that you? Finally! What happened? I have waited all day. After all you promised to call me this morning. I haven't moved away from the phone. **Have you no conscience?**
Q6	А. Со́нечка, дай хоть сло́во вста́вить! У меня́ несча́стье ! В. У́жас ! Что́-то случи́лось? Так я и зна́ла. Я пря́мо как чу́вствовала. **Ты заболе́ла?** Или что-то с детьми́? (2nd dialogue continues…)	A. Sonechka, day khot' slovo vstavit'! U menya neschast'ye! B. Uzhas! Chto-to sluchilos'? Tak ya i znala. Ya pryamo kak chustvovala.**Ty zabolela?** Ili chto-to s det'mi?	A. Sonechka, let me get a word in! I have had a bit of bad luck! B. That's awful! Did something happen? I thought it must have. I almost felt it. **Were you taken ill?** Or was it something to do with the children? (2nd dialogue continues…)

2. Material and methods

2.1. Speech data

Six native Finnish female university students (Fi1-Fi6) (aged 19-25) and six native Russian female university students (Ru1-Ru6) (aged 19-26) were recorded reading aloud two Russian dialogues in pairs. The Finnish subjects had studied Russian for 3 years prior to university entry and for one year at university as their major. They started their second year with a 3.5-month stay in Russia. This was the longest stay in Russia any of these students had experienced. The dialogues were drawn from Russian as a foreign language teaching material on telephone conversations (texts 46 and 100) [7]. These dialogues form a data set used in other research by the present author and thus they were not chosen on the basis of their phonetic content. The Finns were recorded twice during their second year at university: once during (T1) and once after (T2) their stay in Russia. The recordings of the Russian speakers and Finnish speakers at T1 were made with a Sony TDC-D3 DAT recorder and Roland Edirol 24-bit Wave/MP3 R-09 digital recorder with a Sony ECM-959A microphone in a quiet hotel room. The T2 recordings were done with a computer (program Adobe Audition 1.0 and 2.0) and AKG GN30 microphones. All yes/no questions in the data that can be understood as declaratives if a different pitch contour is used were analysed in this study (Table 1).

2.2. Listening experiment data

To evaluate the successfulness of the learners' productions, the L2 speech data described above were played to a total of 40 native Russian listeners in two groups (one group for the T1 and one for the T2 recordings). The stimuli were presented to the listeners only once in the same randomized order with an 11-second interstimulus interval. After each stimulus the judges were to define whether they heard a question or not. Thus, a recognition rate (RR) for each stimulus was obtained (percentage of positive ratings was calculated). In general, only 57% of the utterances intended as questions by Finns were recognized as such by the Russian speakers. The RR varied across the different questions so that learner productions of Q1, Q2 and Q7 were only recognized as questions in less that 40% of cases, whereas the RR for Q5 and Q6 was nearly 60% and for Q3 and Q4 over 80%. (see [8] for a more detailed description of the task and analysis of the results).

2.3. Methods of acoustic analysis

The annotation and analysis were performed using the Praat program [9]. Pitch calculations were done with Praat's autocorrelation method [10] for the entire sound file of the speaker and checked manually. Before that, the other speaker's voice was removed from the file. The same settings were used for all the speakers (Time step automatic, pitch floor 75 Hz, ceiling 600 Hz). The calculations were influenced by creaky voice, which has previously been found to be a typical feature of Finnish [11] (but not Russian). Creaky voice was manually corrected in PitchObject, if the correction was unmistakably possible on the basis of the spectrogram. In most cases it was not possible and the pitch was unvoiced for that part of the file in which case creaky voice was annotated in the TextGrid.

Pitch was measured in semitones (ST). A script was written by the author to measure mean and maximum pitch, pitch range and standard deviation, and mean absolute slope in pitch of each question as well as automatically to draw a picture of the pitch contour for further analysis. Pitch contours were compared both visually and through calculation in SPSS. Furthermore, the RR for each stimulus obtained in the question evaluation task (see 2.2.) was compared against the pitch measurements.

3. Results

3.1. Shape of the L1 and L2 pitch contours

First, the shape of the pitch contours was studied question by question. In Q1 (*U tebya sovyest yest'?*), all the L1 speakers produced a contour similar to that in Figure 1 (peak and a final rise on *yest'*). None of the L2 speakers showed such a contour. Instead most of them produced a peak on the word *sovyest* (on the first, or more often, on the second syllable) together with a fall or creaky voice on *yest'* (Figure 2). However, in the three L2 contours that were clearly recognized as questions by most natives there was a rise in *yest'* whereas the other L2 patterns lacked this. To summarize, in Q1, it was the rise on the final word that signalled a question to the native ear.

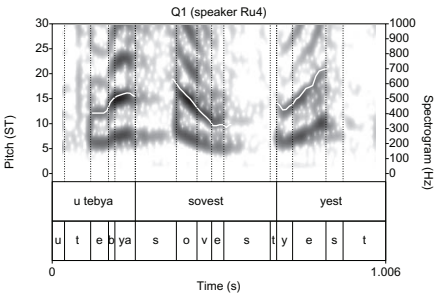


Figure 1: *Example of a native contour in Q1.*

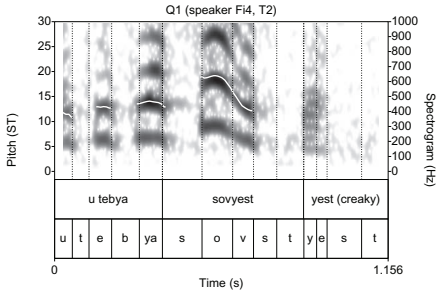


Figure 2: *Example of a non-native contour in Q1 (RR=10%).*

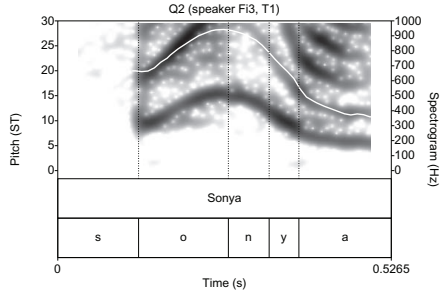


Figure 3: *Example of a non-native contour in Q2 (RR 21%).*

Q2 (*Sonya?*) was produced with three different pitch contours by the L1 speakers. There was a sharp or very flat peak or a rise on the last syllable. It was somewhat unexpected that the change occurred during the last syllable, as it is traditionally believed that it should be on the stressed syllable (i.e. here the first one). In the L2 speech, the final rising contour was the only one that was recognized as a question by the majority of the native listeners. The utterances where the L2 speaker had a peak on the first syllable were usually not recognized as questions (Figure 3). Hence, for L2 speakers it is not enough to produce a high peak, but it also needs to be sharp enough and placed on a correct syllable.

