English monophthong vowels as produced by Finnish and Finland-Swedish ninth-graders

Master's thesis

Valtteri Nyyssönen

University of Jyväskylä
Department of Language and Communication Studies
English

September 2020

JYVÄSKYLÄN YLIOPISTO

Tiedekunta – Faculty Humanistis-yhteiskuntatieteellinen tiedekunta	**Laitos – Department** Kieli- ja viestintätieteen laitos
Tekijä – Author Valtteri Nyyssönen	
Työn nimi – Title English monophthong vowels as produced by Finnish and Finland-Swedish ninth-graders	
Oppiaine – Subject Englannin kieli	**Työn laji – Level** Pro gradu -tutkielma
Aika – Month and year Toukokuu 2020	**Sivumäärä – Number of pages** 78 + 1 liite

Tiivistelmä – Abstract

Tämän pro gradu -tutkielman aiheena on suomalaisten ja suomenruotsalaisten yhdeksäsluokkalaisten englannin vokaaliääntäminen, jota tutkittiin akustisen analyysin menetelmin.

Ääntäminen on merkittävä kielitaidon osuus. Aiempi tutkimus osoittaa, että ääntämisellä ja erityisesti prosodisilla piirteillä on vaikutusta puhujan ymmärrettävyyteen. Vokaaliääntämyksellä on suora vaikutus puhujan vieraan korostuksen vahvuuteen, mutta ainakin englannin kielen ollessa kyseessä se voi myös epäsuorasti vaikuttaa puheen rytmiin ja sujuvuuteen.

Tutkimuksen ensisijainen tarkoitus oli tutkia suomalaisten (n=5) ja suomenruotsalaisten (n=4) yhdeksäsluokkalaisten puhuman englannin akustisia ominaisuuksia. Tutkimushenkilöt olivat Keski-Suomesta ja Uudeltamaalta. Toissijaisena tavoitteena oli verrata tutkittavien ääntämistä toisiinsa sekä brittienglannin ja amerikanenglannin viitearvoihin. Tutkimus on näin ollen deskriptiivinen foneettinen tutkimus, jolla on myös rajallisia implikaatiota kieltenopetuksen ja soveltavan kielitieteen näkökulmasta. Tutkimusaineistona oli nauhoitettua yksittäisiä sanoja ja fraaseja lukupuhuntana. Sanoista eristettiin vokaalit, joista mitattiin formanttitaajuudet ja kesto.

Tulokset olivat melko samanlaisia molemmilla puhujaryhmillä. Molemmat ryhmät äänsivät etu- ja keskivokaalit hyvin, kun taas takavokaalit erosivat vähän, mutta havaittavasti natiivipuhujien viitearvoista. Kummassakin ryhmässä esiintyi pitkä-lyhyt-vokaaliparien laadullista samankaltaistumista ja puuttuvaa vokaalireduktiota, joskin ilmiöt olivat selkeämpiä suomalaisten ääntämisessä. Kumpikaan tutkimusryhmä ei tuottanut riittävää kestoeroa soinnillista ja soinnitonta konsonanttia edeltävien pitkien vokaalien välille. Lisäksi suomenruotsalaiset äänsivät englannin kielen u-vokaalit etisempänä kuin suomalaiset ja olivat ääntämykseltään roottisempia.

Koska tutkimuksen otos oli pieni, tuloksia ei voida pitää yleistettävinä. Tutkimuksen tuloksia voidaan sen sijaan pitää alustavina havaintoina jatkotutkimusta varten. Esimerkiksi suomalaisten ja suomenruotsalaisten puhuman englannin prosodiset erot ovat aihe, jota voisi tutkia, ja josta saattaisi löytyä merkittävämpiä eroavaisuuksia kieliryhmien välillä.

Asiasanat – Keywords englannin kieli, fonetiikka, ääntäminen, akustinen analyysi

Säilytyspaikka – Depository JYX

Muita tietoja – Additional information

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

1. INTRODUCTION

There are three major ways to assess pronunciation, as presented by Munro & Derwing (1999): *comprehensibility, intelligibility* and *foreign accent*. Comprehensibility is the listener's subjective perception of how easy it is to understand the speech. This is usually measured on a numerical scale of 1–9. Intelligibility, on the other hand, is a more concrete and objective way of assessing pronunciation. It can be done by having listeners transcribe what they hear and comparing the transcription to what the speaker intended to say. Last, foreign accent (i.e. accentedness) is the listener's subjective perception of how strong a foreign accent the speaker has. Speakers may receive a 100% score in intelligibility while getting a less than perfect comprehensibility and accentedness ratings (Munro & Derwing 1999). The explanation for this is that speech may be fully intelligible but requires conscious effort to follow, which leads to harsher comprehensibility ratings. In a previous study of theirs, Munro & Derwing (1995) found that lower comprehensibility ratings correlated with longer processing time. They also found that accentedness correlates best with the amount of segmental and suprasegmental deviations from native-like pronunciation, while also having some correlation with comprehensibility and little with intelligibility (Munro & Derwing 1999). A more recent study by Trofimovich & Isaacs (2012) had similar results: English learners' phonological errors were mostly affecting their accentedness ratings, whereas comprehensibility was affected by errors in grammar and vocabulary.

There are, however, segmental errors that may compromise the intelligibility of speech. Traditionally, it is believed that English consonants cause the most trouble to Finnish speakers (Morris-Wilson 2003: 4), which has resulted in consonants getting the most attention when teaching English segmentals to Finns. The same kind of information for Finland-Swedish speakers is not available, since there are virtually no studies about Finland-Swedish-accented English—presumably because Finland-Swedish is such a small language group. There is some knowledge on Sweden-Swedish speakers' typical problems (e.g. Davidsen-Nielsen & Harder 2001) but its relevance is disputable, since the two national varieties of Swedish are phonologically quite different, as Finnish has had a strong influence on Finland-Swedish.

Comprehensibility, on the other hand, could be affected by the rhythm and flow of speech. Unnatural, staccato-like rhythm is a common feature in learner speech. Morris-Wilson (2003: 183, 194) claims that speaking with a "jerky" rhythm draws the listener's attention away from the message itself, when the rhythm is odd and the stress is on the wrong words. This may lead to increased processing cost, although there are no studies to prove that. Saito, Trofimovich & Isaacs (2016) found that prosody is crucial for comprehensibility in English on all levels of language proficiency, whereas high segmental accuracy is significant for advanced speakers' comprehensibility. In addition, recent studies of Swedish have shown that prosodic features, such as sentence-level stress, are quite crucial in native listeners' judgments of L2 pronunciation (Kuronen & Tergujeff 2017; Kautonen 2018). According to Morris-Wilson (2003: 196–197), odd or missing stress patterns are caused mainly by not being able to "think ahead" when speaking in a foreign language, which implies that there is a required level of competence in the language before natural rhythm can be reached.

In English, stress placement is tightly linked with vowel quality: unstressed syllables always have a reduced vowel and stressed syllables cannot have a reduced vowel. A popular example of this is so-called weak forms (i.e. the unstressed forms) of small and common words such as *have* or *to*, which are pronounced with the neutral vowel /ə/ when unstressed. Morris-Wilson (2003: 197) goes as far as say that not using the unstressed weak forms makes the acquisition of natural flow and rhythm "downright impossible." This kind of systematic vowel reduction is not a feature of Finnish or Finland-Swedish, and learning it could aid in producing natural speech rhythm because they are so closely related. It is safe to say that in English, using reduced vowels often enough goes hand in hand with speaking with a natural and flowing rhythm.

Although looking at vowel pronunciation probably does not tell a lot about intelligibility and comprehensibility, which rightfully are the more important goals when learning to speak in a foreign language (see Levis 2005), it does affect the accentedness of speech. Although accentedness has not been found to correlate with processing time or intelligibility (Derwing & Munro 1995; 1999), it has been shown that people tend to react negatively to foreign accents. For example, Morris-Wilson (1999: 276) found that traits associated with status and competence are judged negatively when a person has a strong Finnish accent. However, a clear connection between foreign accent and traits pertaining to solidarity was not found. A large meta-analysis conducted by Fuertes et al. (2011) concluded that speaking with a non-standard accent can have substantial consequences on how the speaker is viewed by other people. The effect was particularly strong when comparing non-standard accents to General American, which makes this particularly relevant for the present study.

