

THE MAKING OF RALLIENGLANTI:
Cross-linguistic influence in Finnish speakers' pronunciation of
English

Master's Thesis
Anssi Kivistö

University of Jyväskylä
Department of Languages
English
May 2016

JYVÄSKYLÄN YLIOPISTO

Tiedekunta – Faculty Humanistinen tiedekunta	Laitos – Department Kielten laitos
Tekijä – Author Anssi Kivistö	
Työn nimi – Title THE MAKING OF RALLIENGLANTI: Cross-linguistic influence in Finnish speakers' pronunciation of English	
Oppiaine – Subject Englannin kieli	Työn laji – Level Pro Gradu -tutkielma
Aika – Month and year Toukokuu 2016	Sivumäärä – Number of pages 69 + 2 liitettä
<p>Tiivistelmä – Abstract</p> <p>Kielten välistä siirtovaikutusta, eli transferia on tutkittu melko paljon, mutta suomen ja englannin kielten suhde, varsinkin fonetiikan osalta, on jäänyt melko vähälle huomiolle. Tämän tutkielman tarkoituksena oli selvittää erilaisia tapoja, joilla suomalaisten englanninääntämys voi erota ns. malliääntämyksestä, sekä selvittää, miten suomen kielen foneettinen järjestelmä vaikuttaa näihin eroihin. Kvalitatiiviseen tapaustutkimukseen valittiin erilaisia englannin kielen ääntämysjärjestelmän osa-alueita, kuten konsonanttiäänteitä, sekä prosodisia seikkoja, kuten painotus ja intonaatio. Tavoitteena oli löytää kattavasti erilaisia esimerkkejä tutkimuksen osallistujien puhenäytteistä ja kuvailla niitä ja niiden suhdetta suomen kielen fonetiikkaan ja sen vaikutukseen.</p> <p>Tutkimuksessa oli kaksi osallistujaa, jotka tekivät ääntämistestin. Testi äänitettiin, minkä jälkeen äänitteistä etsittiin erityyppisiä ääntämisesimerkkejä. Näytteistä löydettyt esimerkit jaoteltiin neljään ryhmään sen perusteella, kuinka paljon ne erosivat englannin kielen malliääntämyksestä. Tämän jälkeen esimerkkejä analysoitiin Praat-ohjelmalla aaltoäänikuvien ja erilaisten käyrien avulla, sekä vertailtiin englannin kielen ääntämissääntöihin. Lisäksi esimerkeistä etsittiin mahdollista suomen kielen vaikutusta.</p> <p>Tutkimuksessa löytyi kattavasti esimerkkejä englannin kielen standardista eroavasta ääntämyksestä, ja monessa tapauksessa mahdollinen syy poikkeavaan ääntämiseen voitiin jäljittää suomen kielen siirtovaikutukseen. Tutkielman tarkoitus ei ollut tuottaa yleistettäviä johtopäätöksiä, vaan kuvata normista eroavaa ääntämystä käytännössä. Löydettyjä esimerkkejä voidaan hyödyntää esimerkiksi ääntämisen opetusta suunniteltaessa.</p>	
Asiasanat – Keywords: pronunciation, language transfer, Finnish language, English language	
Säilytyspaikka – Depository: JYX	
Muita tietoja – Additional information	

Table of Contents

1. Introduction	1
2. Explaining terminology	4
3. Teaching pronunciation	5
3.1 History of pronunciation teaching	5
3.2 English pronunciation in Finnish education and the National Core Curriculum ...	7
3.3 The problem of defining standard pronunciation	9
3.4 Phonological core	9
4. Language transfer in phonetics	12
4.1 Cross-linguistic influence	12
4.2 Phonological transfer	13
5. English phonetics	15
5.1 Consonants	16
5.1.1 Plosives	17
5.1.2 Fricatives	17
5.1.3 Affricates	18
5.1.4 Consonant clusters	18
5.2 Prosody	19
5.2.1 Stress	19
5.2.2 Intonation	20
6. Finnish phonetics	21
6.1 Standard Spoken Finnish	21
6.2 Finnish consonants	22
6.2.1 Group 1 – /p/, /t/, /k/, /s/, /h/, /m/, /n/, /l/, /r/, /v/ and /j/	23
6.2.2 Group 3 – /d/	24
6.2.3 Group 4 – /f/	25
6.2.4 Group 5 – /b/, /g/ and /ʃ/	25
6.2.4 Consonant clusters	26
6.3 Finnish prosody	26
6.3.1 Stress	27
6.3.2 Intonation	28
6.4 Major differences between English and Finnish sound systems	28
7. Voicing in English and Finnish stops (Suomi 1980)	29
7.1 The Study	29
7.2 The English stops produced by the Finnish speakers	30
7.2.1 Word-initial and word-medial position	30

7.2.2 Word-final position	31
7.3 Categorizing the samples given by the Finnish learners.....	31
7.4 Evaluation	32
8. The present study	33
8.1 Data and methods.....	33
8.2 The participants.....	37
9. Results – the samples analyzed according to the framework.....	39
9.1 Group 1	40
9.2 Group 2	46
9.3 Group 3	49
9.4 Group 4	54
9.5 The participants’ own views about the test.....	60
10. Discussion/Conclusion	61
Bibliography.....	67
Appendix	70
Appendix 1. The pronunciation test.....	70
Appendix 2. The waveform images.	71

Table of Figures:

Figure 1. The IPA chart of consonants.....	16
Figure 2. Changes in stressing.....	20
Figure 3. The division of the consonant sounds based on their occurrence in the Finnish sound system.....	22
Figure 4. <i>few/view</i>	41
Figure 5. <i>coat</i>	42
Figure 6. <i>mat</i>	43
Figure 7. <i>mad</i>	43
Figure 8. <i>blueberries</i>	44
Figure 9. <i>Is it funny?</i>	45
Figure 10. <i>astonishing</i>	45
Figure 11. <i>beach</i>	46
Figure 12. <i>car</i>	47
Figure 13. <i>this/thing</i>	49
Figure 14. <i>record</i>	50
Figure 15. <i>this time</i>	51
Figure 16. <i>project</i>	52
Figure 17. <i>Is that car red?</i>	53
Figure 18. <i>drag race</i>	54
Figure 19. <i>desire</i>	55
Figure 20. <i>judeful</i>	56
Figure 21. <i>and</i>	57
Figure 22. <i>vile</i>	58
Figure 23. <i>Say vantering loudly</i>	59

1. INTRODUCTION

A speaker of a foreign language often sounds different compared to a native speaker; the sound system of the speaker's native language affects the foreign language performance, which leads to a foreign-sounding accent (Jarvis & Pavlenko 2008: 62). A rather common, caricaturizing term used of Finnish speakers' pronunciation of English is 'rallienglanti' – 'rally driver English', made famous by Finnish athletes, and especially rally and Formula 1 drivers, as presented by MTV (2011). Even a more negative term 'apinaenglanti' – 'monkey English' has been mentioned (MTV 2009). This stereotypically Finnish pronunciation of English has also been used in commercials in Finland in order to bring a comical touch to them (e.g. Elisa 2016). Although the language skills of Finns are in general good, pronunciation has been left for less attention in education, especially when examining school textbooks (YLE 2013, Tergujeff 2013).

This thesis aims at showing how a native Finnish speaker's pronunciation of English can differ from the target language pronunciation rules, and also at examining how the speaker's native language affects it. Finding the different ways that Finnish affects competence in English benefits teaching and learning by helping to understand which issues should be given more consideration in education.

English has become a *lingua franca*, a global means of communication between people with different mother tongues, and it is nowadays used all over the world and it serves various purposes (Crystal 2003: 86-113). Therefore a certain level of proficiency is needed also from Finnish speakers in order to keep up with the globalization, and the need for English skills is not likely to decrease. Next we will take a brief look at the global situation of the English language, and in addition, the term "nativity" concerning pronunciation will be addressed.

McKay (2002: 15-17) mentions several reasons for the spread of English, as suggested by Crystal (1997, in McKay 2002: 15-17). First of all, the migration of English-speaking individuals to other areas was very rapid during the nineteenth and twentieth centuries. In addition, Britain had become the world's leading industry and trading nation by the beginning of the nineteenth century, with most innovations originating from Britain. The terminology of the new technological advances was therefore in English, which led to the need of speaking and understanding the language. By the end of the nineteenth century the United States had become the fastest growing economy. During the last decades English has taken its place as the language of global culture, especially among young people. This global culture includes for example motion pictures, popular music, international tourism and publications.

Moreover, the widespread use of English in political and intellectual areas makes it imperative for any country wishing to access the global community for economic development to have access to it. English has nowadays a variety of specific purposes, and competence in it is necessary for accessing many discourses at a global level. (McKay 2002: 15-17).

The spread of English has led to a situation in which defining a “native” English speaker has become very difficult, and, in fact, more non-natives use the language than natives. The international status of the language been widely discussed during the recent years. The term *English as a Lingua Franca* (ELF) has been used rather commonly and studied by, for example, Mauranen & Ranta (2009) and Seidlhofer (e.g. 2001, 2011). Sharifian (2009: 2-3) proposes another term, *English as an International Language* (EIL), which is also differentiated from *International English*. The use of the term *International English* suggests a particular variety of English, similar to, for example, American English or Chinese English, whereas EIL emphasizes the view that English, and its varieties, is a “language of international, and therefore *intercultural*, communication”. What is interesting is that EIL disregards Kachru’s (1986, 1992, in Sharifian 2009) circle-model that divides the role and use of English in different countries into three circles: Inner Circle (English used as the primary language), Outer Circle (English used as a second language) and Expanding Circle (English used as a foreign language). The term *World Englishes* has often been used to refer to the countries of the Expanding Circle, but according to Sharifian (2009: 3) the term should cover all Englishes of all circles; in EIL contexts English is used by speakers from various cultural and national backgrounds.

Sharifian’s view of the fading borders between Kachru’s circles seems rational. In a globalized world people are moving from one country to another and spreading their cultures and languages so rapidly that a view of dividing countries into circles based on the status of English is essentially outdated. In a globalizing world all kinds of borders, including ones based on language, between people become vague. In addition, also Canagarajah (2006, in Sharifian 2009: 3) states that world Englishes should not only be viewed through the three circles. One reason is that the Outer-Circle and the Expanding-Circle Englishes are spreading to the Inner-Circle countries, and subsequently native speakers of English are exposed to World Englishes. Furthermore, McKay (2002: 9) mentions that in many countries from the Expanding Circle the number of bilingual English-speakers is greater than in countries of the Outer Circle where English has an official status.

As it was mentioned above, defining a native speaker of English has become difficult, and if we consider the oral skills it obviously leads to the question of which dialect or accent to teach to learners and how precisely certain models of English should be followed. Considering the status of the English language we can make the assumption that proficiency in the language is needed now and in the future, and therefore also pronunciation must be taught to learners. Although a native-like accent is not needed (since even the definition is problematic) one must be intelligible in his or her pronunciation. Understanding on which topics to concentrate is beneficial in the learning and teaching process. Jenkins (2000) has presented a Lingua Franca Core, i.e. a theory of phonetic features in English that are necessary for intelligibility. The LFC will be presented in section 3.4.

The phenomenon of knowledge in one language (e.g. the learner's native language) affecting the performance in another is called cross-linguistic influence, or language transfer, and it also applies to learning the phonetics and pronunciation of a language. Language transfer will be examined in greater detail in chapter 4. In order to distinguish what kinds of issues in the Finnish phonetics are likely to cause negative influence on Finnish speaker's pronunciation of English we must be aware of the phonetic systems of the two languages. Cross-linguistic influence has been studied widely, but the relation between English and Finnish, especially in the field of pronunciation, needs more research. The purpose of this qualitative case study is to shed some light on the issue and provide authentic and illustrative examples of phonetic structures that are problematic to Finnish learners. Therefore no generalizable conclusions can be drawn from the results. It must be noted that the purpose is not to pinpoint errors or give instructions on pronunciation, but the examples and findings can be used to benefit learning and teaching of English pronunciation.

First, in chapter 2 some crucial terminology will be explained. Chapter 3 will focus on pronunciation teaching, covering the main views during the last decades and also how pronunciation is nowadays dealt with in the Finnish National Core Curriculum. In addition, a view of a phonological core, i.e. the necessary issues for intelligibility in English language, will be addressed. In chapter 4 I will present the phenomenon of cross-linguistic influence, and chapters 5 and 6 will focus on the phonetic systems of English and Finnish, respectively. In the seventh chapter I will present a previous study by Kari Suomi, which serves as the main framework in the present study. The main study will be introduced in chapter 8, the findings are presented and analyzed in chapter 9, and the results are discussed and evaluated in chapter 10.

It can be summarized that this study aims to provide various examples of pronunciation that differs from the norms of the English phonology and phonetics. An important part is also to distinguish which of the differences can be traced to the Finnish phonetic system. The study focuses on various phonetic issues such as certain types of consonants and prosodic issues such as stress and intonation.

2. EXPLAINING TERMINOLOGY

In this section some basic terms of phonetics and phonology will be defined. First of all, the difference between *phonology* and *phonetics* has to be clarified. As Gut (2009: 6-7) explains it, they describe and analyze speech from differing perspectives. Phonetics concentrates on the actual sounds produced by humans, and it can be further divided into *articulatory phonetics* (examines how the organs and muscles are used when producing speech), *acoustic phonetics* (focuses on how the sound travels in air from the speaker's mouth to the listener's ear) and *auditory phonetics* (analyzes the effects that the sound has in the listener's ear and brain). Phonology, on the other hand, is interested in how the speech sounds in a language are in theory used to form patterns. Phonology is further divided into *segmental* and *suprasegmental* phonology; segmental phonology focuses on individual sounds and suprasegmental on larger units, such as syllables, words and intonation phrases.

The use of square brackets and slashes in the transcriptions relates to this as well; the phonological model of pronunciation is presented in square brackets (e.g. [b]), i.e. how the word or sound should be pronounced according to the target language rules, whereas the slashes represent the phonetic realization of the utterance, i.e. how it is actually pronounced (e.g. /p/).

A *phoneme* is the simplest unit of sound in a language. If two sounds in a language are in a complementary distribution, meaning that they never occur in the same environment, we can call them different *allophones* of a same phoneme, i.e. the aspirated [p^h] and the unaspirated [p] are both allophones of the same phoneme [p]. (Jensen 1993: 2)

Voicing and *aspiration* are two crucial terms used in this thesis. *Voicing* refers to how the vocal folds act while producing the sound. According to Ladefoged (2005: 3), while producing *voiced* sounds the vocal folds are vibrating, whereas in *voiceless* sounds they are apart. The term *aspiration*, as Ladefoged (2005: 56) describes, refers to a burst of air during a voiceless period after the release of a plosive, i.e. stop consonant.

The basic terminology was explained here very briefly, and for more information on various terms regarding phonetics and phonology I recommend consulting e.g. Ladefoged (2005) which is a thorough but concise and simple textbook on the subject.

3. TEACHING PRONUNCIATION

In this section I will briefly present the views on pronunciation teaching over the last decades, and then take a look at the situation in Finland today based on the National Core Curriculum of comprehensive education in order to see how pronunciation is viewed in it. After that the issue of choosing a specific standard for pronunciation teaching will be addressed, and last, I will introduce a view of a phonological core, i.e. a system of core features in English phonology and phonetics that are vital for intelligibility.

3.1 History of pronunciation teaching

According to Celce-Murcia et al. (2010: 1-3) there are two general approaches to pronunciation teaching; *Intuitive-Imitative Approaches* and *Analytic-Linguistic Approaches*. The *Intuitive-Imitative Approaches* rely on the learner's ability to listen to and imitate the sounds and rhythms of the foreign language without explicit information. The *Analytic-Linguistic Approaches* take advantage of explicit tools and information, such as phonetic alphabet, articulatory descriptions and charts of the vocal apparatus, in addition to listening, imitation and production. Celce-Murcia et al. (2010: 3-12) divide the approaches further into a number of methods of pronunciation teaching that have been used during the last couple of centuries, and in order to give a general picture of the history of pronunciation teaching I will very briefly present them here.

The *Direct Method* of foreign language instruction was popular in the beginning of the 20th century. In *Direct Method* pronunciation was taught through intuition and imitation – students tried to imitate a model provided by the teacher or a recording. This method was based on natural learning environments, such as children learning their first language or adults learning a language in non-instructional settings. The *Direct Method* was succeeded by so-called naturalistic methods, including Asher's (1977, in Celce-Murcia et al. 2010: 3) *Total Physical Response* and Krashen and Terrell's (1983 in Celce-Murcia et al. 2010: 3) *Natural Approach*. (Ibid.)

In the 1890s the first analytic contribution to the teaching of pronunciation arose as part of the *Reform Movement* in language teaching. Almost at the same time (1886) the

International Phonetic Association was formed and also the *International Phonetic Alphabet* (IPA) was developed as a result of establishing phonetics as a science describing and analyzing the sound systems of languages. The phoneticians in the *International Phonetic Association* influenced modern language teaching by promoting practices such as regarding the spoken form of language as primary, applying phonetic findings to language teaching and giving teachers and learners phonetic training. (Ibid.)

According to many historians of language teaching the *Reform Movement* had a role in developing *Audiolingualism* in the US and the *Oral Approach* in the UK during the 1940s and 1950s. These approaches consider pronunciation very important and it is taught explicitly from the start. The teacher uses information from phonetics to demonstrate the articulation of sounds. The teacher also often applies contrastive techniques such as minimal pair drills. (Celce-Murcia et al. 2010: 4)

The *Cognitive Approach* of the 1960s viewed language as “rule-governed behavior rather than habit formation” (Celce-Murcia 2010: 5) and favored grammar and vocabulary, claiming, for example, that native-like pronunciation could not be achieved and learners should instead focus on grammatical structures and words. Therefore pronunciation was not really focused on in the cognitive approach.

In the 1970s methods such as the *Silent Way* and *Community Language Learning* came to attention. The *Silent Way* is in many ways similar to *Audiolingualism*; the learners are taught to produce sounds and structures accurately from the initial stage of instruction. In the *Silent Way* they should also focus on how words are combined in phrases concentrating on how, for example, stress and intonation affect the utterance. The name of the method comes from the fact that the teacher speaks as little as possible and uses gestures, such as tapping out rhythmic patterns or pointing the number of needed syllables with fingers, to indicate what students should do. The *Community Language Learning* lessons consists of various phases in the classroom, including learning to pronounce idiomatic expressions and recording them, then transcribing the utterances and translating them, and then repeating and practicing the utterances. This lets the learners focus on *how* the utterances were said and decide for themselves which utterances and how much they want to repeat until they get as close to the target pronunciation as they want to. (Celce-Murcia et al. 2010: 5-8)

Communicative Language Teaching, also called the *Communicative Approaches*, has been the dominant method in language teaching from the 1980s and still holds its position. The main idea in Communicative Language Teaching is that communication should be the main focus in language instruction. The goal is not to make learners sound native, and such

goal would in many cases be unrealistic. A more realistic goal is to enable the learners to reach a level of establishing intelligible pronunciation so that their speaking skills will not hinder their ability to communicate. In addition, the pronunciation instruction is moving from segmental/suprasegmental debate to a more balanced understanding – the aim is to recognize the most important aspects of both sides of phonetics and incorporate them to create instruction suitable to all learners. (Celce-Murcia et al. 2010: 8-12)

3.2 English pronunciation in Finnish education and the National Core Curriculum

Although Finnish and Swedish are the two official languages of Finland, many in the Finnish-speaking majority feel English to be stronger than Swedish. A possible reason for this is the dominance of English in the Finnish media; for example English-language TV shows are never dubbed but instead presented with their original soundtrack and Finnish subtitles. Additionally, many large companies in Finland use English as their official language of business. (Ranta 2010: 159)

In primary education 89.5% of Finnish pupils chose English as their first foreign language in 2005, and the percentage has long been around 90%. British English was the preferred variant in the curricula through the 1960s to the 1980s, but the curricula from the 1990s onwards have not specified any certain variety. American English was given an equal status to British English in the curricula in 1985. The later curricula also emphasize hearing native speech on tapes or from authentic speakers instead of relying only on the model provided by a non-native teacher. In addition, the lingua franca status of English has been stressed in the curricula from the 1960s onwards. (Ranta 2010: 159-160)

In this section I will briefly present how pronunciation of English is addressed in the latest Finnish National Core Curriculum (NCC) (Perusopetuksen opetussuunnitelman perusteet 2014: 243-247, 398-402). Although the two informants in the present study have both graduated from upper secondary school the focus in this section is on the core curriculum for comprehensive school because it applies to all learners in Finland. The upper secondary education core curriculum will be omitted here since it is applied to only approximately half of the Finnish learners of English, and the CC of the comprehensive school provides a more general view at the way English is taught to all learners in Finland.

The Finnish National Core Curriculum does not really emphasize the significance of pronunciation; it is principally regarded as part of text producing skills. According to the

NCC in the elementary level (3rd to 6th grade) the teacher should provide opportunities to produce speech while paying attention to the basic rules of pronunciation (Perusopetuksen opetussuunnitelman perusteet 2014: 244). The NCC also describes the aims of pronunciation teaching in practice:

“Havainnoidaan ja harjoitellaan runsaasti ääntämistä sekä sana- ja lausepainoa, puherytmiä ja intonaatiota. Harjoitellaan tunnistamaan englannin kielen foneettisen tarkekirjoituksen merkkejä.”

“[The learners will] observe and practice plenty of pronunciation as well as word and sentence stress, speech rhythm and intonation. [They will also] practice to recognize the symbols of the phonetic alphabet of English.”

Perusopetuksen opetussuunnitelman perusteet 2014: 244

Similar view of pronunciation as a part of text producing skills can be seen in the evaluation of learners, and it seems that conveying meaning is more important than correct pronunciation. The goal for the learners, similarly to the goals for teaching, is to learn to produce speech while paying attention on the basic rules of pronunciation. (Perusopetuksen opetussuunnitelman perusteet 2014: 245-247)

The same views are applied also to the lower secondary school level (grades 7-9). According to the NCC (Ibid.: 398-400), pronunciation is a part of the text producing skills, and the teacher’s responsibility is to guide the learners to produce text to different purposes while directing them towards good pronunciation. It is also stated that English should be used as much as possible, but pronunciation is not underlined. The evaluation is very similar to elementary level with the exception of slightly wider criteria, including the ability to apply the text producing skills in different situations (Ibid.: 400-402). Furthermore, it must be noted that the criteria for evaluation does not really separate speaking and writing, either in elementary or lower secondary level; they are both essentially considered text producing skills.

We can see that pronunciation is not really a top priority in the National Core Curriculum in basic education, but instead the criteria and guidelines emphasize communicativity and text production. Looking back at section 3.1 it is clear that the NCC is in line with the *Communicative Language Teaching* method, i.e. the present view of language teaching. Furthermore, the NCC does not really separate spoken and written text. Communicativity is without doubt a very important part of language learning and teaching, and the growing input of spoken English in our everyday lives gives us plenty of models of pronunciation, but still explicit pronunciation teaching in English classes of Finland seems, based on the NCC at

least, to be slightly forgotten. Naturally the point of explicit pronunciation teaching can be questioned when implicit models are provided everywhere, and also the NCC leaves it up to the teachers to decide how thoroughly they want to cover the topic in their classrooms.

3.3 The problem of defining standard pronunciation

The question of choosing a specific variation of English to be a model of pronunciation is at the least problematic. Teaching pronunciation is often based on a specific standard of a variety of English; British English was the preferred variation in the Finnish curricula till the 1980s, as it was mentioned in the previous section. Preston (2005: 38) notes that such models do not in fact exist. It is essentially impossible that a person would speak English following a standard model, e.g. Received Pronunciation, perfectly. Furthermore, Seidlhofer (2005: 59) notes that the majority of users of English do not have it as their first language, and that native speakers are often considered providers of acceptable usage models. However, the problem of choosing an alternative still persists since there are regional and individual differences between native speakers as well. Therefore a specific standard to follow would solve the problem of the alternatives, and, furthermore, defining nativity would not be necessary. In view of globalization, and especially how the status of English has changed during the last decades as it was reviewed in section 1, the concept of nativity can be considered almost irrelevant nevertheless.

As McKay (2002: 49-53) notes, the issue of standards persists in all languages. There is no official organization to uphold the standards in the English language, and there has been worry that other languages used in many Outer Circle countries would influence English to a degree in which it would vary so much that mutual intelligibility between English speakers would be lost. Standard English has been given quite a few definitions, but often they are limited to the written form of the language.

In Finnish education the model for pronunciation, i.e. a standard used for the English language, was, as mentioned earlier, for a rather long period of time British English, and only the recent curricula have taken also other variations into account, mainly American English. The issue of standard pronunciation will be discussed in next section concerning a phonological core for intelligible pronunciation.

3.4 Phonological core

The previous section discussed the problems in finding a specific model for pronunciation teaching. Jenkins (2000: 99-121) discusses finding a phonological core, i.e. the essential

features of a phonological system of a language that are essential for intelligibility, and the advantages in doing so by using language transfer. Seidlhofer (2002: 67-68) also presents Jenkins' model and describes it as "groundbreaking". Jenkins (2000: 119) states that "phonological transfer is deep-rooted and can be of benefit to the learners". This is backed by the idea that universal difficulties in learning a specific feature of L2 English pronunciation could be a hint towards an "unteachable" item of language when considering learning in a classroom environment. Whether these unteachable items should be tried to be taught at all, Jenkins suggests that "the outcome depends on whether the item is needed in order to preserve intelligibility". She also discusses the relevance of getting rid of L1 transfer: if a learner does not experience problems in intelligibility of their speech despite using a particular phonological transfer (for example not distinguishing voiced /ð/ and voiceless /θ/), they might not be motivated to replace it in order to only have their accent sound more native-like, nor is it even necessary to do so.

Jenkins (2000: 123-164) discusses the contents of the LFC, a Lingua Franca Core, i.e. the items which are necessary for intelligible pronunciation in English language. Based on research, she has found four general areas (Ibid.: 132) that will be presented in this section: 1. Most consonant sounds, 2. Appropriate consonant cluster simplification, 3. Vowel length distinctions and 4. Nuclear stress.

Jenkins (2000: 137-138) suggests based on her data that when assessing consonants there are two possible sounds that could be omitted from the core without causing unintelligibility; the dental fricatives /ð/ (voiced) and /θ/ (voiceless). The dental plosives /d/ and /t/ occur in many variants of English, which suggests that their pronunciation is simpler and therefore they are often used to substitute /ð/ and /θ/. The LFC does not directly suggest that the /ð/ and /θ/ sounds should be replaced with /t/ and /d/, but only excludes them from being necessary for phonological intelligibility.

Other suggestions for consonant voices relate to differences between the General American accent (GA) and Received Pronunciation (RP). With /r/ Jenkins chooses the GA rhotic variant [ɹ] (retroflex approximant) which is indicated orthographically in post-vocalic situations, e.g. in the word *four*. The GA alternative of pronouncing the *r* is simpler than RP in which it is pronounced only when the following sound is a vowel, e.g. *four eggs*. On the contrary, the RP variation is included in the LFC regarding the consonant /t/; this leads to the omission of the intervocal /t/ becoming the voiced flap /ɾ/ in GA. (Ibid. 139-140)

According to Jenkins (2000: 140) the LFC also includes aspiration, which helps distinguishing the voiceless /ptk/ sounds from the voiced /bdg/ sounds in initial position in

stressed syllables. In addition, lengthening the vowel that is followed by a voiced consonant will be included in the LFC since it benefits the distinction of voiced and voiceless plosives in syllable-final position. This phenomenon of lengthening the vowel will be examined in greater detail in section 5.1 regarding English consonants.

With consonant clusters, analyzed further in section 5.1.4, in LFC the addition of extra vowel sounds (e.g. Japanese learners pronouncing *product* as /pə'rɒdʌkɒtə/) is preferable than deletion (mostly in word-initial clusters, e.g. a Taiwanese learner pronouncing *product* as /'pɒdʌk/). Word-medial or word-final clusters are sometimes simplified in English, e.g. the omission of /t/ in words like *castle* and *facts*.

The vowel sounds are not focused on in the present study, but their inclusion in Jenkins' Lingua Franca Core will be very briefly presented here. Generally speaking, according to Jenkins (2000: 144), the LFC prioritizes quantity over quality; vowel quantity is rather stable across different variations of English, whereas there is more variation in quality (e.g. the pronunciation of *dog* as /dɒg/ in RP and /dag/ in GA). The length of the vowel depends on the following consonant, and will be presented in section 5.1. Jenkins (2000: 146) also takes the quality into account in the LFC in relation to the sound /ɜ:/; for example pronouncing the word *birthplace*, /'bɜ:θpleɪs/ as *bathplace*, /'bɑ:θpleɪs/, might result in non-comprehension.

Jenkins (2000: 150-151) considers word stress rules somewhat unteachable referring to their complexity. However, there are some guidelines that can be provided to the learners; for example, that the majority of two-syllable nouns have stress on the first syllable and two-syllable verbs on the second, and that suffixes such as *-ee* and *-ese* tend to be stressed. Most of all the learners should be made aware that there are a number of exceptions to the stress rules and taught to identify the word stress in the dictionary.

The issue of intonation is slightly problematic. Tone is often connected to certain grammatical structures, i.e. yes/no questions having a rising intonation (Ladefoged 2005: 124). However, Jenkins (2000: 152) refers to a few opposing views (Kreidler 1989: 182-183 and McCarthy 1991: 106, in Jenkins 2000: 151-152) that claim that a specific type of melody might not necessarily be related to any certain grammatical rules. Instead she regards the *nuclear stress* as more important concerning the LFC. In an utterance nuclear stress “highlights the most salient part of the message, indicating where the listener should pay particular attention” (Jenkins 2000: 153).

Jenkins (2000: 158-160) concludes that the LFC is not a model of pronunciation and neither a restricted simplified core, but instead it allows individual freedom for the speakers to express their identity and to adjust to their receivers. It provides guidelines on what

requires pedagogic focus, and the issues excluded from it are the ones that should be regarded as regional accent variation. The issues included in it are the ones that have potential for errors; in other words, a phonological error is defined in relation to its effect on intelligibility.

4. LANGUAGE TRANSFER IN PHONETICS

Language transfer, or *cross-linguistic influence*, according to Jarvis & Pavlenko (2008: 3) refers to a person's knowledge of one language affecting knowledge of another language. CLI, can occur in many areas of language use, e.g. phonology, orthography, lexis, semantics, morphology, syntax, discourse, pragmatics and sociolinguistics. (Jarvis and Pavlenko 2008). The following sections will focus on its effect on language learning and pronunciation.

4.1 Cross-linguistic influence

A speaker of a foreign language rarely sounds like a native speaker, and naturally we must also remember the difficulties in defining a native speaker, which was discussed earlier in sections 1 and 3.3. Foreign accents are an example of *cross-linguistic influence (CLI)*, or *language transfer*, which, as defined by Jarvis & Pavlenko (2008: 3) refers to a person's knowledge of one language affecting knowledge of another language.

Jarvis & Pavlenko (2008: 3) and Odlin (1989: 25) state that the term to describe the phenomenon of knowledge of one language influencing the knowledge of another has been argued about by many researchers. According to Jarvis & Pavlenko (2008: 3) the term *interference* has rather negative connotations, whereas the phenomenon can also have a positive effect on learning another language. Transfer (or interference) has been described as falling back on old knowledge, but as stated by Odlin (1989: 27), this ignores that fact that some learners have "a head start" in getting to know a new language. As an example he mentions similarities in vocabulary and writing systems that reduce the amount of new linguistic issues in English to Spanish speakers, compared to e.g. Arabic speakers. Odlin (1989: 26) states that despite the negative associations the term *interference* is still used, and that it seems applicable e.g. with "phonetic inaccuracies that resemble sounds in the learner's native language".

The term *transfer* has been considered unsuitable because of its association to behaviorist notion of *skill transfer*. (Jarvis & Pavlenko 2008: 3). The term *Cross-linguistic influence*, proposed by Kellerman and Sharwood Smith (1986, in Jarvis & Pavlenko 2008: 3) is considered appropriate since it is a theory-neutral term and it takes into account the wide

range of ways in which knowledge of one language can affect another. Since then this term has gained acceptance among researchers, but still also the terms *transfer* and *interference* have been used synonymously. On top of all this Odlin (2008: 436) also mentions terms *the role of mother tongue*, *native language influence* and *language mixing*. In the present paper I will use the terms *transfer* and *cross-linguistic influence*, which can refer to both positive and negative transfer. The term *interference* will also be used, but only when referring to negative transfer.

4.2 Phonological transfer

The term *phonological transfer* used in this thesis, as defined by Jarvis & Pavlenko (2008: 62), refers to “the ways in which a person’s knowledge of the sound system of one language can affect that person’s perception and production of speech sounds in another language”. They also differentiate the effect of CLI to phonetics, i.e. the actual sounds that L2 users perceive and produce and to phonology, i.e. how the users organize, structure and categorize the sounds. Phonological transfer is a useful cover term for both phenomena.

The most noticeable effect of phonological transfer at the segmental level, according to Jarvis & Pavlenko (2008: 63) is the difficulty of differentiating two sounds in the second language that do not have a distinct phonemic contrast in the native language (e.g. the difference between the English sounds /i/ (*sheep*) and /ɪ/ (*ship*) to Spanish speakers or the Finnish short (*tuli*) and long (*tuuli*) vowel to English speakers. Other examples of transfer effects in receptive phonological skills include spectral qualities (or formant frequencies, i.e. vowels), duration, voicing and aspiration.