Overall in Q3 and Q4 the L2 speakers' pitch contours were very often recognized as questions. In Q3 (*Chaynyy ili stolovyy?*) all the L1 speakers had a sharp peak on the first syllable of *chaynyy*, while 4/6 then had a rise on the final syllable of the last word and 2/6 had a fall. The L2 speakers, on the other hand, rarely produced a rising contour on *stolovyy*. The most important cue for perceiving Q3 as a question would seem to be a peak or a rising contour on the word *chaynyy*. In Q4, most of the L1 speakers produced a contour similar to Figure 4. Most L2 speakers also had a very similar contour and Q4 utterances were mostly recognized as questions by native speakers. However, most L1 speakers had the peak on *za menya* whereas most L2 speakers preferred to place the peak on *rada.*

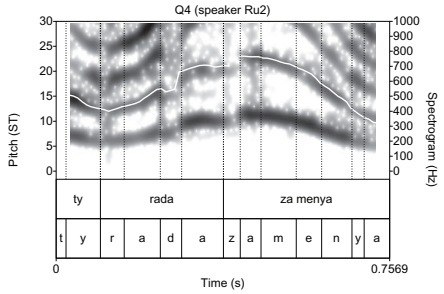


Figure 4: *Example of a native contour in Q4.*

In Q5 (Da?) 4/6 of the L1 speakers had a rise and 2/6 a peak. Most of the successful L2 speakers also had a rise. The utterances that were not recognized as questions by the majority had e.g. a flat tone, a small peak or a small rise.

In Q6 (*Ty zabolela?*), most of the L1 speakers had a flat contour with a peak on the stressed syllable *lel* (Figure 5). When the L2 speakers produced a similar contour, it was recognized as a question. Also, when the L2 speakers had a peak on the unstressed syllable *bo* or a final rise on *la*, they were recognized as questions by the majority of the L1 speakers, providing the peak was high enough. Figure 6 gives an example of a rather native-like pitch contour in L2 speech, which, however, received a rather low RR. In this example the peak is slightly flatter, lower and earlier than in the L1 utterances, which, perhaps, resulted in the mediocre RR.

Finally, Q7 (*Sudya po torzhestvennomu tonu, ty khochesh' soobshchit' mne nechto vazhnoye?*) is interesting, because it is a rather long utterance. Here, the L2 speakers' productions were rather poorly recognized as questions. Most of the L1 speakers had two sharp peaks in this utterance, on *tonu* and *vazhnoe* while the rest was rather flat. None of the L2 speakers had a similar contour. The L2 speaker whose utterances were best recognised as questions (RR=89%) had a peak on *tonu* but a final rise in *vazhnoe.* The rest of the successful L2 utterances had peaks on the two words (but sometimes on unstressed syllables) and in addition to that, often, on some other word of the utterance. The unsuccessful L2 realizations had peaks e.g. on *sudya* or *khochesh*. Hence, the L2 speakers had difficulties in deciding which word to accentuate and sometimes accentuated almost every word.

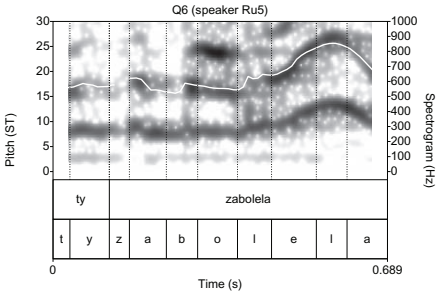


Figure 5: *Example of a native contour in Q6.*

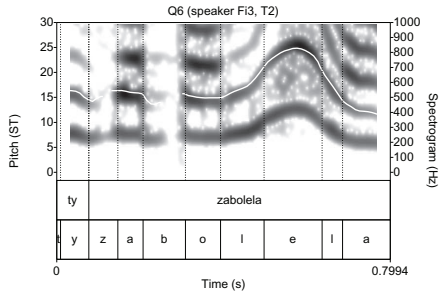


Figure 6: *Example of a non-native contour in Q6 (RR 55%)*

3.2. L1 and L2 pitch measurements

Second, pitch values (mean, maximum, std, mean absolute slope and range) were measured in ST and compared between L1 and L2. Also, L2 values were compared to the RR.

Table 2. *The mean pitch values (ST).*

Pitch (ST)	L1 (n=42)	L2 (n=84)	t-test results
Mean	14.53	13.28	t (124)=-6.371, p=0.0001
Max	22.90	21.37	t (124)=-2.412, p=0.017
Std	3.62	3.46	t (124)=-0.685, p=0.495
Range	13.13	12.47	t (124)=-0.705, p=0.482
Slope	53.65	43.34	t (124)=-2.700, p=0.008

Table 2 shows the pitch measurements for the L1 and L2 speakers. All the measured values are higher for the L1 speakers. The L1 speakers speak with a higher pitch, have more variation, a wider range and deeper slopes than the L2 speakers. The statistical significance of the differences was tested by independent samples t-test which yielded the result that mean and maximum pitch and mean absolute slope were significantly different between the two groups. The statistical significances were also tested between T1 and T2 for the L2 speakers, but although mean pitch and mean absolute slope were closer to the L1 values in T2, no significant difference was found.

The correlation between the RR and pitch measurements was calculated for the L2 speakers and verified in scatter plot graphs. Pearson's correlation coefficients (n=84) yielded the following correlations: RR&mean pitch=0.150 (p=0.172),

RR&max pitch=0.221 (p=0.042), RR&slope=0.055 (p=0.621), RR&range=0.127 (p=0.250), RR&std=0.359 (p=0.001). The only statistically significant correlations were between RR and std, and RR and max pitch, but they were very weak. Hence, it can be concluded that no single pitch measurement used here contributes significantly to the recognition of the utterance as a successful interrogative.