However, it must be noted that many of the studies that Fuertes et al. (2011) and Derwing & Munro (1999) have listed as proof for the negative effect of foreign accent are from the 60s and 70s. The world has changed tremendously since then in terms of globalization and contact to people from cultures and languages. Recent studies such as Dewaele & McCloskey (2015) show that people tend to react less negatively to foreign accents if they have experiences of living abroad or working in ethnically diverse environments. International and multilingual working environments are much more common today than fifty years ago. Considering this, the consequences of having a strong foreign accent are not probably as large as they were decades ago, when these studies were conducted. However, it is the author's belief that there is still something to benefit from losing a strong foreign accent. First, the negative effect may be diminished from that of fifty years ago, but not vanquished. The consequences of having a strong foreign accent are especially strong on an English as a Second Language (ESL) speaker, because a foreign accent immediately reveals the speaker’s status as an immigrant, “foreigner” or an “outsider” among native English speakers. The accent itself is not frowned upon, but negative stereotypes linked to (especially, but not necessarily limited to, non-white) immigrants etc. may be evoked by the accent and lead to discrimination. (Derwing & Munro 2015: 17–18) Because of this, ESL learners often want to lose their foreign accent. For example, Derwing (2003) found that 95% of the ESL learners that she interviewed in Canada would have wanted to pronounce English with a native accent. The interviewees felt that native speakers did not pay attention to their message or treated them rudely because of their accent. The participants of the present study are learning English as a Foreign Language (EFL), which is why they are not subject to everyday discrimination based on their English skills like ESL learners often are. Still, many people are quite self-conscious about their foreign accent (Dewaele & McCloskey 2015), which is why reducing accentedness might increase one's language confidence. In addition, a strong foreign accent may prove burdensome in an international career in business, for example. In summary, improving one's pronunciation need not—or rather, should not—stop after reaching full intelligibility.

The present study is essentially a descriptive phonetic study. The study aims to describe the acoustic (as well as articulatory) qualities of English vowels as pronounced by Finnish and Finland-Swedish ninth-graders. Additionally, this study aims to preliminary knowledge of how accurately Finnish and Finland-Swedish intermediate learners pronounce English vowels. This is why the participants' productions are compared to the two native varieties of English: General British (GB) and General American (GA). These two were chosen because British English and American English are the two primary native varieties in English teaching. Furthermore, I chose GB and GA because both are relatively unmarked standard varieties of English. There has been a long debate among linguists over the standard variety of British English, in particular, and what it should be called. Cruttenden (2014: 80) has chosen to use General British, as the old term Received Pronunciation has become obsolete and other terms, such as Standard Southern British English, are not as neutral. GB also parallels the name of the other regional standard, GA. As for Munro & Derwing's (1999) three aspects of pronunciation, I am not going to analyze the intelligibility, comprehensibility or accentedness of the participants' pronunciation per se. However, author's notions about intelligibility, comprehensibility and accentedness are given when summarising the participants' differences to native speakers of English, because looking at the acoustic signal alone does not give very useful results from the applied linguist's point of view.

This study will be one of the first ones to study Finland-Swedish speakers’ English pronunciation and compare it with Finnish speakers. Tergujeff's ongoing project *Intelligibility, comprehensibility and accentedness of English spoken by Finns* (ICASEF, 2018–2021), for which the author collected data alongside the present study, is also aimed at this obvious gap in research. There is not too much research on English vowel pronunciation of either language group with Finland-Swedish learners being a practically unexplored area. This is probably because vowel pronunciation rarely causes serious problems, at least for Finns (Morris-Wilson 2003: 4). There are some pedagogical implications to be derived from the results of this study, mainly what vowels and features of vowel pronunciation ninth-graders have already learned and what needs reinforcement. Last, Finnish and Finland-Swedish provide an interesting ground for comparison because the two languages are almost similar on a segmental phonetic level but otherwise quite different.

The study is divided into three parts. First, I will cover the theoretical background in Section 2.1 by exploring the basics of acoustic analysis and theoretical models of L2 pronunciation learning. The vowel systems of the languages in question are presented and compared in Section 2.2. Before moving on to the present study, some previous studies are presented in Section 2.3. Second, the research questions, methods and participants of this study are presented in Section 3. Last, the results are presented in Section 4, while the discussion and concluding remarks are reserved for Sections 5 and 6 respectively.

2 BACKGROUND

2.1 Theoretical background

Before discussing actual speech sounds, some basic principles of acoustics are to be clarified. The present study and its results are phonetic in nature, and the purpose of this section is to explain why it is worthwhile to study the acoustic properties of speech sounds. The acoustics of speech is thoroughly discussed in Suomi's (1990) book, which is aimed for aspiring students and/or researchers interested in phonetics and which is also the basis for the majority of Sections 2.1.1 and 2.1.2. Studying speech through acoustics is called **acoustic analysis**, because the object of interest is the speech signal and its properties. The benefit of acoustic analysis over **articulatory analysis** is that it is non-invasive, whereas articulatory analysis often uses a palatograph or a camera inserted into the mouth. Articulatory analysis may also use non-invasive medical imaging, such as X-ray imaging, but they require very expensive equipment, whereas basic acoustic analysis can be done with inexpensive equipment and free software.

2.1.1 The acoustics of speech sounds

A **simple** sound consists of only one sound wave. Such sounds are also called sine waves. However, almost all natural (i.e. not synthesized) sounds that we hear – including all speech sounds – are **complex**, which means that they consist of numerous different and simultaneous sine waves. The French mathematician Joseph Fourier first introduced the theorem that all complex waves can be seen as a group of individual sine waves in 1822. The individual sound waves of a complex sound cannot be distinguished from each other by looking at the waveform alone, which is why a complex sound must be decomposed into its sine components with **Fourier analysis**, named after the French mathematician (Suomi 1990: 27). Then, the individual sound waves, along with their frequencies, amplitudes and phases, can be observed in isolation. A wave's frequency measures how quickly it vibrates whereas amplitude measures the magnitude of the vibration. The phase of a wave describes its position in relation to other sound waves. These sine waves are called the component frequencies of a complex sound. Fourier analysis uses a mathematical formula to decompose a complex periodic wave, which used to be a laborious process and too complex to understand for most linguists. However, modern speech processing software, such as Praat (Boersma & Weenink 2019), does it automatically, which has made it tremendously more accessible for linguists.

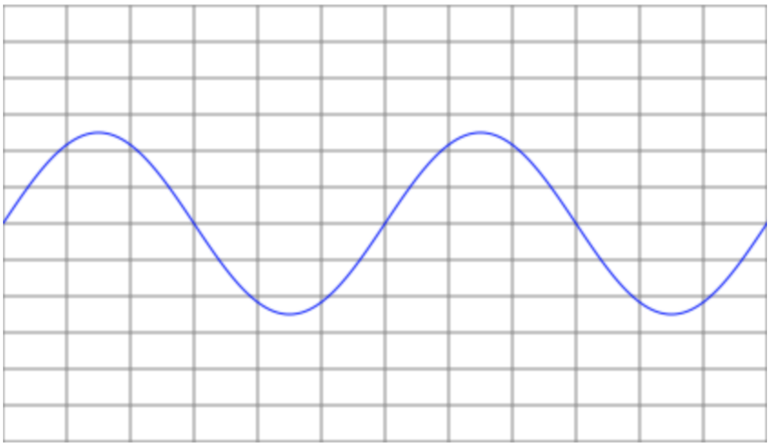

Figure 1. A simple 100Hz sine wave, two full cycles.

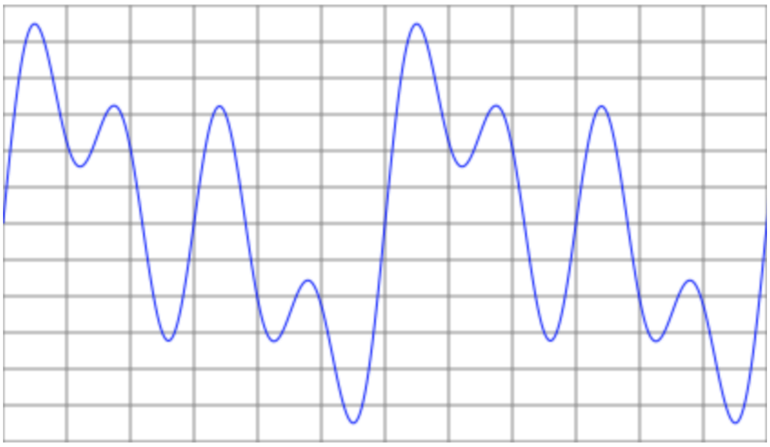

Figure 2. A periodic complex wave with 100Hz, 200Hz and 400Hz components with the same amplitude and phase, two full cycles.

Furthermore, complex sounds can be either **periodic** or **aperiodic**. Periodic sounds have a regular repeating waveform, which is why they have a perceivable and measurable pitch (Suomi 1990: 37). In turn, aperiodic sounds consist of irregular sound waves, which is why they do not have a clearly perceivable pitch. However, they do have a vague sense of pitch. For example, /ʃ/ sounds somewhat "darker" than /s/, because the energy is concentrated on lower frequencies in /ʃ/ (Suomi 1990: 41–42). In human speech, sonorants (vowels and consonants that are produced with a continuous non-turbulent airflow) are periodic. On the other hand, all obstruents, i.e. sounds that are produced by obstructing the airflow completely (plosives) or partially (fricatives), are aperiodic. As an exception, voiced obstruents are simultaneously periodic and aperiodic: the sound created in the place of articulation is aperiodic and the sound coming from the glottis is periodic. For example, one can sing a melody with [ʒ] despite it is an obstruent, although not as easily as with a vowel. Vowels, which are in the focus of this study, are complex periodic sounds.