Odlin (1989) presents Moulton’s (1962a, in Odlin 1989: 115-117) further categorization of segmental errors into phonemic, phonetic, allophonic and distributional errors. Phonemic errors can occur when there are differences between the phonemic inventories of two languages (i.e. the German distinction between the voiceless velar fricative /x/ and the voiceless velar stop /k/, which might cause difficulties to e.g. native English speakers since such distinction does not exist in the English phonological system). Phonetic errors can occur between, for example, the German uvular /ʁ/ and the English retroflex /ɻ/, which are phonologically similar but phonetically different. Allophonic errors may arise when a certain allophone of a sound is appropriate in one language but not in another, for example the American English flap /ɾ/ which would not be used as an allophone of /t/ in e.g. German. Distributional errors relate to phonotactics; Odlin (1989: 117) presents as an example the

word-initial consonant cluster /ts/, which is common in German but not in English and could therefore cause problems to English-speaking learners of German.

In addition, Jarvis & Pavlenko (2008: 67-68) also mention phonotactics and state that the most problematic issue to learners is often consonant clusters (more in section 5.1). What kinds of problems the learners have and how they cope with them depends on their L1, and even on their L1 dialect.

Research concerning productive skills has concentrated on studying why and how often learners produce specific segmental substitutions. According to Jarvis & Pavlenko (2008: 64) it has been found that using a certain substitution depends on several factors, e.g. language universals and the phonetic environment of the substitution. L1 influence is strongly connected to these factors in determining which sound is used and in which contexts in the substitution. However, substitutions do not stem only from following the constraints of the native language, but also from concerns about meaning; Jarvis & Pavlenko (2008: 65) present an example of Japanese speakers pronouncing *think* as /sɪŋk/ and trying to maintain the contrast between *think* and *sink* by pronouncing *sink* as /fɪŋk/.

Another area of research concentrates on producing the right phonemes with wrong segmental properties, e.g. voicing and aspiration; transfer is present in the tendency to pronounce word-final voiced obstruents, i.e. stop consonants, as voiceless, e.g. pronouncing *bag* as /bæk/, and, in addition, it may cause differences in voice onset time (VOT), i.e. the time that passes from releasing the stop consonant to the beginning of the following vowel sound. According to Jarvis & Pavlenko (2008: 65), VOT is closely associated with the views about a speaker's nativeness or accentedness.

Phonological transfer affects also suprasegmental issues, i.e. stress, tone, rhythm and other factors. According to Odlin (1989: 117), especially stress patterns are very important since they are used to distinguish words from one another and therefore play an important role in listeners' recognition of words. This will be discussed in greater detail in sections 5.2.1 (stressing in the English system) and 6.3.1 (stressing in the Finnish system). Also intonation can be affected by transfer; in some languages pitch levels are used to differentiate words and in others the pitch does not have such function. Intonation will be discussed further in sections 5.2.2 (regarding English) and 6.3.2 (regarding Finnish).

As Jarvis & Pavlenko (2008: 66) note, transfer can affect a speaker's skills also in reverse direction, i.e. from L2 to L1. In addition, a transfer from a learner's L2 to the learner's another L2 is possible. Since the present study only focuses on forward transfer, these phenomena will not be discussed further here.

According to Major (2008: 81), because of the variability of speech it is often difficult to distinguish what constitutes transfer and what does not. In order to distinguish transfer one must have a thorough description of the L1. Without it the identification of transfer is impossible (Ibid.: 82). The researcher should be native or near-native speaker of the participants' L1, or rather have "a conscious knowledge" of the L1. However, as Major notes, this is often unrealistic and restricting towards the research and the research should be limited to the issues that one has L1 descriptions for. This also confirms the fact that spotting examples of transfer is always a matter of subjective opinion and that it is difficult to distinguish issues that stem completely from transfer. One can only make good guesses based on the knowledge of L1 phonetics.

5. ENGLISH PHONETICS

This section of the thesis will focus on the phonetics and phonology of the English language. The relevant consonants of English phonetics are presented and then I will briefly discuss prosodic issues, such as stress and intonation. The following phonetic issues are chosen according to their relevance, and a number of interesting topics had to be left out due to the scope of the study. For example, the consonant sounds chosen for this study are fricatives, plosives and affricates since they are the ones that contain issues that are more likely to cause difficulties to native speakers of Finnish (e.g. voicing and aspiration). In addition, the vowel sounds were not chosen for the analysis because there is more variance between them in the different alternatives of English, especially concerning their quality, for example the word 'dog' pronounced likely as /dɑg/ in GA and as /dɒg/ in RP, and choosing the correct target language model of pronunciation is more controversial. The consonants have more qualities that are shared with the many variations of English, for example the voiced/voiceless contrast. Moreover, because of the scope of this study something had to be omitted, and by examining the different phonetic issues I came to the conclusion that certain types of consonants and suprasegmental properties would possibly provide more material for analysis and would also be more effectively analyzed in the waveforms and by listening.

This section of the thesis is based on a few sources, mainly Ladefoged (2005) and Odisho (2003) since the issues covered here are rather elementary concepts of phonetic sciences. The sources mentioned, especially Ladefoged (2005), provide a great deal of information on the subject in a concise form.

5.1 Consonants

According to Deterding (2005: 23) there are 24 consonant sounds in the English language, (although the number can be questioned, see section 5.1.3): six plosives [p b t d k g], nine fricatives [f v θ ð s z ʃ ʒ h], two affricates [tʃ dʒ], three nasals [m n ŋ], one lateral-approximant [l] and three approximants [w j r]. The consonants are all visible in the IPA chart (Figure 1).

THE INTERNATIONAL PHONETIC ALPHABET (revised to 2015)

CONSONANTS (PULMONIC) © 2015 IPA

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d		ʈ ɖ	c ɟ	k ɡ	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	ɴ		
Trill	ʙ			r					ʀ		
Tap or Flap		ⱱ		ɾ		ɽ					
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Lateral fricative				ɬ ɮ							
Approximant		ʋ		ɹ		ɻ	j	ɰ			
Lateral approximant				l		ɭ	ʎ	ʟ			

Symbols to the right in a cell are voiced, to the left are voiceless. Shaded areas denote articulations judged impossible.

Figure 1. The IPA chart of consonants. International Phonetic Association (2015).

<https://www.internationalphoneticassociation.org/content/full-ipa-chart>

Consonants are commonly categorized using the three parameters; phonatory status, place of articulation and manner of articulation. For example, the consonant sound [b] can be described by using the three parameters as *voiced* (phonatory status), *bilabial* (place of articulation) and *plosive*, i.e. *stop* (manner of articulation). However, two of the parameters are occasionally sufficient in making the categorization since certain types of consonants, e.g. nasals, can only be voiced. In addition, sometimes the three parameters are not enough, e.g. with plosives, which can also be aspirated or unaspirated (which will be discussed in greater detail later in this section). (Odisho 2003: 35).

The most common phonatory contrast, voiced and voiceless, according to Odisho (2003: 35-36), is considered an "absolute universal", i.e. a feature that can be found in all the languages. Additionally, between the extremes, *voiced* and *voiceless*, there are language-specific glottal gestures such as *creaky*, *whispery* and *breathy*, but having little relevance with the subject of this paper they will not be discussed further. Yet it must be remembered that the voiced-voiceless division is not actually so absolute in actual speech, but more like a generalization, since a voiced sound can also be partially voiced or more voiced initially or finally depending on, for example, the speaker and the context in which it occurs.

5.1.1 Plosives

As Ladefoged (2005: 13) defines plosives (/p/, /b/, /t/, /d/, /k/ and /g/), i.e. stop consonants, a closure is formed in the articulators, and they are called stop consonants because the air flow is stopped completely for a moment. Stop consonants can be divided into oral stops and nasal stops, depending on how the obstruction is formed. When articulating the oral stops the soft palate in the back of the mouth is raised so that the nasal tract is blocked and the airstream is completely stopped. At the same time the articulators block the air in the mouth and when they come apart a stop sound is formed, as in words like *pie* and *buy* (bilabial closure), *tie* and *dye* (alveolar closure), and *kye* and *guy* (velar closure). Nasal stops are formed when the airstream through the mouth is blocked but the soft palate is down, so that the air can escape through the nasal cavity. Nasal stops can be heard in words like *my* (bilabial closure), *nye* (alveolar closure) and at the end of the word *sang* (velar closure). The closures are similar in both oral and nasal stops, but the term *stop* is practically always used to describe oral stops. Nasal stops are instead referred to with the term *nasal*.

Ladefoged (2005: 13) explains that stops, as other consonants as well, can be either voiced (vibrating vocal folds) or voiceless (non-vibrating vocal folds). Voiced stops in the English language are the bilabial /b/, the alveolar /d/ and the velar /g/, and the voiceless stops are the bilabial /p/, alveolar /t/ and the velar /k/. Voicing is one aspect that differentiates English sounds from one another, yet however, it is not enough as such. Ladefoged (2005: 56) notes that in most people's speech the voicing is rather weak in words such as *buy*, and even voiced sounds can in certain phonetic environments be practically voiceless. What helps making the distinction between, for example, the voiced [b] and the voiceless [p] is *aspiration*, a small puff of air immediately after the release of the stop. Therefore *pie* is pronounced as /p^hai/. (Ibid.)

The voicing is rather weak also in the stops /b/, /d/ and /g/ in a word-final position in words like *nab*, *mad* and *nag*. What helps distinguishing them from *nap*, *mat* and *knack* is the fact that a voiced sound lengthens the preceding vowel, i.e. the [æ] sound in *nab* is slightly longer than in *nap*. (Ibid.)

5.1.2 Fricatives

In the English phonetic system there are nine fricatives, which can, similarly to stops, be either voiced (the former) or voiceless (the latter): the labiodental /f/ and /v/, the dental /θ/ and /ð/, the alveolar /s/ and /z/, the postalveolar /ʃ/ and /ʒ/ and the voiceless glottal fricative

/h/. With fricatives the sound is made by forcing the air through a narrow gap. There is less variation among the fricatives of English compared to the stops, but still the major allophonic variations are more or less parallel to those of the stops. Similarly to the stops, the voicing in the fricative also affects the length of the preceding vowel in a way that the vowel sound is longer in words *strive* and *rise*, and shorter in *strife* and *rice*, respectively. Stops and fricatives (and affricates, depending on whether we consider them a group of their own) are the only consonant sounds that can be either voiced or voiceless, and therefore we can state that vowels are shorter before all of the voiceless consonants and longer before all of the voiced consonants. (Ladefoged 2005: 64)

Also similarly to stops, a voiceless fricative at the end of a word is longer than its voiced counterpart (Ibid.: 65). Consequently, the voiceless /f/ is longer in *safe* than the voiced /v/ in *save*.

The voicing that occurs in a fricative that is at the end of a word does not in most pronunciations last throughout the whole sound, but instead changes to a voiceless sound. However, if the first sound in the next word is a voiced sound, the fricative is voiced throughout, compare for example the word *prove* by itself and in a phrase *prove it*. This also happens with stops. Fricatives and stops are also common in a way that they involve obstruction of the airflow through the mouth, and therefore they are referred to as *obstruents*. (Ladefoged 2005: 65)

5.1.3 Affricates

According to the International Phonetic Association (2015) there are two affricates in the English sound system; the voiceless palato-alveolar /tʃ/ and the voiced post-alveolar /dʒ/. The categorization of the affricates is slightly problematic. They could be regarded as clusters of two separate sounds, such as /ts/ in *cats* or /ps/ in *lapse*. However, according to Ladefoged (2005: 66) they are the only affricates that can occur at both the beginning and the end of words, and from this point of view they are, also in this study, considered single, individual phonemes.

5.1.4 Consonant clusters

A consonant cluster (or a *consonant blend*) is, according to Odisho (2003: 43), a combination of consonants in a close-knit sequence within a syllable, e.g. *str* in the word *stress*. The occurrence of consonant clusters in a language is an issue causing possible problems in cross-language situations, for example with Japanese or Finnish people pronouncing English.

Odisho (2003: 43-44) states that there can be clusters in initial, medial and final positions in English words, which makes the language very rich in clusters. The most common clusters have two consonants, such as *play* and *inspire*, but combinations of three (e.g. *spring*) or four consonants (e.g. *attempts* or *sixths*) are possible as well. These complicated structures are problematic also to native speakers; often in casual speech the pronunciation of, for example, *attempts* is closer to [ətemts] and *sixths* may be simplified to [sɪks]. According to Odisho (2003: 44) non-native speakers of English often either delete consonants or insert vowels (in the beginning, middle or the end of a cluster) and restructure words when facing pronunciation problems with consonant clusters.

5.2 Prosody

5.2.1 Stress

According to Ladefoged (2005: 110) a stressed syllable is produced by increasing the amount of air that is being pushed out of the lungs compared to unstressed syllables. A stressed syllable is often louder than an unstressed syllable, and usually it has also a higher pitch. In addition, a stressed vowel is longer than the same vowel would be in an unstressed position. (Ibid.)

Stress is used in speech to give emphasis on a word or contrast one word from another. In addition, another important function of stress is to show syntactic relationships between words. There are many noun-verb oppositions that are distinguished by stress, e.g. *an 'insult – to in'sult* and *an 'increase – to in'crease*. Similarly compound words can be distinguished from phrasal verbs, e.g. *a 'walkout – to 'walk 'out* and *a 'pushover – to 'push 'over*. In addition, compound nouns can be differentiated from adjective-noun combinations with stressing, e.g. *a 'hot dog – a 'hot 'dog*. (Ibid. 111)

Sentence stress modifies the stress of individual words, namely by dropping some of the stresses. For example, a sentence like *The big brown bear bit ten white mice* would not have stress on every word, but instead, according to Ladefoged (2005: 115) people would in ordinary speech say *The 'big brown 'bear bit 'ten white 'mice*. In addition, the stress in a word can shift from a syllable to another if a stressed syllable in a neighboring word is too close, e.g. in sentences *He had a 'clarinet 'solo* and *He 'plays the clari'net*. Stresses also tend to recur at regular intervals in spoken words, i.e. in a specific rhythm. However, not all sentences are regular; the sound pattern of English does not require the adjustment of syllable lengths to force the sound into a regular rhythm. The stress intervals are affected by the

number of stressed syllables in a stress group, the type and number of vowels and consonants in each syllable and also the variations in emphasis given to each word. (Ibid. 115-116)

In order to clarify the irregularity of stress in the English language Figure 2 is presented from Ladefoged (2005):

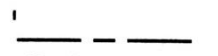
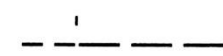
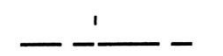
 diplomat photograph monotone	 diplomacy photography monotony	 diplomatic photographic monotonic
--	--	---

Figure 2. Changes in stressing. The stress (marked with ') changes from syllable to another depending on the inflections of the words. (Ladefoged 2005: 111)

Paananen-Porkka (2007: 20-21) distinguishes *stress* and *accent (sentence stress)*. *Stress* refers to stressing certain syllables in a word; in Finnish the stress is practically always on the first syllable of a word, whereas in English the place of the stress depends on word structure. *Accent*, or *sentence stress*, refers to textual level and depends on pragmatic and semantic factors, such as emphasizing or contrasting certain parts of the text. This view of accent is rather close to what Ladefoged describes as *intonation*, which will be covered below. How these are used in Finnish spoken language will be covered in section 6.3 about Finnish prosody.

5.2.2 Intonation

Intonation, according to Ladefoged (2005: 116) means the changes that happen in the pitch in a sentence. In English intonation has effect on meaning on sentence and phrase level rather than on individual words. In normal statements the pitch of the sentence is descending. There is variation in the intonation between different kinds of questions. Questions such as *Where are you going?* have a normal descending intonation. However, questions such as *Are you going home?* that require either *yes* or *no* for an answer usually have an ascending intonation (Ladefoged 2005: 124). However, Jenkins (2000: 151), as mentioned earlier in section 3.4, presents a differing point of view stating that the pitch does not necessarily relate to any specific grammatical structure, meaning, for example, that yes/no questions do not always have a rising tone. Ladefoged (2005: 124) also mentions that intonation depends on individual variations and is affected, for example, by the speaker's mood and attitude.

6. FINNISH PHONETICS

In this chapter I will briefly explain the term Standard Spoken Finnish and then present the features of the Finnish phonetics and phonological system that are relevant considering this thesis. Finnish is an interesting language, especially phonetically, and there are some aspects that must be left out (for example the morphophonological alternation, presented briefly in Suomi et al. 2008: 8-10) for the sake of relevance to the present paper. This chapter relies strongly on the work of Suomi et al. (2006 and 2008) since they provide a thorough view on the issue, and in need of more information consultation of their books is recommended (2006 in Finnish and 2008 in English).

6.1 Standard Spoken Finnish

Suomi et al. (2008: 7) define the Standard Spoken Finnish (SSF) as a form of speech used in education and media across Finland. It was originally based on Standard Written Finnish, which consecutively was created as a compromise between several dialects. What is interesting is that SSF was not based on the dialect spoken by the people in power at that time, but instead it was no one's native dialect and it was used by a small number of educated people – Swedish was the language of the upper class and the most educated people.

Finnish language has a number of regional dialects, and most people in Finland learn a native dialect first and then SSF. Most Finns can use both their native dialect and the SSF, the former in informal situations and the latter in formal. However, Suomi et al. (2006, 2008) do not take into account the possibility that some speakers simply may not have any specific dialectal features in their speech, i.e. their regional background cannot be distinguished based on their pronunciation of Finnish. It is also possible that a person never uses the SSF. The SSF also has some local features, “colourings”, considering e.g. prosody, and they are tolerated also in formal situations. In addition, there are informal forms of SSF emerging, which can be seen (or heard) as deletions of certain segments or replacing foreign phonemes with native ones. (Suomi et al. 2008: 7-8). In short, the SSF is not a (never-changing) monolith, but instead it also goes through gradual changes. However, it serves well as a standard for reference in the present study, which is also backed by the fact that in my BA thesis (Kivistö 2013) there were no observable major effects on the results by the participants' native dialects.

6.2 Finnish consonants

Defining the overall number of Finnish consonant phonemes is rather problematic – it varies between different dialectal variants. Suomi et al. (2006: 156-177) present a division of consonants into five groups based on their occurrence in the Finnish variants proposed by Karlsson (1983: 65-66, in Suomi et al. 2006). Group 1 has the consonants that occur in all Finnish dialects, and is therefore a core group, and group 5 has consonants occurring only in the speech of some Finnish speakers, and not even in theirs in all situations. Group 1 occurs in all variations, and usually a variant that has the consonants from highest numbered group also includes the ones in the lower groups. The graph depicting the division is included below in Figure 3.

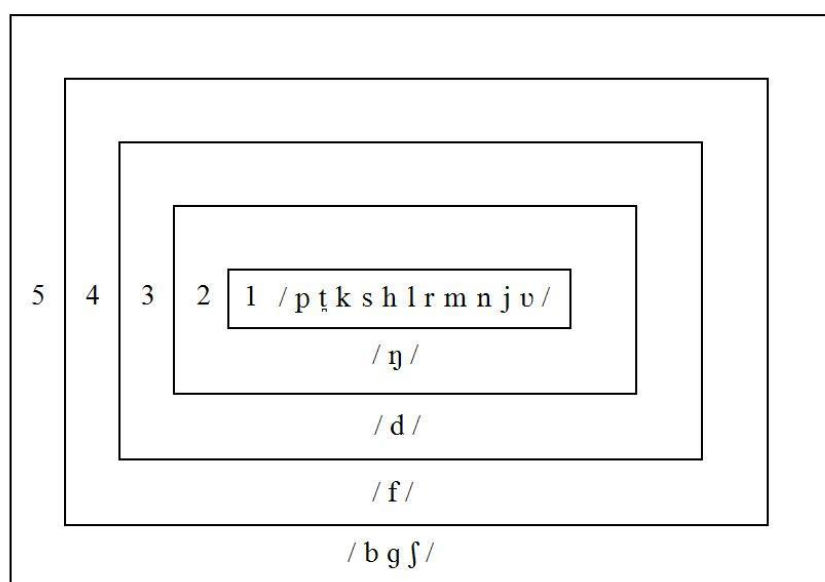


Figure 3. The division of the consonant sounds based on their occurrence in the Finnish sound system. (Suomi et al 2008: 25)

As an example of a dialect that includes the consonants in all of the five groups we can think of the Standard Spoken Finnish. As it was mentioned above, some Finns might never use the SSF, and similarly dialects with consonants from group 5 are rather uncommon.

According to Suomi et al. (2008: 26 and 2006: 157-159) the Finnish consonant phonemes have at least two allophones: rounded (i.e. pronounced with rounded lips, marked with [ʷ]) and unrounded. The rounded allophones occur near rounded vowels ([y], [u], [ø] and [o]), e.g. [ʷ] in *luumu* ‘a plum’, and the unrounded near unrounded vowels, e.g. [l] in *liima* ‘glue’, respectively. Whether the consonant is rounded or not can be predicted from the context, and

hence will not be concentrated upon in greater detail. For the same reason the consonants are not marked with the diacritical mark [ʷ] in the phonetically transcribed example words.

I will now present the groups briefly and concentrate only on the relevant phonemes, i.e. plosives and fricatives, since they are the ones with the most notable and interesting differences compared to the English phonetic system. For this reason Group 2 will be excluded.

6.2.1 Group 1 – /p/, /t/, /k/, /s/, /h/, /m/, /n/, /l/, /r/, /v/ and /j/

Plosives /p/, /t/ and /k/

The Finnish plosives /p/, /t/ and /k/ are generally voiceless and unaspirated. Therefore they do not realize as the voiced sounds [b], [d] and [g] or with aspiration as [p^h], [t^h] and [k^h] in careful pronunciation. Yet it is possible to come across the voiced realizations in rapid speech. In addition, being voiceless and unaspirated they do not differ greatly from the plosives in most languages in the world. (Suomi et al. 2006: 159)

Since we are not focusing on the roundedness of the consonants, we can see that the only relevant allophone of /p/ in Finnish is a voiceless, unaspirated bilabial plosive [p]. The main allophone of /t/ is a voiceless dental plosive [t̪]. However, it is more precise to define the Finnish [t̪] as a laminal dentalveolar consonant, the term laminal referring to the blade of the tongue touching the alveolar ridge. In addition to the laminal dentalveolar variation the Finnish /t/ also has a laminal alveolar allophone [t] that is pronounced further back in the mouth without the tip of the tongue touching the upper front teeth. This allophone occurs after [l], [s] and [r] (e.g. in words *alta* ‘from under something’, *aste* ‘a degree’ and *varten* ‘for’). The place of articulation with /k/ varies depending on its surrounding vowels; [k] and [k̟] are voiceless and unaspirated, but [k̟] is pre-velar and it occurs only before frontal vowels (compare for example the pronunciation of words *kissa* ‘a cat’ and *kassa* ‘a cash register’/‘a counter’). (Suomi et al. 2006: 159-160).

Sibilant /s/

The Finnish /s/ differs from the [s] in the IPA chart, which is slightly sharper. The most usual allophone of the Finnish /s/ is somewhere between the [s] and the [ʃ] of the IPA chart. Since the IPA chart does not include a symbol suitable for the Finnish sibilant, /s/ is widely used to represent it. In most variants of Finnish /s/ is the only sibilant, and also the only fricative if we count /h/ as a glottal. Therefore, according to Suomi et al. (2006: 161-162) it has a lot of “phonetic space for itself without any danger of perceptual confusion”. In other words, the

various possible allophones ranging from /z/ to /ʃ/ produced by individual speakers are not usually misunderstood since they all represent the same phoneme [s].

Central approximant /v/

The central approximant (or the labiodental approximant, according to the International Phonetic Association (2015)) /v/ can only appear in Finnish in the syllable onset position, i.e. at the beginning of a syllable. In addition, after diphthongs that end with /u/ it can realize as its allophone /w/ in words such as *sauva*, /sauwa/ ‘a staff’/‘a wand’, or *rouva*, /rouwa/ ‘a lady’. (Suomi et al. 2008: 31).

What is interesting is that, as stated by Suomi et al. (2006: 77), the voiced labiodental fricative /v/ does not occur in Finnish at all while it is included in the English sound system. The Finnish equivalent, the central approximant /v/ is slightly closer to the labial-velar approximant /w/. This might have some connection with the fact that, according to Korpela (2011: 38) the ‘w’ was earlier used as the grapheme for the /v/ sound. In addition, Korpela (2011: 110) mentions that even earlier in the old written Finnish the graphemes *u*, *v* and *w* were used rather interchangeably, for example the first Finnish translation of the New Testament was named as *Se Wsi Testamenti*, written nowadays also without the definitive article *se* as *Uusi testamentti*. Although the ‘v’ grapheme was established in the 20th century, *w* is still used in several proper nouns, e.g. a number of surnames, such as *Wirtanen* (Ibid. 38). However, the *v* and *w* are not differentiated in speech; *w* is usually pronounced as /v/, also in words that have an English origin (Ibid. 113).

However, the difference in pronunciation between /v/ and /v/ is minor and it is not likely to cause ambiguousness, and therefore it will not be discussed further here, although we will be coming back to this phenomenon briefly in the present study in section 9.4.

6.2.2 Group 3 – /d/

Plosive /d/

The symbol /d/ refers to a voiced alveolar plosive sound, but the Finnish /d/ is not, according to Suomi et al. (2006: 169-170), a proper plosive. The Finnish sound is apical alveolar, and the occlusion is very short – it could even be considered a flap consonant. In other words, the Finnish /d/ differs from the sound that is represented by the symbol ‘d’ (a voiced alveolar plosive) in the International Phonetic Alphabet (International Phonetic Association 2015). Since the IPA chart does not have a symbol more fitting to represent the Finnish sound, ‘d’ is commonly used. (Ibid.)

In many dialects the /d/ is substituted with another sound. For example, in the Ostrobothnian area it is common to use /r/ or /r̥/ instead, and in eastern, northern and some older Tavastian dialects it can be replaced with /j/, /v/, /w/, /h/ and /t/. Sometimes it can also be completely omitted. The /d/ has an interesting status in the Finnish phonetic system; it can appear in original Finnish words, but only word-medially, but in newer loanwords it can be also in word-initial and -final positions. (Suomi et al. 2006: 170)

6.2.3 Group 4 – /f/

Fricative /f/

In the standard spoken Finnish the voiceless labiodental fricative /f/ occurs only in rather new loanwords, such as *filmi* ‘a film’ *fasismi* ‘fascism’, and *toffee* ‘toffee’/‘fudge’. In older loanwords the /f/ in word-initial position has been replaced with /v/, e.g. in *vaari* (‘grandfather’, Swedish *far* = ‘father’) and *vaara* (‘danger’, Swedish *fara*). The /f/ sound in word-medial position has often been replaced with the sequence /hv/, e.g. in words like *kahvi* (‘coffee’, Swedish *kaffe*). Many variants of Finnish lack the /f/ completely, in which cases it is replaced with /v/, e.g. in *farkut* ‘jeans’ - [varkut]. (Suomi et al. 2006: 171-172).

6.2.4 Group 5 – /b/, /g/ and /ʃ/

Plosives /b/ and /g/, and fricative /ʃ/

The consonants in group 5 have entered Finnish in recent loanwords. Plosives [b] and [g] sometimes occur in rapid and careless speech as allophones of /p/ and /k/. However, the real issue is to determine whether they occur systematically also in careful speech, and whether they can be distinguished from [p] and [k]. There are many loanwords such as *baari* ‘a bar’, *baletti* ‘ballet’ and *galleria* ‘a gallery’, and in many varieties of Finnish these words are pronounced with a voiceless plosive, although they are written with the voiced equivalent. Therefore different words like *bussi* ‘a bus’ and *pussi* ‘a bag’ are both pronounced as [pus:i] and Finns have only learned to write phonetically identical words with different letters in order to distinguish the meaning. (Suomi et al. 2006: 172-173). With some speakers it is possible that /s/ realizes as [ʃ], but, as with /b/ and /g/, it is important to examine the systematicity of the occurrence in order to determine whether [ʃ] is used as an allophone of /s/ or as an individual phoneme /ʃ/. These phonemes belong to the rarest group of consonants in the Finnish phonetics, but there are speakers who use them in their speech as separate phonemes. Reasons for this could be, as Suomi et al. (2006: 173) suggest, that the speakers have learned foreign languages in which [b], [g] and [ʃ] are separate phonemes, that the

speaker is possibly rather young, educated and lives in an urban area, or that the speaker speaks slowly and carefully. In addition, the speaker could be using a formal register, for example when giving a presentation, and therefore it must be noted that the systematic use of the phonemes in question depends often on the register.

6.2.4 Consonant clusters

Words that are completely of Finnish origin can only contain one consonant sound in word-initial position. Also loanwords used to be modified to follow this rule; for example, the Swedish words *strand* ‘a beach’ became *ranta* and *spel* ‘a game’ became *pele*. Nowadays, however, there are a great number of loanwords in Finnish that do have a word-initial consonant cluster with more than one sound, but in spoken language and dialects they are often simplified (e.g. *stressi* ‘stress’, becomes *ressi*). (Suomi et al. 2006: 193-194)

There are some rules concerning word-medial consonant clusters that have to be addressed here because of their relation to English phonetics. According to Suomi et al. (2006: 194-196), clusters of two consonants can contain the /d/ sound only in the cluster /hd/ and the /ŋ/ sound only in the cluster /ŋk/, except in some newer loanwords such as *kandidaatti* ‘a candidate’, and *kognitio* ‘a cognition’. In addition, a plosive cannot be clustered with a nasal (except in loans like *hypnoosi*, ‘hypnosis’), and a nasal cannot be combined with an approximant or a trill (except in *genre* and *englanti* ‘English’). Also bilabial plosives cannot be combined with non-bilabial plosives (except in some loanwords, e.g. *apteekki* ‘a pharmacy’, and *kapteeni* ‘a captain’). Finally, in originally Finnish words a velar plosive cannot be in a cluster with a dental plosive, but however, loanwords with such combinations are rather common (e.g. *traktori* ‘a tractor’, and *aktiivinen* ‘active’).

Word-medial clusters of three consonants are common in Finnish (e.g. *helppo* ‘easy’, and *kurssi* ‘a course’). Clusters of four consonants are rare, and they can only be found in loanwords (e.g. *abstrakti* ‘abstract’). Word-final clusters do not generally exist in Finnish words, but they can be found in some onomatopoeic expressions (e.g. *poks* and *plumps*) and loanwords (a greeting *morjens* and *preesens*, the Finnish word for the present tense). (Ibid.: 196-197)

6.3 Finnish prosody

As it was mentioned in section 2, prosody includes phenomena in speech that relate to the pitch and the intensity of the sounds, and prosody can be examined on word level and

utterance level. In this section I will briefly review issues such as stress and intonation focusing on the Finnish phonetic system.

Suomi et al. (2006: 219) state that while some languages have word-level prosodic features that have to be marked in the lexicon, Finnish does not have such requirements, excluding situations where quantity affects the meaning of the word. For example, the primary stress falls on the first syllable, and therefore there is no need to mark it explicitly in the lexicon.

6.3.1 Stress

There are three degrees of word stress in Finnish: primarily stressed, secondarily stressed and unstressed. Primary stress is fixed on the first syllable (with a few exceptions). Secondary stress occurs in words longer than three syllables, and it is not entirely predictable; usually it falls on the third or the fourth syllable, and after that on every other syllable, but however, never on the last syllable. The place of the secondary stress depends on the segmental and morphological structure of the word. (Ibid.: 220). Based on our knowledge of the rather fixed word-level stressing in the Finnish language, we can make an assumption that this is one issue causing difficulties to Finns speaking English. The English stressing was covered in section 5.2.1.

Utterance-level stressing has an important effect on pointing out the information structure of the sentence in Finnish, and also in languages in general. Sentence stress, i.e. accent, falls on the most prominent words on sentence level. Suomi et al. (2006: 236) present a division to three sentence-level accents made by Iivonen (1998, in Suomi 2006: 236): rhematic, contrastive and emphatic. Rhematic stress refers to new information in the utterance, contrastive means there is a contrast between a word in the sentence and information presented earlier, and emphatic stress emphasizes a specifically important word. Suomi et al. (2006: 236) provide an example utterance on the issue: *Minähän inhoan Idols-kisaa* ‘But I hate the Idols competition’. This kind of an expression occurs probably during a conversation about the Idols competition, which means that the sentence accent is most likely on the word *inhoan* ‘(I) hate’. This kind of prosodic action is supposedly contrastive since the speech situation concerns the partakers’ opinions about the competition. If the speaker wishes to underline his or her hatred towards the show the accent is emphatic and the prosody regarding the word *inhoan* is even more dynamic. In contrast, the rhematic accent is used to bring up new information neutrally without certain emphasis or contrast, as in an utterance *Idols-kisa oli hyvä* ‘The Idols competition was good’. In this utterance the rhematic accent would fall on

the adjective *hyvä* ‘good’, considering that the concept of Idols competition has already been addressed in the conversation and the speaker’s opinion about it is the only piece of new information. Contrastive and emphatic accent can occur anywhere in the utterance depending on what the speaker wishes to highlight in the utterance, but the rhematic accent falls usually on the last meaningful word if the utterance provides new information. In general, the accents in Finnish are used in order to emphasize new information in different ways.

6.3.2 Intonation

Generally speaking, the intonation in neutrally uttered statements in Finnish is smoothly descending; the first syllable is uttered at or slightly above the center of the speaker’s vocal range, and the last syllable has a very low pitch. This descending pitch has led many authors to describe the Finnish intonation as “flat” and “monotonous”. However, in questioning and commanding utterances the pitch is higher than in statements. This, according to Suomi et al. (2008: 115) reflects a language universal; across languages, questions tend to have a higher intonation than statements. In general, the higher intonation in Finnish relates to the distributions of accents in the utterance, which was covered in the previous section.