4. Discussion and Conclusions

This paper set out to determine what kinds of pitch contours Finns use in Russian yes/no questions as compared to L1 speakers. However, the study also showed that the native speakers sometimes produced pitch contours in a way that did not fully correspond to the norms presented in Bryzgunova's [2] theory of Russian intonation patterns. According to the theory, in IK-3 constructions, the peak should always be on the stressed syllable. However, this study showed that in L1 Russian the most important pitch movements are often realized on the unstressed syllable following the stressed one. This phenomenon has been described in Gussenhoven's [12] theory of Biological Codes as a delayed peak. According to him, this can function as an equally important cue for interrogativity as raised F0, because late peak is perceived as having a higher pitch. Peak delay has also been observed previously in Russian in IK-3 [13].

Consistent with previous studies [5,6] this study has also shown that Finnish L2 speakers of Russian often fail to produce a pitch contour similar to that of L1 speakers. This results sometimes in very low RRs. One of the new findings to emerge from this study is that the peak position seems to be the most important cue for perceiving yes/no questions produced by L2 speakers. It was most difficult for L2 speakers to produce in Q1. Other important cues were the sharpness of the peak and its height. These results cannot be applied to all languages, as e.g. Toivanen [14] found that Finnish students did not encounter difficulties in peak placement in English, despite the fact that it differed from peak placement in their L1.

The second major finding was that L1 and L2 yes/no questions differ in mean pitch and mean absolute slope. A statistically significant difference in mean pitch has been established earlier [15], but mean absolute slope reveals the differences in the variability of L1 and L2 pitch contours. It needs to be pointed out, however, that there was a lot of variation in the L2 group. There were cases where the L2 speaker produced a near-native-like contour, which was, consequently, recognized as a question by all of the L1 listeners. Furthermore, in other cases the L2 pitch contour was rather close to the L1 contour, yet the utterance received a low RR (likely due to peak placement and the height of the peak).

Although it is known that other prosodic factors, such as intensity, also contribute to the perception of question, in this study the pitch contours were used as the sole acoustic indicator of interrogativity. The research could be further supplemented by more detailed analysis of peak placement. Furthermore, it would be useful to study intensity, speech rate and pausing and their relationship to the perception of a question. For example, van Heuven & van Zanten [16] found that, in addition to higher pitch, questions differ from statements by their faster speech rate, whereas House [17] in turn found that pausing affected question perception.

These findings enhance our understanding of L2 prosody in Russian. Furthermore, they indicate a specific theme L2 teachers of Russian could usefully focus on in order to improve their students' speech and thus raise their competence in Russian oral interaction.

5. Acknowledgements

The author would like to acknowledge the help of Docent Nina B. Volskaya from St. Petersburg State University with the data collection for the listening experiment.

6. References

[1] Anttila, Hanna. 2008. *The effect of interrogative function on intonation in spontaneous and read Finnish.* Master's thesis. Department of speech sciences, University of Helsinki. http://urn.fi/URN:NBN:fi-fe200807301740 (10 Feb 2010)

[2] Bryzgunova, Elena A. 1977. *Zvuki i intonaciya russkoy rechi*. Moscow: Russkiy yazyk.

[3] Svetozarova, Natalia D. 1998. Intonation in Russian. In Hirst, Daniel & Albert Di Cristo (eds.), *Intonation systems: A survey of twenty languages*, 261-274. England: Cambridge University Press.

[4] Volskaya, Nina B. 2009. Aspects of Russian intonation. In de Silva, Viola & Riikka Ullakonoja (eds.), *Phonetics of Russian and Finnish: General introduction, spontaneous and read-aloud speech,* 133-143. Frankfurt: Peter Lang.

[5] Kuosmanen, Anne & de Silva, Viola. 2007. Problemy finskih studentov pri ovladeniy russkoy voprositel'noy intonaciey. In Kasatkina Rozaliya F. (ed.), *Problemy Fonetiki V,* 297-307. Moscow: Nauka.

[6] Kuosmanen, Anne & de Silva, Viola. 2003. Why don't Russians answer my questions? Finnish students' problems in producing Russian interrogative intonation. In Solé Maria-Josep & Daniel Recasens & Joaquín Romero (eds.), *Proceedings of the 15th ICPhS, Barcelona 3-9 August 2003,* 523-526.

[7] Shilova, Klara A. & Usmanova Elena E. 1990. *100 dialogov po telefonu*, Moscow: Russkiy yazyk.

[8] Ullakonoja, Riikka. (to appear) How do native Russian speakers evaluate yes/no questions produced by Finnish L2 learners? *Rice Working Papers in Linguistics,* Vol. 2.

[9] Boersma, Paul & Weenik, David. *Praat: doing phonetics by computer* (Version 5.1) http://www.praat.org, (27 Feb 2009).

[10] Boersma, Paul. 1993. Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound, *IFA Proceedings 17*, 97-110. http://www.fon.hum.uva.nl/paul/papers/Proceedings_1993.pdf, (12 Nov 2009).

[11] Ogden, Richard. 2003. Voice quality as resource for the management of turn-taking in Finnish talk-in-interaction. In Solé Maria-Josep & Daniel Recasens & Joaquín Romero (eds.), *Proceedings of the 15th ICPhS, Barcelona 3-9 August 2003*, 123-126.

[12] Gussenhoven, Carlos. 2002. Intonation and interpretation: phonetics and phonology. In Bel, Bernard & Isabelle Marlien (eds.), *Proceedings of the Speech Prosody 2002*, Aix-en-Provence, 47-57.

[13] Volskaya, Nina B. 2007. Variativnost' intonacionnykh modeley v spontannoy rechi i chtenii: intonacionny center. In Vinyaminovna, Mirra (ed.), *Materialy XXXVI Mezhdunarodnoy filologicheskoy konferencii, 12-27 Marta 2007*, Vypusk 17, Fonetika, 16-25.

[14] Toivanen, Juhani H. 2001. Perspectives on intonation: English, Finnish and English spoken by Finns. Frankfurt am Main: Peter Lang.

[15] Ullakonoja, Riikka. 2007. Comparison of pitch range in Finnish (L1) and Russian (L2). In Trouvain, Jürgen & William J. Barry (eds.), *Proceedings of the 16th ICPhS, Saarbrücken, 6-10 August 2007*, 1701-1704.