Furthermore, all sonorants are actually quasi-periodic. This means that while there are miniscule fluctuations in cycle length, they can be considered periodic for the purpose of phonetic analysis. For example, a perfectly periodic 100Hz sound has a constant cycle length of 10ms. Its cycle length is 10ms because it vibrates 100 times in a second and it is constant because the sound is periodic. In turn, the cycle length in a 100Hz speech sound is approximately 10ms and most often marginally longer or shorter. Thus, sonorant speech sounds are not periodic per se, but the fluctuation in cycle length (also called jitter) is so small that they can be treated as periodic in acoustic analysis. A large amount of jitter makes a person's speech sound distorted, which usually occurs in the pathologic speech, i.e. the speech of people that have injured or deformed speech organs.

The perceived pitch of a complex sound comes from its fundamental frequency (F_0), which is the frequency of the slowest (i.e. lowest) component frequency (Suomi 1990: 29). All components above the F_0 are in a harmonic relation to the F_0 in a periodic sound, which means that they are divisible by the F_0. The components of periodic complex sounds are called harmonic partials. This means that the frequencies of harmonic partials can only be between increments that are the size of the F_0 (Suomi 1990: 38). For example, in a complex periodic sound with the F_0 of 100Hz, the next components would be 200Hz, 300Hz, etc. In addition to frequency, every component has its own amplitude and phase. The amplitudes of the components are the cause for differences in sound quality, which has been proven by experiments with sound synthesis. Two sounds can have exactly the same component frequencies while sounding different in quality, if the amplitudes of the component waves are different. For example, we can consider a 100Hz complex periodic sound that has harmonic partials between 100Hz increments up to 1kHz. If the third partial, i.e. the one with a frequency of $3*F_0 = 300$Hz, is relatively strong, it will sound different when compared to an otherwise identical sound that has a strong sixth partial ($8*F_0 = 800$Hz) in turn. In this case, the former will sound somewhat "darker" than the latter, because there is more acoustic energy on lower frequencies. The role of phase in sound perception has been found largely irrelevant, which is why it is not taken into account in acoustic analysis of speech (Suomi 1990: 32).

As a more concrete example, let us consider the General British vowels /aː/ and /iː/, which are both pronounced with the fundamental frequency of 100Hz by a male speaker. The sounds have the same pitch, but different sound quality, much like the difference between the same note played on two different musical instruments. As stated in the previous section, vowels are complex periodic sounds. In both sounds, the F_0 is 100Hz and the harmonic partials above that are 200Hz, 300Hz etc. However, in /aː/, the partials at around 700Hz and 1100Hz are relatively strong when looking at the spectrum of the sound. The vowel /iː/, in turn, has relatively strong partials at around 300Hz and 2200Hz. (Cruttenden 2014: 104.) In other words, two frequency peaks can be observed in the spectra of both sounds, but the peaks are on different frequencies. These frequency peaks are characteristic to each vowel, and do not change when producing the vowels with different F_0. The frequency peaks are called the **formant frequencies** of vowels, which are usually referred to as formants. The study of vowel quality has revolved around formants ever since the invention of the spectrograph in the 1940s, and vowel formants are also in the focus of this study.

In the study of vowels, no more than the first four formants (F_1–F_4) are usually taken into account in the analysis. There are formant frequencies above F_4, but they are so close to each other that they blend together in human hearing. They are also relatively weak when compared to F_1 and F_2, which is why they are most often left out of the analysis. The actual distances between higher formants are not smaller, but due to the logarithmic nature of human hearing, the same distance in Hz becomes effectively smaller on higher frequencies (Suomi 1990: 180). Vowel formants are especially useful because they do not only describe the sounds acoustically but they can also predict the position of the tongue and lips with some accuracy. For example, F_1 correlates negatively with vowel height (also called vowel closeness). In turn, F_2 correlates positively with vowel advancement and negatively with vowel rounding. (Suomi 1990: 147.) For example, F_1 is higher in /ɑ/ than in /i/ because the tongue is lower in the mouth. Furthermore, F_2 is lower in /ɑ/ than /æ/ because the tongue is further back in the mouth. Last, F_2 is also lower in /y/ than /i/ because the lips are rounded. However, the articulation of a vowel cannot be described by looking at formants alone; auditory and articulatory information is also needed to describe a vowel accurately.

2.1.2 The perception of speech sounds

Fant's (1960) **source-filter theory** explains how different vowel sounds are produced. The basic principle of his theory is that the differences between vowels are formed when the sound travels through the oral (and in sometimes, nasal) cavity, which acts as a filter. The effects of the filter are independent of the sound source (Fant 1960: 20). It means that the oral cavity always filters the same frequencies regardless of the sound that is projected through it. This is why two different vowels can be produced with the same F_0 and the same vowel can be produced with varying F_0.

The source sound is produced in the glottis and is called the glottal pulse. The glottal pulse is a complex sound with a strong F_0 component and a very large number of harmonic components that gradually decrease in their amplitudes with each component (Suomi 1990: 70). In other words, the spectrum of the glottal pulse resembles a downward slope, which does not have any frequency peaks. When producing different vowels at the same fundamental frequency, the glottis pulse is identical, because the sounds have the same F_0, which in turn determines the component frequencies of a periodic sound. However, the air column in the oral cavity has certain resonant frequencies that are determined by the length and shape of the oral cavity. This means that the air column starts to vibrate strongly when it is excited by sound that includes the resonant frequencies of the air column, i.e. it starts to resonate. (Suomi 1990: 54.) The resonant frequencies of the oral cavity can be changed by moving the speech organs, such as the tongue and lips, which changes the shape and length of the oral cavity. The frequencies of the glottal pulse that are near the resonant frequencies of the oral cavity "pass through" with less dampening than the others, which results in the frequency peaks that we call vowel formants. To put it very simply, the glottal pulse determines the frequencies of the components i.e. pitch, subglottal pressure (i.e. the pressure generated by squeezing the lungs with the diaphragm) determines their absolute amplitudes i.e. volume, and the shape of the oral cavity determines their relative amplitudes i.e. quality. (Suomi 1990: 80.)

Fletcher's (1940) **critical band theory** made it possible to assess what differences are perceivable and what are not. This is an important part of evaluating the findings of a phonetic study, because there is no point in looking at imperceptible differences even from a descriptive viewpoint, let alone a pedagogical one. The human ear is quite accurate in distinguishing sounds that have different fundamental frequencies, but at the same time, it is substantially less accurate in distinguishing component frequencies in sounds. There are so called critical bands in human hearing, which means that the ear measures the total amount of acoustic energy inside a certain bandwidth at a time (Suomi 1990: 180). The increments between the frequencies of harmonic partials are usually so small that several partials fit inside the critical band. This means that component frequencies that fit inside the critical band are heard as one. In addition, the width of the critical band increases along with frequency, which is one reason to not include higher individual formants in the analysis.

Zwicker (1961) divided the bandwidth of human hearing (20–20000Hz) into 24 critical bands. The unit he used for the bandwidth of one critical band is 1 Bark. The Bark scale is used in phonetic studies to evaluate the perceptive significance of differences between speech sounds. However, the critical bands are not fixed and they also overlap each other (Suomi 1990: 180). In other words, the bandwidth of human hearing is 24 Bark wide but the number of critical bands is greater. For example, between 0–6kHz, which is the most important frequency range for speech sounds, there are about

20 critical bands (Kuronen 2000: 42). An important implication of this is that if a formant shifts a distance that is less than 1 Bark, the difference is not audible, and therefore not significant. For example, if a language learner pronounces a target language vowel so that both F_1 and F_2 are less than 1 Bark away from the formants of a native speaker, the difference is inaudible and the production is therefore accurate. In turn, if either formant is more than 1 Bark away from native formant values, there is an audible difference between the vowels.

The latest advancement in this area of perceptual phonetics is the ERB_N scale, which also the one used in this study. In principle, it works similarly to the Bark scale (Zwicker 1961) but it has been found to represent human perception better (Moore 2010: 459, Iivonen 2012). The formula used for the conversion is from Glasberg & Moore (1990, see Section 3.4). Although it used to be a common view that the difference threshold between vowel formants is 1 Bark (1.3ERB), there does not seem to be a consensus anymore – not among the users of the ERB scale, at least. Iivonen (2012) argues that there is no absolute difference threshold for vowel formants but it can be approximated by examining minimal formant distances in languages with large vowel inventories. If quantity differences are taken into account, the average minimal distance between two vowels is 1.06ERB. If quantity is left out, the average minimal distance between two vowels is 1.4ERB. As English is a language with quantitatively different vowel pairs, 1.06ERB is going to be used as the limit for a just-noticeable-difference (JND), which is also Iivonen's (2012) suggestion. In other words, a vowel has to have a 1.06ERB difference in F_1 and/or F_2 when compared to another vowel for the two to be qualitatively distinguishable from each other to most listeners.