Unlike in English (see section 5.2.2) the rising final intonation in an utterance does not usually occur in Finnish. A so-called *high-rise* in questions that indicate shock and amazement is possible, for example in an utterance *Hän sai siis synttärilahjaksi mitä?* ‘He/She got for birthday **what?**’ In this kind of a question, it must be noted, the final rise in *mitä* has a more emotional than syntactic function. Another example of a rising intonation can be found in the utterance *Anteeksi?* ‘Excuse me?’ in the meaning of *Mitä sanoit(te)? En kuulut.* ‘What did you say? I didn’t hear you.’ (Suomi et al. 2006: 242-243)

However, Suomi et al. (2006: 243-245) state that the rising utterance-final intonation is becoming more common in contemporary spoken Finnish, both in statements and questions and especially among young people from the Helsinki metropolitan area. In questions the rising intonation does have a syntactic function, but it also seems to indicate “uncertainty and openness”, they mention (2006: 244). For more information on the issue I suggest consulting Suomi et al. (2006: 244-255) or Routarinne (2003).

6.4 Major differences between English and Finnish sound systems

When we compare the phonetic systems of the two languages in question presented above we can see that the most notable differences relate to aspiration of plosive sounds and voicing in consonants in general; the Finnish system lacks aspiration, and although the voiced

consonants are included in the system, they do not usually realize as such, but in normal speech are replaced with their voiceless counterparts. Another difference is the occurrence of consonant clusters; they are rather usual in English words, but in Finnish they have traditionally been modified into single sounds it is also possible that in spoken language and dialects they are simplified. In addition, stress and intonation, i.e. the suprasegmentals, are an issue that contains significant differences between Finnish and English. Finnish is suprasegmentally a rather unsurprising language; the primary stress is always on the first syllable and the intonation is usually slowly descending, whereas in English the stress can vary and the intonation is rising in yes/no questions.

These are the main differences between English and Finnish phonetics regarding the issues that were presented in the previous sections, and therefore they can also be seen in the hypotheses that will be covered later in section 8.1. In the next section, however, I will present briefly a study by Kari Suomi (1980) that focuses on the pronunciation of English plosives by Finnish learners.

7. VOICING IN ENGLISH AND FINNISH STOPS (SUOMI 1980)

Kari Suomi (1980) conducted a study on voicing in English and Finnish stops in January 1978, and it was published in 1980. Although a rather long period of time has passed since the study was published, it is still relevant, especially because of its categorization of pronunciation errors, but also because it is still the only study on Finnish speakers' pronunciation of English that fits in the framework of the present paper. In this section I will briefly present the study and its main findings concentrating on a section of English produced by native speakers of Finnish. The aim of this section of Suomi's study was to give a description of the learner language in relation to their native language Finnish and the target language English.

7.1 The Study

Suomi (1980: 54-55) examined the speech of three different groups of participants in the study: native English speakers, native Finnish speakers with secondary school level training of English and native Finnish speakers with university level training of English. There were a total of fifteen participants in the study, five in each group, and all of them were male. The

participants in the native speakers' group were British English speakers, the more advanced group of Finnish speakers had had about three years of English training at university, and the less advanced Finnish group were secondary school students in their last form.

According to Suomi (1980: 56), the participants read out loud words set in constant English and Finnish frame sentences: *Say ___ loudly* and *Joko ___ luettiin*. The purpose of using frame sentences was to have the sample words in a similar phonetic environment, both in English and Finnish. There were 206 different simple mono- or disyllabic English words used in the study with plosive consonants in initial, medial and final positions, and they were easy to pronounce. In disyllabic words the stress was always on the first syllable. The total number of sample words provided by the 15 participants was 3090, and there were about 4410 stops altogether. The English words were read out loud and recorded by all of the participants, and in addition, there were 72 Finnish words read by the Finnish participants. As a result, there were three types of data: English samples provided by native English speakers, Finnish samples provided by native Finnish speakers, and English samples provided by Finnish learners of English. (Suomi 1980: 56-57)

In the following section we will focus on the English stops produced by the Finnish learners for the sake of relevance considering the present thesis.

7.2 The English stops produced by the Finnish speakers

7.2.1 Word-initial and word-medial position

In general, Suomi (1980: 113) presents, the voiced /bdg/ sounds in the study were pronounced as voiceless unaspirated, moderately voiced and extensively voiced, and the voiceless /ptk/ sounds as voiceless aspirated or voiceless unaspirated. Furthermore, based on the findings it seems that producing the voiceless /ptk/ was easier for the Finnish learners than the voiced /bdg/ (Ibid. 114). In addition, a few instances of /ptk/ realizing as voiced and /bdg/ as voiceless aspirated were found from the samples of the Finnish participants. The native English speakers did not make such errors. The reasons for these kinds of errors are unknown; they might be due to the learners' internalized rules of English pronunciation or simply just slips of tongue. In the study there were a fairly large number of voiceless unaspirated realizations found from the samples, which could be explained by native language interference or by phonological simplicity compared to voiced or aspirated sounds. It is also possible that the two factors, interference and simplicity, emphasize each other and thus increase the occurrence of the voiceless unaspirated realizations.

7.2.2 Word-final position

As stated by Suomi (1980: 140-141), categorizing the word-final plosives to voiced and voiceless sounds was not accomplished meaningfully in the study since most of the informants did not produce an adequate sample of voice onset time, i.e. the sound for analysis, in word-final position. However, the duration of the vowels preceding the word-final plosives is in English a more significant issue than the actual release of the plosives. In other words, the release of the plosive has a rather low functional load, which may have an effect on this matter. In addition, there was variation in the voicing of the word-final /bdg/ sounds, and therefore, instead of the plosive sounds themselves, it is more relevant to concentrate on the preceding vowels.

It was found that in the samples of seven of the ten Finnish informants there was no distinction between how /ptk/ and /bdg/ affected the duration of the preceding vowel. In Finnish the quantity (i.e. duration) of vowels has a meaning distinguishing function, but it is interesting how Finns in the study could not distinguish differences in duration when their linguistic function is different from the function of vowel duration in Finnish. It could be suggested that in this case the phonological system of the native language acts as a filter on the perception of this phonological phenomena of the target language. The filter effect is a type of interference from the native language. (Suomi 1980: 150-151)

7.3 Categorizing the samples given by the Finnish learners

Suomi (1980: 151) divides the sample sounds provided by the ten Finnish participants into four groups:

- (1) learners use a correct target language rule,
 - (2) learners use a rule that is faulty in terms of the target language but bears an obvious resemblance to and is readily identifiable as an attempt at the correct target language rule,
 - (3) learners obviously lack a target language rule, and
 - (4) learners are in possession of a rule that is definitely not used by native speakers.
- (Suomi 1980: 151)

These groups can be analyzed further by examining whether similar rules, *i.e. phonetic phenomena*, can be found in the learners' native language. Suomi (1980: 152) states that if a learner uses a rule, i.e. a phonetic form, which cannot be found in the target language but occurs in the native language we may assume that interference has occurred. If, however, there are similar rules in the foreign and the native language it is not possible to determine whether using the correct form stems from positive transfer or simply learning the correct

rule. Certainly we can argue that positive transfer helps learning the rules, and therefore the relevance of examining positive transfer and learning separately can be questioned. As Suomi (1980: 140) summarizes, “our mother tongue determines for us what is easy and what is difficult to pronounce and to perceive”.

Suomi (1980: 152) gives an example of group 1, the correct use of target language rules; the Finnish participants varied the duration of the occlusion in the plosive sounds even though similar variation does not occur in Finnish (e.g. pronouncing the [æ] sound in *mad* as slightly longer). For the second group Suomi (Ibid.: 152-153) mentions overlapping of voiced and voiceless stops, i.e. regression to the native language rules (e.g. substituting [z] with /s/ in the plural form *peas* – /p^hi:s/). For the third group (Ibid.: 153-154), some learners failed to lengthen the vowel before voiced plosives in word-final position (e.g. pronouncing the [æ] sound in *mad* as a short vowel, as in *mat*), which could stem from ignoring the rule, lacking the ability to differentiate durational patterns in a foreign language, or being exposed to an incorrect pronunciation model from their native Finnish teacher, i.e. “transfer of training”. An example of the fourth and final group (Ibid.: 154) is slips of the tongue that do not seem to arise from either the native or the target language, but instead they might relate to ease of articulation and the anatomy of the human vocal system.

7.4 Evaluation

The data of Suomi’s study was gathered in January 1978, which naturally has an impact on how it should be viewed. The significance of English in the world and in Finland has greatly increased since, and it is likely to have affected on how English is being taught in schools nowadays. In addition, as it was discussed in chapter 1 of this paper, the amount of authentic input of English has increased. Suomi (1980) describes as an example of the third group of the study a phenomenon that he calls “transfer of training”. Making errors in pronunciation only because the teacher makes similar errors is still possible but less likely since it is unavoidable to hear the language from native speakers via, for example, television, the internet and also games. Native Finnish speaking English teacher is no longer the only or even the main model for learners. For these same reasons it is fairly safe to make the assumption that the Finnish learners’ skills, especially oral, in English have increased in average during the last few decades.

In addition, the number of participants, fifteen, is rather small, and furthermore, they were all male. The participants did provide an ample of data for analysis, but still, drawing

any generalizable conclusions based on the speech of 15 people is slightly problematic. However, the findings serve as an excellent framework for further studies.

The same study was used as an outline for analysis in my BA thesis (Kivistö 2013), and therefore it will be more than suitable to be used similarly also in this MA thesis. Although Suomi concentrated on analyzing only plosive sounds, the criteria he used for categorizing the errors is applicable to any other phonetic phenomena as well. One of the limitations in the study is that he does not provide concrete examples of his findings, and that is a gap that the present thesis aims to fill. Describing the different ways a speaker's output differs from the target language rules and finding reasons for it gives important, practical insight in the matter, and can also be applied to teaching or further studies.

8. THE PRESENT STUDY

In this chapter I will introduce the study conducted for this MA thesis, i.e. present the data, methods and participants. After that, in chapter 9, the results of the study are presented and analyzed, and lastly, in chapter 10, the results are discussed and the findings will be summarized and evaluated.

8.1 Data and methods

The aim of this thesis is to provide concrete examples and thorough analysis of the different ways the pronunciation of native Finnish speakers can differ from the English language rules. Another important goal is to find possible reasons for those differences and determine how strong an effect language transfer has in them. Studies regarding cross-linguistic influence and phonetics have been conducted earlier, but the work of Kari Suomi is essentially the only material available on the relation of Finnish and English. Consequently there is need for more research on the subject, and especially regarding the two languages in question.

The research question in this thesis is in two parts:

1. How does the pronunciation of the participants differ from the target language rules?
2. What kind of an effect does language transfer have in those differences?

Two participants were chosen to take part in a pronunciation test that was recorded, and then their samples were analyzed according to the target language rules. The differences between the Finnish and English phonetic systems served as an outline for initial analysis,

and the four groups by Suomi (1980) that were presented in the previous section were used as a framework when categorizing individual sounds from the samples. The aim was to provide concrete examples and thorough analysis of the different ways the pronunciation of native Finnish speakers can differ from the English language rules.

The participants took a pronunciation test that was created for the purpose of this study only. Creating the test solely for this purpose gave the chance to control the sounds in the samples so that there would be as much analyzable content as possible. The test consisted of a longer text (Part 1), word pairs (Part 2), a short dialogue (Part 3) and words in a frame sentence (Part 4). The longer text was chosen to be used because it would present the sounds in a normal phonetic environment. In addition, when reading the longer text the participants would not pay too much attention in single sounds. Word pairs were chosen to see if the participant could differentiate sounds in minimal pairs, but also to give them a chance for more careful pronunciation. The short dialogue was included in order to see if the participants would apply different intonation to, for example, different kinds of questions, and the words in a frame sentence were used to present the words in a similar phonetic environment so that the surrounding sounds would not affect the pronunciation. The frame sentence was the same that was used by Kari Suomi (1980) in his study (*Say ___ loudly*). Furthermore, after the initial test there was a brief conversational section in English, in which the participants answered a couple of questions regarding their everyday life. This part was included in order to gather data also from normal, natural speech.

Part 4 contained also pseudowords, i.e. words that can be pronounced according to the English language rules but are not real words. It must be noted, however, that the pseudowords must be treated with certain cautiousness in the analysis since there are no established models of pronunciation for them and, in addition, the rules of pronunciation in English are rather irregular in general. The test used in the study is included in Appendix 1.

Before taking the test the participants had an opportunity to read it through once. Each section of the test was read out loud twice before moving on to the next one, which provided a chance to be more familiar with the section on the second reading. Re-reading it immediately also helped remembering the material in the section and hence the second run was slightly more fluent. The participants were instructed to read the test as naturally as possible and not pay excessive attention to their pronunciation, and they were also informed that the point was not to give examples of perfect articulation but instead to pronounce the words like they would normally do. They were also given the instruction not to cling to difficult words but instead to go on. However, they both tended to repeat words that they felt

they pronounced incorrectly. The test was similar to the one used in my BA thesis (Kivistö 2013) with the exceptions of covering more phonetic issues and including a free conversational section. The test was not piloted since a similar test proved to be a success in the BA thesis, and provided more than enough data for analysis.

There were also short interview sections in Finnish before and after the test. The interview before the test covered briefly the participants' backgrounds and their views on their pronunciation of English. After the test the participants had a chance to reflect on their performance and give feedback on the test. The interviews were recorded and transcribed, and they were done in the participants' native language Finnish so that it would be as easy as possible for them to express their thoughts.

The data was gathered in March 2016, and the test was recorded with a Roland R-09 Wave/mp3 recorder. Before the actual analysis the stereo recordings were converted to mono format and a noise reduction was applied to them in Audacity, a free open-source audio recording and editing software. The noise reduction removes all extra background sounds that might interfere with the analysis. There is no risk of losing important sound information since the user can determine which frequencies should be removed. After that the recordings were carefully examined in order to find samples of interest and relevance for further analysis. The samples were then studied by using the Praat software developed by Paul Boersma and David Weenink from the University of Amsterdam. Praat shows the sound files in waveform images and lets the user select sections of sounds and create images that contain, in addition to the waveform, e.g. pitch and intensity curves. An equally important method of analysis was simply listening to the samples. There were instances in which the software's pitch analysis could not, for example, distinguish actual voicing in a sound from the slight echo in the room during the recording situation, which was clearly distinct by human ear. Furthermore, especially in some words in sentence-final position the pitch descended rather low and the voice of the participant distorted slightly, and the pitch curve could not identify the voicing. Therefore it was important to listen to the sounds carefully. Any possible irregularities in the pitch curves are pointed out in the analysis and images when relevant. The headphones used for listening were Klipsch Image S4 in-ear noise-isolating headphones with frequency response from 10 Hz to 19 kHz. The frequency response of the headphones was more than sufficient considering that the human hearing range is from 20 Hz to 20 kHz (National Physical Laboratory 2016) and the average human voice is around 200 Hz for adult females and 125 Hz for adult males (The National Center for Voice and Speech 2016).

The analysis of the samples took place in April 2016. Analyzing cross-linguistic influence has its difficulties, as Jarvis & Pavlenko (2007: 27) point out. CLI cannot always be detected by looking at the data if it is very subtle or "obscured by other factors", in which cases it may be difficult to differentiate it from the other factors. It is difficult to isolate, identify and measure the effects of CLI and it is also possible that the researcher sees something that seems to be CLI but in fact is not. It has to be remembered that also in this study the analysis is based strongly on subjective findings and must be treated as such.

The methodology, according to Jarvis & Pavlenko (2007: 28), used to identify the instances of transfer depends on the scope of the study. The relevant method of collecting data for analysis can involve e.g. conducting surveys and questionnaires, consulting grammars of the languages in question, gathering empirical observations or compiling a corpus of contemporary language use from a cross-section of speakers. After collection the data is then analyzed either quantitatively or qualitatively. It is also possible to analyze the frequency of borrowed structures.

In the present study the method used to collect the data could be described as empirical observation and elicitation. The samples were carefully chosen from the recordings in a way that they would give clear and interesting representations of the issues under scrutiny. Because the goal was not making generalizable conclusions based on the data, the chosen samples were analyzed qualitatively based on the previous research. These methods were chosen while keeping in mind the purpose of the study.

The main goal of the study was to find concrete examples of sounds and issues in a non-native speaker's output that differ from the target language rules and to see whether the differences can be considered cross-linguistic influence from the phonetic system of the participants' native language Finnish. The data was collected by using elicitation technique, i.e. recording the needed information directly from the participants. With direct recording it was possible to give the participants clear instructions, and they could also ask for more information and help in any unclear situations if it was needed.

The hypotheses in this study related closely to the differences between the English and Finnish sound systems, which were presented in section 6.4. The most prominent differences concern issues such as voicing, aspiration and consonant clusters, and also suprasegmental features like stress and intonation. Consequently, it was expected that, in addition to pronunciation that follows the target language rules, examples of different types of "errors" in pronunciation could be found in the samples relating to the phonetic features listed above.

8.2 The participants

In this section the participants of the study are presented and their learning history of English and their own views of their pronunciation will be briefly described.

Participant 1 was female, 25 years old and at the time of the recording lived in the Central Finland area. She was a university student but not majoring in any language. She had received very good grades in English at school, and she had always felt very confident about her skills. Apart from school she had learned the language by watching TV-series and movies in English. She remembered that in upper secondary school she had had a very good English teacher who was very proficient also in pronunciation. They had quite a lot of practice in pronunciation and the teacher was very instructive and well oriented for it. The participant felt the amount of explicit instruction adequate in school and did not feel like she would have needed more of it.

She had lived in England for a year six years before the recordings. She felt that the year spent in England benefited her pronunciation of English and that it still had some effect on it. A couple of years before the recording she had also lived a year in Norway, which, on the contrary, had a negative influence on her pronunciation because she spoke English mostly with other exchange students.

She described that she did not speak English regularly, although she read a great deal of books in English regarding her studies at university. Earlier she did also speak English more with her MA thesis instructor who was a native English speaker.

She felt English pronunciation easy, especially after spending the year in England, but she also mentioned that she sometimes makes “the usual Finnish mistakes”, e.g. not distinguishing the voiced /b/ from the voiceless /p/. She described that she does not pay very much attention to her pronunciation when she speaks English.

In her own words she has always aimed to as correct pronunciation as possible, and that she has considered British English as a desirable model. She also mentioned that she likes different accents and does not think that everyone should sound the same. On the contrary, she stated that she does not really like the way Finns pronounce English and described it as “funny”.

She did not feel that there are any specific phonetic issues in general that are usually more difficult for her, apart from the voiced/voiceless distinction, but that she instead sometimes notices certain words that she has problems with and could not distinguish what it is that causes them.

Participant 2 was also female and 25 years old at the time of the recording. She lived in the Helsinki metropolitan area. She had finished a Bachelor's degree at university (not in any language) and after that switched to vocational school and studied a degree in care business. Her English grades had, contrary to participant 1, never been good, and she had had some problems with learning the language. She is dyslexic, which has probably influenced her learning. She described that she has always felt pronunciation unpleasant and difficult. In her English lessons pronunciation was taught explicitly by repetition and reading out loud, and she felt that the teaching methods did not meet her needs at all and she was not motivated; she would probably have benefited more from implicit learning, i.e. learning the pronunciation with the help of discussion and conversational exercises. Furthermore, the teachers at that time, especially in elementary school, did not really take her dyslexia into account but instead just deemed her to be "dumb".

She stated that pronouncing English is still rather difficult for her, but that her skills in English in general have improved after her school years. She mentioned that she uses English sometimes at work with patients who do not understand Finnish. She could not say whether her pronunciation skills had improved or not, but she stated that she considers communicativity more important than perfect pronunciation. In her opinion her communication in English is usually successful, and she does not really pay attention to pronunciation when she speaks.

The skills and language learning history of the two participants is very different, which brings an interesting view into this study. Naturally the meaning here is not to compare their skills or the errors they make, but these differences and their effects are briefly addressed in section 10. One of the main reasons for choosing these two participants was that despite their differences in learning history and language skills they both are rather fluent speakers of English and they are not nervous or anxious about using a foreign language, and neither about recording their speech. Therefore the recording situations were very pleasant and relaxed, and there were no signs of speaker anxiety in either of the samples. In addition, a female voice is easier to analyze considering the pitch of the voice; the program picks up a higher pitch more easily compared to a lower male pitch. The speakers' native dialects proved to have no connection to the learner's oral output in English on my BA thesis, and consequently they were not taken into account here.

9. RESULTS – THE SAMPLES ANALYZED ACCORDING TO THE FRAMEWORK

The analyzed samples were divided into the four groups by Suomi (1980), which were presented in section 7.3 and they are listed here as well:

- (1) learners use a correct target language rule,
 - (2) learners use a rule that is faulty in terms of the target language but bears an obvious resemblance to and is readily identifiable as an attempt at the correct target language rule,
 - (3) learners obviously lack a target language rule, and
 - (4) learners are in possession of a rule that is definitely not used by native speakers.
- (Suomi 1980: 151)

In this section I will give six examples of each group, and every example is explained and analyzed. The examples were chosen from the recordings of each part of the test in a way that they would represent various types of differing pronunciation as comprehensively as possible. The purpose was to include various types of phonetic issues, i.e. segmental (e.g. different types of consonants) and suprasegmental (i.e. prosodic) features. With groups 2 and 3 it is relevant to assess whether the examples can be regarded as cross-linguistic influence from the speakers' native language Finnish. In group 1 the examples represent correct pronunciation and in those cases determining the positive transfer from Finnish is difficult and irrelevant. Moreover, the examples in group 4 cannot be considered to have been affected by language transfer since, as Suomi (1980: 154) stated, the samples in the last group are rather just "slips of the tongue that do not seem to arise from either the native or the target language".

The waveform image and/or the pitch and intensity curves accompany most of examples, and all the rest of them are found in Appendix 2. The thinner line in the curve picture represents the pitch curve and the thicker line the intensity curve. In some examples there was slight distortion in the pitch curve which related to, for example, the slight echo in the recording situation or to the low pitch of the sound in sentence-final positions. The waveforms and curves of all sample sounds mentioned here are also included in Appendix 2. After the examples the participants' own views of the test and their pronunciation will be presented and discussed, and in the conclusive chapter the findings are analyzed further.

9.1 Group 1

The first group included samples of the participants using a correct target language rule.

One example of a correct pronunciation is the distinction of voiceless /f/ in *few* - /'fju:/ and voiced /v/ in *view* - /'vju:/ in the second recording of part 2 of the test from participant 1. When we look at the pitch curve (Figure 4, lower picture) we can see that in *few* the /f/ is clearly voiceless – the voicing does not begin until at the /j/ sound. On the contrary, in *view* the /v/ sound is clearly voiced since the pitch curve begins at the beginning of the sound. Although there is slight disturbance in the pitch curve, the voicing does last throughout the whole utterance. In addition, the waveforms look different; the voiced /v/ grows more steadily and evenly compared to the voiceless /f/ – in *few* the /f/ sound is more quiet, and the loudness of the sound grows more rapidly at the /j/ and /u/ sounds. The image is included also in Appendix 2 (Image 1).

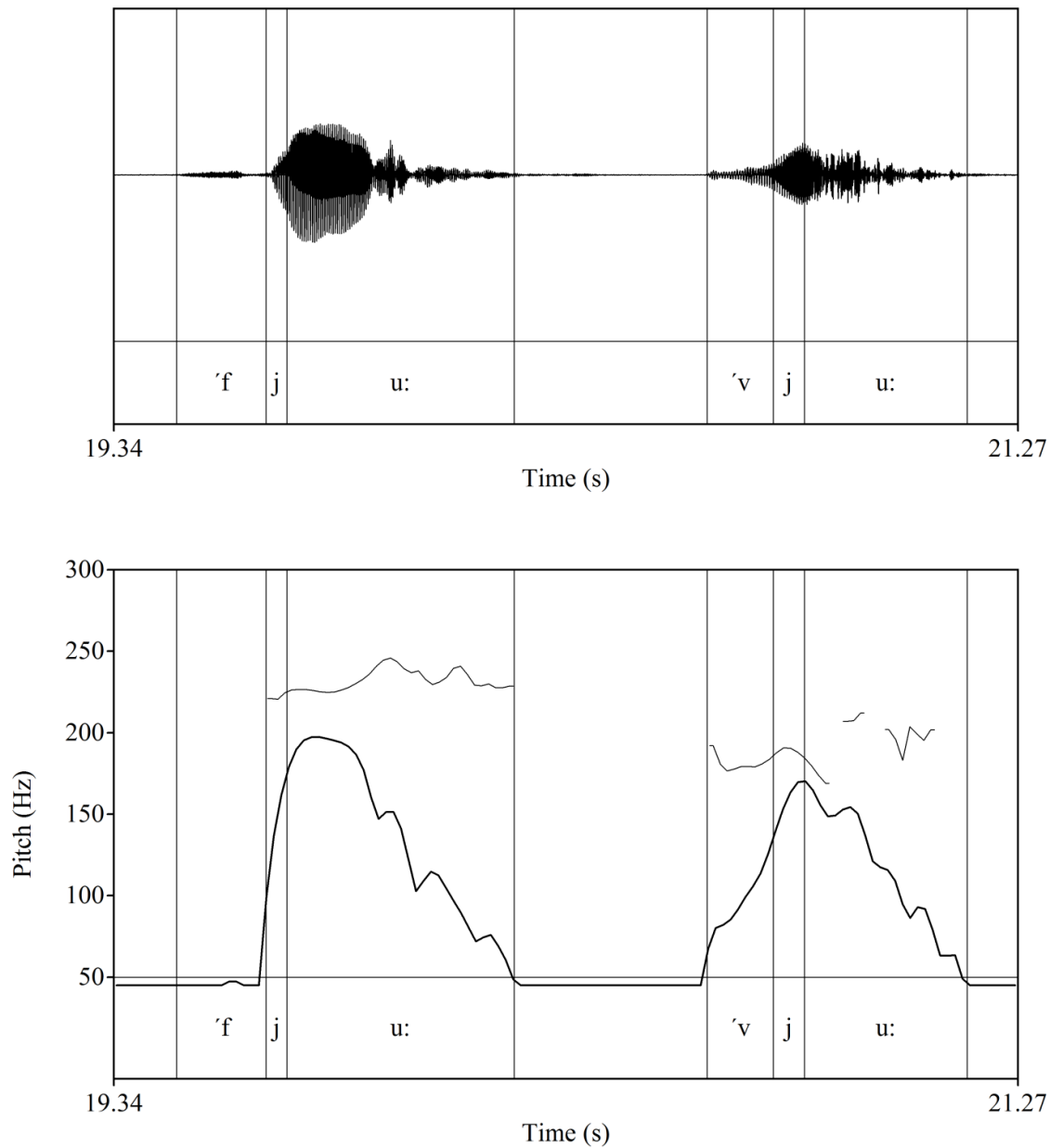


Figure 4. *few/view*. Participant 1, part 2 of the test, second recording. Intensity: thick line, pitch: thin, dotted line.

It was presented in section 5.1.1 that aspiration is one significant phenomenon that differentiates voiced and voiceless stop consonants, and the word *coat* from participant 2 (part 1 of the test, first recording) is a good example of this. When we take a look at the waveform of the sound (Figure 5) we can see the puff of air right before the vowel sound (the diphthong /*ɔʊ*/). In addition, the pitch curve is visible only during the vowel. The image with the pitch and intensity curves is included in the Appendix 2 (Image 2). The participant produces, for some reason, a slight aspiration also on the /*t*/ at the end of the word, which can

be seen in the waveform as well. This kind of overpronunciation of the plosive can be a mere slip of the tongue, and therefore it could serve as an example of the fourth group, presented later in section 9.4.

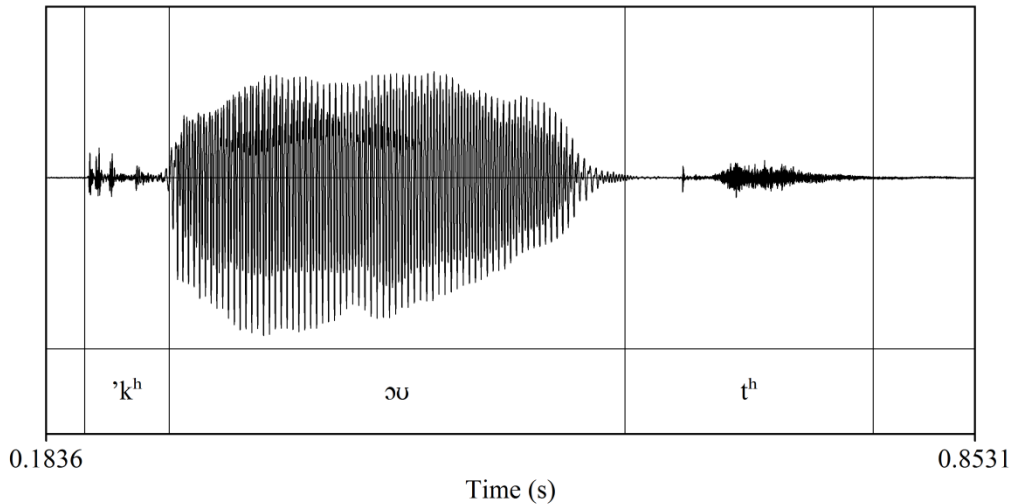


Figure 5. *coat*. Participant 2, part 1 of the test, first recording.

An example of lengthening the vowel that precedes a voiced vowel can be seen in the way Participant 1 distinguishes *mat* (part 1 of the test, second recording, Figure 6) from *mad* (part 4 of the test, second recording, Figure 7). As we compare the lengths of the sounds we can see that the /æ/ in *mad* is significantly longer in duration, almost twice as long, although the words are otherwise relatively as long. This stems from the duration of the occlusion in the voiceless /t/ in *mat*; the occlusion before the burst of the sound is even longer than the vowel /æ/ preceding it, as in *mat* the burst of the plosive /d/ comes immediately after the vowel. The word *mat* is included also in Appendix 2 (Image 3).

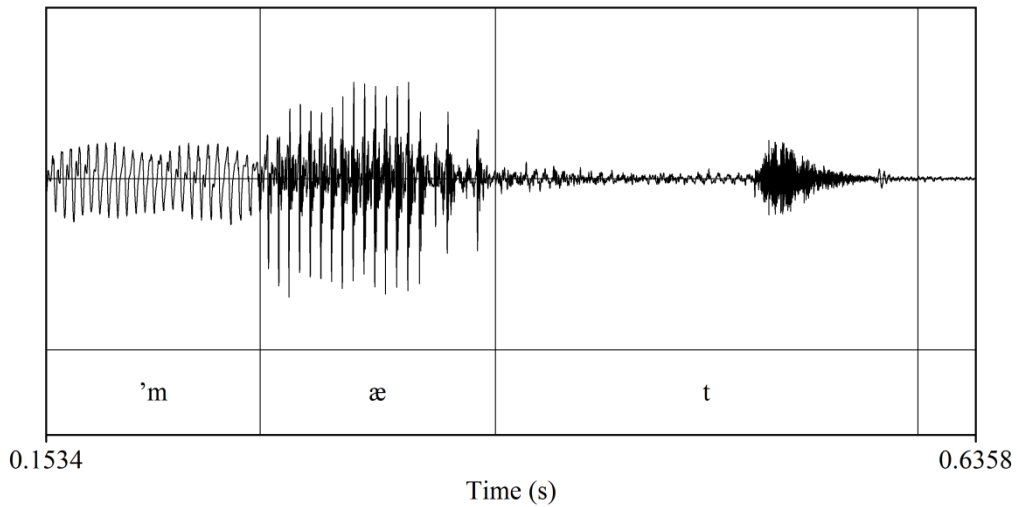


Figure 6. *mat*. Participant 1, part 1 of the test, second recording.

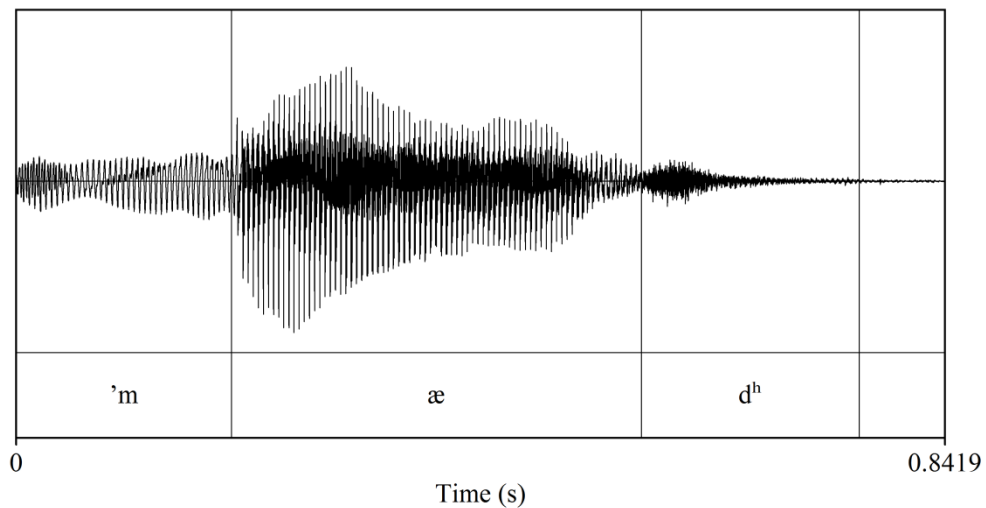


Figure 7. *mad*. Participant 1, part 4 of the test, second recording.

Correct pronunciation regarding voiced plosives can be seen in participant 2's sample of word *blueberries* from the free discussion section of the test (Appendix 2, Image 4); both of the /b/ sounds are clearly voiced (the pitch curve in Figure 8 continues almost throughout the whole word). Furthermore, also the /i/ sound is lengthened since it follows a sound that is supposed to be voiced but which, however, realizes in the sample as only moderately voiced, and is thus transcribed as /s/. This represents the fact discussed in section 5.1.1 that although the supposedly voiced consonant might not always in normal speech realize as voiced, the length of the preceding vowel differentiates it from its voiceless counterpart.