[16] van Heuven, Vincent J. & van Zanten, Ellen. 2005. Speech rate as a secondary prosodic characteristic of polarity questions in three languages, *Speech Communication* 47(1-2), 87-99.

[17] House, David. 2003. Perceiving question intonation: the role of pre-focal pause and delayed focal peak. In Solé Maria-Josep & Daniel Recasens & Joaquín Romero (eds.), *Proceedings of the 15th ICPhS, Barcelona 3-9 August 2003,* 755-758.

JYVÄSKYLÄ STUDIES IN HUMANITIES

1 KOSTIAINEN, EMMA, Viestintä ammattiosaamisen ulottuvuutena. - Communication as a dimension of vocational competence. 305 p. Summary 4 p. 2003.

2 SEPPÄLÄ, ANTTI, Todellisuutta kuvaamassa – todellisuutta tuottamassa. Työ ja koti televisio ja vähän radionkin uutisissa. - Describing reality – producing reality. Discourses of work and home in television and on a small scale in radio news. 211 p. Summary 3 p. 2003.

3 GERLANDER, MAIJA, Jännitteet lääkärin ja potilaan välisessä viestintäsuhteessa. - Tensions in the doctor-patient communication and relationship. 228 p. Summary 6 p. 2003.

4 LEHIKOINEN, TAISTO, Religious media theory - Understanding mediated faith and christian applications of modern media. - Uskonnollinen mediateoria: Modernin median kristilliset sovellukset. 341 p. Summary 5 p. 2003.

5 JARVA, VESA, Venäläisperäisyys ja ekspressiivisyys suomen murteiden sanastossa. - Russian influence and expressivity in the lexicon of Finnish dialects. 215 p. 6 p. 2003.

6 USKALI, TURO, "Älä kirjoita itseäsi ulos" Suomalaisen Moskovan-kirjeenvaihtajuuden alkutaival 1957–1975. - "Do not write yourself out" The beginning of the Finnish Moscow-correspondency in 1957–1975. 484 p. Summary 4 p. 2003.

7 VALKONEN, TARJA, Puheviestintätaitojen arviointi. Näkökulmia lukioikäisten esiintymis- ja ryhmätaitoihin. - Assessing speech communication skills. Perspectives on presentation and group communication skills among upper secondary school students. 310 p. Summary 7 p. 2003.

8 TAMPERE, KAJA, Public relations in a transition society 1989-2002. Using a stakeholder approach in organisational communications and relation analyses. 137 p. 2003.

9 EEROLA, TUOMAS, The dynamics of musical expectancy. Cross-cultural and statistical approaches to melodic expectations. - Musiikillisten odotusten tarkastelu kulttuurienvälisten vertailujen ja tilastollisten mallien avulla. 84 p. (277 p.) Yhteenveto 2 p. 2003.

10 PAANANEN, PIRKKO, Monta polkua musiikkiin. Tonaalisen musiikin perusrakenteiden kehittyminen musiikin tuottamis- ja improvisaatiotehtävissä ikävuosina 6–11. - Many paths to music. The development of basic structures of tonal music in music production and improvisation at the age of 6–11 years. 235 p. Summary 4 p. 2003.

11 LAAKSAMO, JOUKO, Musiikillisten karakterien metamorfoosi. Transformaatio- ja metamorfoosiprosessit Usko Meriläisen tuotannossa vuosina 1963–86. - "Metamorphosis of musical characters". Transformation and metamorphosis processes in the works of Usko Meriläinen during 1963–86. 307 p. Summary 3 p. 2004.

12 RAUTIO, RIITTA, *Fortspinnungstypus* Revisited. Schemata and prototypical features in J. S. Bach's Minor-Key Cantata Aria Introductions. - Uusi katsaus kehitysmuotoon. Skeemat ja prototyyppiset piirteet J. S. Bachin kantaattien molliaarioiden alkusoitoissa. 238 p. Yhteenveto 3 p. 2004.

13 MÄNTYLÄ, KATJA, Idioms and language users: the effect of the characteristics of idioms on their recognition and interpretation by native and non-native speakers of English. - Idiomien ominaisuuksien vaikutus englannin idiomien ymmärtämiseen ja tulkintaan syntyperäisten ja suomea äidinkielenään puhuvien näkökulmasta. 239 p. Yhteenveto 3 p. 2004.

14 MIKKONEN, YRJÖ, On conceptualization of music. Applying systemic approach to musicological concepts, with practical examples of music theory and analysis. - Musiikin käsitteellistämisestä. Systeemisen tarkastelutavan soveltaminen musikologisiin käsitteisiin sekä käytännön esimerkkejä musiikin teoriasta ja analyysistä. 294 p. Yhteenveto 10 p. 2004.

15 HOLM, JAN-MARKUS, Virtual violin in the digital domain. Physical modeling and model-based sound synthesis of violin and its interactive application in virtual environment. - Virtuaalinen viulu digitaalisella alueella. Viulun fysikaalinen mallintaminen ja mallipohjainen äänisynteesi sekä sen vuorovaikutteinen soveltaminen virtuaalitodellisuus ympäristössä. 74 p. (123 p.) Yhteenveto 1 p. 2004.

16 KEMP, CHRIS, Towards the holistic interpretation of musical genre classification. - Kohti musiikin genreluokituksen kokonaisvaltaista tulkintaa. 302 p. Yhteenveto 1 p. 2004.

17 LEINONEN, KARI, Finlandssvenskt sje-, tje- och s-ljud i kontrastiv belysning. 274 p. Yhteenveto 4 p. 2004.

18 MÄKINEN, EEVA, Pianisti cembalistina. Cembalotekniikka cembalonsoittoa aloittavan pianistin ongelmana. - The Pianist as cembalist. Adapting to harpsichord technique as a problem for pianists beginning on the harpsichord. 189 p. Summary 4 p. 2004.

19 KINNUNEN, MAURI, Herätysliike kahden kulttuurin rajalla. Lestadiolaisuus Karjalassa 1870-1939. - The Conviction on the boundary of two cultures. Laestadianism in Karelia in 1870-1939. 591 p. Summary 9 p. 2004.