Although the study of formants is ubiquitous in acoustic vowel studies, some criticism against it has also surfaced. One common view is that describing the acoustics of vowels with formant frequencies alone is simplifying, and that the overall shape of the spectrum is a more accurate way to do it. However, the overall shapes of the vowel spectra are harder to compare with each other than formant frequencies. In addition, it has been found that two or more formants are heard as one if they fit inside 3–3.5 Bark (Kuronen 2000: 45). In some vowels, F_2–F_4 fit inside this range and are thus perceived as one formant. In this case, looking at F_2 alone would be invalid, especially if one aims to describe what a vowel sounds like.

However, F_1–F_4 can be used to calculate the effective second formant (F_2'), which gives a more accurate representation of how the vowel is perceived (Suomi 1990: 148–149). Unfortunately, there is no standard formula for calculating F_2', which is why it is often left out of the analysis altogether (Kuronen 2000: 44). Fant (1959) was the first to introduce a F_2' formula. However, it was rather simple and did not take F_4 into account, which yielded unreliable results. Bladon & Fant (1978) created a more sophisticated formula, which is the one used in this study (see Section 3.4). This formula, in turn, takes F_1–F_4 and their relative distances into account. After testing the formula with the cardinal vowels, they concluded that the formula predicts the measured F_2' accurately enough for all cardinal vowels except for /ʉ/, where the error is just above the difference limen i.e. the threshold of noticeability. There are also differing views on how the formants should be interpreted when describing the articulatory and perceptional features of vowels. For example, Aaltonen (1985) found that between Finnish /i/ and /y/, the contrastive feature is the distance between F_2 and F_3 rather than the frequency of F_2. Still, formant analysis remains the most common method in vowel studies. The reason for this might well be that formant analysis is relatively simple to carry out and it produces easily comparable results. Vowels cannot be described exhaustively by their F_1 and F_2, but nevertheless, they provide enough information to recognize and distinguish different vowels (Suomi 1990: 147).

2.1.3 Theoretical models of L2 pronunciation learning

Cross-linguistic influence can be defined as "the influence of a person's knowledge of one language on that person's knowledge or use of another language" (Jarvis & Pavlenko 2008: 1). Cross-linguistic influence affects all components of language, but it is discussed mainly from a phonological perspective in this section. The cross-linguistic influence that is most relevant for this study is L1 influence in L2 pronunciation, although L2 influence in L1 (also called L1 attrition) is also a known phenomenon. The most common problem that this influence causes is that phonemes that are not phonemically contrastive in the L1 are heard as one and the same phoneme (Jarvis & Pavlenko 2008: 63). The effect gets stronger if the vowels are phonetically close to each other. This is supported by both by Flege's (1995) Speech Learning Model (SLM) and Best's (1995) Perceptual Assimilation Model (PAM), which are going to be discussed next. In addition, there are other factors that contribute to the learnability of a vowel sound, which include phonotactics of both the L2 sound and its possible L1 counterpart, for example (Jarvis & Pavlenko 2008: 63).

The results of the present study are presented within the framework of Flege's (1995) **Speech Learning Model**. The basic assumption is that a foreign accent is caused by the learner's inability to perceive sounds accurately. The most important of Flege's postulates is that speech sounds are represented as phonetic categories in the human mind, and that even L1 phonetic categories are subject to change over a person's lifespan. He argues that a person comes attuned to perceive the contrastive sounds of his/her L1 (Flege 1995: 238). This means that if there is an L2 sound that is phonetically different from an L1 sound—but not in a way that is contrastive in the speaker's L1—the attunement reduces the speaker's ability to make the distinction between the sounds. Humans are able to understand even fairly disturbed speech, which can be seen as an example of this phonological conditioning. This is why L2 sounds that resemble L1 sounds assimilate to L1 phonological categories: otherwise we would not be able to understand even slightly foreign-accented speech.

For example, a Swedish learner of English might hear the English /ɜː/ as the Swedish /øː/ because 1) /ɜː/ is not a phoneme of Swedish and 2) /øː/ is an L1 phoneme that is phonetically quite near to the English /ɜː/. For these two sounds to be perceived and produced differently, separate phonetic categories must be established for both sounds in the learner's mind. This is possible only if the speaker can perceive at least some of the phonetic differences between the sounds. The likelihood of this happening depends on the sounds' phonetic similarity. The more similar the two sounds are, the less likely it is that the speaker establishes separate phonetic categories for both sounds. Last, Flege (1995) makes the hypothesis that the production of a vowel sound always reflects its phonetic category. This implies that accurate perception always precedes the accurate production of a sound.

Best's (1995) **Perceptual Assimilation Model** is another influential theory of L2 speech perception. It is mostly compatible with SLM, although it does not fully support Flege's (1995) assumption of phonetic categories. Best & Tyler (2007) argue that speech sounds are also represented in a more abstract phonological level in the human mind. Evidence for this is that although the English /r/ and the French /ʁ/ are phonetically very different from each other, English learners of French put them into the same phonological category of /r/ (Best & Tyler 2007: 28). This might also be the result of orthography, because the sounds are represented by the same grapheme in both languages. The phonetic categories of the SLM would then be subcategories for the phonological category.

Best (1995: 194) posits that non-native sounds can be perceived in three ways. The first option is that the sound is heard as a more or less acceptable exemplar of an L1 phonological category and is assimilated to it. This leads to inaccurate perception and production of the L2 sound, given that there is a phonetic difference between the L1 and L2 sounds. An L2 sound can also be recognized as a speech sound that does not fall into any L1 category. In this case, it is likely that the person establishes a new phonetic category for the sound, which in turn predicts accuracy in perception and production. Last, there is a possibility that the sound is not recognized as a speech sound at all. An example of this, from a Western point of view, could be the click consonants of African languages. However, this is a practically impossible scenario with the languages in the present study, and is thus left outside this study.

Best (1995: 195) also describes different kinds of assimilation between two L2 sounds. Distinguishing two L2 phonemes is expected to be poorest if both sounds fall into the same L1 category of which both are heard as equally good or bad exemplars (Single-Category Assimilation). If only one of them is seen as a less acceptable exemplar of the L1 category (Category-Goodness Difference), the chances of discrimination are slightly better. If both L2 sounds fall into different L1 categories (Two-Category Assimilation), discrimination is expected to be excellent. This is also the case if one sound is assimilated into an L1 category and the other one is not (Uncategorized vs. Categorized). Last, if both sounds are uncategorizable to an L1 category (Both Uncategorizable), their discrimination depends on their phonetic similarity. The SLM can be referred to here: the further away the sounds are from each other phonetically, the more likely it is that the speaker creates separate phonetic categories for the sounds and achieves accuracy.

In conclusion, both theories suggest that in order to produce an L2 sound accurately, it either has to be identical to an L1 sound or perceivably different from its closest L1 equivalent. In the first case, assimilation inherently does not cause accentedness. In other cases, the speaker needs to hear a difference and establish a new phonetic category for the sound, which the speaker might not able to do on his/her own. One factor that further complicates accurate perception of sounds is orthography, which is especially relevant for Finnish and Finland-Swedish learners. Both groups are expected to assimilate English long-short vowel pairs into their long-short L1 categories (i.e. pronounce them with similar quality but different duration), because

a) long and short allophones of the same vowel are qualitatively similar in both L1s
b) they are qualitatively not very different in English either
c) they are represented by the same grapheme in both L1s
d) English orthography, while usually having different spellings for the short and long counterparts (e.g. /iː/ is often spelled <ee> or <ea> while /ɪ/ is often spelled <i>) does not intuitively guide the learner towards the correct pronunciation.

Although orthography is not a matter of speech perception per se, Jarvis & Pavlenko (2008: 70) argue that its impact can be so strong that it overrides a person's ability to perceive speech, which is when it becomes a serious hindrance for learning pronunciation. This is especially relevant in a foreign language acquisition context, where the learner is not surrounded by the language in his/her everyday life and the language is learned in a formal context. This is the learning context for both groups of English learners in the present study, although its impact is somewhat lessened by the fact that English is prevalent in the lives of Finnish and Finland-Swedish youth through media and entertainment. However, English is not so common in Finland that it would resemble an ESL context. Last, learners also tend to pronounce L2 words with L1 sound-to-letter correspondences (Jarvis & Pavlenko: 70). This, however, is not an accent feature – it is not knowing the correct pronunciation and reverting to one's L1 grapheme-to-phoneme correspondences to figure out the pronunciation of a word.

2.2 Vowel inventories of the languages

In this section, I present the vowel inventories of the two varieties of English and the participants' native languages, Finnish and Finland-Swedish. The latter are presented mostly in relation to their differences to English.

2.3.1 General British English vowel system

The vowels of GB and example words are presented in Table 1. The vowels' formant frequencies can be found in Figure 3.