Although the different *r*-sounds were not an intended issue of analysis in the thesis I wish to address one interesting detail in this sample. In many examples from the recordings from participant 2, mainly sections 1-4 in which she read the pre-written text from the paper, pronounced the *r*-sounds as the trill variant /r̄/, whereas in this example, and during the free conversation in general, she used the approximant /ɹ/ that is more common in e.g. RP and American accent. This issue will not be discussed further because the *r*-sounds were not intended to be taken into account in the study. In the transcriptions the sounds are all transcribed with /r/ for the sake of clarity.

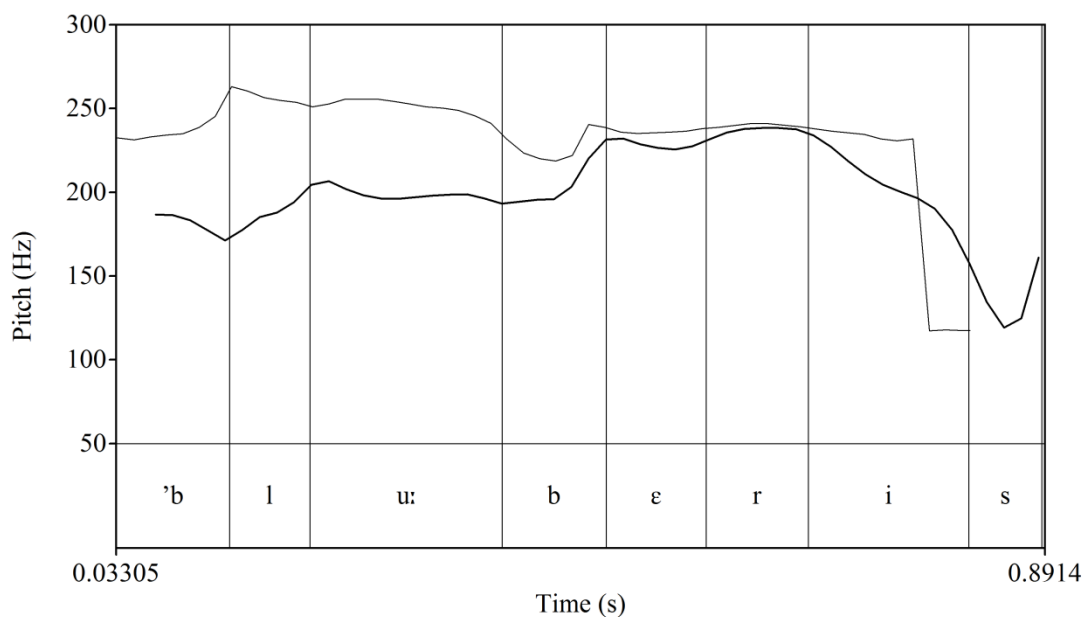


Figure 8. *blueberries*. Participant 2, free discussion. Intensity: thick line, pitch: thin, dotted line.

The question *Is it funny?* from participant 1 (part 1 of the test, second recording) is a good example of group 1, using a correct target language rule. As it was mentioned in section 5.2.2, the intonation in yes/no questions is rising in the English language, and so is the intonation in this sample. The pitch curve (thin, dotted line, Figure 9) rises clearly at the /ʌ/ sound. This sample can also be found in Appendix 2 (Image 5).

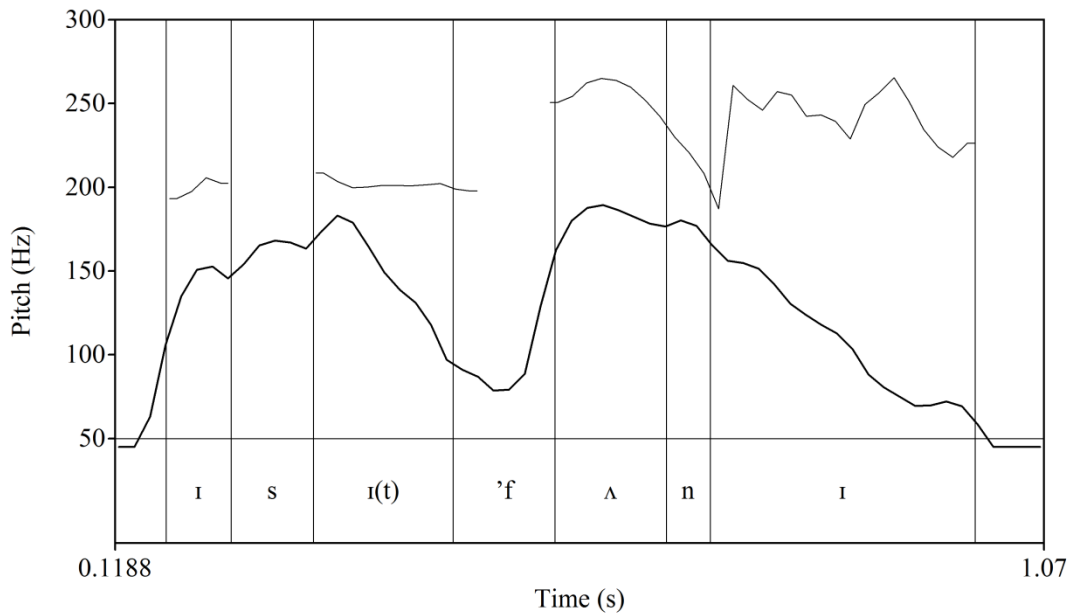


Figure 9. *Is it funny?*. Participant 1, part 1 of the test, second recording. Intensity: thick line, pitch: thin, dotted line.

The word *astonishing* (from participant 2's first recording of part 1 of the test) fits in the first group especially because of the stressing in it. The stress in the word is on the second syllable, which is visible also in the highest point of the intensity curve (thick line in Figure 10); the curve starts rising in the end of the /t/ sound and is at its peak in the vowel /ɔ/. In addition, also the plosive (/t/) and the fricative sounds (/s/ in comparison with /ʃ/) follow the correct target language rules. The waveform is included in Appendix 2 (Image 6).

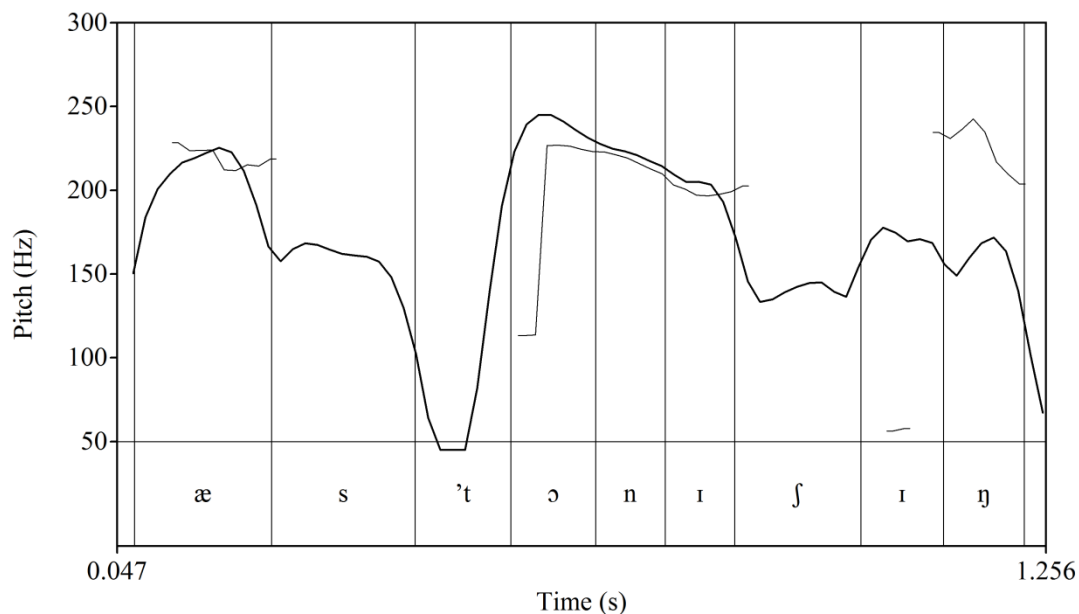


Figure 10. *astonishing*. Participant 2, part 1 of the test, first recording. Intensity: thick line, pitch: thin, dotted line.

9.2 Group 2

The second group has examples of learners using a faulty target language rule which still resembles the correct rule. In section 4.2 it was mentioned that transfer can often be observed as producing the right phonemes with wrong segmental properties, e.g. voicing and aspiration (Jarvis & Pavlenko 2008: 65), and many of the examples in this group present this phenomenon.

An example of the second group is the word *exposed* – [ɪk'spəʊzd] – in the second recording of the first part of the test from participant 1, which is pronounced as /ɪk'spəʊsd/. The waveform image can be found in Appendix 2 (Image 7). The sound that is marked with the grapheme *s* at the end of the word should be pronounced as a voiced [z], especially since the following sound [d] is voiced, and also the following word, *in*, begins with a voiced sound (the vowel /ɪ/). However, the sound in the sample is clearly pronounced as the voiceless sibilant /s/. The fact that the sound used to substitute the correct [z] sound is its voiced counterpart, and phonetically very close to it, can be regarded as an attempt at the correct target language rule. In addition, in the /d/ sound at the end there is no observable action in the vocal cords, but still it can, however, be distinguished from the voiceless [t] because of differences in the place of articulation between [t] and [d]. Furthermore, it could be argued that since the /p/ at the middle of the word is in a stressed position, it could be somewhat more intensive, possibly even with a slight aspiration.

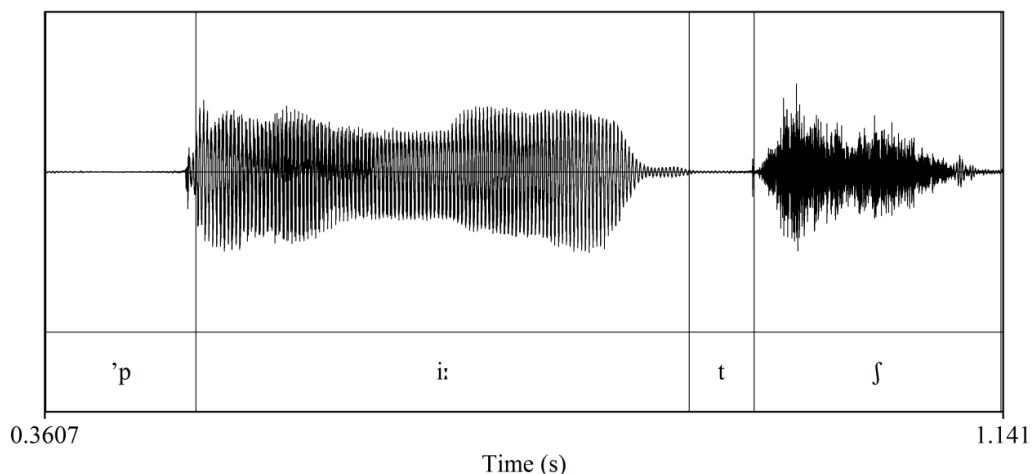


Figure 11. *beach*. Participant 2, free conversation.

Missing voicing in the word *beach* (Appendix 2, Image 8) from participant 2's free conversational section is the reason why the word fits in the second group. When we take a

look at the waveform in Figure 11 we can see that the bilabial consonant at the beginning of the word is voiceless. Voicing in the consonant would be visible in the waveform, as it is for example in the two /b/ sounds in the word *blueberries* analyzed in section 9.1 (Appendix 2, image 4). However, categorizing this example in group 2 is not completely unproblematic; as mentioned in section 5.1.1, according to Ladefoged (2005: 56) the voicing in voiced sounds can sometimes be rather weak, and in this case the word is understandable albeit the missing voicing. Nevertheless, if we compare this sample to the word *blueberries* we see a clear difference in the waves, the main difference between the waveforms of the voiced /b/ and voiceless /p/ being the absence of wave in /p/ during the closure of the plosive. The wave does not begin until the burst of the sound, whereas in /b/ it looks rather stable and even. Although the word is understandable, the initial plosive clearly differs from the target language rule.

The word *car* from participant 2 (part 3 of the test, first recording) fits in the second group because it is missing the aspiration in the /k/ sound. In the waveform in Figure 12 the aspiration should be clearly visible as a combination of a peak at the release of the sound and a smaller wave following it before the beginning of the vowel /a/. In other words, the VOT (voice onset time, presented in section 4.2) should be distinctly longer as it is in this example. The word is included in Appendix 2 (Image 9).

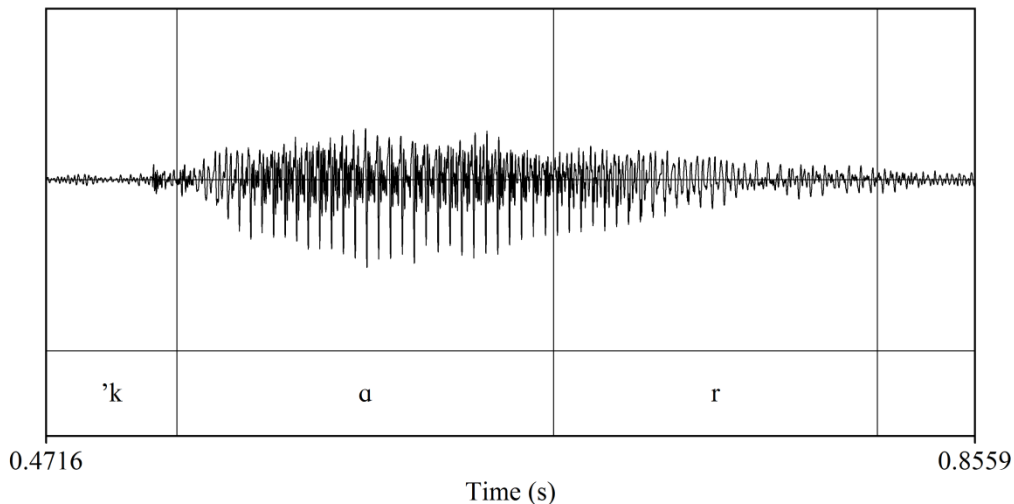


Figure 12. *car*. Participant 2, part 3 of the test, first recording.

The sample of participant 2 pronouncing the phrase *perfect bird* (first recording of part 1 of the test, Image 10 in Appendix 2) fits well in the second group of Suomi's classification. The main issue of interest here is the lack of aspiration in both of the voiceless bilabial /p/

sounds. In stressed (in this case word-initial) position the sound should be aspirated in order to distinguish it from its voiced counterpart, /b/. However, as we can see in the /p/ sounds in waveform image, the following vowel sound starts immediately after the burst of the plosive, and there is no time for the aspiration to take place. In other words, the voice onset time is very short. The lack of the aspiration is even clearer when we compare this waveform to the waveform in the word *coat* in Figure 3, which was analyzed earlier. Another interesting issue in this sample is the slight aspiration in the last sound, /t/, which realizes as /t^h/, even though the intended sound is in fact the voiced /d/. Although the voicing can be rather weak or even nonexistent in such word-final positions, and therefore a voiceless /t/ could possibly be considered an acceptable allophone of /d/, the aspiration is definitely against the target language rules. Similar examples will be analyzed in greater detail in section 9.4 regarding group 4 of Suomi's classification.

As it was mentioned in section 6.2.5, in consonant clusters some sounds are often dropped in order to ease the pronunciation. Pronunciation of the word *attempts* from participant 1 (part 1 of the test, second recording, Image 11 in Appendix 2) is an attempt towards the target language rule, and a great example of the second group; the participant attempts the pronunciation twice and drops the /t/ sound for the sake of easier pronunciation on her first try. However, she then corrects her utterance and instead drops the /p/ sound. Dropping the /p/ is more natural and logical since there is already a bilabial sound, the nasal /m/, which in a way substitutes the /p/. The release of the bilabial nasal /m/ creates an illusion that the /p/ would actually be there. If the alveolar plosive /t/ is dropped from the sound the illusion will not occur since the /p/ is already there, but instead the word sounds, and is, incomplete without the /t/.

The last example from the second group is the comparison between words *this* and *thing* from participant 1's second recording of part 2 of the test (Appendix 2, Image 12). The dental fricatives [ð] and [θ] differ from each other only by their phonatory status, meaning that [ð] in *this* [ðis] is voiced and [θ] in *thing* [θiŋ] is voiceless. In this sample both words have in word-initial position the /θ/ sound, i.e. also *this* is pronounced with a voiceless fricative: /θis/. The voicelessness is visible in the pitch curve, but also in the waveform; a voiced /ð/ would be visible as a rather long and even wave before the vowel. Instead when we look at the waveform in Figure 13 we can see that the fricative sound is very short and practically identical in both words. Since English does not have a word that would be pronounced that way there is essentially little chance for misinterpretation, and the message would probably get through anyway. We can still say that this pronunciation does not follow the target

language rule, but is definitely an attempt at the right direction since the dental fricative is correct with the exception of the phonatory status.

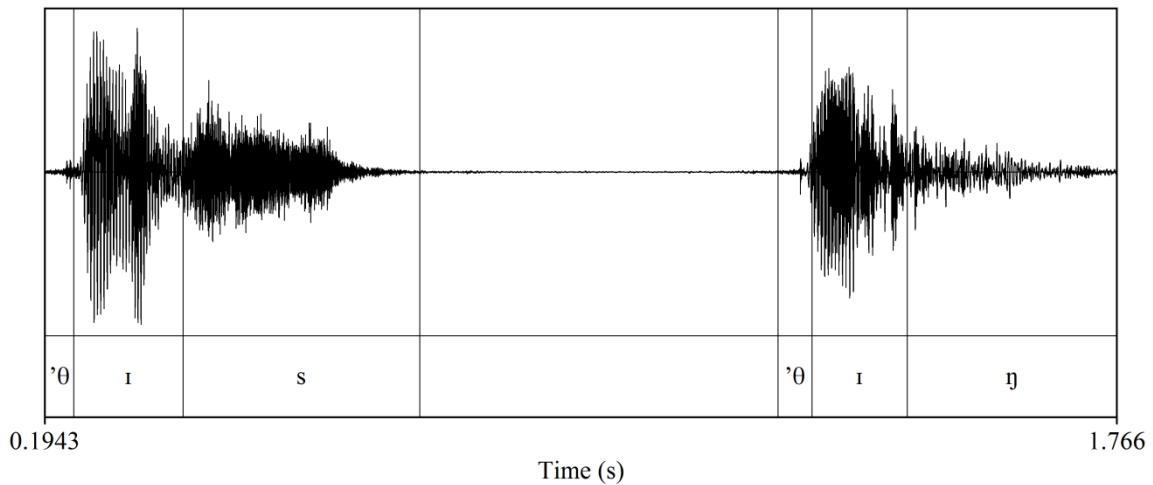


Figure 13. *this/thing*. Participant 1, part 2 of the test, second recording.

9.3 Group 3

Group 3 of Suomi's (1980) categorization contains examples in which the “learners obviously lack a target language rule”, i.e. they use a rule that a native speaker would not use.

The sample of the noun *record* from participant 1 (part 1 of the test, second recording, Image 13 in Appendix 2) is the first example from the third group because of the wrong stressing of the word. The noun *record* should have stress on the first syllable, [ˈrɛkə(r)d], but instead in the sample the stress is on the second, /rɪˈkʰɔrd/, which is more or less how the verb, (to) *record*, would be pronounced. This is clear also in the intensity curve (Figure 14), which rises on the second syllable. In addition, there is clear aspiration on the voiceless /k/ sound, visible in the waveform, and the sound combination /ɔr/ is significantly longer in duration than the vowel /ɪ/ in the first syllable. Therefore we can state that the sound is a good example of the participant lacking the target language rule of distinguishing the verb and noun with the same orthographical appearance from each other by stress. We can also assume that the faulty rule used by the participant stems from Finnish since in Finnish the primary stress is fixed on the first syllable. However, the impact of transfer is not completely unambiguous because she stresses the second syllable, which is practically very rarely the case in Finnish.

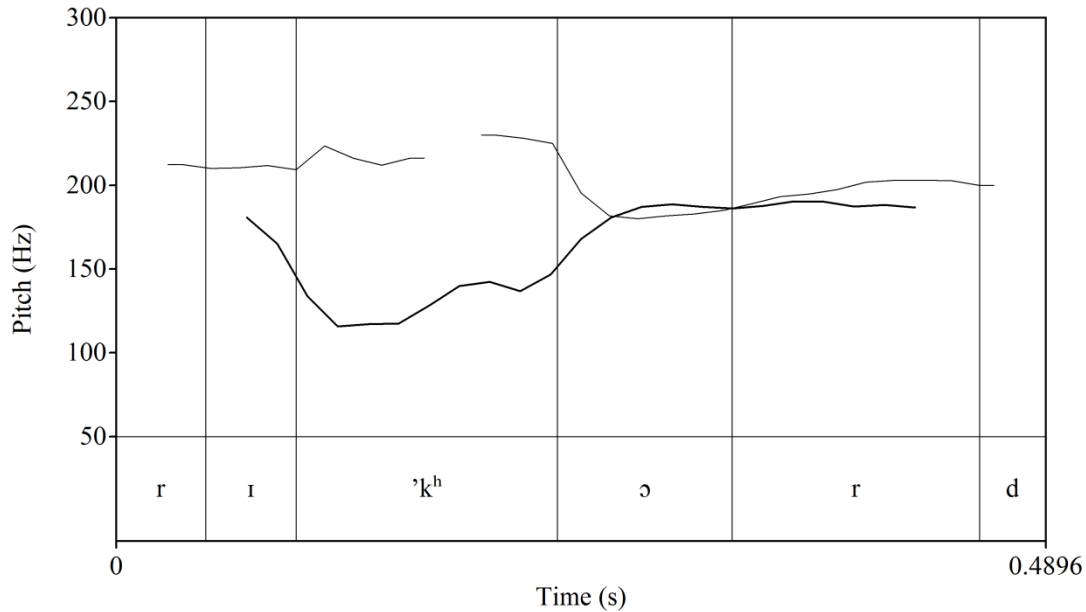


Figure 14. *record*. Participant 1, part 1 of the test, second recording. Intensity: thick line, pitch: thin, dotted line.

The word *ridge* from participant 2's sample (part 1 of the test, first recording) provides another good example of the third group. The waveform and the curve image can be found in Appendix 2 (Image 14). The issue of interest here is the affricate /dʒ/ that realizes in the sample as /ts/. The /d/ sound voiceless, but what makes this example fitting in the third group is the post-alveolar /ʒ/ sound differing from the target language rule by its place of articulation. The voiced post-alveolar /ʒ/ is substituted with the voiceless alveolar /s/, i.e. the most common, and in many dialects the only, sibilant found in the Finnish language. Therefore it is safe to assume in this case that the /ts/ substitution of /dʒ/ stems directly from the Finnish sound system, and that the participant lacks the target language rule that would be used in this sample by native speakers.

Participant 2's pronunciation of the phrase *this time* from the conversational section of the test serves as another example from the third group. The waveform of the phrase can be seen in Figure 15, and the pitch and intensity curves are included in Appendix 2 (Image 15). According to the target language rules the first sound of the phrase is the voiced dental fricative [ð]. In this sample the participant substitutes the sound with the dentalveolar /t/, which is most likely due to the influence of the Finnish language. The Finnish phonetic system lacks both the dental fricatives, which results in many Finns replacing them with sounds that are easier for them to pronounce, mainly /t/. It was mentioned in section 7.3 that our mother tongue determines which sounds are easier and which more difficult for us. The ease of articulation results simply from the fact that the dentalveolar /t/ is the closest

equivalent to the dental fricatives considering the place articulation. In the waveform we can see that the release of the first /t/ sound that substitutes the [ð] is even slightly stronger than the second /t/ in 'time'. The stronger stressing of the first /t/ in the phrase is a result from the participant stressing the word *this*; "Not **this** time--". In short, this kind of substitution of the voiced dental fricative [ð] with the voiceless dentalalveolar plosive /t/ is an example of the speaker lacking the target language rule. Although the example here does not follow the target language rules, it would probably not affect intelligibility. As it was mentioned in section 3.4 about the Lingua Franca Core, the dental fricatives could be omitted and substituted with another sounds.

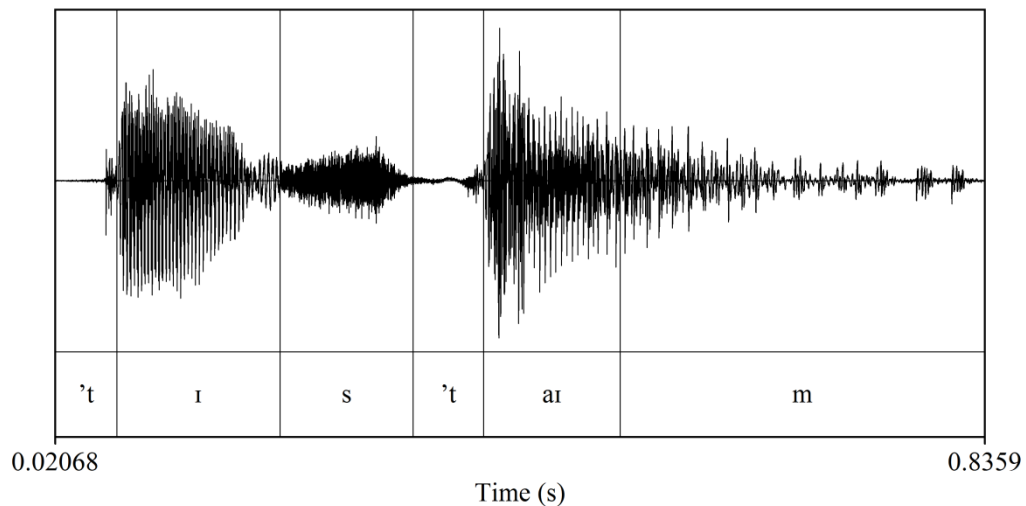


Figure 15. *this time*. Participant 2, conversational section of the test.

Another example of lacking the target language rule is in the verb (*to*) *project* from participant 1's first recording of part 1 of the test (Figure 16, Appendix 2: image 16). According to the target language rules the verb *project* should have stress on the second syllable, i.e. [prə'dʒɛkt]. However, the participant stresses the first syllable, as in the noun *a project*, and pronounces the word as /'prəʊdʒɛkt/. This is clearly visible on the intensity curve, which is significantly higher on the first syllable. This is most likely due to the interference of the Finnish phonetic system, which requires the primary stress to always fall on the first syllable of the word, as it was mentioned in section 6.3. Furthermore, the participant pronounces the first vowel sound of the word, [ə], as a diphthong /əʊ/. This probably relates to stressing the wrong syllable, or this kind of "overpronunciation" can also stem from the fact that the grapheme 'o' is often pronounced as /oʊ/ or /əʊ/ (e.g. *hold*, *cope* and *rose*), which cannot be considered transfer from the Finnish system. In addition, the voiced post-

alveolar fricative [ʒ] realizes as its voiceless counterpart /ʃ/, which serves as an example of group 2 of Suomi's categorization.

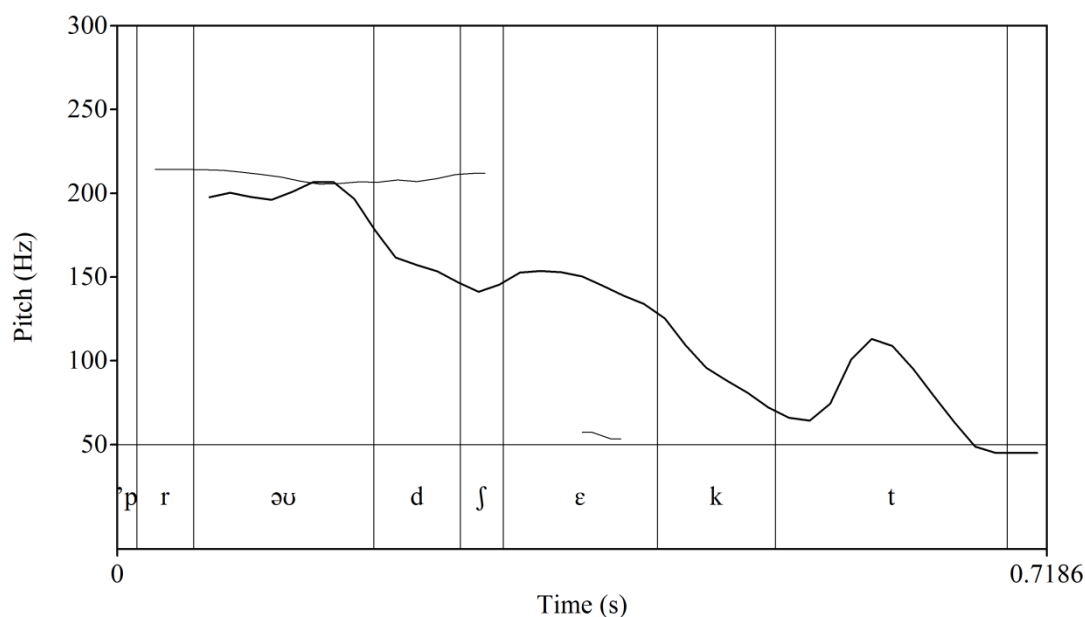


Figure 16. *project*. Participant 1, part 1 of the test, first recording. Intensity: thick line, pitch: thin, dotted line.

The question *Is that car red?* in section 3 of the test is a simple yes/no question which, as mentioned in section 5.2.2, requires a rising intonation. However, in this sample from participant 2 (first recording of section 3, Appendix 2: Image 17) the intonation is steadily falling and thus very similar to the normal Finnish intonation, which was covered in section 6.3.2. As such the sample fits in the third group of Suomi's categorization. However, Jenkins (2000), as presented in section 3.4, does not consider this kind of mistake in stressing dangerous to intelligibility. Therefore categorizing this sample in the third group could be questioned. But if we take into account Ladefoged's (2005: 124) presentation of the rules this sample can be included in the third group. The falling intonation can be seen in the pitch curve in Figure 17 although there is some disturbance in the curve, it is still generally descending. The problem with this sample is that the question was read out loud from the paper, which might result in a less natural intonation. However, participant 1 pronounced the same question with a rising intonation on the second try, and both participants had a correct intonation in another question, *Is the red car John's?*, in the same section on both recordings. The stress is not marked in the transcription because the participant stresses all of the words with practically equal strength, which is visible in the intensity curve and also audible in the sample. Moreover, the voiced fricative /ð/ realizes as the voiceless plosive /t/ similarly as in

participant 2's sample of the phrase *this time* analyzed above, which is another example that belongs to the third group. In addition to this and the lack of rising intonation there are some other issues with e.g. voicing and aspiration, which should be noted briefly although they might fit better in other groups; The /s/ sound in *is* is completely voiceless, the word *car* has no aspiration in the voiceless velar plosive /k/, and the /ɛ/ in *red* is very short even though it should be lengthened before a voiced sound. These all would fit well in the second group of Suomi's classification. Furthermore, the final /d/ sound in the sentence is voiceless and aspirated. The /d/ can be distinguished as a /d/ instead of a /t/ by listening carefully since they differ slightly by the place of articulation, even though it is clearly voiceless. A similar aspirated /d/ in the word *and* will be discussed further in section 9.4 concentrating on the fourth group.

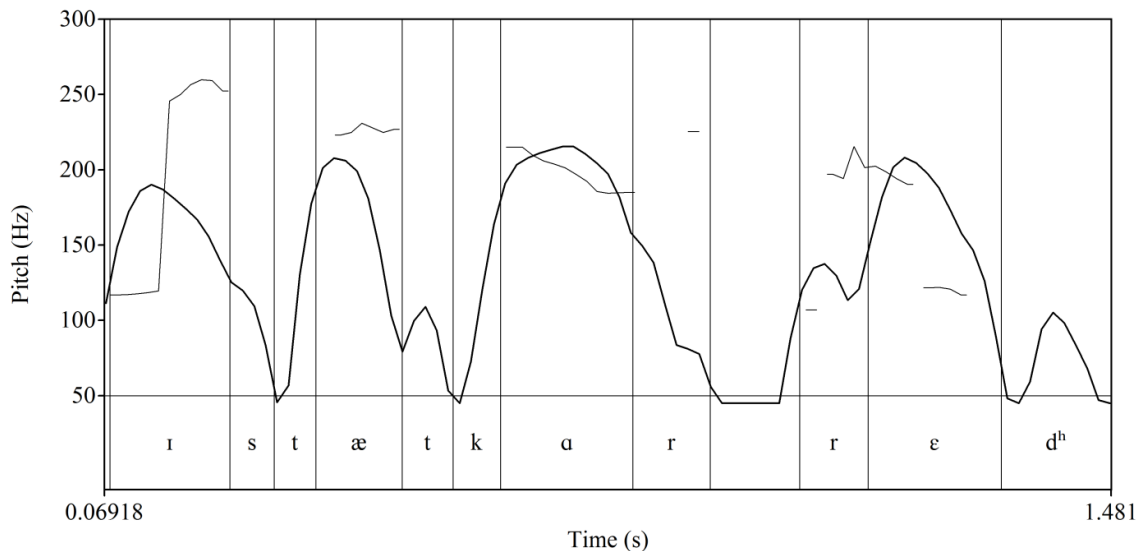


Figure 17. *Is that car red?*. Participant 2, part 3 of the test, first recording. Intensity: thick line, pitch: thin, dotted line.

The last example in the third group is the phrase *drag race* from the first recording of part 3 of the test from participant 1 (Image 18 in Appendix 2). The reason why this sample fits in the third group is that the participant does not apply the target language rule of lengthening the vowel that precedes a voiced consonant, in this case the vowel /æ/ before the voiced velar plosive /g/. We can see from the waveform (Figure 18) that the /æ/ sound (117 milliseconds) is only about half the length of the diphthong /eɪ/ (200 milliseconds), while it should be at least the same length or even longer since the diphthong is followed by the voiceless alveolar sibilant /s/. Therefore we can say that in this sample the participant lacks

the target language rule. What must also be noted here is that the /g/ sound in the sample is only moderately voiced (the pitch curve is affected by the slight echo in the recording situation). By listening to the sample it is possible to hear the voicing disappear for a very brief moment before and during the release of the plosive /g/ while the echo of the preceding vowel is vaguely audible in the background. Yet the sound is voiced almost throughout its duration and the release burst of the plosive is rather weak, and thus it can be described as moderately voiced transcribed with /g/ instead of /k/. It is possible that there is a connection between the short vowel preceding the plosive and the moderate voicing.