20 Лилия Сиберг, "БЕЛЫЕ ЛИЛИИ". ГЕНЕЗИС ФИНСКОГО МИФА В БОЛГАРИИ. РОЛЬ РУССКОГО ФЕННОИЛЬСТВА. ФИНСКО-БОЛГАРСКИЕ КОНТАКТЫ И ПОСРЕДНИКИ С КОНЦА XIX ДО КОНЦА XX ВЕКА. 284 с. - "Belye lilii". Genezis finskogo mifa v Bolgarii. Rol' russkogo fennoil'stva. Finsko-bolgarskie kontakty i posredniki s konca XIX do konca XX veka. 284 p. Yhteenveto 2 p. 2004.

21 FUCHS, BERTOLD, Phonetische Aspekte einer Didaktik der Finnischen Gebärdensprache als Fremdsprache. - Suomalainen viittomakieli vieraana kielenä. Didaktinen fonetiikka. 476 p. Yhteenveto 14 p. 2004.

22 JÄÄSKELÄINEN, PETRI, Instrumentatiivisuus ja nykysuomen verbinjohto. Semanttinen tutkimus. - Instrumentality and verb derivation in Finnish. A semantic study. 504 p. Summary 5 p. 2004.

23 MERTANEN TOMI, Kahdentoista markan kapina? Vuoden 1956 yleislakko Suomessa. - A Rebellion for Twelve Marks? – The General Strike of 1956 in Finland. 399 p. Summary 10 p. 2004.

24 MALHERBE, JEAN-YVES, L'œuvre de fiction en prose de Marcel Thiry : une lecture d'inaboutissements. 353 p. Yhteenveto 1 p. 2004.

25 KUHNA, MATTI, Kahden maailman välissä. Marko Tapion *Arktinen hysteria* Väinö Linnan haastajana. - Between two worlds. Marko Tapio's Arktinen hysteria as a challenger to Väinö Linna. 307p. Summary 2 p. 2004.

26 VALTONEN, HELI, Minäkuvat, arvot ja mentaliteetit. Tutkimus 1900-luvun alussa syntyneiden toimihenkilönaisten omaelämäkerroista. - Self-images, values and mentalities. An autobiographical study of white collar women in twentieth century Finland. 272 p. Summary 6 p. 2004.

27 PUSZTAI, BERTALAN, Religious tourists. Constructing authentic experiences in late modern hungarian catholicism. - Uskontoturistit. Autenttisen elämyksen rakentaminen myöhäismodernissa unkarilaisessa katolisuudessa. 256 p. Yhteenveto 9 p. Summary in Hungarian 9 p. 2004.

28 PÄÄJOKI, TARJA, Taide kulttuurisena kohtaamispaikkana taidekavatuksessa. - The arts as a place of cultural encounters in arts education. 125 p. Summary 3 p. 2004.

29 JUPPI, PIRITA, "Keitä me olemme? Mitä me haluamme?" Eläinoikeusliike määrittelykamppailun, marginalisoinnin ja moraalisen paniikin kohteena suomalaisessa sanomalehdistössä. - "Who are we? What do we want?" The animal rights movement as an object of discursive struggle, marginalization and moral panic in Finnish newspapers. 315 p. Summary 6 p. 2004.

30 HOLMBERG, JUKKA, Etusivun politiikkaa. Yhteiskunnallisten toimijoiden representointi suomalaisissa sanomalehtiuutisissa 1987–2003. - Front page politics. Representation of societal actors in Finnish newspapers' news articles in 1987-2003. 291 p. Summary 2 p. 2004.

31 LAGERBLOM, KIMMO, Kaukana Kainuussa, valtaväylän varrella. Etnologinen tutkimus Kontiomäen rautatieläisyhteisön elinkaaresta 1950 – 1972. - Far, far away, nearby a main passage. An ethnological study of the life spans of Kontiomäki railtown 1950 – 1972. 407 p. Summary 2 p. 2004.

32 HAKAMÄKI, LEENA, Scaffolded assistance provided by an EFL teacher during whole-class interaction. - Vieraan kielen opettajan antama oikea-aikainen tuki luokkahuoneessa. 331 p. Yhteenveto 7 p. 2005.

33 VIERGUTZ, GUDRUN, Beiträge zur Geschichte des Musikunterrichts an den Gelehrtenschulen der östlichen Ostseeregion im 16. und 17. Jahrhundert. - Latinankoulujen musiikinopetuksen historiasta itäisen Itämeren rannikkokaupungeissa 1500- ja 1600-luvuilla. 211 p. Yhteenveto 9 p. 2005.

34 NIKULA, KAISU, Zur Umsetzung deutscher Lyrik in finnische Musik am Beispiel Rainer Maria Rilke und Einojuhani Rautavaara. - Saksalainen runous suomalaisessa musiikissa, esimerkkinä Rainer Maria Rilke ja Einojuhani Rautavaara. 304 p. Yhteenveto 6 p. 2005.

35 SYVÄNEN, KARI, Vastatunteiden dynamiikka musiikkiterapiassa. - Counter emotions dynamics in music therapy. 186 p. Summary 4 p. 2005.

36 ELORANTA, JARI & OJALA, JARI (eds), East-West trade and the cold war. 235 p. 2005.

37 HILTUNEN, KAISA, Images of time, thought and emotions: Narration and the spectator's experience in Krzysztof Kieslowski 's late fiction films. - Ajan, ajattelun ja tunteiden kuvia. Kerronta ja katsojan kokemus Krzysztof Kieslowskin myöhäisfiktiossa. 203 p. Yhteenveto 5 p. 2005.

38 AHONEN, KALEVI, From sugar triangle to cotton triangle. Trade and shipping between America and Baltic Russia, 1783-1860. 572 p. Yhteenveto 9 p. 2005.

39 UTRIAINEN, JAANA, A gestalt music analysis. Philosophical theory, method, and analysis of Iegor Reznikoff's compositions. - Hahmoperustainen musiikkianalyysi. Hahmofilosofinen teoria, metodi ja musiikkianalyysi Iégor Reznikoffin sävellyksistä. 222 p. Yhteenveto 3 p. 2005.