Table 1. The short and long vowels of General British English with example words.

ɪ	bit
ʊ	put
e	bet
ɒ	bot
æ	bat
ʌ	cut
ə	phonetics

iː	beat
uː	boot
ɜː	bird
ɔː	board
ɑː	bard

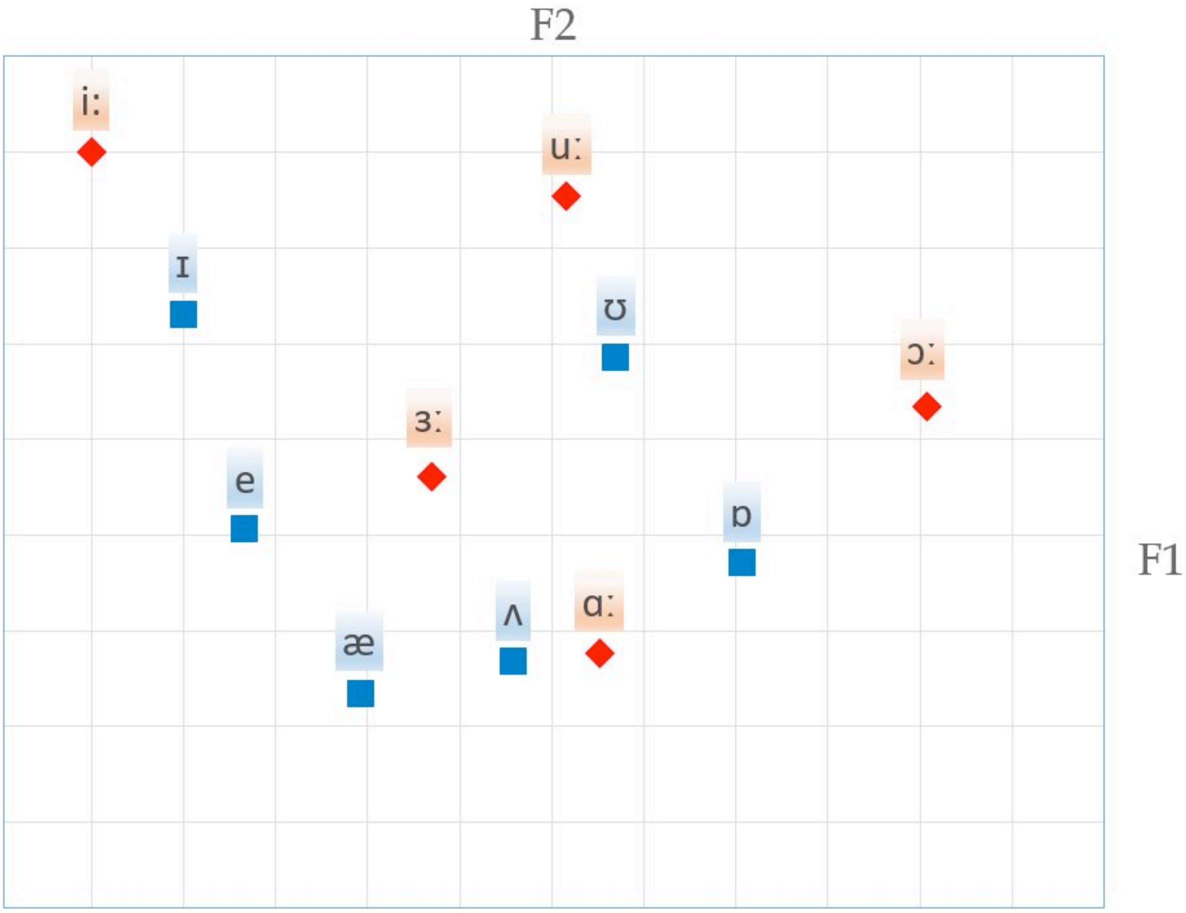

Figure 3. General British short and long vowels (normalized ERB, see 3.4 for details). Formant values taken from Cruttenden (2014: 104). The grid spacing is 1.06ERB.

General British (GB) has five long and seven short (relatively) pure vowels i.e. monophthongs. The long sounds are /iː/, /uː/, /ɜː/, /ɔː/ and /ɑː/ and the short sounds are /ɪ/, /ʊ/, /e/, /ɒ/, /æ/, /ʌ/ and /ə/. Although the English /æ/ is traditionally short and treated as such in the literature, it is often pronounced with significantly longer duration than the other short vowels, especially when preceding /b, d, g, dʒ, m, n/. It also behaves like a short vowel in English phonotactics, because it cannot occur word-finally, which is why it is considered short (Cruttenden 2014: 98). In addition, the long pure vowels /iː/ and /uː/ are often not entirely pure in modern English pronunciation (Cruttenden 2014: 112, 134). Both sounds are often pronounced with an upwards glide that starts slightly lower than the pure vowel. The diphthongized variations can be written /ɪi/ and /ʊu/ respectively, although there is a lot of variation in how the diphthongization is produced and transcribed. When producing English vowels, the position of the tongue is extremely important, because the position of the lips varies only slightly between vowels, if at all (Morris-Wilson 2003: 139). There are no vowels in English that require extreme rounding or spreading of the lips. Some vowels, such as /uː/, are sometimes even pronounced with virtually no lip-rounding (Cruttenden 2014: 134). All long-short vowel pairs, such as /iː/ and /ɪ/ are different in terms of both quality and quantity, hence the different symbols (Morris-Wilson 2003: 136). The symbol set used in this study is the one by A. C. Gimson (Cruttenden 2014: 104). It is particularly good for learners that are not accustomed to the quality differences between English long and short vowels, such as Finns. The long vowels are most often more peripheral than their short counterparts, i.e. farther from the centre of the mouth. For example, the tongue position in /iː/ is higher than it is in /ɪ/. This difference is partly necessitated by vowel duration: when pronouncing a long vowel, the speaker has more time to move their speech organs which allows for more peripheral vowels.

The duration of an English vowel is determined by its length and the sound that it precedes. The latter makes English vowel duration a complicated matter. Wiik (1965: 116) found that a vowel preceding a voiced consonant can be 64–100% longer than the same vowel preceding a voiceless consonant (e.g. *bead* and *beat*). Morris-Wilson (2003: 155) claims that this feature of vowel duration is crucial for understanding, because otherwise it would be harder to distinguish word-final voiced and voiceless consonants. In fact, he sees the vowel duration between *bead* and *beat* as the primary distinction. One reason for this is that utterance-final devoicing is a linguistic universal, which makes it hard to recognize the consonant based on its voicing, or the lack of it. Furthermore, utterance-final stops are often unreleased in informal English speech, which makes it impossible to use aspiration as a cue. Therefore, the only noticeable difference between *bead* and *beat* can be the duration of the vowel in some cases. Furthermore, Morris-Wilson (2003: 158) argues that mastering this is a prerequisite for learning a natural and flowing rhythm when speaking English.

English vowel distribution is determined by stress. Syllables with primary or secondary stress can include all vowels except for /ə/, which is a reduced vowel. In turn, unstressed syllables can only include the vowels /ə/ and /ɪ/, although there are some exceptions. For example, the English word *investigation* /ɪnˌves.tɪˈɡeɪ.ʃən/ has the primary stress on the fourth syllable and the secondary stress on the second syllable, both of which include so-called full vowels. However, the rest of the syllables only include the reduced vowels /ə/ and /ɪ/. The pronunciation of the reduced central vowel /ə/, which is also called schwa, varies greatly and is largely dependent on its phonetic environment (Morris-Wilson 2003: 141). Because the vowel only appears in unstressed syllables, its duration is often very short, which makes it more susceptible to coarticulation.

2.3.2 General American English vowel system and comparison with GB

The vowels of GA and example words are presented in Table 2. The vowels' formant frequencies can be found in Figure 4.

Table 2. The short and long vowels of General American English with example words.