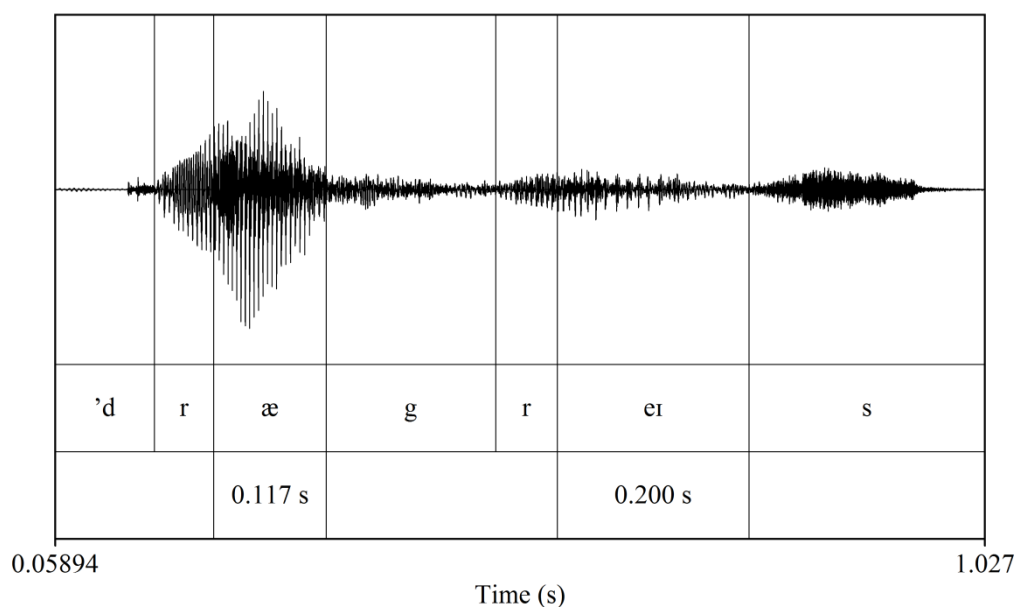


Figure 18. *drag race*. Participant 1, part 3 of the test, first recording.

9.4 Group 4

The fourth and final group in Suomi's (1980) classification was described as learners using "a rule that is definitely not used by native speakers" and that the rules used seem to stem neither from the target nor the learners' first language.

The word *desire* (participant 2, part 2 of the test, first recording) is a good example of group 4 since the pronunciation here stems from participant 2's learning style of pronunciation and not from her native language, nor the target language, and a native speaker would definitely use such pronunciation. In the interview she described that as an auditive learner she learns pronunciation and English words in general by listening and that because of her dyslexia she often has problems connecting written words with their pronunciation. In other words she knows how a word is pronounced and used, but cannot always recognize

some words that she sees as the ones that she already knows and can pronounce. In this example she pronounces the word *desire*, [dɪ'zʌɪə(r)], as /'desire:/, more or less in a way the name *Désirée* would be pronounced. The waveform can be seen in Figure 19 and in Appendix 2: Image 19. After the test she commented on her pronunciation that the name was the first association she had when she read the word on the paper, and that she pronounced the word according to that association on both recordings. It must be noted that all of the words, including *desire* was written capitalized. After the test she heard the correct model of pronunciation, recognized the word and pronounced it correctly.

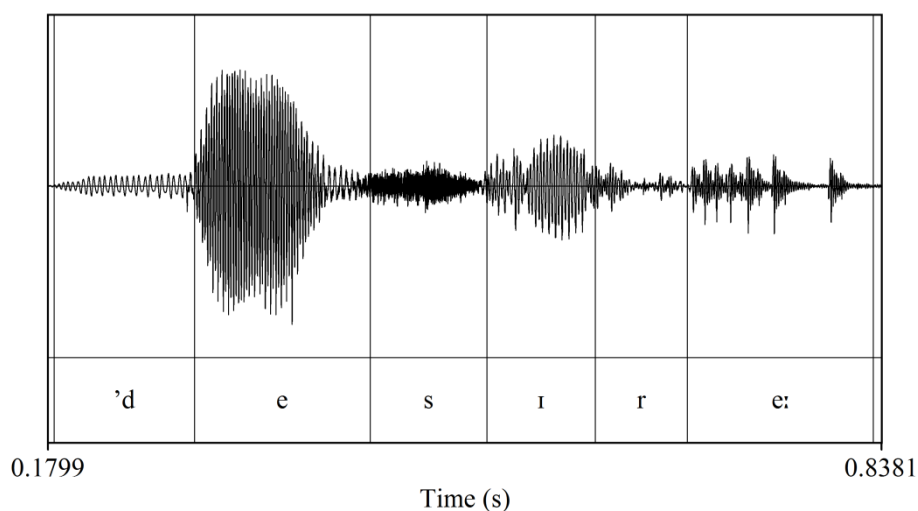


Figure 19. *desire*. Participant 2, part 2 of the test, first recording.

The pseudoword *judeful* from participant 2's first recording of part 4 of the test is another example of how her learning style of pronunciation affects her oral skills. As mentioned above, she is an auditory learner regarding her language learning, and also with this instance she seems to connect the written word to a pronunciation she already knows and relies strongly on her first association of a word. As a result the pronunciation of the word that would be, according to target language rules, something like [dʒu:dfʊl] realizes as /tʃatsfʊl/. This implies that the participant aimed for the word *judge* with the suffix *-ful*, although the voiced consonants [dʒ] and [d] are voiceless, /tʃ/ and /ts/. This is also observable in the pitch curve (the thin dotted line, Figure 20). The waveform is included in Appendix 2 (Image 20).

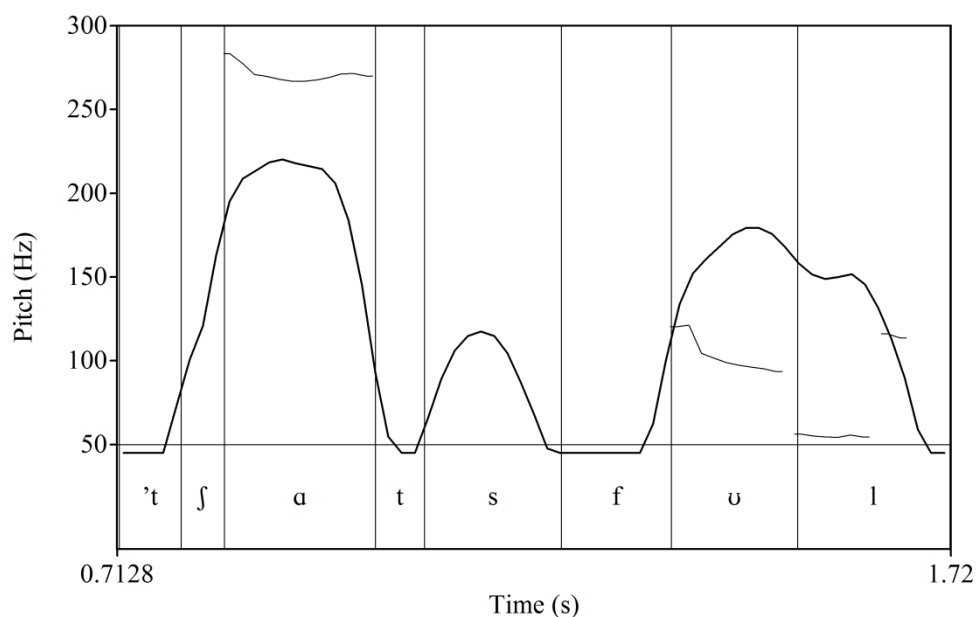


Figure 20. *judeful*. Participant 2, part 4 of the test, first recording. Intensity: thick line, pitch: thin, dotted line.

The kind of aspirated voiceless alveolar plosive /d/ that is presented next in the fourth group was also briefly introduced earlier in section 9.3 in the phrase *Is that car red?*. This /d/ in question here is the final sound in the word *and* from participant 2's conversational part of the test (Appendix 2: Image 21). The aspiration is visible in the waveform (Figure 23); the noise of the aspiration continues for a moment after the release of the /d/. The /d/ is completely voiceless, but as it was mentioned earlier in section 9.3 it can be differentiated from /t/ since the place of articulation is slightly different, /t/ being almost dentalveolar, and /d/ being almost postalveolar. This kind of aspirated /d/ is very interesting since it does not seem to stem from either the speaker's native language Finnish or the target language English. The voicelessness could be considered negative transfer from Finnish, but the Finnish sound system does not explain the aspiration since, as it was discussed in section 6.2, it does not normally occur in Finnish at all. The rather lengthy /æ/ is due to the fact that the sample is from free conversation and the participant was thinking and stretched the word to some extent.

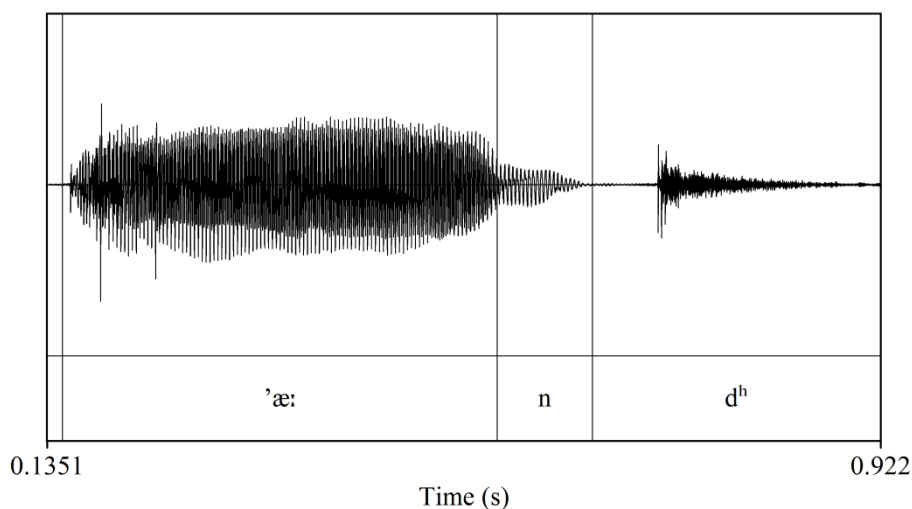


Figure 21. *and*. Participant 2, conversational section of the test.

A similar aspirated /d/ occurs also in participant 1's speech. This example is the word *mad* from her second recording of part 4 (Appendix 2: Image 22). The same word was also mentioned earlier in group 1 as a good example of lengthening the vowel before a voiced consonant. Here we focus on the aspirated /d/, which is visible in the waveform (Figure 7). The burst of the plosive occurs immediately after the vowel /æ/, and it continues for a rather long period of time until it is no longer visible in the waveform image. In fact, since the aspiration occurs after an alveolar sound it sounds almost like a sibilant, a very frontal /ʃ/, to be exact. This will be discussed further in section 10. Nonetheless, this is a good example of using a rule that is not used by native speakers.

The next example of the fourth group is the word *vile* from the first recording of part 1 of the test from participant 2 (Appendix 2: Image 23). According to the target language rules the word begins with a voiced labiodental fricative [v]. In the sample the word is pronounced as /waɪl/, i.e. as the word *while* – the voiced /w/ is very clearly visible in the waveform (Figure 22). Substituting the [v] with the voiced labial-velar approximant /w/ is a very interesting phenomenon, since it does not occur in the Finnish sound system and is rather different from [v] by its manner and place of articulation. As mentioned in section 6.2 regarding Finnish phonetics, the Finnish system lacks the voiced labiodental fricative [v] and instead has the central approximant [ʋ], which is very close to [v] but still slightly towards [w]. If we try to examine this from the point of view of language transfer this could be one reason for confusion between the sounds, but it still does not quite explain why the speaker moves from the Finnish /v/ towards the /w/ instead of the /ʋ/ which would be closer to the Finnish sound.

The opposite substitution of [w] with /v/ would be more likely considering the Finnish system. However, this substitution is a good example of a sound that would not be used by a native speaker in such context.

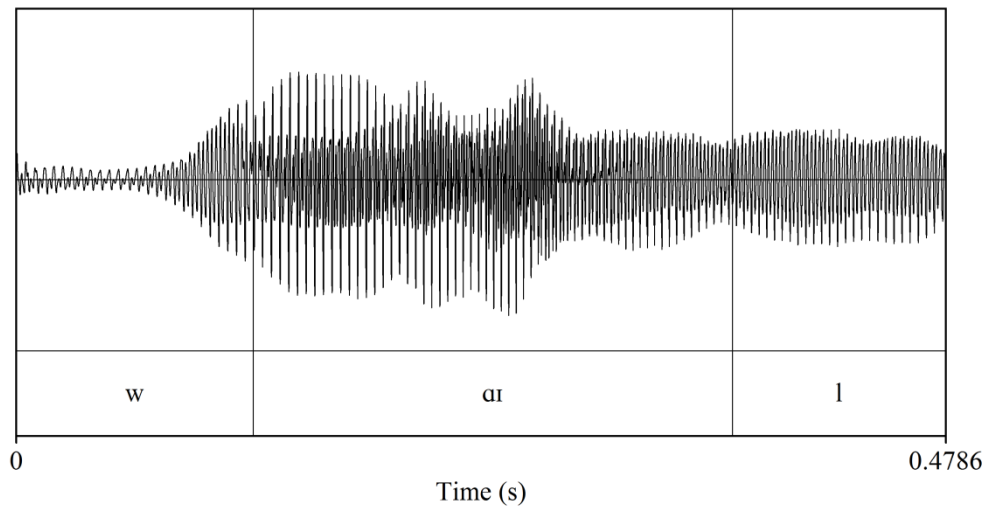


Figure 22. *vile*. Participant 2, part 1 of the test, first recording.

The last sample in the fourth group is Participant 1 pronouncing the phrase *Say vantering loudly* (part 4 of the test, second recording, Appendix 2: Image 24). *Vantering* is a pseudoword, a progressive verb form, to be precise. It could be pronounced as, for example, [vɔ̃ntərɪŋ] (based on the pronunciation of *wandering*) or [væntərɪŋ] (based on *vantage*). The correct pronunciation of the first vowel in the word is not relevant, however. In this sample we will concentrate on the initial [v] sound in *vantering*, and also on an additional nasal /n/ and an aspirated /d/ in *loudly*.

The [v] in *vantering* is a voiced labiodental fricative, but in this sample participant 1 replaces it with the labial-velar approximant /w/. A similar substitution occurred also with participant 2 in the word *vile* and it was analyzed earlier in this section with some possible reasons given for it as well. It is interesting how a sound is replaced with another sound that does not have very much in common with it regarding both their place and manner of pronunciation. Although presented as separate graphemes in both Finnish and English alphabet *v* and *w* are orthographically similar and have been used rather interchangeably in Finnish, which might result to this kind of “overpronunciation” of the [v] sound.

Another brief example of a sound that would not be used by a native speaker is the additional alveolar nasal sound /n/ that appears in the middle of the word *loudly* - /laʊnd^hli/, also visible in the waveform (Figure 23). This is most likely only a slip of the tongue since

there is no reason in the Finnish phonetic system that would explain it. On the other hand, the additional /n/ occurs right before the alveolar stop /d/, which is the sound that is formed if the nasal tract is blocked while pronouncing /n/. It seems that the participant keeps nasal tract open during the closure of the /d/ and does not close it until right before the release of the stop. This way the air flows through the nasal tract, which results in the alveolar nasal /n/ sound.

The third example in this sample from participant 1 that fits in the fourth group is a voiceless aspirated /d/, similar to the one presented earlier in this section in the word *and* from participant 2. The difference here is that the aspirated /d/ occurs word-medially, and there is a clear voiceless moment of aspiration, visible also in the waveform, before the following sound, the alveolar lateral approximant /l/. It is interesting that the following sound is also voiced, as are all the sounds in the word *loudly*, but still the voicing cuts away word-medially. Phonetically it would be easier to pronounce the whole word with voicing instead of stopping the vibrating of the vocal folds and then beginning it again. Furthermore, the word *loudly* is repeated after each word in the fourth part of the test and the participant has the same voiceless moment in most instances of them on both recordings, and a number of times she adds the aspiration there as well, or at least has a rather strong release of the plosive. This seems to be a speaker-specific phenomenon, which does not have any effect on understandability. Nevertheless, it would not be used by a native speaker, and neither does it seem to stem from either Finnish or English phonetics.

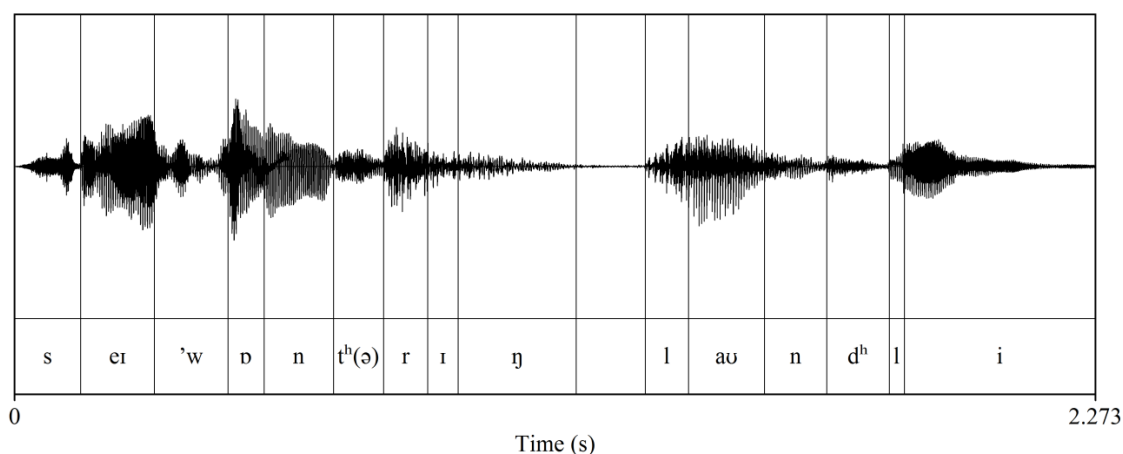


Figure 23. *Say vantering loudly*. Participant 1, part 4 of the test, second recording.

9.5 The participants' own views about the test

Neither one of the participants had any major difficulties with the test. Participant 1 stated that in general the test was easy but, however, there were some words that were more difficult than others. She mentioned that several years have passed since she has received explicit instruction in English, and she has therefore forgotten many of the rules of pronunciation. She was sure that she had learned many rules in school, but probably makes mistakes just because she does not remember them anymore. It is interesting how she still after spending a year in England and learning pronunciation implicitly focuses on forgetting rules when she reflects on her mistakes in pronunciation.

Usually most words or phrases come to Participant 1 automatically, but at some places during the test she noticed that she stopped to think of the pronunciation of a word, especially when the words were represented separately in the test and not in longer sentences. She stated that it is easier for her to pronounce words in the middle of sentences, i.e. in natural context, because that way she does not pay too much attention to the pronunciation; “You don’t stop to think about it too much, you just let it come out. Once you’ve learned it in the past it comes out right.” She mentioned also that when she read the words on paper she focused more on her pronunciation, compared to speaking freely, when she pays more attention to structures and vocabulary instead.

As an example of separate individual words she pointed out the word *attempts*, which was analyzed in section 9.2: “I’ve never thought of it as a difficult word, but suddenly when it’s there it feels somehow really tricky”. She thought that the difficulties were due to the consonant cluster at the end of the word.

In addition, she mentioned noticing some difficulties with distinguishing voiced and voiceless consonants, both plosives and fricatives: “I can’t hear the difference if I say *peas* or *peace*”, she said.

According to her, there were also several easier words, such as *closed* and *mad*, which she thought are probably easier because they are familiar to her, or because they lack the sound combinations that she feels are more difficult to pronounce.

Participant 2 did not feel that the test was very difficult, but she felt it slightly unpleasant. However, she stated that it helped her to know that she was not supposed to give correct pronunciation but instead she could pronounce the words like she would normally do.

Because of her dyslexia she normally feels reading more difficult than informal conversation, and this was the case in the test as well. Usually speaking becomes more fluent

when she gets herself “to a specific mode”, which did not quite happen in the conversational section of the test. Yet she still did feel it was easier than reading out loud.

She mentioned that recognizing the word certainly helps pronunciation, meaning that words that she knows beforehand are easier to pronounce. The problem is, however, that when reading she does not always recognize the written form of the word, which was the case in the word *desire*, for example. In other words, it is easier for her to recognize words auditively than visually, and it takes a moment until she can make the connection between a word’s written and spoken realizations. This is probably due to her dyslexia. As an example of this she mentioned the word *dangerous* in part 1 of the test; when she recognizes the word she can pronounce it correctly. Furthermore, language learning is easier for her by listening and doing instead of reading. She also identifies the transfer of Finnish in her speech; according to her there are some words that she cannot pronounce even though she theoretically is familiar with the correct rule of pronunciation.

10. DISCUSSION/CONCLUSION

The research question in this thesis, as presented in section 8.1, was in two parts:

1. How does the pronunciation of the participants differ from the target language rules?
2. What kind of an effect does language transfer have in those differences?

It can be seen from the questions that the purpose was to present how pronunciation can differ from the target language rules in the speech of the two participants based on the four groups and to examine the different ways in which the phonetic system of Finnish language can affect the English pronunciation of Finnish speakers. Choosing to begin the questions with the word ‘how’ and the phrase ‘what kind of’ aimed the research towards finding and describing examples, and the questions can be considered successful. The examples presented in this study include various types of phonetic issues and, furthermore, they also give some insight on how the participants’ differing backgrounds as speakers of English as a foreign language affect their pronunciation skills. Thus it can also be assumed in general that the purpose of this thesis was fulfilled and the questions were answered thoroughly.

The results found and analyzed in the previous section include various examples from all four groups of Suomi’s (1980) classification. The groups were the following:

- (1) learners use a correct target language rule,
 - (2) learners use a rule that is faulty in terms of the target language but bears an obvious resemblance to and is readily identifiable as an attempt at the correct target language rule,
 - (3) learners obviously lack a target language rule, and
 - (4) learners are in possession of a rule that is definitely not used by native speakers.
- (Suomi 1980: 151)

It was easiest to find examples to group 1, i.e. pronunciation that followed the target language rules, which was expected. For groups 2 and 3 there were a number of possible examples on top of the ones that were chosen for analysis. The least examples were found for group 4, but there were also a couple of samples that could have been analyzed in this group as well but were omitted. The hypotheses were met rather well; as it was expected, the most common issues related to voicing, or the lack of it, to be accurate, and aspiration. There were also a few examples regarding prosody, e.g. stressing and intonation, that differed from the target language rules.

The issues included as examples of groups 2 and 3, e.g. the changing phonatory status (i.e. voicing), aspiration, varying stress and certain types of intonation, are all phenomena that are more or less absent from the Finnish sound system, and therefore it is clear that cross-linguistic influence plays an important part in those examples. The examples in group 4 represented individual differences in pronunciation that could not be traced to the phonetic systems of Finnish or English, but instead some of them (e.g. pronouncing *loudly* as /laʊnd^hli/) were found to be related to the physical vocal system of forming different sounds and sound combinations, and some are just slips of tongue. It is interesting, however, that both participants produced similar individual slips of the tongue, which suggests that there are logical reasons for them. There were a couple of those kinds of examples in group 4 that we will take a closer look at here. The sounds cannot be regarded as examples of transfer from Finnish, but they certainly differ from the target language rules, which makes them worthy of slightly deeper examination.

First we must look at the aspirated /d/ that appeared in a couple of examples in the fourth group. The International Phonetic Alphabet does not have a separate symbol for a voiceless **postalveolar** /d/, although it clearly differs from the alveolar /d/; in postalveolar /d/ the tongue touches the alveolar ridge further back in the mouth, whereas in the alveolar /d/ the contact occurs very close to the upper teeth. This results in slightly differing sounds, which is clearly audible, especially if the two sounds are voiceless; a voiceless alveolar /d/ is very close to, and practically indistinguishable from the (denti)alveolar /t/ sound, whereas the postalveolar

/d/ can be recognized as a /d/, even if it is voiceless. The /d/ pronounced by the participants is postalveolar, and thus easily differentiated from /t/. This leads us to the aspirated /d/, i.e. /d^h/, analyzed in section 9.4. The /d^h/ could in many instances be transcribed even as /dʃ/. The aspiration is so strong and the closure occurs rather back in the mouth that the release sound is essentially a sharp alveolar fricative /ʃ/. In other words, the sounds /d/ and /ʃ/ have the same place of articulation. The effect is very strong especially in the word *loudly* that was analyzed last in section 9.4; although the word had an additional alveolar nasal /n/ in the middle, the vowel sound preceding the consonant combination /nd^h/ is /ʊ/, a near-close near-back rounded vowel. The rounding is still more or less present in the release of the plosive, and this leads to the rounded allophone of /ʃ/, which sounds significantly darker and stronger compared to, for example if the /dʃ/ sound was instead of /loodʃ/ in a phonetic environment like /li:dʃ/. This can also be heard between words like *pooch*, /p^hu:tʃ/ and *peach*, /p^hi:tʃ/. The aspirated /d^h/ did not occur at all in word-initial position; in word-final position it could be explained as relaxation of the vocal apparatus after the release, which would clarify why it does not appear word-initially, but that does not quite fit for its word-medial occurrence, e.g. in the word *loudly*. Another interesting issue here is that the aspirated /d^h/ occurred in the speech both of the participants and in several occasions, which suggests that there are phonological reasons for it, as was just examined above. The occurrence in the samples of both of the participants could also result from the fact that when reading from the paper the speaker tends to overemphasize his/her pronunciation, which might lead to stronger bursts in the release of the plosive /d/. What makes /d/ so prone to this kind of aspiration compared to the other plosives is the way it differs from its voiceless counterpart /t/; a voiceless /g/ is essentially a /k/, and a voiceless /b/ is a /p/.

The vowel sounds in the samples must also be addressed here briefly, although they were not focused on in the research. The two participants tended to pronounce the vowels more carefully and purely than a native speaker would probably have done. In English non-stressed vowels tend to be weakened (Jenkins 2000: 146-148), which in general leads to them being pronounced as the mid-central vowel *schwa*, /ə/. Generally both the participants lacked the weakening and “overpronounced” also non-stressed vowels. Furthermore, especially participant 2 tended to mix short and long vowel sounds – not by their quantity, i.e. duration, but their sound quality. For example, she often substituted the near-close near-back rounded vowel /ʊ/ (as in *look*) with the close back rounded vowel /u/ (as in *Luke*), and the near-close near-front unrounded vowel /ɪ/ (as in *sit*) with the close front unrounded vowel /i/ (as in *seat*). This is most likely transfer from Finnish, but it does not affect eligibility.

The two participants were very different regarding their skills and learning history. However, they both provided examples for each group, and comparing their success in the test would be irrelevant. It can be stated that participant 1 had trouble mostly with the voiced and voiceless distinction, both with plosives and fricatives, as she also herself mentioned: “I can’t hear the difference if I say *peas* or *peace*”. Participant 2 provided more general examples of how pronunciation can differ from the target language rules. Furthermore, participant 2 had more examples in group 4, i.e. individual issues that could not be traced to either the target language or the native language. What was common with the two participants was that they both seem to learn the correct pronunciation by hearing the model and not by concentrating on explicit pronunciation rules, which was emphasized with participant 2 (e.g. *desire* in group 4), but was also noticeable with participant 1 (e.g. she lengthened the vowel in *mad* but not in *drag*). In addition, participant 1 mentioned in the interview that although she has probably learned the rules in school she has already forgotten them.

The language test used as a means of collecting data was successful. It covered several phonetic issues and provided a great deal of material to choose for the analysis. Most examples were found in part 1, the longer text. This was probably due to the fact that when reading a longer text the participants could not focus too much on single words and pronounce them too carefully. The downside was that there were instances of phrases that are more difficult to pronounce than phrases that one would encounter in normal language use. These kinds of tongue twisters might have resulted in unnatural pronunciation. Furthermore, participant 1 mentioned in the interview that part 1 was difficult also because it “did not make sense”. In other words, the text was not cohesive which made it difficult for the reader to connect it to the real world and therefore it is also impossible to try to predict what is coming next in the text. Least mistakes in pronunciation were found in part 2, word pairs, because the participants could focus more on the pronunciation of single words. Especially minimal pairs steer the participant to pay more attention to the issue of interest since there is only one phonological difference between the two words. However, there were, for example, instances of the missing voiced/voiceless distinction in the word pairs. Part 3, the dialog, provided some examples, but its relevance could be questioned. A dialog is supposed to be free conversation, and reading it from paper does not result in very natural pronunciation. However, there were words and examples of intonation that functioned well in the analysis. Part 4, words in a frame sentence, proved to be a success in the analysis. The frame sentence provided a useful phonological environment for the words, and there were some minimal pairs but they were scattered in different places in the list, and thus the participants did not

focus too much on them. Additionally, there were also pseudowords, i.e. nonsense words that sound English but were in fact created for this purpose only. They provided interesting insight on whether the participants had learned to connect a certain pronunciation to certain words, or whether they had internalized the rules and could apply them also to words that are unfamiliar to them. There were instances of both cases, the most interesting of the former probably being *vantering*, which the participants tended to connect to either ‘*wondering*’ or ‘*wandering*’. This is an example of learning the pronunciation of whole words instead of single sounds. The last part, i.e. the free conversation was successful as well; it gave a sample of the participants’ pronunciation in a natural form and provided good examples in several of the four groups, some of which were included in the analysis. This part could have been replaced with doing the interview sections in English, but separating them from the actual data collection and letting the speakers answer the questions in their native language Finnish was beneficial since they could most likely reflect on their learning and their views more naturally than they would have in a foreign language.

There are always some shortcomings in research, and I would like to address some issues here. The recording situation with participant 1 could have been better; there was a faint humming noise from the air conditioning system observable in the recordings, and the room echoed slightly. The location was chosen simply because of practical logistic reasons. However, the humming was removed from the recordings by the noise reduction and it did not affect the analysis. The echo, however, caused some interference in the pitch curve, but it was easy to distinguish actual voicing from the echo by listening. In addition, the participants were very similar, apart from the differences in their skills and learning history. A wider sample of participants from, for example, different social backgrounds or age groups would have provided an interesting addition to the analysis. However, in that case the number of participants should be larger since trying to pinpoint differences in pronunciation based on the participants’ age or social situation would be irrelevant if there were only two people involved. The samples could have also been compared to a sample given by a native speaker. Yet, as discussed in chapter 1, defining a native speaker or their accent is difficult, and also native speakers’ pronunciation probably differs from the so-called standard. Hence the samples were compared to a model based on the general rules of English pronunciation, and the issues chosen for analysis were not dependent on the specific accent used but instead were applicable at least to GA and RP. In addition, the receptive skills of the participants could have been examined. This would have further helped to make the distinction of whether the

participants had learned the rule of pronunciation, i.e. if they could distinguish different sounds by listening, or whether they had learned the pronunciation for the certain words.

An important thing that must also be taken into account is that analyzing phonetic content and language transfer is essentially based on subjective opinions and decisions, and even though a researcher might have good knowledge on the speaker's L1 everything has to be based on good guesses. This, however, does not diminish the relevance of the examples found in the study; although they can always be discussed or questioned they still provide important insight on the issues, especially since the aim was not to make generalizable conclusions but instead to provide descriptive examples.

A quantitative study on this subject with a larger number of participants and a wider array of phonetic issues to focus on would be one way to continue this field of research. In addition, a longitudinal study that would also consider the development of a learner's pronunciation skills would be very interesting to conduct. Furthermore, participant 1 mentioned in the interview that spending a year in Norway and speaking English only with other exchange students had a negative influence on her pronunciation skills. Therefore the effect of faulty input is also one topic that could be investigated.

The present thesis aimed at finding reasons for pronunciation that differs from the target language rules. It is good to acknowledge said reasons and identify the effect of the native language on foreign language pronunciation. This can also benefit teaching and learning of pronunciation. Nevertheless, perfect pronunciation that follows the rules by the book is not a goal that should, or even could be accomplished. An analysis on the subject often disregards the real-life situations in which the pronunciation and communication occur; in everyday conversation between speakers with different L1's there is always the possibility for clarifications and negotiation of meaning, let alone using gestures and body language, i.e. all the other tools that complement pronunciation. All these together influence the intelligibility and understanding. Every speaker is different and communication situations vary from one another. It is good to know the basics also on pronunciation (i.e. Jenkins' LFC), but it is not a bad thing if a speaker's cultural background can be heard in their speech. After all, it is an important part of our identity.