40 MURTORINNE, ANNAMARI, *Tuskan hauskaa*! Tavoitteena tiedostava kirjoittaminen. Kirjoittamisprosessi peruskoulun yhdeksännellä luokalla. - Painfully fun! Towards reflective writing process. 338 p. 2005.

41 TUNTURI, ANNA-RIITTA, Der Pikareske Roman als Katalysator in Geschichtlichen Abläufen. Erzählerische Kommunikationsmodelle in *Das Leben des Lazarillo von Tormes*, bei Thomas Mann und in EinigenFinnischen Romanen. 183 p. 2005.

42 LUOMA-AHO, VILMA, Faith-holders as Social Capital of Finnish Public Organisations. - Luottojoukot – Suomalaisten julkisten organisaatioiden sosiaalista pääomaa. 368 p. Yhteenveto 8 p. 2005.

63 RUUSKANEN, ESA, Viholliskuviin ja viranomaisiin vetoamalla vaiennetut työväentalot. Kuinka Pohjois-Savon Lapuan liike sai nimismiehet ja maaherran sulkemaan 59 kommunistista työväentaloa Pohjois-Savossa vuosina 1930–1932. - The workers' halls closed by scare-mongering and the use of special powers by the authorities. 248 p. Summary 5 p. 2006.

64 VARDJA, MERIKE, Tegelaskategooriad ja tegelase kujutamise vahendid Väinö Linna romaanis "Tundmatu sõdur". - Character categories and the means of character representation in Väinö Linna's Novel *The Unknown Soldier*. 208 p. Summary 3 p. 2006.

65 TAKÁTS, JÓZSEF, Módszertani berek. Írások az irodalomtörténet-írásról. - The Grove of Methodology. Writings on Literary Historiography. 164 p. Summary 3 p. 2006.

66 MIKKOLA, LEENA, Tuen merkitykset potilaan ja hoitajan vuorovaikutuksessa. - Meanings of social support in patient-nurse interaction. 260 p. Summary 3 p. 2006.

67 SAARIKALLIO, SUVI, Music as mood regulation in adolescence. - Musiikki nuorten tunteiden säätelynä. 46 p. (119 p.) Yhteenveto 2 p. 2007.

68 HUJANEN, ERKKI, Lukijakunnan rajamailla. Sanomalehden muuttuvat merkitykset arjessa. - On the fringes of readership. The changing meanings of newspaper in everyday life. 296 p. Summary 4 p. 2007.

69 TUOKKO, EEVA, Mille tasolle perusopetuksen englannin opiskelussa päästään? Perusopetuksen päättövaiheen kansallisen arvioinnin 1999 eurooppalaisen viitekehyksen taitotasoihin linkitetyt tulokset. - What level do pupils reach in English at the end of the comprehensive school? National assessment results linked to the common European framework. 338 p. Summary 7 p. Sammanfattning 1 p. Tiivistelmä 1 p. 2007.

70 TUIKKA, TIMO, "Kekkosen konstit". Urho Kekkosen historia- ja politiikkakäsitykset teoriasta käytäntöön 1933–1981. - "Kekkonen´s way". Urho Kekkonen's conceptions of history and politics from theory to practice, 1933–1981 413 p. Summary 3 p. 2007.

71 Humanistista kirjoa. 145 s. 2007.

72 NIEMINEN, LEA, A complex case: a morphosyntactic approach to complexity in early child language. 296 p. Tiivistelmä 7 p. 2007.

73 TORVELAINEN, PÄIVI, Kaksivuotiaiden lasten fonologisen kehityksen variaatio. Puheen ymmärrettävyyden sekä sananmuotojen tavoittelun ja tuottamisen tarkastelu. - Variation in phonological development of two-year-old Finnish children. A study of speech intelligibility and attempting and production of words. 220 p. Summary 10 p. 2007.

74 SIITONEN, MARKO, Social interaction in online multiplayer communities. - Vuorovaikutus verkkopeliyhteisöissä. 235 p. Yhteenveto 5 p. 2007.

75 STJERNVALL-JÄRVI, BIRGITTA, Kartanoarkkitehtuuri osana Tandefelt-suvun elämäntapaa. - Manor house architecture as part of the Tandefelt family´s lifestyle. 231 p. 2007.

76 SULKUNEN, SARI, Text authenticity in international reading literacy assessment. Focusing on PISA 2000. - Tekstien autenttisuus kansainvälisissä lukutaidon arviointitutkimuksissa: PISA 2000. 227 p. Tiivistelmä 6 p. 2007.

77 KŐSZEGHY, PÉTER, Magyar Alkibiadés. Balassi Bálint élete. - The Hungarian Alcibiades. The life of Bálint Balass. 270 p. Summary 6 p. 2007.

78 MIKKONEN, SIMO, State composers and the red courtiers - Music, ideology, and politics in the Soviet 1930s - Valtion säveltäjiä ja punaisia hoviherroja. Musiikki, ideologia ja politiikka 1930-luvun Neuvostoliitossa. 336 p. Yhteenveto 4 p. 2007.

79 SIVUNEN, ANU, Vuorovaikutus, viestintäteknologia ja identifioituminen hajautetuissa tiimeissä. - Social interaction, communication technology and identification in virtual teams. 251 p. Summary 6 p. 2007.

80 LAPPI, TIINA-RIITTA, Neuvottelu tilan tulkinnoista. Etnologinen tutkimus sosiaalisen ja materiaalisen ympäristön vuorovaikutuksesta jyväskyläläisissä kaupunkipuhunnoissa. - Negotiating urban spatiality. An ethnological study on the interplay of social and material environment in urban narrations on Jyväskylä. 231 p. Summary 4 p. 2007.

81 HUHTAMÄKI, ULLA, "Heittäydy vapauteen". Avantgarde ja Kauko Lehtisen taiteen murros 1961–1965. - "Fling yourself into freedom!" The Avant-Garde and the artistic transition of Kauko Lehtinen over the period 1961–1965. 287 p. Summary 4 p. 2007.

82 KELA, MARIA, *Jumalan kasvot* suomeksi. Metaforisaatio ja erään uskonnollisen ilmauksen synty. - God's face in Finnish. Metaphorisation and the emergence of a religious expression. 275 p. Summary 5 p. 2007.