ɪ	bit
ʊ	put
ɛ	bet
ɑ	bot, bard
æ	bat
ʌ	cut
ə	phonetics

iː	beat
uː	boot
ɔː	caught*

*not a part of the vowel inventory for all Americans, see below

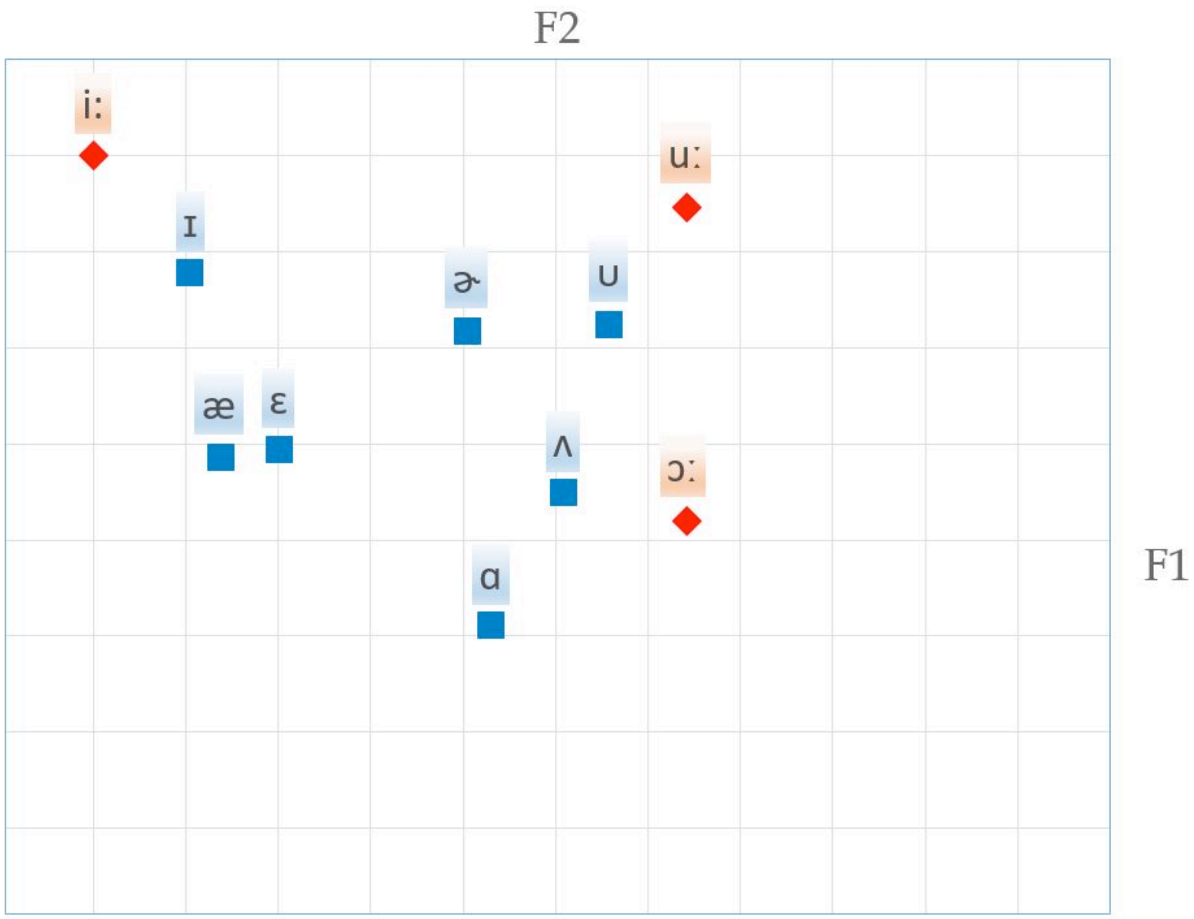

Figure 4. General American short and long vowels (normalized ERB). Formant values taken from Hillenbrand et al. (1995). The grid spacing is 1.06ERB.

General American (GA) vowel system is quite similar to the GB vowel system, which is why only differences relevant to this study are discussed in this section. Although there are of course countless small differences between these two varieties of English, the most relevant in this study are ones between the vowel qualities. The most noticeable difference is in the long vowels; more specifically, the lack of the mid-central vowel /ɜː/ and the long /ɑː/. This is caused by rhoticity. GB, or rather its precursor from the early modern era, lost it during the 17th century (Cruttenden 2014: 70). In rhotic accents, all r-sounds are pronounced, whereas in non-rhotic accents, only prevocalic /r/ is pronounced. In other words, an /r/ that is followed by a consonant, silence or a pause is pronouned only in rhotic accents (Cruttenden 2014: 87). For example, whereas the word *bird* is pronounced /bɜːd/ in GB, it is pronounced with a rhotic vowel in GA /bɚd/. Furthermore, the word *car* is pronounced /kɑː/ in GB and /kɑr/ in GA.

Interestingly enough, the GA /ɛ/ is indeed closer than the GB /e/, although the symbols tell the opposite. By the same token, the GA /æ/ is less than 1.06ERB away from /ɛ/, which means that it is significantly closer than the GB vowel. Hillenbrand et al. (1995) also noted this in their study, explaining that while there is significant F_1/F_2 overlap between the GA /ɛ/ and /æ/, they are still systematically identified in listening tests. This is because both vowels have different spectral change patterns. In other words, they are diphthongized to an extent, and their formants move in different directions. To put it very simply, /æ/ moves to a more open position (i.e. higher F_1) whereas /ɛ/ moves to a more central position (i.e. lower F_2).

On the other hand, some GA vowels are more open than their GB counterparts, namely /ɒ/ and /ɔː/. The most obvious difference is in the pronunciation of words like *cot*, where there is a mid vowel in GB /kɒt/ while the GA pronunciation has an open vowel /kɑt/. However, it is slightly inaccurate to say that the GA short "o-sound" is lower than its GB counterpart, because the two sounds are also represented differently in the orthography. The GB /ɒ/ is usually written with the grapheme <o>. In turn, the GA /ɑ/ is represented by <o> (e.g. *cot*) as well as <a> (e.g. *car*). One interpretation of this would be that the short "o-sound" of GA has become more open and lost its roundedness and thus become merged with the short /ɑ/. In addition to the vowel height differences, the vowel /uː/ is significantly further back and/or more rounded. As for the present study, the most important phonological difference apart from the vowel qualities is the cot-caught merger. Many Americans do not make a difference between the vowels in the words *cot* and *caught* but rather pronounce both words with the vowel /ɑ/. This merger is still in process, i.e. some American speakers still maintain a difference between the vowels in the words by pronouncing *caught* with the vowel /ɔː/. The difference is maintained in three areas: The Inland North, The Mid-Atlantic and The South (Labov, Ash & Boberg 2006: 59). Although not the only one merger of GA, this is the only one that could have any impact on the data. Despite the differences between GB and GA, their features that are expected to be difficult for

Finnish and Finland-Swedish learners are essentially the same, i.e. the quality difference between long and short allophones, the durational differences between vowels preceding voiceless and voiced consonants, and the effects of same-category assimilation (see 2.1.3).

2.3.3 Standard Finnish vowel system and comparison with English

The vowels of Finnish and example words are presented in Table 3. The vowels' formant frequencies can be found in Figure 5.

Table 3. Finnish short and long vowels with example words.

/i/	kivi, kiivi
/e/	elo, eepos
/æ/	käsi, kääpä
/y/	tyvi, tyyni
/ø/	tönö, kööri
/u/	tuli, tuuli
/o/	pomo, poolo
/ɑ/	kala, maa

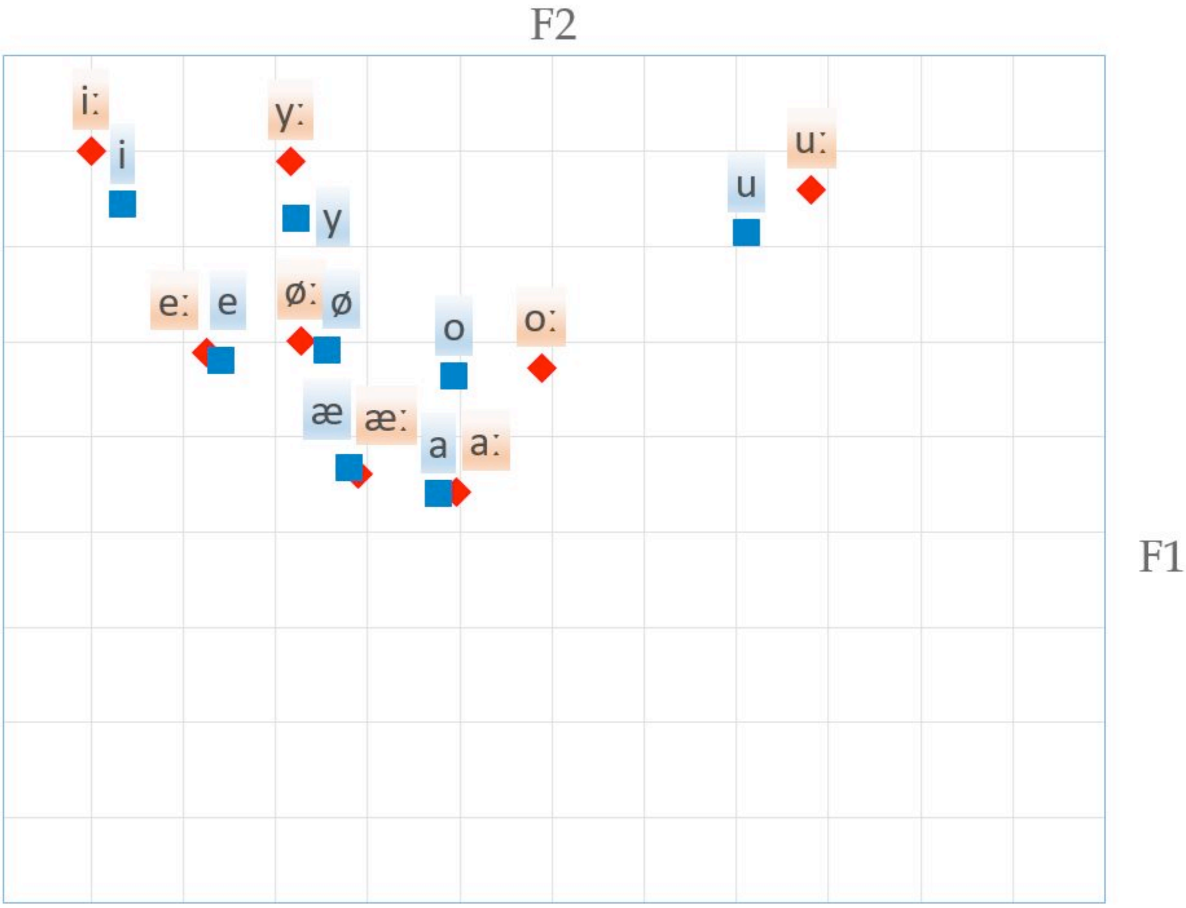

Figure 5. Finnish short and long vowels (normalized ERB). Formant values taken from Kuronen (2000: 166, 170). The grid spacing is 1.06ERB.