BIBLIOGRAPHY

- Celce-Murcia, M., Brinton, D. M. and Goodwin, J. M. (2010). *Teaching pronunciation: A course book and reference guide (2nd ed.)*. New York: Cambridge University Press.
- Crystal, D. (2003). *English as a global language (2nd ed.)*. Cambridge: Cambridge University Press.
- Deterding, D. (2005). How many consonant sounds are there in English?. *STETS Language and Communication Review* 4 (1), 23–25.
- Doughty, C. J., Long, M. H. (Eds.) (2008). *The handbook of second language acquisition*. Malden: Blackwell.
- Dziubalska-Kolaczyk, K. and Przedlacka, J. (Eds.) (2005). *English pronunciation models: A changing scene*. Bern : New York: Peter Lang.
- Edwards, J. G. H. and Zampini, M. L. (2008). *Phonology and second language acquisition*. Amsterdam: John Benjamins.
- Elisa (2016). *Hintasaarnaajat Miamissa* [online video clip]. YouTube, 8 January, 2016. <https://www.youtube.com/watch?v=vyQ4jg8esJ8>. (21 May, 2016).
- Gut, U. (2009). *Introduction to English phonetics and phonology*. Frankfurt: Peter Lang.
- International Phonetic Association 2015: Full IPA Chart. <https://www.internationalphoneticassociation.org/content/full-ipa-chart>. (22 April, 2016).
- Jarvis, S. and Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. New York: Routledge.
- Jenkins, J. (2000). *The phonology of English as an international language: New models, new norms, new goals*. Oxford: Oxford University Press.
- Jensen, J. T. (1993). *English phonology*. Amsterdam ; Philadelphia: John Benjamins.
- Kivistö, A. (2013). *Bet big or pet pig?: The influence of the Finnish phonetics on Finnish speakers' pronunciation of English*. BA Thesis. Jyväskylä: University of Jyväskylä. JYX. <https://jyx.jyu.fi/dspace/handle/123456789/41350>
- Ladefoged, P. (2005). *A course in phonetics. 5th ed.* Boston, MA: Thomson Wadsworth.
- Major, R. C. (2008). Transfer in second language phonology – a review. In Edwards, J. G. H. and Zampini, M. L. (eds.), *Phonology and second language acquisition*. Amsterdam: John Benjamins, 63-94.

- Mauranen, A., & Ranta, E. (2009). *English as a lingua franca: Studies and findings*. Newcastle upon Tyne: Cambridge Scholars.
- McKay, S. L. (2002). *Teaching English as an international language: Rethinking goals and approaches*. Oxford: Oxford University Press.
- MTV 2011: TOP 10: "Rallienglannin" taitajat.
<http://www.mtv.fi/sport/ralli/uutinen/artikkeli/top-10-rallienglannin-taitajat/4160028>. (29 April, 2016).
- MTV 2009: Lehti: Lenita Airisto haukkuu Vanhasen kielitaidon
<http://www.mtv.fi/viihde/muut/artikkeli/lehti-lenita-airisto-haukkuu-vanhasen-kielitaidon/2979672>. (29 April, 2016).
- National Center for Voice and Speech 2016: Factors Influencing Fundamental Frequency.
<http://www.ncvs.org/ncvs/tutorials/voiceprod/tutorial/influence.html>. (22 March, 2016).
- National Physical Laboratory 2016: Acoustics. <http://www.npl.co.uk/educate-explore/factsheets/acoustics/>. (22 March, 2016).
- Odisho, E. Y. (2003). *Techniques of teaching pronunciation in ESL, bilingual & foreign language classes*. München: Lincom Europa.
- Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning*. Cambridge: Cambridge University Press.
- Odlin, T. (2008). Cross-linguistic influence. In Doughty, C. J., Long, M. H. (Eds.). *The handbook of second language acquisition*. Malden: Blackwell, 436-486.
- Perusopetuksen opetussuunnitelman perusteet* (2014). Helsinki: Opetushallitus.
- Paananen-Porkka, M. (2007). *Speech rhythm in an interlanguage perspective : Finnish adolescents speaking English*. Helsinki: University of Helsinki.
- Preston, D. R. (2005). How can you learn a language that isn't there? In Dziubalska-Kolaczyk, K. and Przedlacka, J. (eds.), *English pronunciation models: A changing scene*. Bern : New York: Peter Lang, 37-58.
- Ranta, E. (2010). English in the real world vs. English at school: Finnish English teachers' and students' views. *International Journal of Applied Linguistics* [online] 20 (2), 156-177. <http://onlinelibrary.wiley.com.ezproxy.jyu.fi/doi/10.1111/j.1473-4192.2009.00235.x/epdf>
- Routarinne S. (2003). *Tytöt äänessä: parenteesit ja nouseva sävelkulku kertojien vauorovaikutuskeinoina*. Helsinki: Suomalaisen Kirjallisuuden Seura.

- Seidlhofer, B. (2001). Closing a conceptual gap: The case for a description of English as a Lingua Franca. *International Journal of Applied Linguistics* 11 (2), 133-158.
- Seidlhofer, B. (2005). Language variation and change: The case of English as a Lingua Franca. In Dziubalska-Kolaczyk, K. and Przedlacka, J. (eds.), *English pronunciation models: A changing scene*. Bern : New York: Peter Lang, 59-76.
- Seidlhofer, B. (2011). *Understanding English as a Lingua Franca*. Oxford: Oxford University press.
- Sharifian, F. (2009). *English as an international language: Perspectives and pedagogical issues*. Bristol, UK ; Buffalo, NY: Multilingual Matters.
- Suomi, K. (1980). *Voicing in English and Finnish stops: A typological comparison with an interlanguage study of the two languages in contact*. Turku: Turun yliopisto.
- Suomi, K., Toivanen, J. and Ylitalo, R. (2006). *Fonetiikan ja suomen äänneopin perusteet*. Helsinki: Gaudeamus.
- Suomi, K., Toivanen, J. and Ylitalo, R. (2008). *Finnish Sound Structure: Phonetics, phonology, phonotactics and prosody*. Oulu: Oulu University Press.
- Tergujeff, E. (2013). *English pronunciation teaching in Finland*. Jyväskylä: University of Jyväskylä. https://jyx.jyu.fi/dspace/bitstream/handle/123456789/41900/978-951-39-5322-5_vaitos03082013.pdf?sequence=1.
- YLE 2013: Rallienglanti tarttuu oppikirjoista.
http://yle.fi/uutiset/rallienglanti_tarttuu_oppikirjoista/6758253. (29 April, 2016).

APPENDIX

Appendix 1. The pronunciation test

Part 1 – Text

It would be shocking to see you exposed in a booth to the rich that way. A vile record similar to the one I made just topped up his hash. When has he cut that hedge thing? It must have been a difficult project. I suppose you're just confused this time of the year as usually. Did they take that mashed coat to the tailor like she asked them? Hopefully they remember the hats and the mat, too. That was very kind of them. Have you ever fought in a war? I don't believe peace would make a film that astonishing.

I don't like thick bees because sometimes they sting me. But it could be worse. They saw a wise man with a boot in a stressful situation. The man, whose name was Owen Shaw, made many attempts to catch the perfect bird. Is it funny? Maybe it's not. He should keep in mind that standing on a thermal ridge facing that direction can be dangerous to your gut. Arching over to the other side of the site is easier. I might want to project this image to the wall, and while a sock in New Zealand is warm the exposure to cold objects can pull heads down.

Part 2 – Word pairs

Came – Game

Try – Dry

Expose – Exposure

Harsh – Arch

Joke – Choke

This – Thing

Few – View

Decide – Desire

Part 3 – A dialog

What color is that car? Is that car red?

- No, it's blue.

Is the red car John's?

- I suppose it is.

What car do you drive?

- I drive a black Porsche. I even broke a track record with it once.

You did? When was that?

- Three years ago in a drag race.

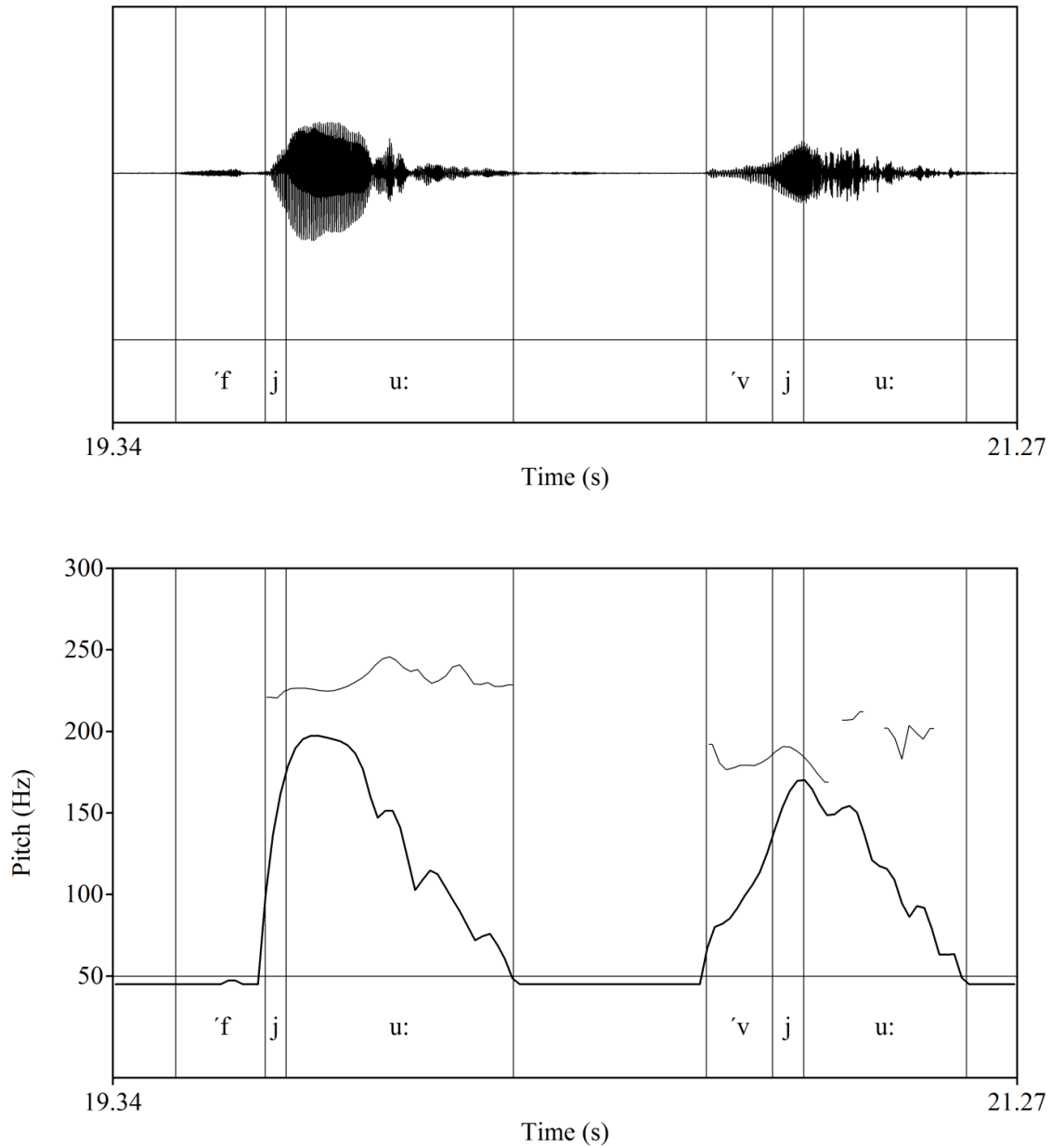
Part 4 – Words in a frame sentence (real and pseudowords)

- Say **wife** loudly
- Say **mad** loudly
- Say **judeful** loudly
- Say **vice** loudly
- Say **heat** loudly
- Say **hetch** loudly
- Say **peas** loudly
- Say **mass** loudly
- Say **comble** loudly
- Say **five** loudly
- Say **breeze** loudly
- Say **dopped** loudly
- Say **wise** loudly
- Say **heed** loudly
- Say **mazz** loudly
- Say **itch** loudly
- Say **vantering** loudly
- Say **closed** loudly
- Say **straffest** loudly
- Say **blows** loudly

Appendix 2. The waveform images.

Thick line: intensity of the sound. Thin, dotted line: pitch of the sound.

Image 1. *few/view*. Participant 1, part 2 of the test, second recording.



There is distortion in the pitch curve; the /u:/ sound in 'view' is completely voiced.

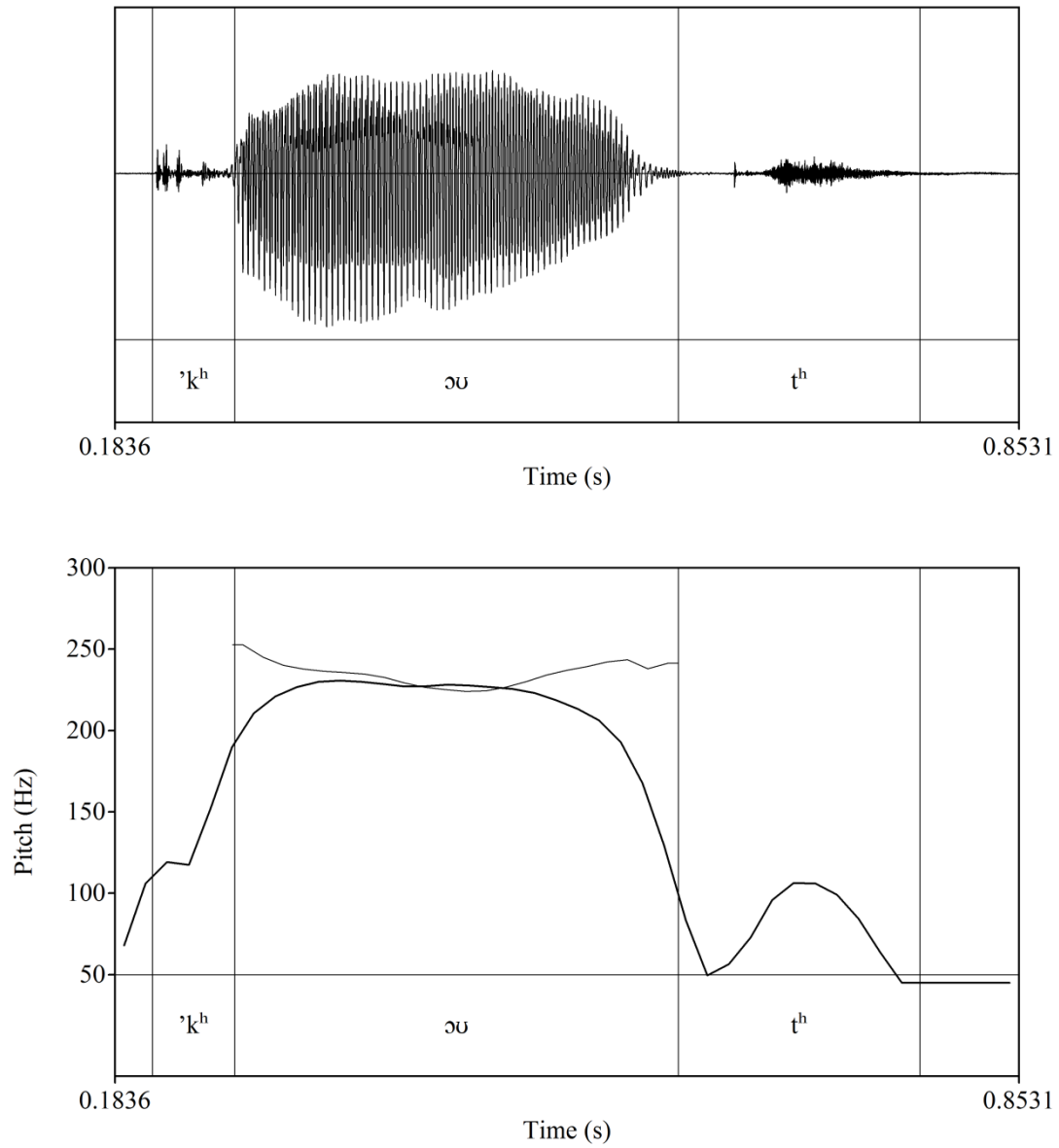
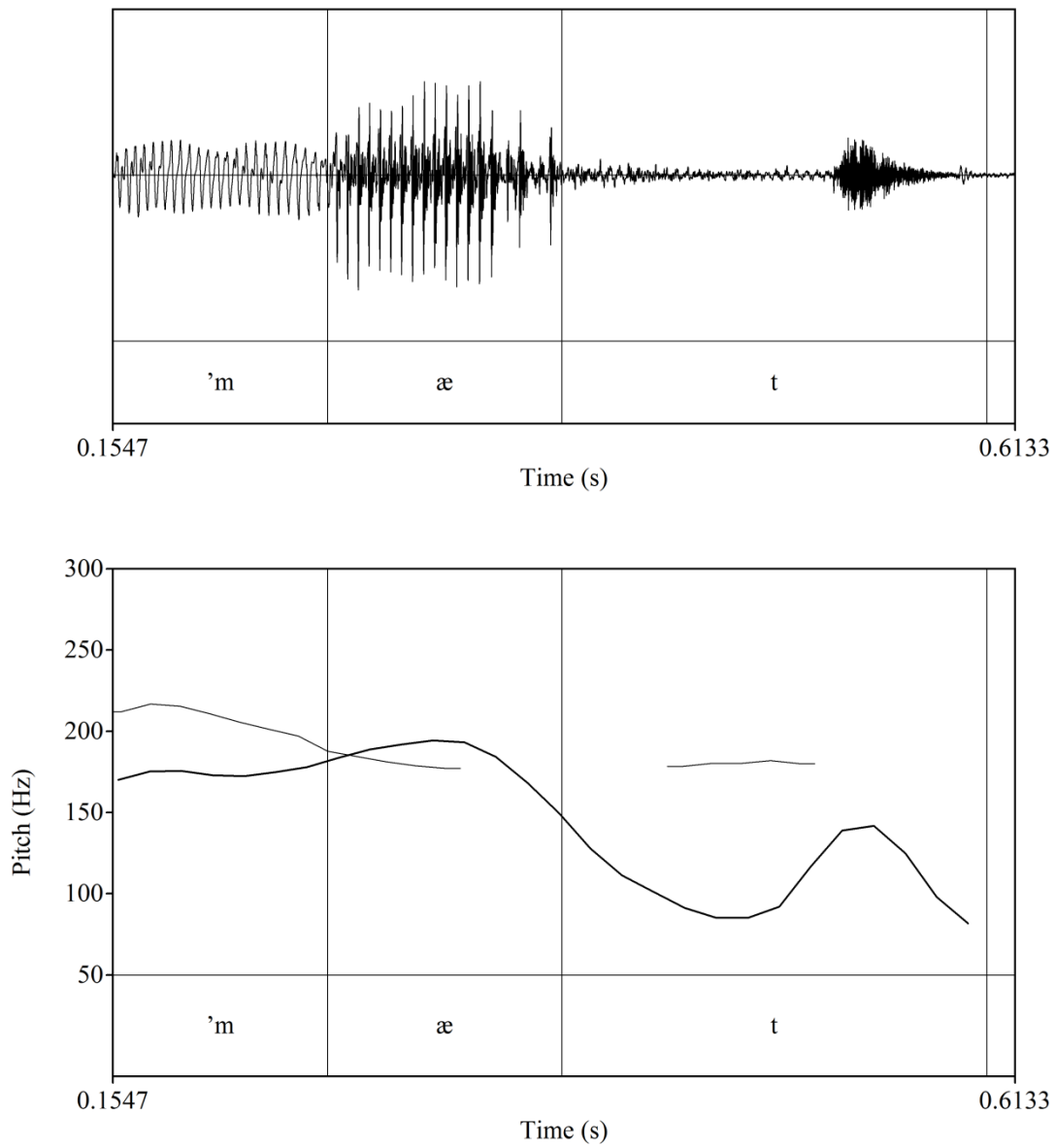
Image 2. *coat*. Participant 2, part 1 of the test, first recording.

Image 3. *mat*. Participant 1, part 1 of the test, second recording.

There is distortion in the pitch curve; the /æ/ sound is completely voiced and the /t/ sound is completely voiceless.

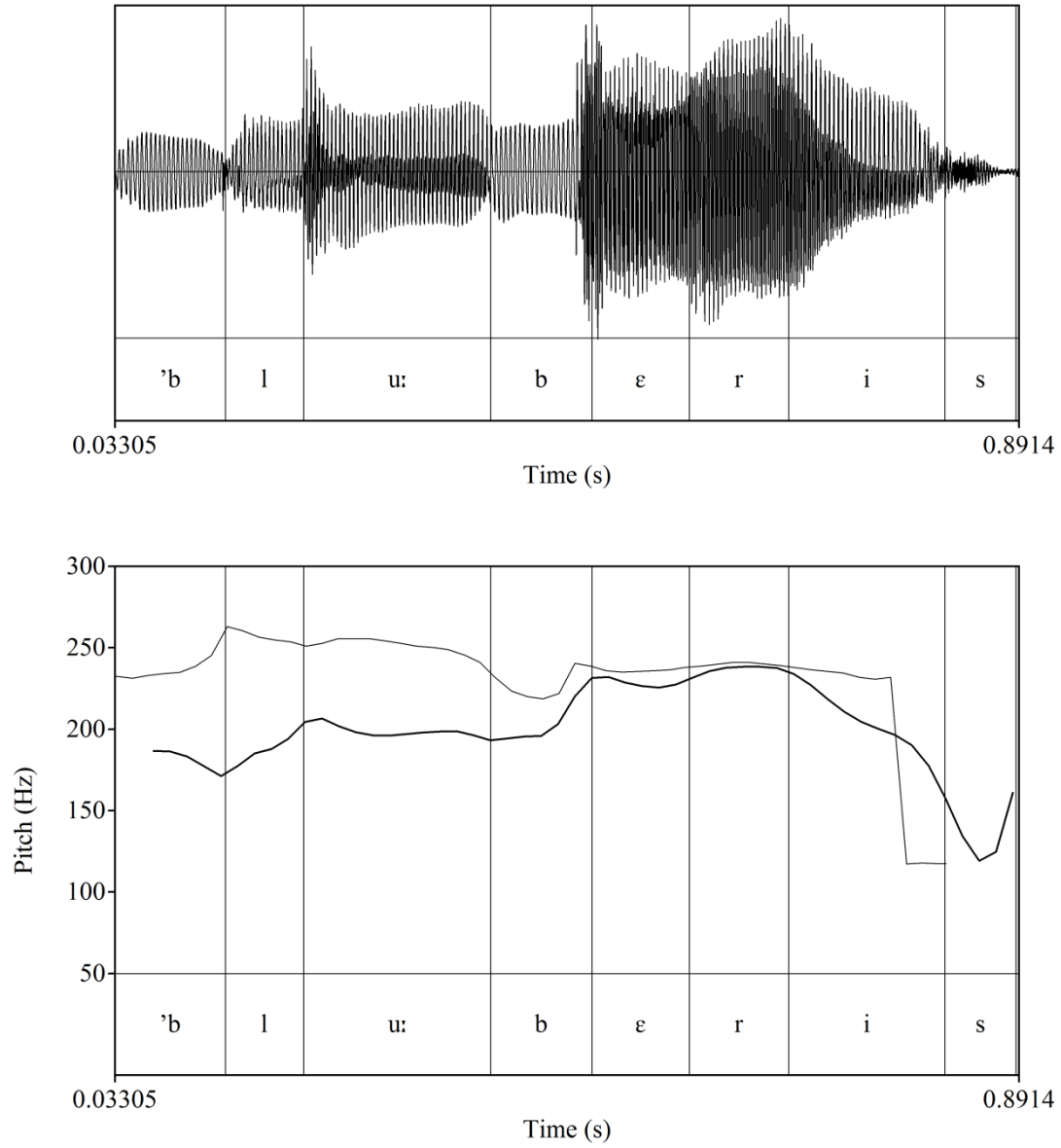
Image 4. *blueberries*. Participant 2, free discussion

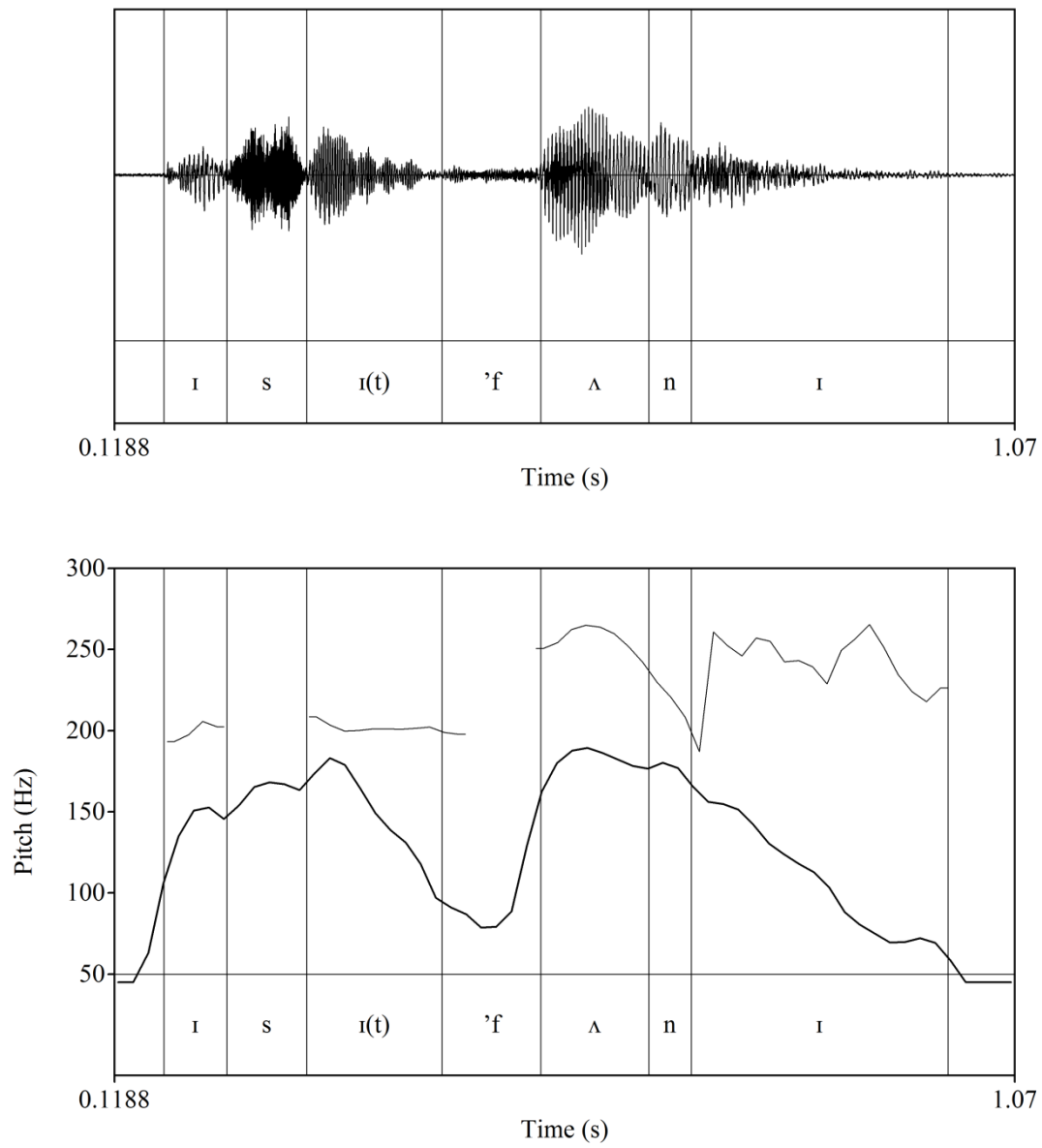
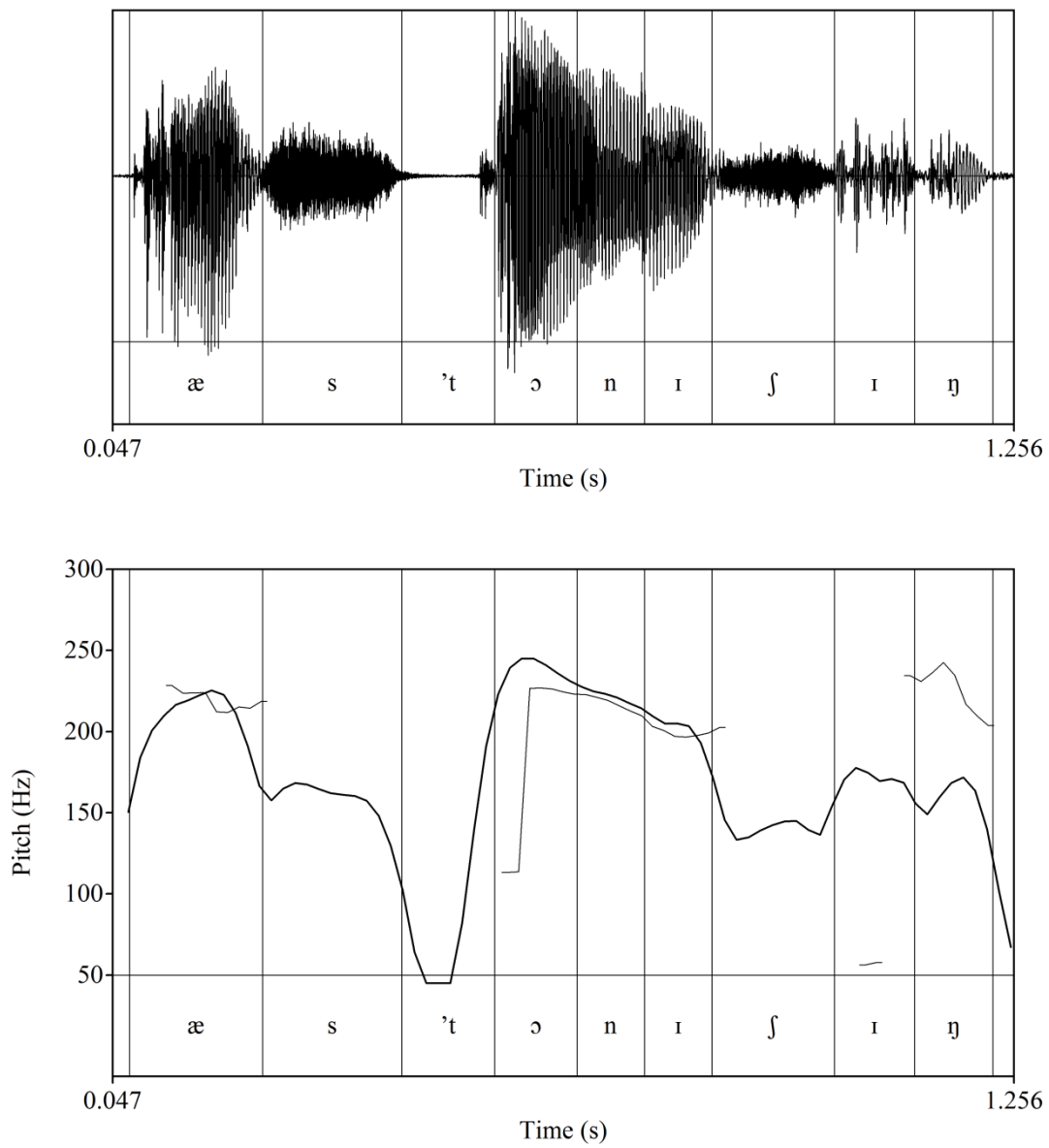
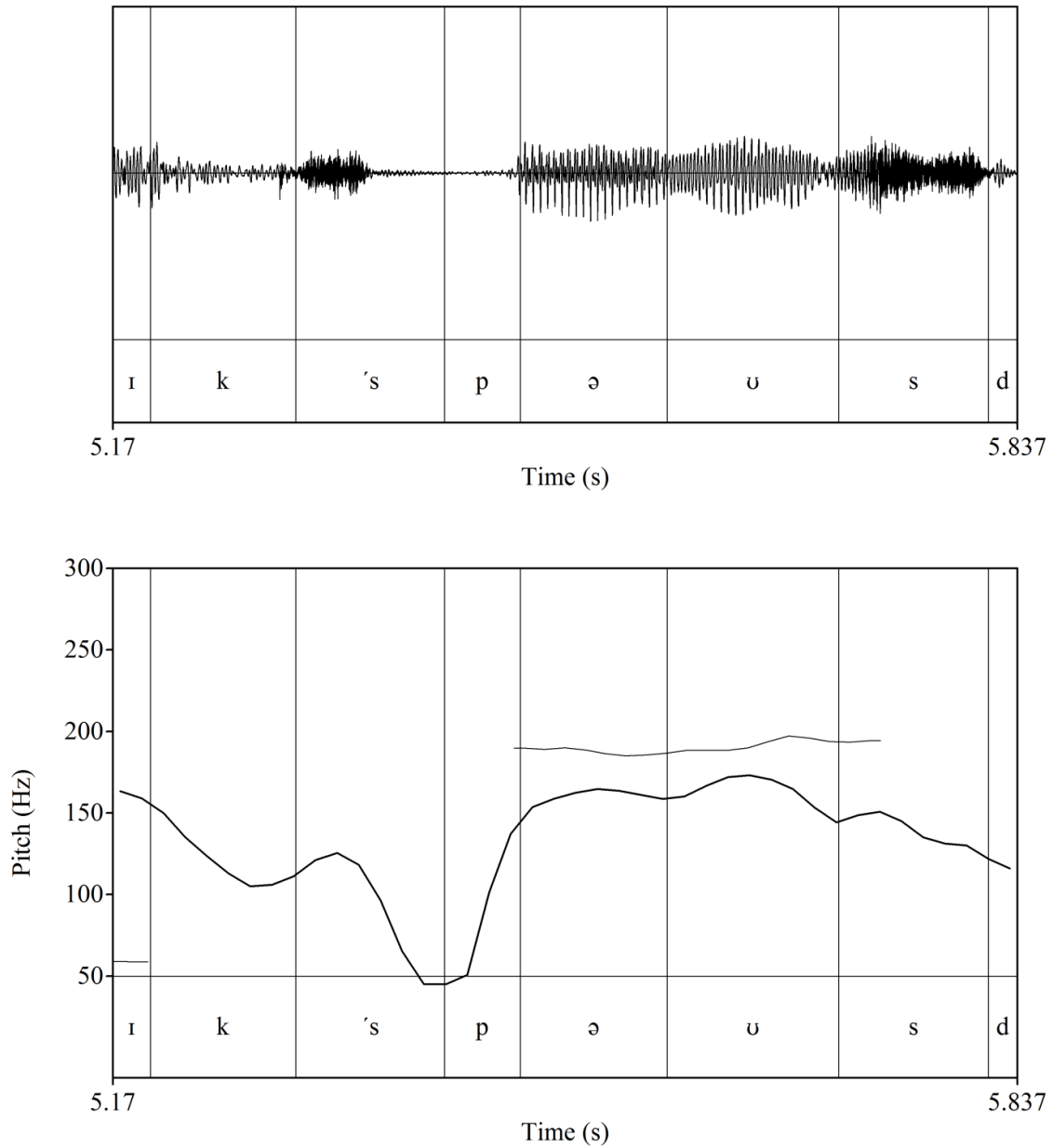
Image 5. *Is it funny?* Participant 1, part 1 of the test, second recording.