83 SAARINEN, TAINA, Quality on the move. Discursive construction of higher education policy from the perspective of quality. - Laatu liikkeessä. Korkeakoulupolitiikan diskursiivinen rakentuminen laadun näkökulmasta. 90 p. (176 p.) Yhteenveto 4 p. 2007.

84 MÄKILÄ, KIMMO, Tuhoa, tehoa ja tuhlausta. Helsingin Sanomien ja New York Timesin ydinaseuutisoinnin tarkastelua diskurssianalyyttisesta näkökulmasta 1945–1998.

- "Powerful, Useful and Wasteful". Discourses of Nuclear Weapons in the New York Times and Helsingin Sanomat 1945–1998. 337 p. Summary 7 p. 2007.

85 KANTANEN, HELENA, Stakeholder dialogue and regional engagement in the context of higher education. - Yliopistojen sidosryhmävuoropuhelu ja alueellinen sitoutuminen. 209 p. Yhteenveto 8 p. 2007.

86 ALMONKARI, MERJA, Jännittäminen opiskelun puheviestintätilanteissa. - Social anxiety in study-related communication situations. 204 p. Summary 4 p. 2007.

87 VALENTINI, CHIARA, Promoting the European Union. Comparative analysis of EU communication strategies in Finland and in Italy. 159 p. (282 p.) 2008.

88 PULKKINEN, HANNU, Uutisten arkkitehtuuri - Sanomalehden ulkoasun rakenteiden järjestys ja jousto. - The Architecture of news. Order and flexibility of newspaper design structures. 280 p. Yhteenveto 5 p. 2008.

89 MERILÄINEN, MERJA, Monenlaiset oppijat englanninkielisessä kielikylpyopetuksessa - rakennusaineita opetusjärjestelyjen tueksi. - Diverse Children in English Immersion: Tools for Supporting Teaching Arrangements. 197 p. 2008.

90 VARES, MARI, The question of Western Hungary/Burgenland, 1918-1923. A territorial question in the context of national and international policy. - Länsi-Unkarin/Burgenlandin kysymys 1918–1923. Aluekysymys kansallisen ja kansainvälisen politiikan kontekstissa. 328 p. Yhteenveto 8 p. 2008.

91 ALA-RUONA, ESA, Alkuarviointi kliinisenä käytäntönä psyykkisesti oireilevien asiakkaiden musiikkiterapiassa – strategioita, menetelmiä ja apukeinoja. – Initial assessment as a clinical procedure in music therapy of clients with mental health problems – strategies, methods and tools. 155 p. 2008.

92 ORAVALA, JUHA, Kohti elokuvallista ajattelua. Virtuaalisen todellisen ontologia Gilles Deleuzen ja Jean-Luc Godardin elokuvakäsityksissä. - Towards cinematic thinking. The ontology of the virtually real in Gilles Deleuze's and Jean-Luc Godard's conceptions of cinema. 184 p. Summary 6 p. 2008.

93 KECSKEMÉTI, ISTVÁN, Papyruksesta megabitteihin. Arkisto- ja valokuvakokoelmien konservoinnin prosessin hallinta. - From papyrus to megabytes: Conservation management of archival and photographic collections. 277 p. 2008.

94 SUNI, MINNA, Toista kieltä vuorovaikutuksessa. Kielellisten resurssien jakaminen toisen kielen omaksumisen alkuvaiheessa. - Second language in interaction: sharing linguistic resources in the early stage of second language acquisition. 251 p. Summary 9 p. 2008.

95 N. PÁL, JÓZSEF, Modernség, progresszió, Ady Endre és az Ady–Rákosi vita. Egy konfliktusos eszmetörténeti pozíció természete és következményei. 203 p. Summary 3 p. 2008.

96 BARTIS, IMRE, „Az igazság ismérve az, hogy igaz". Etika és nemzeti identitás Sütő András Anyám könnyű álmot ígér című művében és annak recepciójában. 173 p. Summary 4 p. 2008.

97 RANTA-MEYER, TUIRE, Nulla dies sine linea. Avauksia Erkki Melartinin vaikutteisiin, verkostoihin ja vastaanottoon henkilö- ja reseptiohistoriallisena tutkimuksena. - *Nulla dies sine linea*: A biographical and reception-historical approach to Finnish composer Erkki Melartin. 68 p. Summary 6 p. 2008.

98 KOIVISTO, KEIJO, Itsenäisen Suomen kanta-aliupseeriston synty, koulutus, rekrytointitausta ja palvelusehdot. - The rise, education, the background of recruitment and conditions of service of the non-commissioned officers in independent Finland. 300 p. Summary 7 p. 2008.

99 KISS, MIKLÓS, Between narrative and cognitive approaches. Film theory of non-linearity applied to Hungarian movies. 198 p. 2008.

100 RUUSUNEN, AIMO, Todeksi uskottua. Kansandemokraattinen Neuvostoliitto-journalismi rajapinnan tulkkina vuosina1964–1973. - Believed to be true. Reporting on the USSR as interpretation of a boundary surface in pro-communist partisan journalism 1964–1973. 311 p. Summary 4 p. 2008.

101 HÄRMÄLÄ, MARITA, Riittääkö *Ett ögonblick* näytöksi merkonomilta edellytetystä kielitaidosta? Kielitaidon arviointi aikuisten näyttötutkinnoissa. – Is *Ett ögonblick* a sufficient demonstration of the language skills required in the qualification of business and administration? Language assessment in competence-based qualifications for adults. 318 p. Summary 4 p. 2008.

102 COELHO, JACQUES, The vision of the cyclops. From painting to video ways of seeing in the 20th century and through the eyes of Man Ray. 538 p. 2008.

103 BREWIS, KIELO, Stress in the multi-ethnic customer contacts of the Finnish civil servants: Developing critical pragmatic intercultural professionals. – Stressin kokemus suomalaisten viranomaisten monietnisissä asiakaskontakteissa: kriittis-pragmaattisen kulttuurienvälisen ammattitaidon kehittäminen. 299 p. Yhteenveto 4 p. 2008.