The Standard Finnish vowel system has eight long and eight short monophthongs. The sounds are /i/, /e/, /æ/, /y/, /ø/, /u/, /o/ and /ɑ/, all of which have qualitatively identical long and short counterparts. Finnish vowel qualities differ from IPA cardinal vowels by being less peripheral. For example, Finnish /ɑ/ is closer than the cardinal vowel with the same symbol and can be characterized as near-open (Suomi, Toivanen & Ylitalo 2008: 21). In addition, the mid series /e/, /ø/ and /o/ is not, like their symbols would suggest, close-mid but rather between close-mid and open-mid (Suomi, Toivanen & Ylitalo 2008: 20). The English vowel sounds that have at least nearly similar counterparts in Finnish are /iː/, /uː/, /e/, /ɔː/ and /æ/, which adds up to a total of 5 out of 12 vowels. In contrary to English, Finnish makes frequent use of rounding as a contrastive feature. For example, the only significant difference between /i/ and /y/ as well as /e/ and /ø/ is that the first one is unrounded and the second one is rounded. Furthermore, Finnish vowels can only be front or back, i.e. there are no central vowels. Vowel closeness also has only three steps: close, mid and open, although with the exception of the near-open /ɑ/ as mentioned before. In short, the qualitative differences in Finnish vowels are therefore clearer than the differences in English vowels that, in turn, use all three degrees of advancement and all four degrees of closeness.

Finnish has phonemic long-short pairs for every vowel, which is not the case in English. For example, the English /æ/ does not have a phonemically long counterpart. This is supported by the fact that the duration of /æ/ is very flexible in English, as pointed in the previous section. However, there are different interpretations, as some researchers seem to pair /ɑː/ and /æ/ and some others treat /ɑː/ and /ʌ/ as pair. However, it always leaves one phoneme without a long counterpart, because there are three a-like vowels in English. Furthermore, Finnish long and short vowels are thought to differ from each other in quantity only, whereas English long and short vowels are different in both quantity and quality. This is supported by the fact that contrastively long phonemes are usually interpreted as sequences of two identical phonemes (Suomi, Toivanen & Ylitalo 2008: 19). However, as the results of Wiik's (1965: 57) study suggest, this is not entirely true. In fact, Finnish short vowels tend to be more central (i.e. relaxed) than their phonemically long counterparts (Wiik 1965: 65). Nevertheless, native speakers perceive short and long Finnish vowels as qualitatively identical, and the centralization of short vowels can be seen as a result of shorter duration rather than an inherent characteristic (Suomi, Toivanen & Ylitalo 2008: 20), which is not the case in English vowels. The duration of Finnish vowels is determined mostly by their phonemic length, whereas in English, the following consonant can have a substantial effect on vowel duration (Wiik 1965: 150). Thus, vowel duration in Finnish is at the same time a simpler but more important feature. This is very likely to cause non-native pronunciation in the English of Finnish learners. Because Finns are used to make the distinction between long and short phonemes with duration differences, it can be assumed that Finns do not produce English long-short vowel pairs differently enough when it comes to vowel quality. However, because the distinction between the vowels has to be maintained (Jarvis & Pavlenko 2008: 65), Finns are likely to overdo the quantity differences. This will also conflict with the effect that the following sound has on a vowel's duration.

Finnish vowel distribution is determined by vowel harmony. Finnish vowels can be put into three categories: front vowels /æ/; /y/ and /ø/, back vowels /u/; /o/ and /a/ and neutral vowels /i/ and /e/. According to Finnish vowel harmony, front vowels and back vowels cannot occur in the same word. Neutral vowels can occur with both front and back vowels. This differs fundamentally from the stress-based vowel distribution of English (see previous section). The effect of vowel harmony was diminishing already in the 1960s because of the influx of loanwords (Wiik 1965: 50) and therefore it is reasonable to assume that English words that contradict Finnish vowel harmony are not a major problem for Finns, especially in these modern times. Instead, adopting the stress-based vowel distribution of English can be problematic for Finns. In conversational English speech, most function words and words with low semantic content are very often unstressed and thus pronounced with reduced vowel quality (Morris-Wilson 2003: 195). The unstressed syllables in content words are also pronounced with reduced vowels in most cases.

Because vowel reduction is not a feature of Finnish and the reduced vowels /ɪ ə/ are not part of the Finnish vowel system, it can prove problematic when learning English pronunciation and cause the substitution of reduced vowels with full vowels. For example, Peltola, Lintunen & Tamminen (2014: 93) describe the English /ɪ/, which is used in unstressed syllables along with /ə/, as maximally difficult for Finnish students. This is because it is very likely to assimilate to the Finnish phonetic category /i/. Consequently, they found that first-year English students in Finnish universities often pronounce it as /i/. However, failure to produce the so-called weak forms of English words (cf. Morris-Wilson 2003: 196) cannot be solely accounted to not being able to perceive and produce the reduced vowels accurately. In order to use the weak forms accurately, the speaker needs to be able to produce a stress pattern for the sentence. Furthermore, in order to produce a native-like stress and intonation patterns, the speaker needs to be able to "think ahead"; the speaker needs to know what he or she is going to say. Last, if the speaker needs to stop and think about word choices and word order in the middle of a sentence, it practically prevents him from producing the prosodic patterns even if the speaker would know them. This suggests that there is a certain threshold in language proficiency that a speaker needs to pass in order to use native-like prosody. Word and grammar choices must be automatic to a certain degree before a speaker can concentrate on prosody in conversational speech.

2.3.4 Standard Finland-Swedish vowel system and comparison with English

The vowels of Finland-Swedish and example words are presented in Table 4. The vowels' formant frequencies can be found in Figure 6.

Table 4. Finland-Swedish short and long vowels with example words.

/i/	vitt, vit
/y/	bytt, by
/e/	bett, be
/ø/	rött, röd
/æ/	kärra, skära
/a/	back, bad
/o/	åtta, råka
/u/	oxe, hota
/ʉ/	hutta, hus