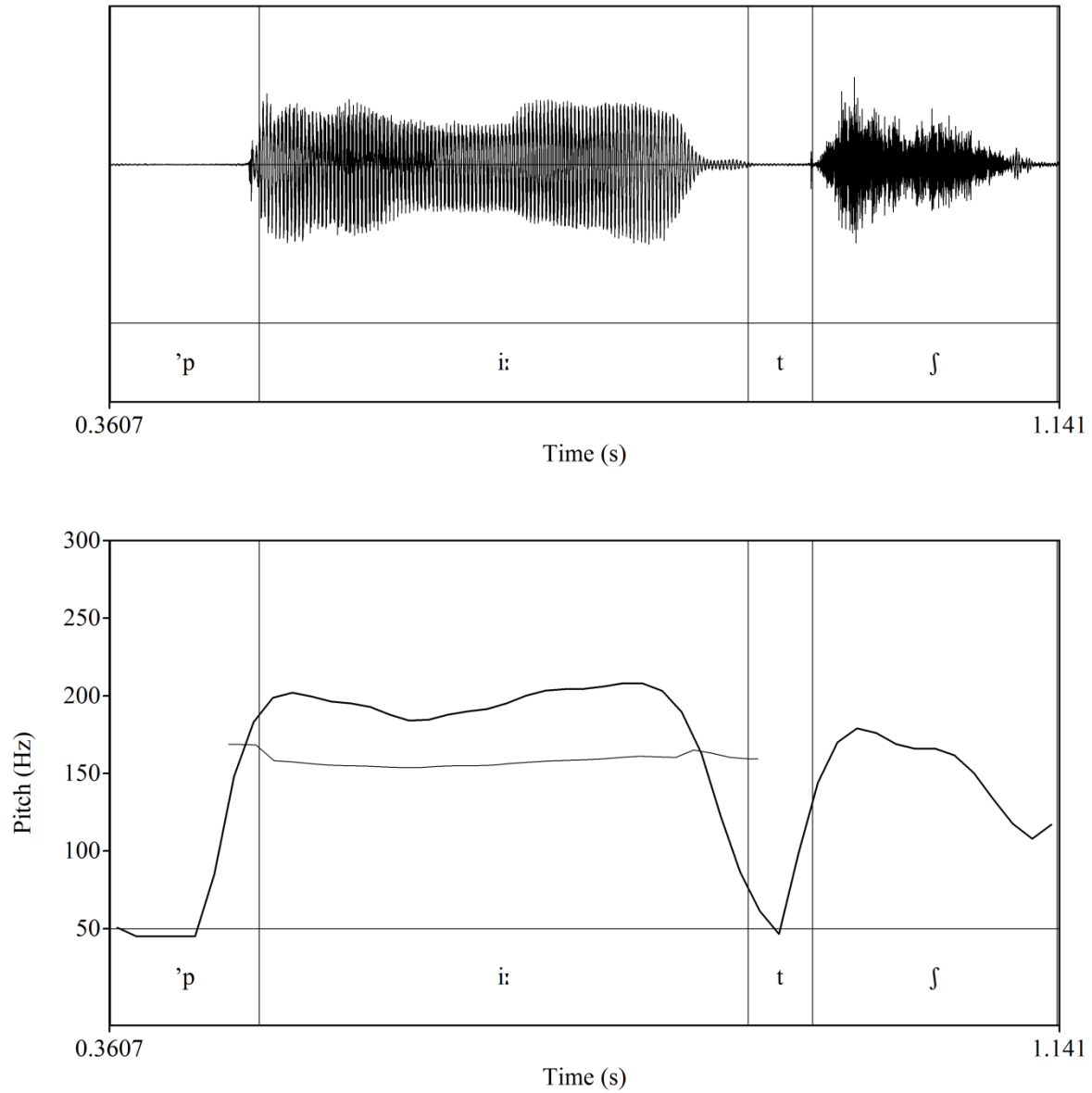
Image 6. *astonishing*. Participant 2, part 1 of the test, first recording.



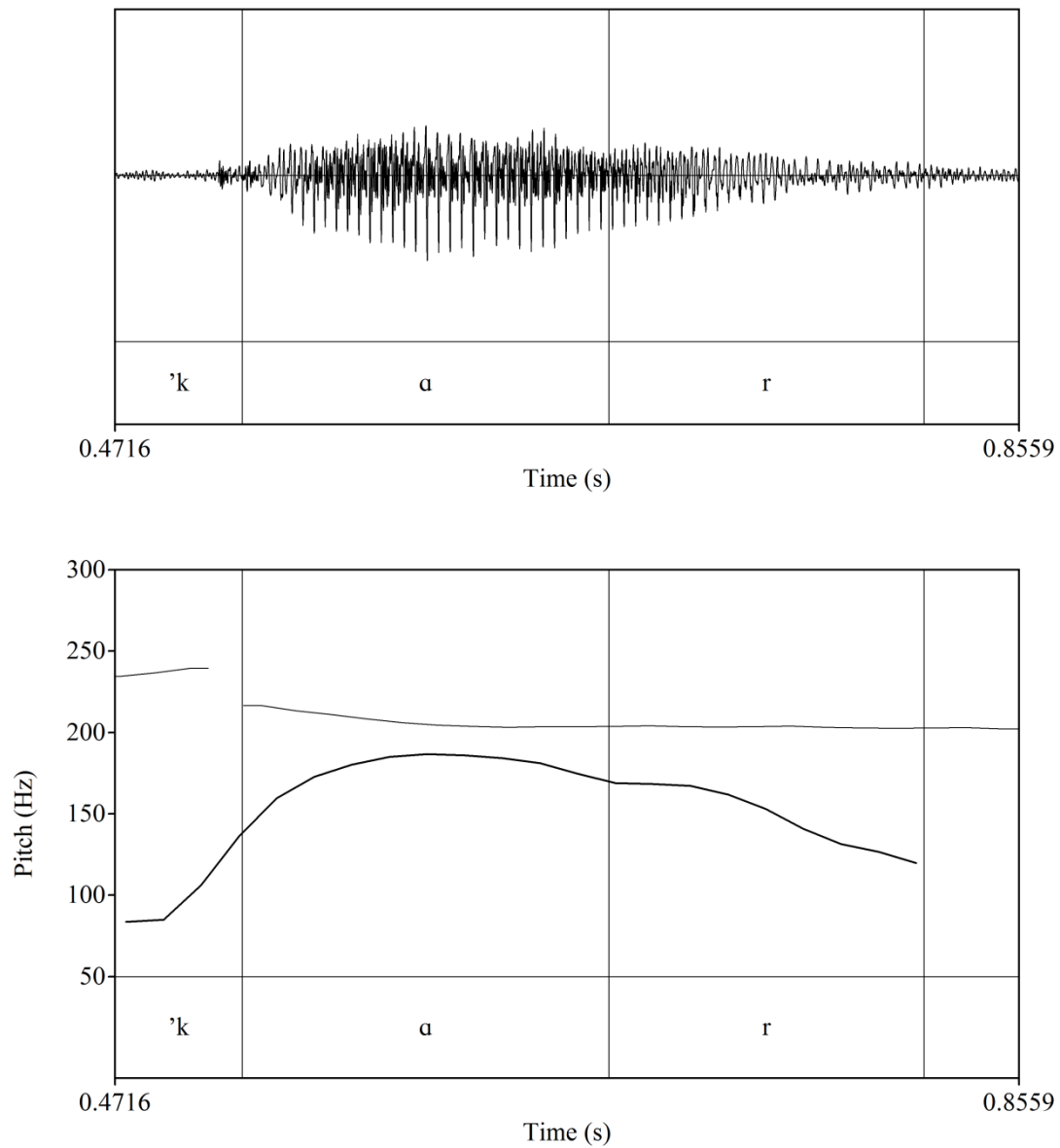
There is distortion in the pitch curve; the curve should travel all the way around 200-250 Hz.

Image 7. *exposed*. Participant 1, part 1 of the test, first recording.

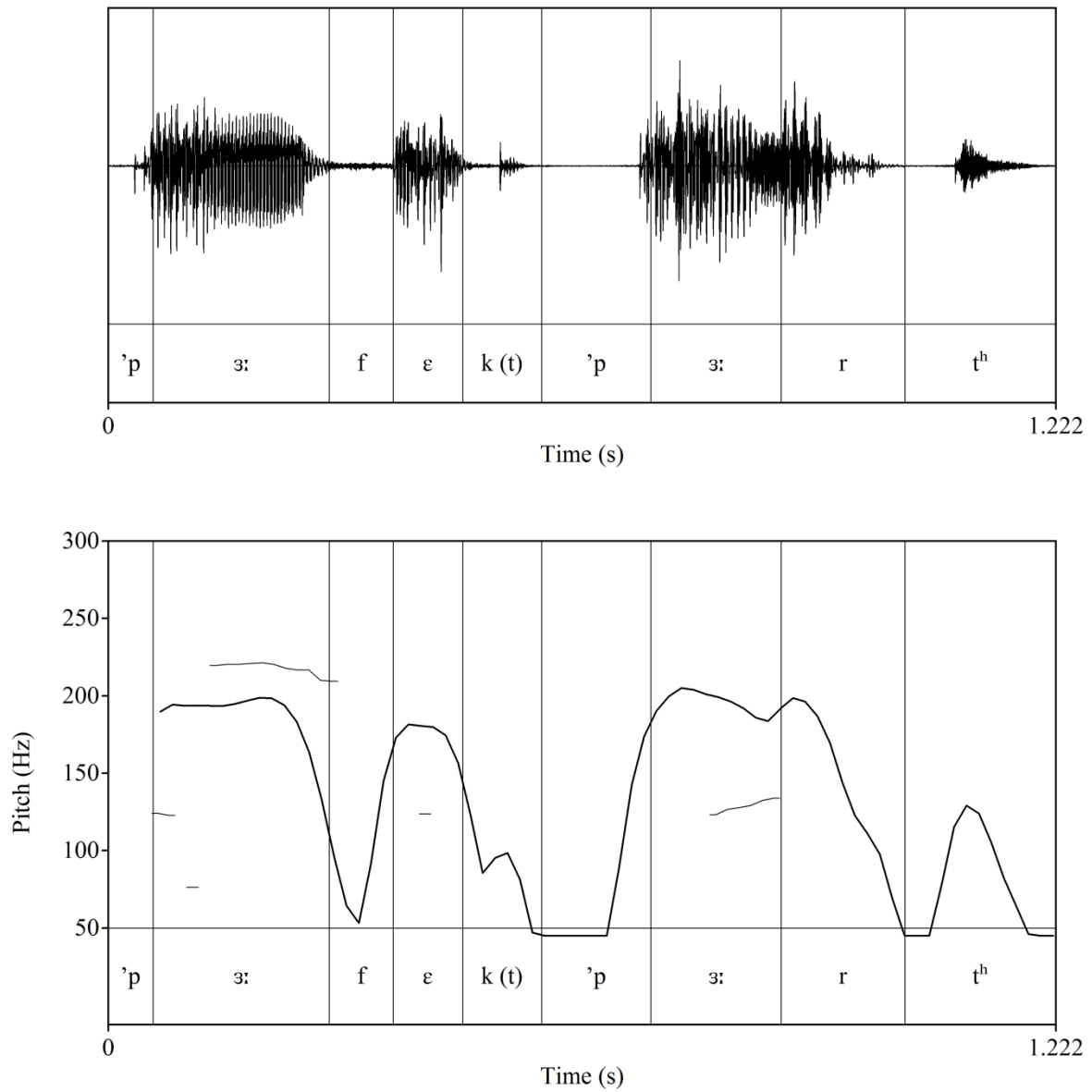
There is distortion in the pitch curve; the voicing in the /s/ is caused by the echo.

Image 8. *beach*. Participant 2, free conversation.

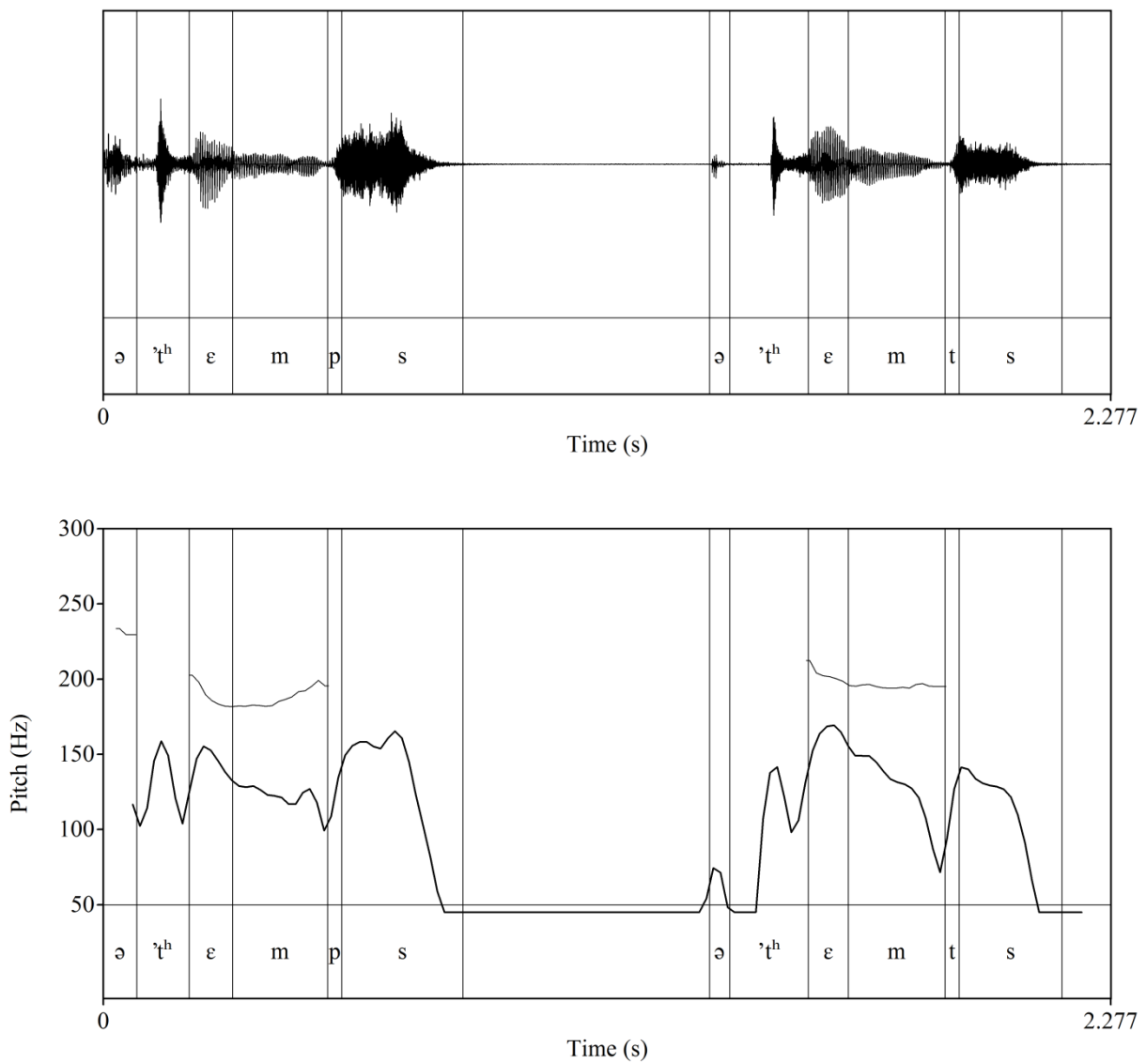
There is distortion in the pitch curve; the voicing does not begin until at the /i:/ sound.

Image 9. *car*. Participant 2, part 3 of the test, first recording.

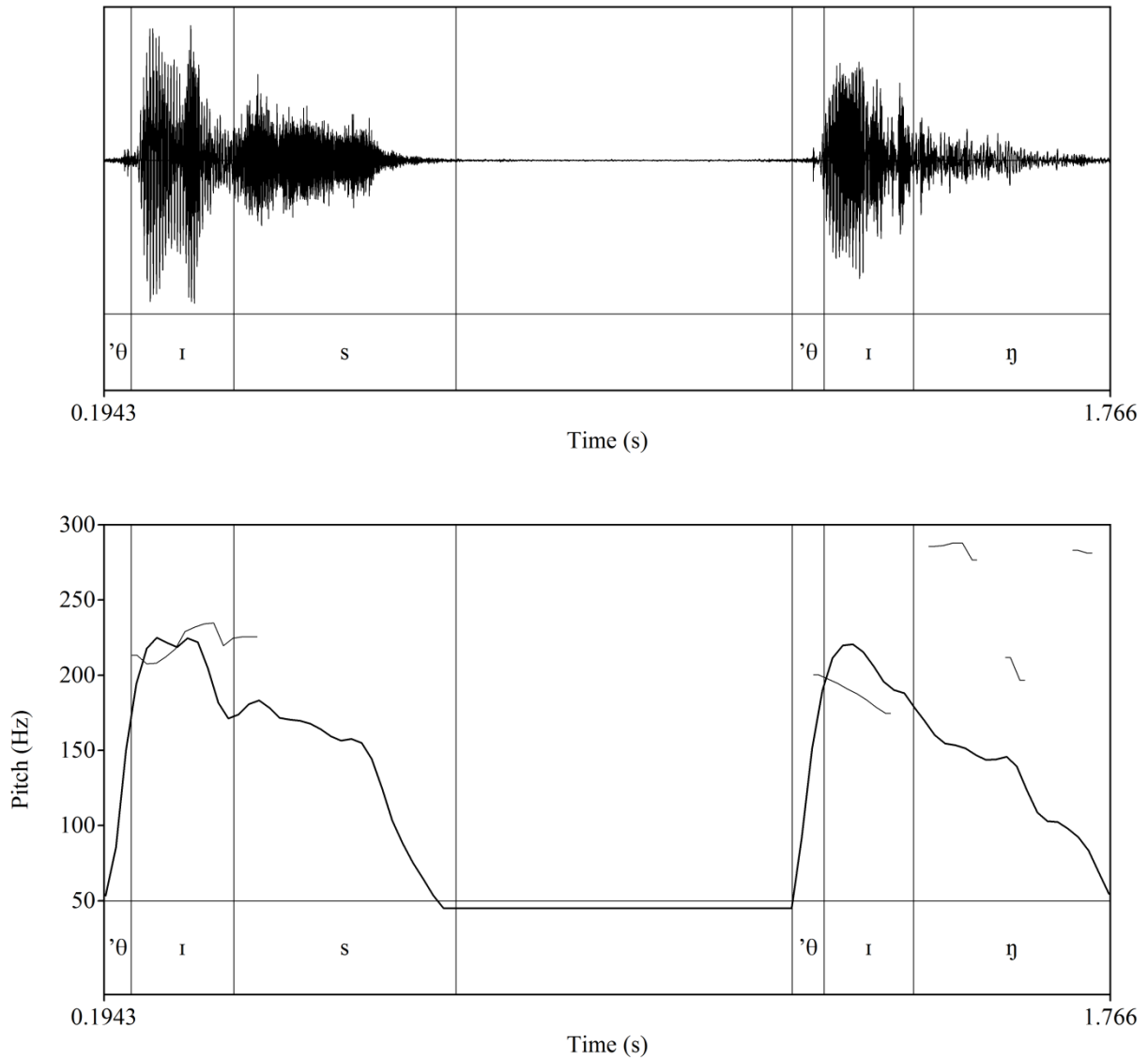
There is distortion in the pitch curve; the /k/ sound is completely voiceless.

Image 10. *perfect bird*. Participant 2, part 1 of the test, first recording.

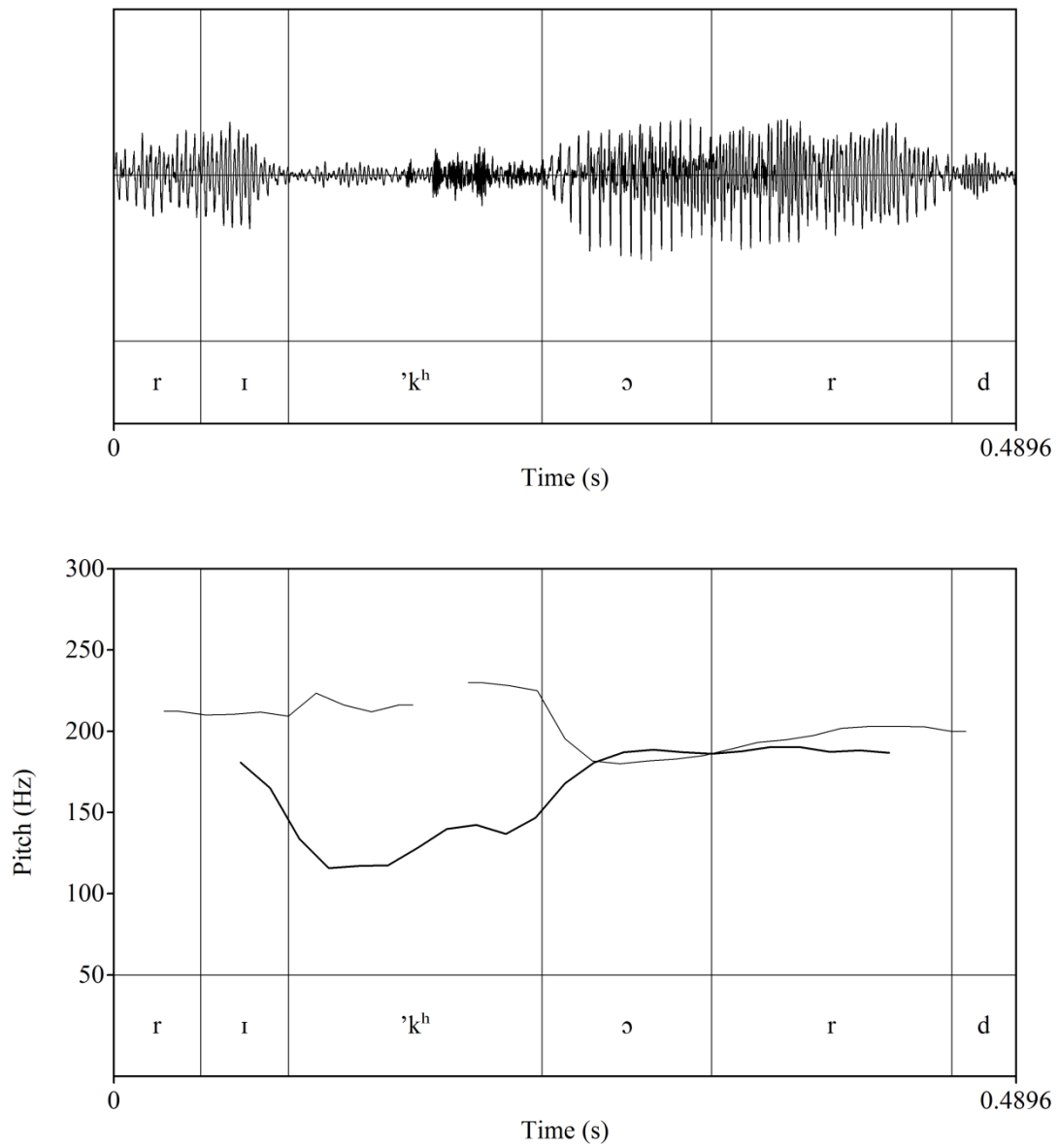
There is distortion in the pitch curve; the /ɜ:/, /ɜ:/ and /r/ sounds are all completely voiced, and the curve should travel between 100 and 250 Hz.

Image 11. *attempts*. Participant 1, part 1 of the test, second recording.

There is distortion in the pitch curve; the short /ə/ sound in the latter word is also voiced.

Image 12. *this/thing*. Participant 1, part 2 of the test, second recording.

There is distortion in the pitch curve; the /ŋ/ sound is completely voiced and the curve in the latter word should be descending from 200 Hz to around 125-150 Hz.

Image 13. *record*. Participant 1, part 1 of the test, second recording.

There is distortion in the pitch curve; the /k^h/ is completely voiceless.

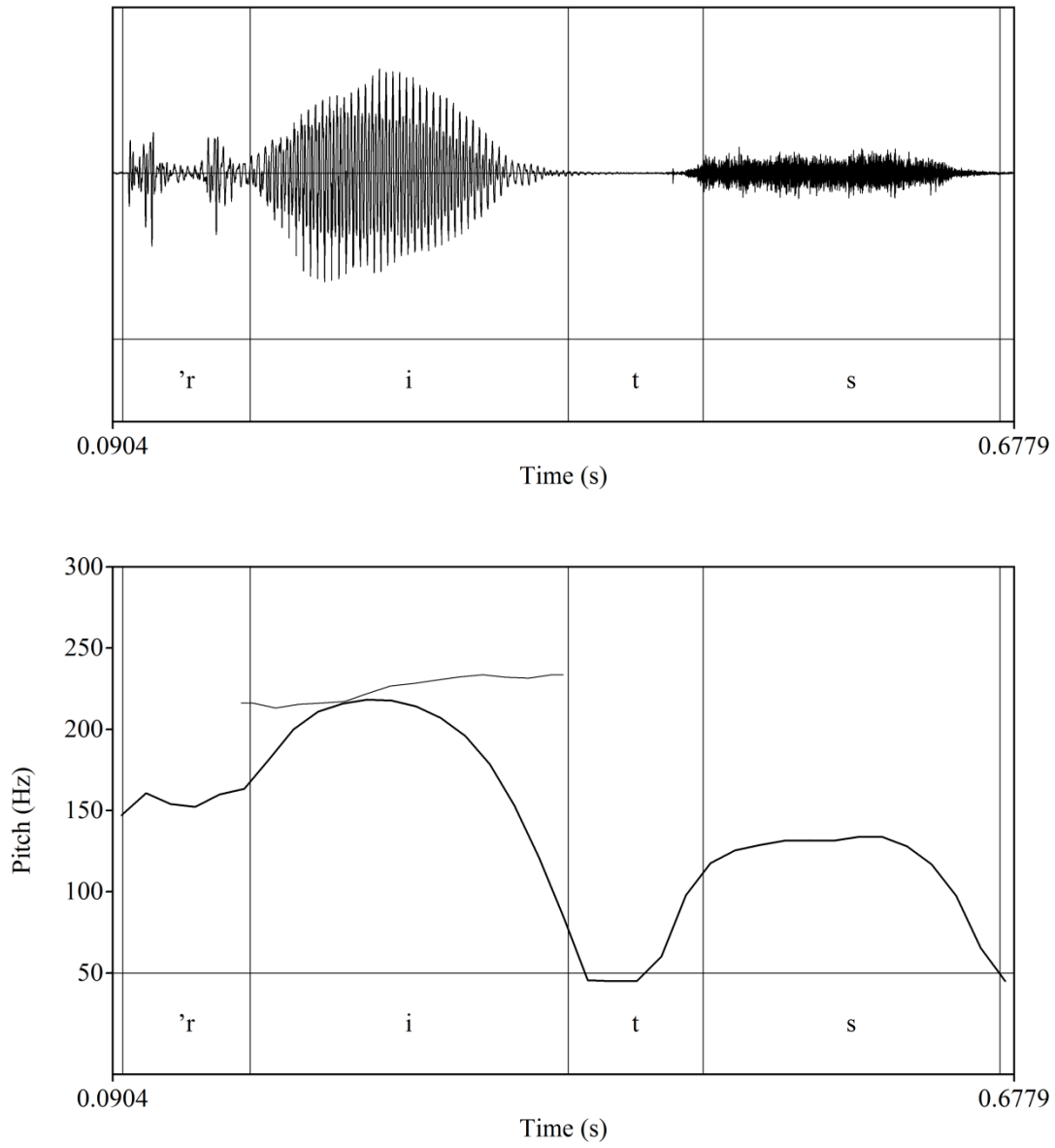
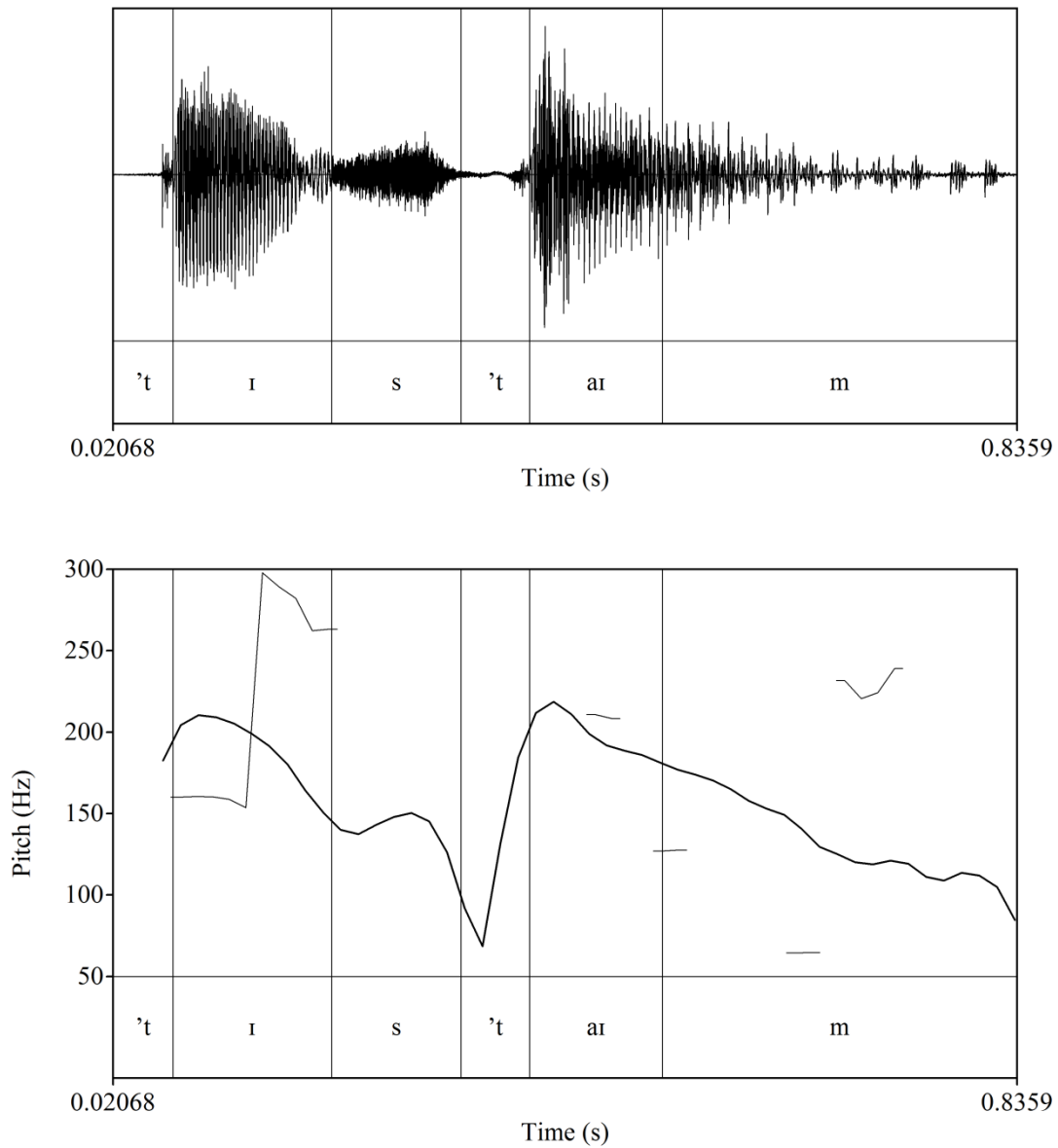
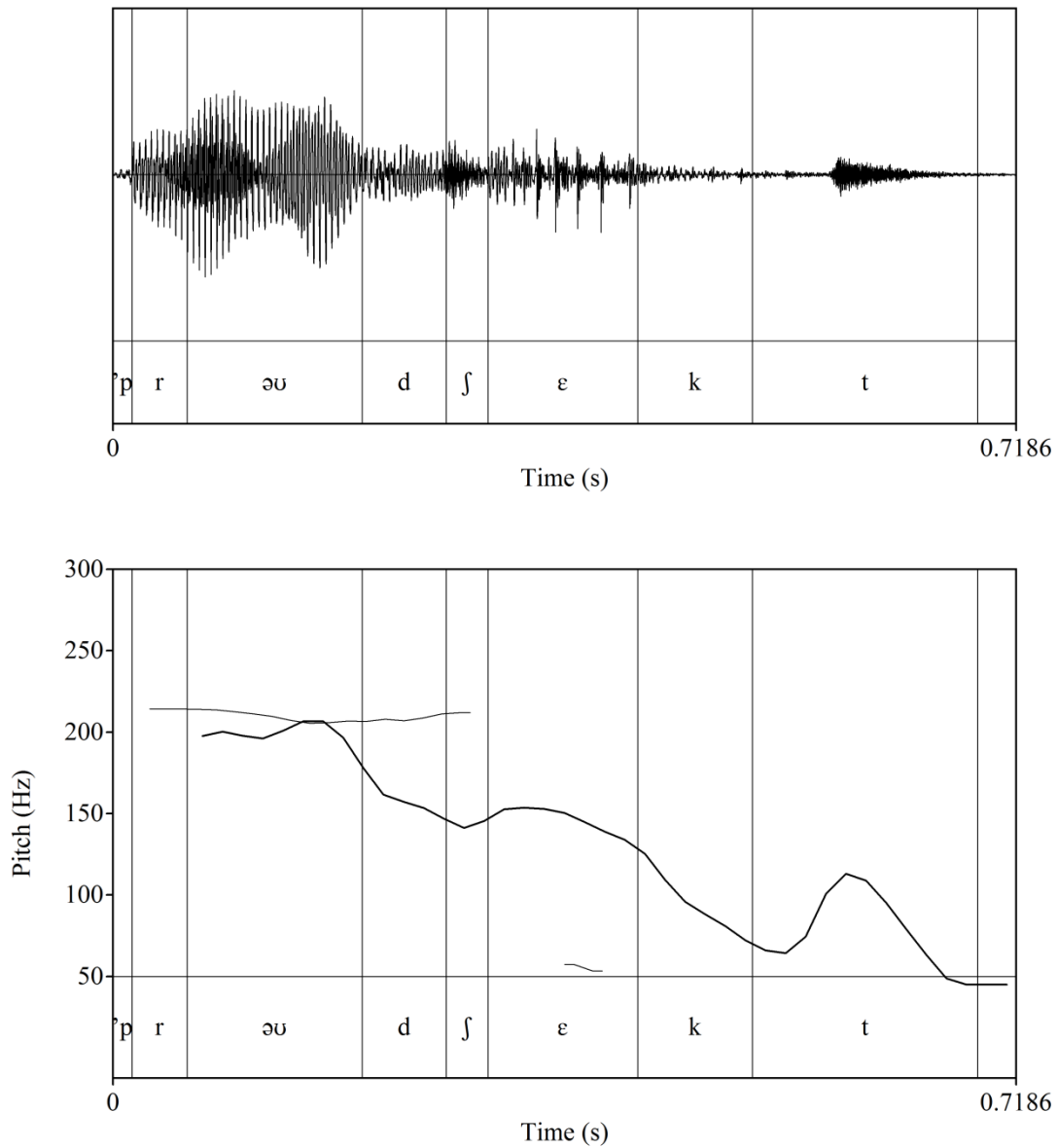
Image 14. *ridge*. Participant 2, part 1 of the test, first recording.