104 BELIK, ZHANNA, The Peshekhonovs' Workshop: The Heritage in Icon Painting. 239 p. [Russian]. Summary 7 p. 2008.

105 MOILANEN, LAURA-KRISTIINA, Talonpoikaisuus, säädyllisyys ja suomalaisuus 1800- ja 1900-lukujen vaihteen suomenkielisen proosan kertomana. – Peasant values, estate society and the Finnish in late nineteenth- and early

and early twentieth-century narrative literature. 208 p. Summary 3 p. 2008.

106 PÄÄRNILÄ, OSSI, Hengen hehkusta tietostrategioihin. Jyväskylän yliopiston humanistisen tiedekunnan viisi vuosikymmentä. 110 p. 2008.

107 KANGASNIEMI, JUKKA, Yksinäisyyden kokemisen avainkomponentit Yleisradion tekstitelevision Nuorten palstan kirjoituksissa. - The key components of the experience of loneliness on the Finnish Broadcasting Company's (YLE) teletext forum for adolescents. 388 p. 2008.

108 GAJDÓ, TAMÁS, Színháztörténeti metszetek a 19. század végétől a 20. század közepéig. - Segments of theatre history from the end of the 19th century to the middle of the 20th century. 246 p. Summary 2 p. 2008.

109 CATANI, JOHANNA, Yritystapahtuma kontekstina ja kulttuurisena kokemuksena. - Corporate event as context and cultural experience. 140 p. Summary 3 p. 2008.

110 MAHLAMÄKI-KAISTINEN, RIIKKA, Mätänevän velhon taidejulistus. Intertekstuaalisen ja -figuraalisen aineiston asema Apollinairen L'Enchanteur pourrissant teoksen tematiikassa ja symboliikassa. - Pamphlet of the rotten sorcerer. The themes and symbols that intertextuality and interfigurality raise in Apollinaire's prose work L'Enchanteur pourrissant. 235 p. Résumé 4 p. 2008.

111 PIETILÄ, JYRKI, Kirjoitus, juttu, tekstielementti. Suomalainen sanomalehtijournalismi juttutyyppien kehityksen valossa printtimedian vuosina 1771-2000. - Written Item, Story, Text Element. Finnish print journalism in the light of the development of journalistic genres during the period 1771-2000. 779 p. Summary 2 p. 2008.

112 SAUKKO, PÄIVI, Musiikkiterapian tavoitteet lapsen kuntoutusprosessissa. - The goals of music therapy in the child's rehabilitation process. 215 p. Summary 2 p. 2008.

113 LASSILA-MERISALO, MARIA, Faktan ja fiktion rajamailla. Kaunokirjallisen journalismin poetiikka suomalaisissa aikakauslehdissä. - On the borderline of fact and fiction. The poetics of literary journalism in Finnish magazines. 238 p. Summary 3 p. 2009.

114 KNUUTINEN, ULLA, Kulttuurihistoriallisten materiaalien menneisyys ja tulevaisuus. Konservoinnin materiaalitutkimuksen heritologiset funktiot. - The heritological functions of materials research of conservation. 157 p. (208 p.) 2009.

115 NIIRANEN, SUSANNA, «Miroir de mérite». Valeurs sociales, rôles et image de la femme dans les textes médiévaux des *trobairitz*. - "Arvokkuuden peili". Sosiaaliset arvot, roolit ja naiskuva keskiaikaisissa *trobairitz*-teksteissä. 267 p. Yhteenveto 4 p. 2009.

116 ARO, MARI, Speakers and doers. Polyphony and agency in children's beliefs about language learning. - Puhujat ja tekijät. Polyfonia ja agentiivisuus lasten kielenoppimiskäsityksissä. 184 p. Yhteenveto 5 p. 2009.

117 JANTUNEN, TOMMI, Tavu ja lause. Tutkimuksia kahden sekventiaalisen perusyksikön olemuksesta suomalaisessa viittomakielessä. - Syllable and sentence. Studies on the nature of two sequential basic units in Finnish Sign Language. 64 p. 2009.

118 SÄRKKÄ, TIMO, Hobson's Imperialism. A Study in Late-Victorian political thought. - J. A. Hobsonin imperialismi. 211 p. Yhteenveto 11 p. 2009.

119 LAIHONEN, PETTERI, Language ideologies in the Romanian Banat. Analysis of interviews and academic writings among the Hungarians and Germans. 51 p. (180 p) Yhteenveto 3 p. 2009.

120 MÁTYÁS, EMESE, Sprachlernspiele im DaF-Unterricht. Einblick in die Spielpraxis des finnischen und ungarischen Deutsch-als-Fremdsprache-Unterrichts in der gymnasialen Oberstufe sowie in die subjektiven Theorien der Lehrenden über den Einsatz von Sprachlernspielen. 399 p. 2009.

121 PARACZKY, ÁGNES, Näkeekö taitava muusikko sen minkä kuulee? Melodiadiktaatin ongelmat suomalaisessa ja unkarilaisessa taidemusiikin ammattikoulutuksessa. - Do accomplished musicians see what they hear? 164 p. Magyar nyelvü összefoglaló 15 p. Summary 4 p. 2009.

122 ELOMAA, EEVA, Oppikirja eläköön! Teoreettisia ja käytännön näkökohtia kielten oppimateriaalien uudistamiseen. - Cheers to the textbook! Theoretical and practical considerations on enchancing foreign language textbook design. 307 p. Zusammanfassung 1 p. 2009.

123 HELLE, ANNA, Jäljet sanoissa. Jälkistrukturalistisen kirjallisuuskäsityksen tulo 1980-luvun Suomeen. - Traces in the words. The advent of the poststructuralist conception of literature to Finland in the 1980s. 272 p. Summary 2 p. 2009.

124 PIMIÄ, TENHO ILARI, Tähtäin idässä. Suomalainen sukukansojen tutkimus toisessa maailmansodassa. - Setting sights on East Karelia: Finnish ethnology during the Second World War. 275 p. Summary 2 p. 2009.

125 VUORIO, KAIJA, Sanoma, lähettäjä, kulttuuri. Lehdistöhistorian tutkimustraditiot Suomessa ja median rakennemuutos. - Message, sender, culture. Traditions of research into the history of the press in Finland and structural change in the media. 107 p. 2009.

126 BENE, ADRIÁN Egyén és közösség. Jean-Paul Sartre *Critique de la raison dialectique* című műve a magyar recepció tükrében. - Individual and community. Jean-Paul Sartre's