Image 15. *this time*. Participant 2, conversational section of the test.

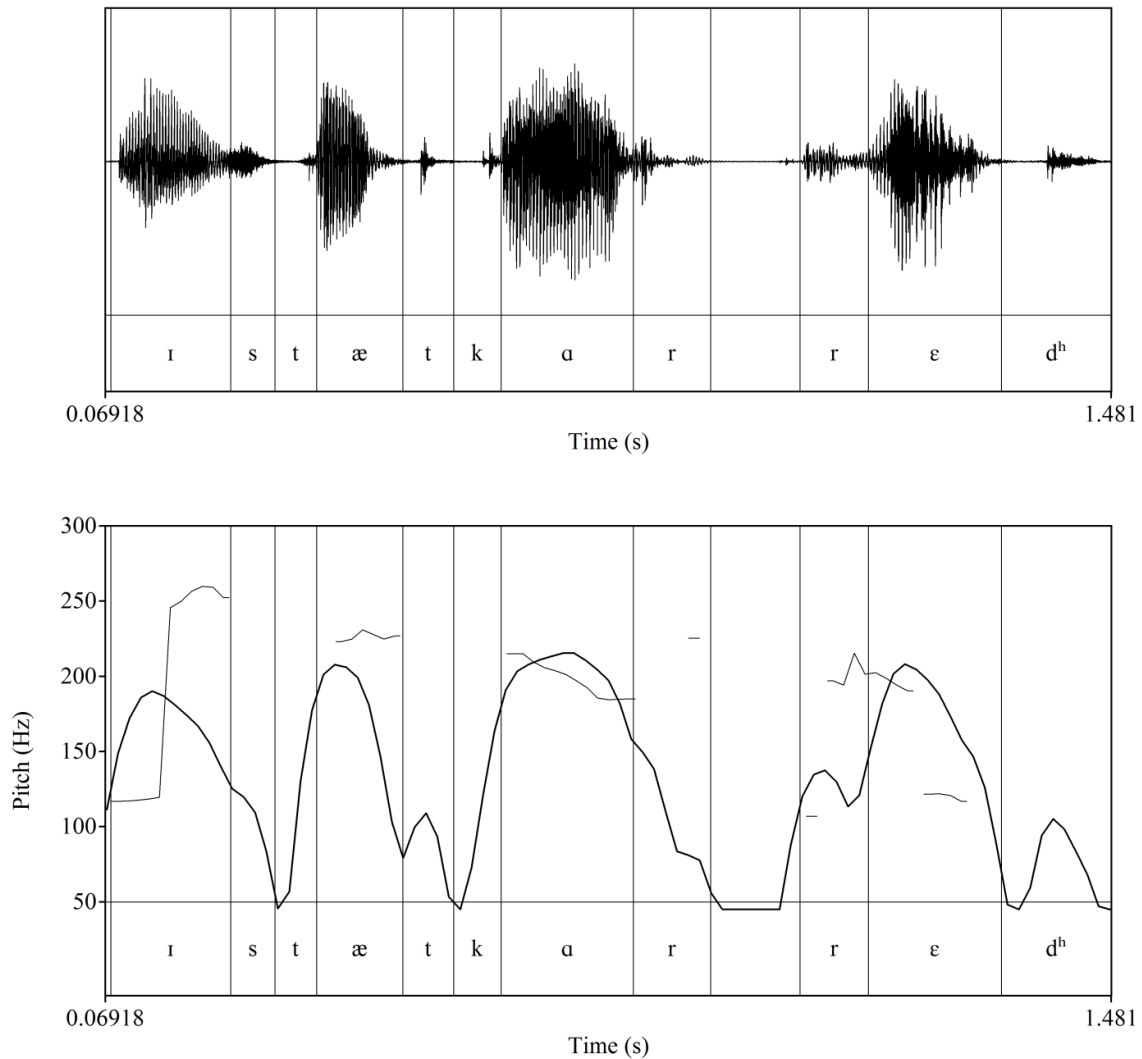


There is distortion in the pitch curve; the end of the word '*time*', /am/, should be voiced throughout, and the pitch descends from around 200 Hz to 150 Hz.

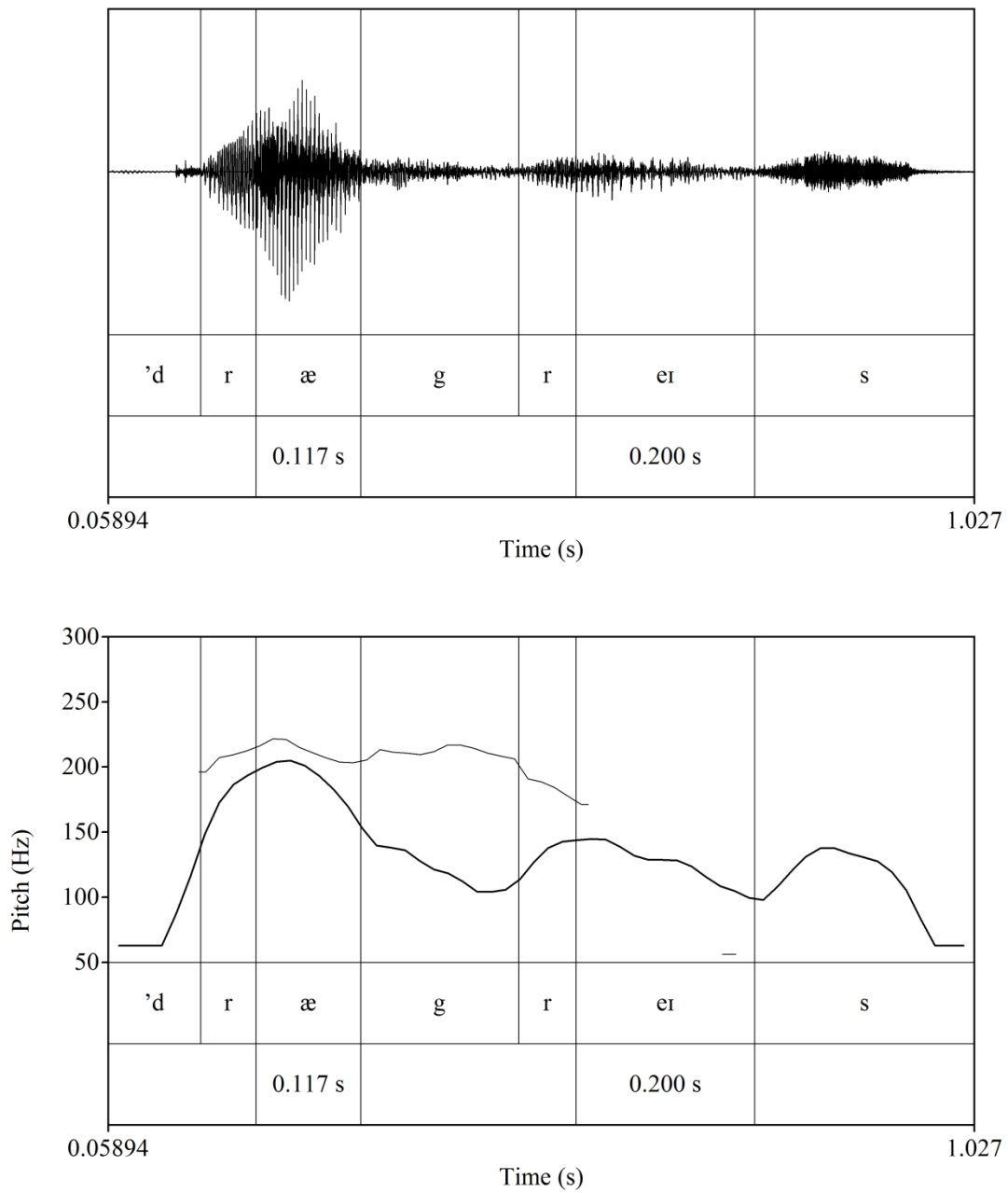
Image 16. *project*. Participant 1, part 1 of the test, first recording.

There is distortion in the pitch curve; the /ʃ/ is voiceless and the /ε/ is voiced.

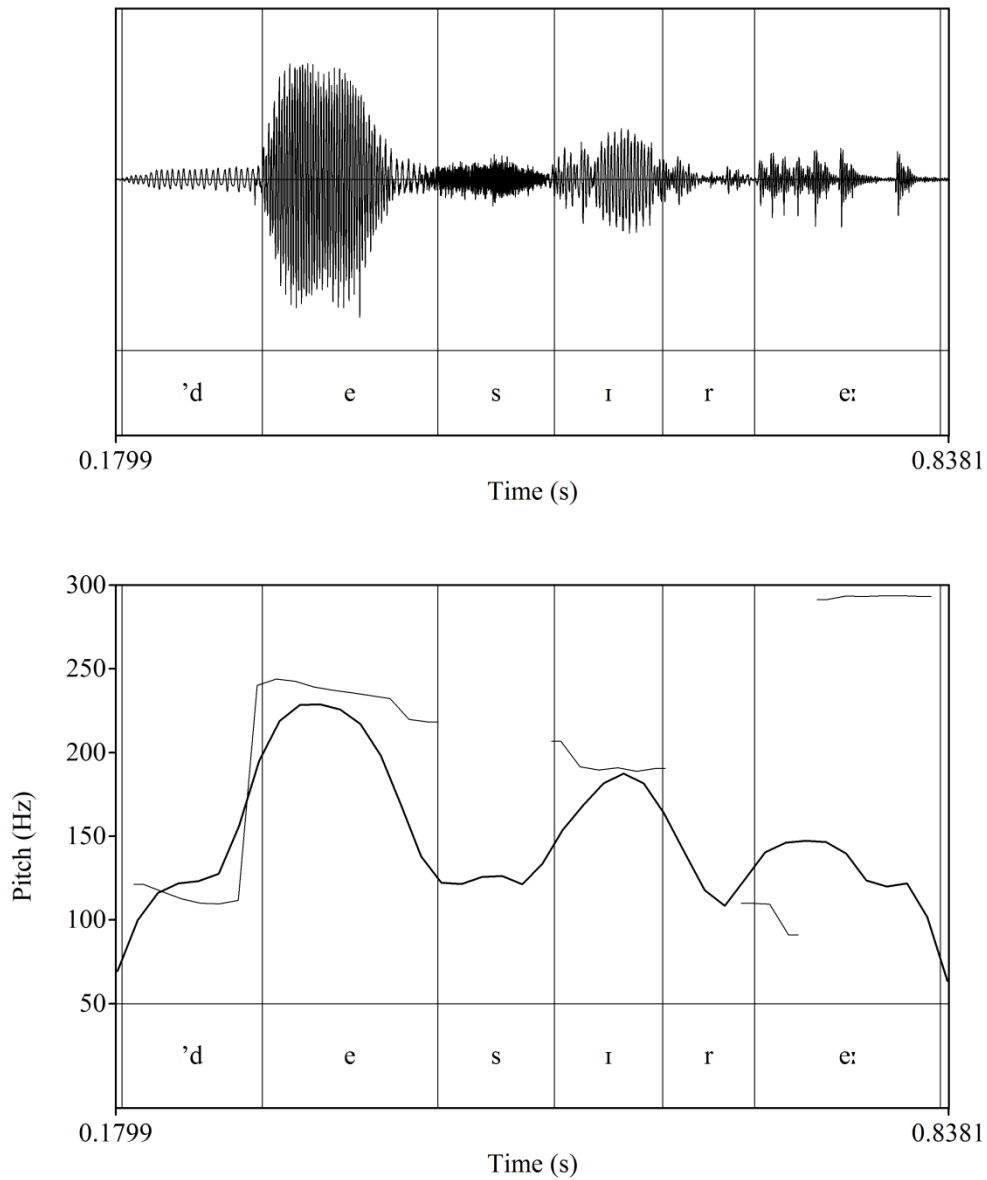
Image 17. *Is that car red?*. Participant 2, part 3 of the test, first recording.



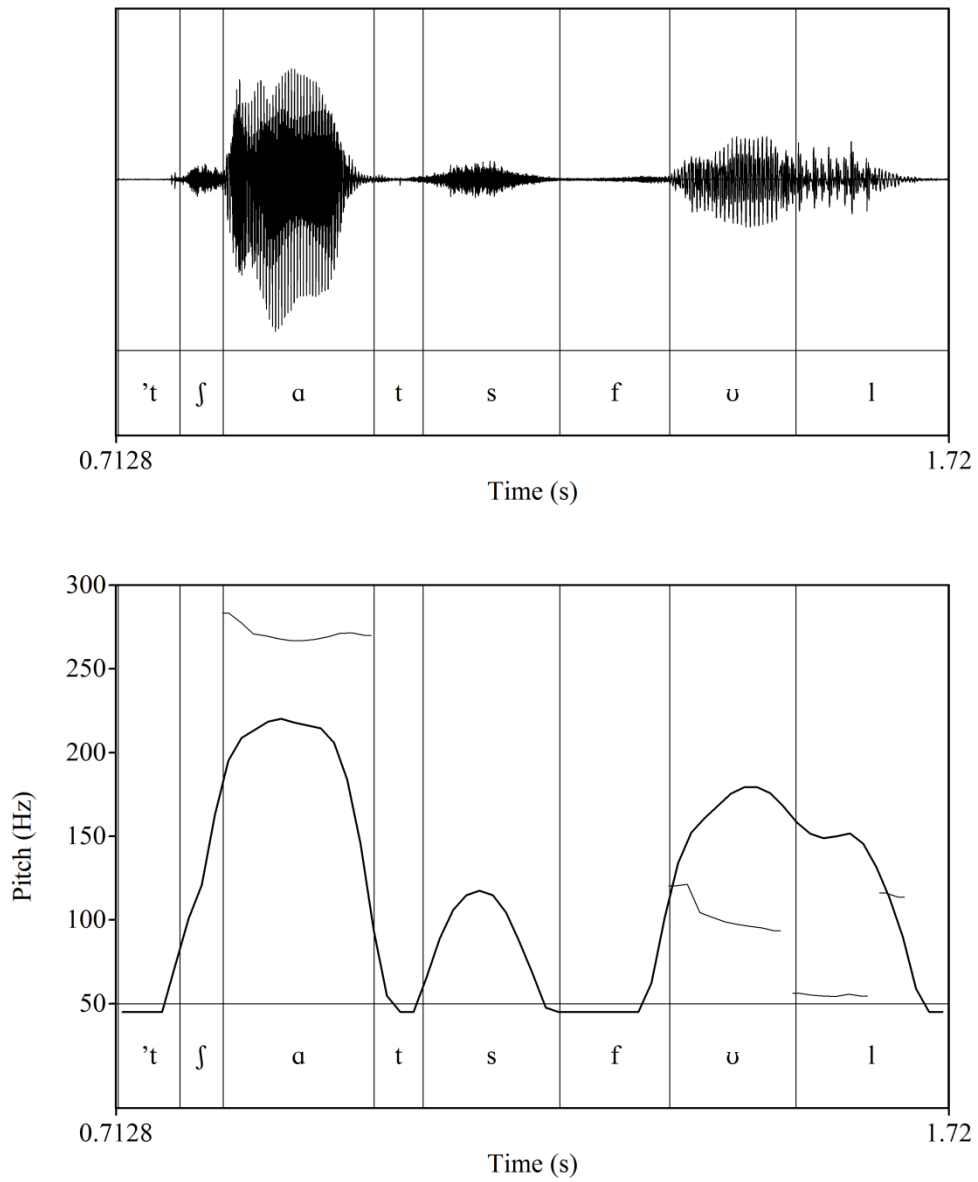
There is distortion in the pitch curve; all the /r/ sounds are voiced, and the pitch curve should be slowly descending from around 250 Hz to 125 Hz.

Image 18. *drag race*. Participant 1, part 3 of the test, first recording.

There is distortion in the pitch curve; also the diphthong /eɪ/ is voiced.

Image 19. *desire*. Participant 2, part 2 of the test, first recording.

There is distortion in the pitch curve; the voicing continues throughout the end of the word, /ire:/.

Image 20. *judeful*. Participant 2, part 4 of the test, first recording.

There is distortion in the pitch curve; the voicing is continuous at the end of the word, /ɔl/.

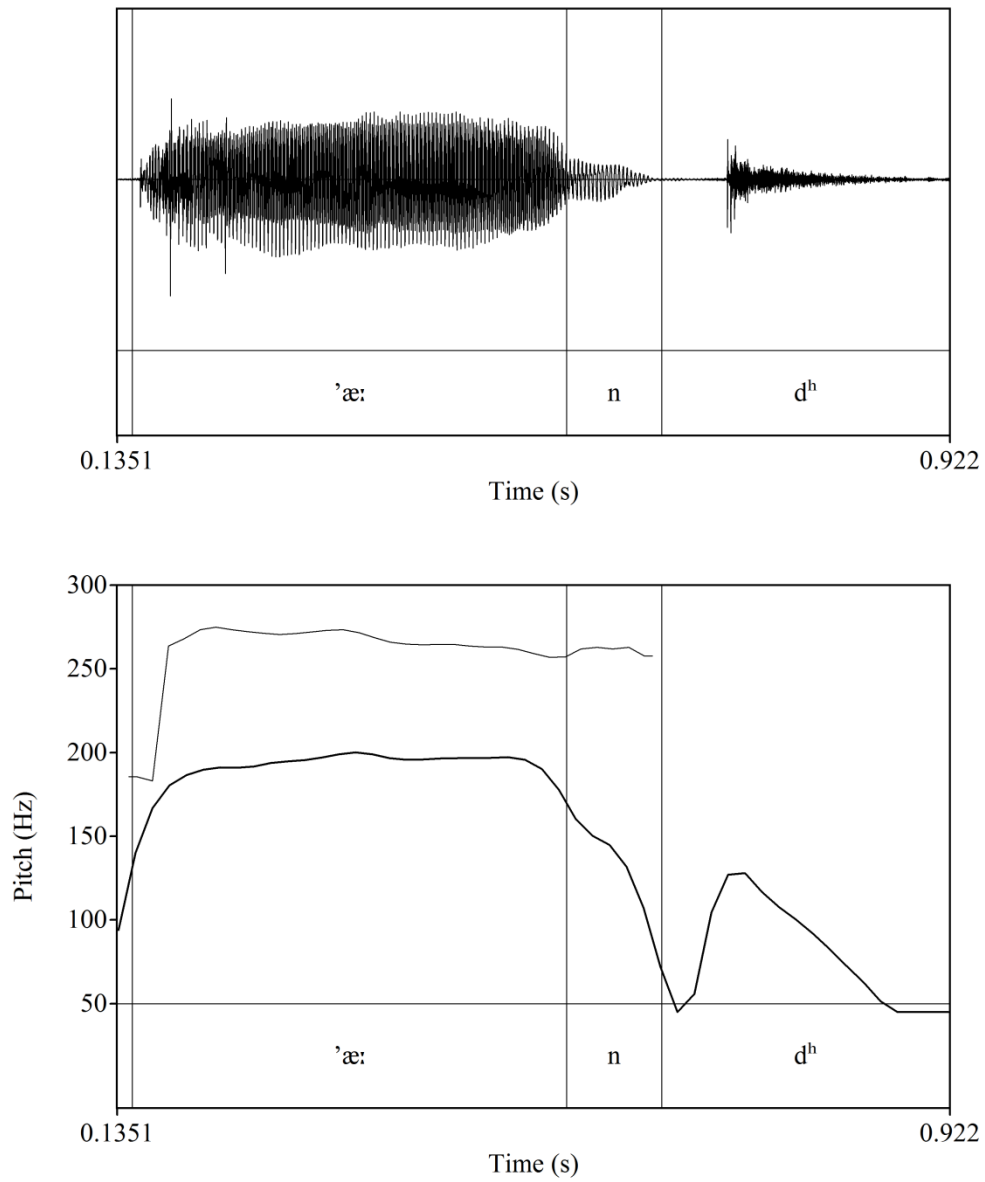
Image 21. *and.* Participant 2, conversational section of the test.

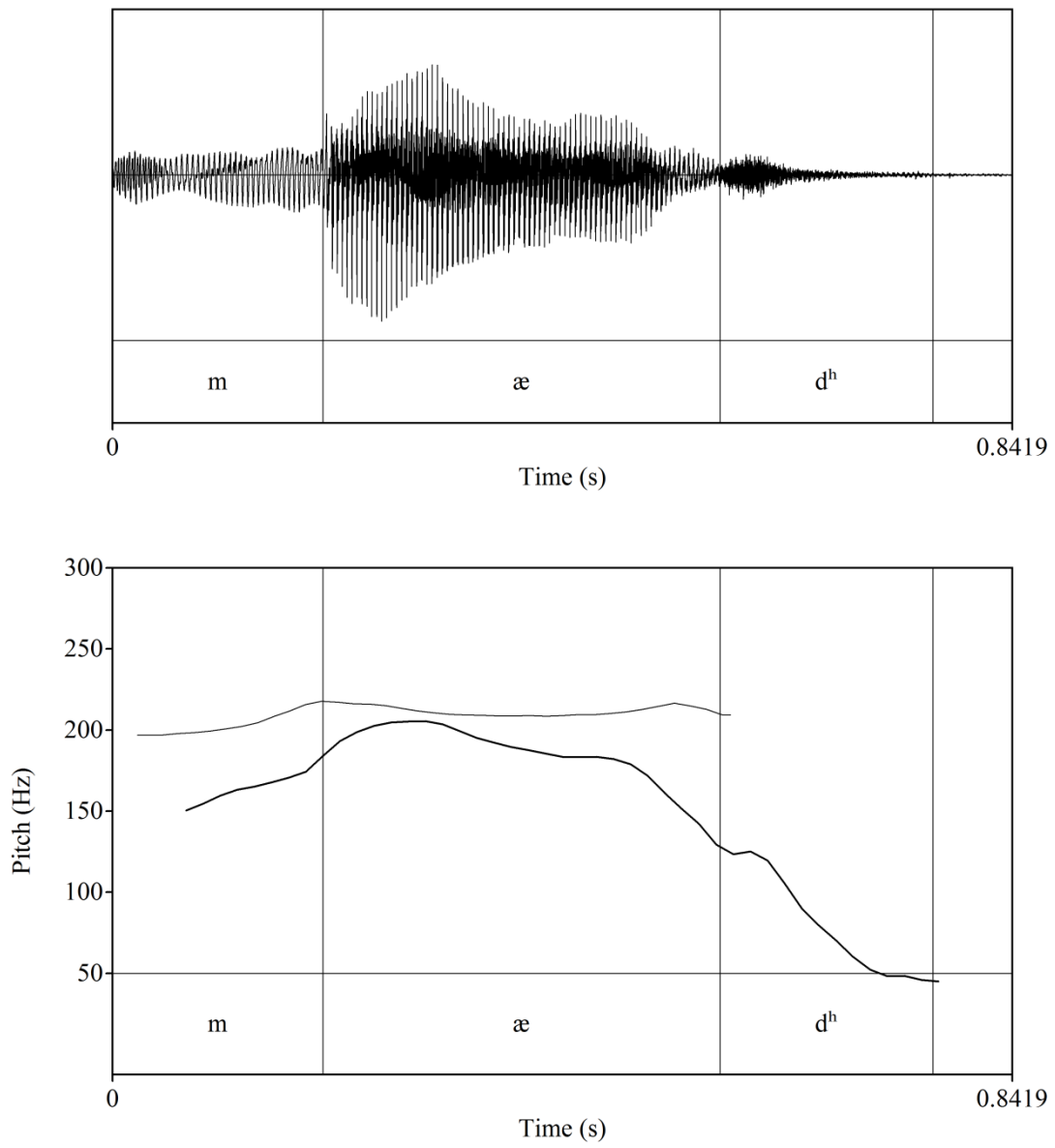
Image 22. *mad*. Participant 1, part 4 of the test, second recording.

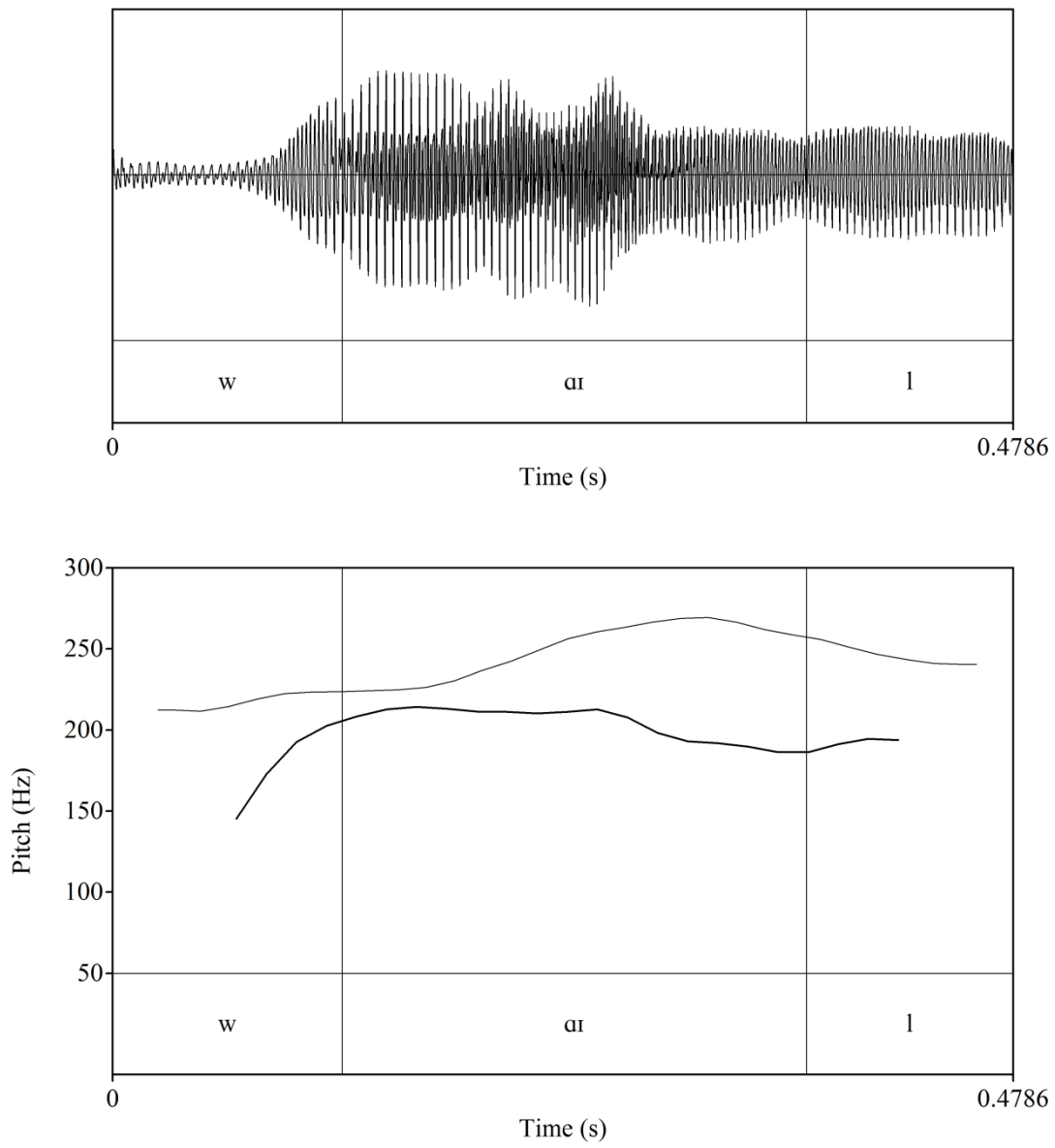
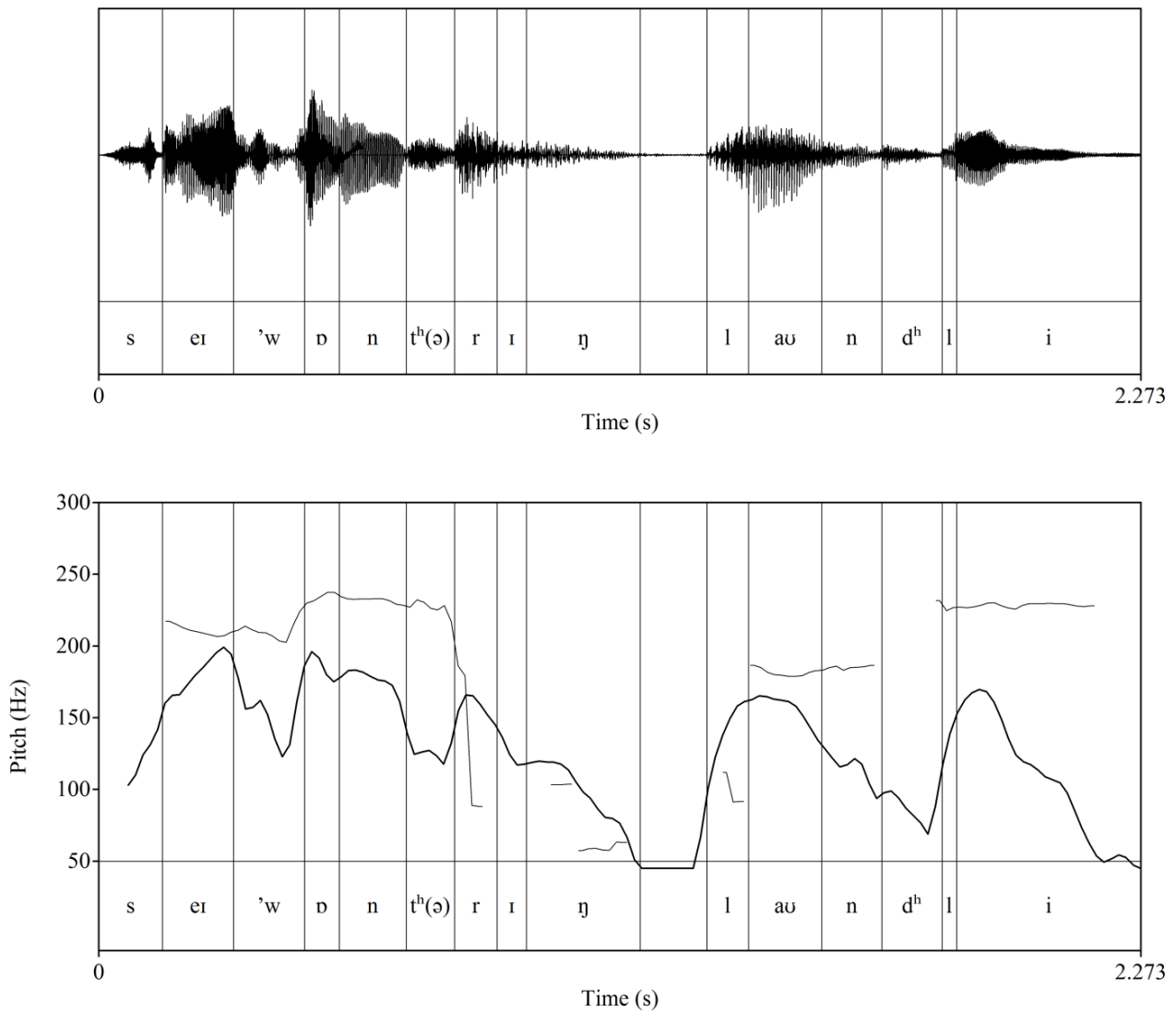
Image 23. *vile*. Participant 2, part 1 of the test, first recording.

Image 24. *Say vantering loudly.* Participant 1, part 4 of the test, second recording.



There is distortion in the pitch curve; the voicing is continuous in the sound combinations /rŋ/ and /laʊn/.