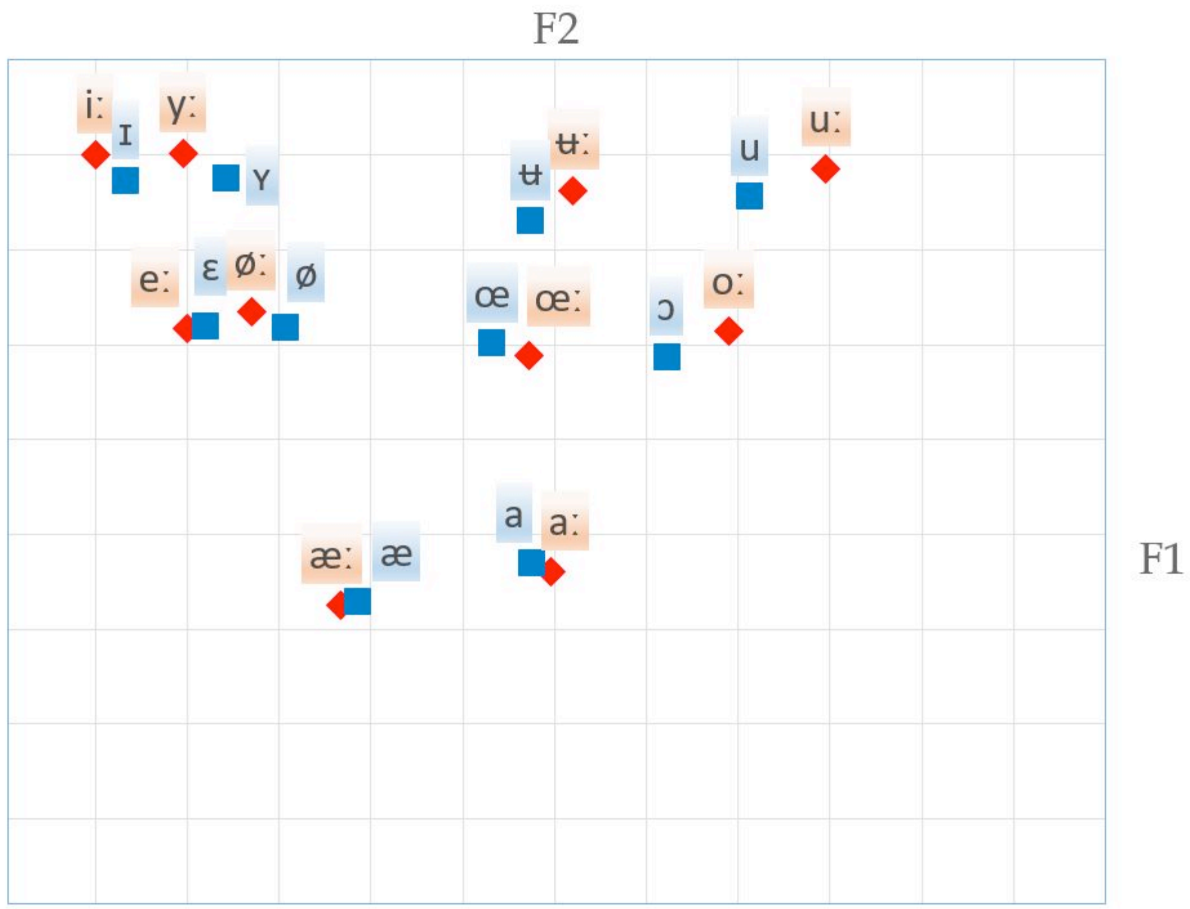

Figure 6. Finland-Swedish short and long vowels (normalized ERB). Values taken from Kuronen (2000: 140, 148). The grid spacing is 1.06ERB.

Standard Finland-Swedish has nine short and long monophthongs. The vowels are /i, y, e, ø, æ, a, o, u, ʉ/, all of which have qualitatively similar long and short counterparts. This is a result of the influence of Finnish, because Sweden-Swedish has some differences between long-short vowel pairs such as /ɑː/ and /a/. Reuter (1971: 246) found that the short mid vowels of Finland-Swedish /e, ø, o/ are most different to their long counterparts by being more central, which is also a feature of both languages discussed earlier in this study. However, a more recent study by Kuronen (2000: 147) found that there is a slight tendency towards centralization in all short vowels. The degree of centralization in short vowels correlates with vowel backness. Kuronen (2000: 147) argues that the different results might be caused by him using connected speech from a speech corpus as material, whereas Reuter (1971) used individual words and phrases in his study. However, none of the short vowels is over 1 Bark away from its long counterpart, which means that there is no significant differences between them. Contrary to English and Finnish, the vowel /ø/ has two allophones that are in complementary distribution. It can be pronounced as [ø] and [œ]: the first is the main allophone and the latter only appears when /ø/ precedes /r/. There is also a similar pattern with the vowel /e/, which is pronounced [æ] when preceding /r/. However, /æ/ is often treated as a separate phoneme in Swedish literature, regardless of the complementary distribution that points toward allophony.

Swedish vowel distribution is determined by stress. Vowel quality is not restricted by stress, but long monophthongs can only occur in syllables with primary stress. Short vowels can occur in stressed and unstressed syllables, but unstressed syllables always have a short vowel. This resembles English vowel distribution, although English unstressed syllables have vowels that are reduced in both quantity and quality. Furthermore, vowel length in stressed syllables is determined by the sound following the vowel: if the syllable ends in a long consonant, the vowel sound is short (e.g. *vitt*) and if the syllable ends with a short consonant or there is no final consonant (e.g. *vit* and *vi*), the vowel is long. This phenomenon is called complementary length in Swedish-language literature (Riad 2014: 10).

For the time being, there are no contrastive phonetic studies between English and Finland-Swedish. Even comparisons of English and Sweden-Swedish are scarce, which makes it difficult to make valid hypotheses about Finland-Swedish speakers pronouncing English. Davidsen-Nielsen & Harder (2001) discuss the usual problems that speakers of Scandinavian languages face when learning English pronunciation in a book aimed for EFL teachers. However, it is problematic to treat all Scandinavian languages as a whole because there are clear differences even between Finland-Swedish and Swedish vowels, let alone between Finland-Swedish and Danish vowels. Davidsen-Nielsen & Harder (2001) make no distinction between Swedish spoken in Sweden and Finland. However, the assumed difficulties seem quite relevant from a Finland-Swedish point of view, although the varieties are quite different.

I will discuss the relevant difficulties suggested by Davidsen-Nielsen & Harder (2001: 22) with regard to Reuter's (1971) and Kuronen's (2000) descriptions of Finland-Swedish vowels. First, Finland-Swedish speakers are expected pronounce the English /ɪ/ as more tense than native speakers, and its quality is not dissimilar to /iː/. The reason for this is most likely the same as with Finnish: in both Finnish and Finland-Swedish, short and long /i/ is pronounced with similar quality and the sole distinctive feature is quantity. Second, the short /ʊ/ is expected to be too close and clearly rounded. This is probably caused by the sound assimilating into the Finland-Swedish vowel category /ʉ/. Third, the vowel /ɜː/ is expected to be pronounced further front and more rounded. The vowel /ɜː/ probably assimilates into the L1 category of /ø/, which would cause such as difference to L1 pronunciation. Last, the vowel /ə/ is not sufficiently reduced, which is caused by the lack of vowel reduction in Swedish. In conclusion, the differences in vowel qualities seem to be that short vowels are more peripheral than native speakers' productions that the neutral vowel /ə/ is replaced with full vowels. These, in turn, are very similar to the differences between Finnish learners and native English speakers.

2.4 Previous research on English learners’ vowel pronunciation

Studies of learners producing English vowels often end up with results that are in line with cross-linguistic theories. The usual findings are that L2 vowels that have nearly similar L1 counterparts are pronounced similarly to their L1 counterparts, but L2 vowels that do not assimilate into L1 categories are pronounced more accurately. These findings support both Flege's (1995) and Best's (1995) models of speech learning and perception. There are no studies of Finnish or Finland-Swedish learners dedicated to their vowel pronunciation, but some information can be gathered from pronunciation teaching experiments. Immonen & Peltola (2018) studied the effects of an early-age language immersion program on vowel production. They studied children aged between 11-13 years. One group had been in an English language immersion program and the control group had studied in a regular Finnish school. They found greatest differences to native pronunciation in the English vowels /ɪ/, /ɒ/, /ɔ/ and /ɜː/, which are not parts of the Finnish vowel inventory. In addition, Peltola, Lintunen & Tamminen (2014) found that Finnish first-year English students pronounce /ɪ/ as more tense than native speakers.

An interesting point of reference from a Finnish point of view is Thai, which has a similar vowel system in the way that it has qualitatively identical long-short vowel pairs. A study by Pillai & Salaemae (2012) suggests that Thai learners of English produce long and short English vowels with similar quality, which is also one of the hypotheses about Finnish speakers in the present study. In addition, Sarmah et al. (2009) found that Thai speakers produced English long-short vowel pairs with greater duration differences than a native speaker. This is also a feature that Thai and Finnish learners of English should have in common, as Wiik (1965: 113) that the duration differences are greater in Finnish long-short vowel pairs. This is also in line with Jarvis & Pavlenko's (2008: 65) claim that learners strive to maintain contrastive differences between L2 vowels, sometimes even in a non-native manner.

However, some studies have had results that do not follow the usual learning patterns. An example of a study similar to the present study is the one by Hunter & Kebede (2012). They measured F_1 and F_2 values of English vowels produced by native speakers of Farsi using the same stimulus words as was used in the present study and in numerous other vowel studies. Although the F_2 of English /uː/ is over 2 Bark higher than in its Farsi counterpart, Farsi speakers still failed to establish a new phonetic category for the English sound and produced it near to the L1 equivalent. One possible reason for this is limited exposure to native English speech. In addition, a study of English vowels as produced by Turkish-English bilinguals found no difference between vowels that have or do not have an L1 category to assimilate into; significant differences to native pronunciation were found in both types of vowels (Ng, Chen & Sadaka 2008).

3 THE PRESENT STUDY

3.1 Research questions and hypotheses

The problem that the present study aims to solve is that we do not know what features of English vowel pronunciation the Finnish and Finland-Swedish pupils in our schools have learned and what they still need to work on, and whether there are differences between the two groups. The aim is to solve this problem by recording, analyzing and describing both groups' English vowel pronunciation. With the knowledge gained from the present study, we will know where pronunciation learning has been effective and what areas of pronunciation need more attention. In addition, comparisons between Finnish and Finland-Swedish learners' pronunciation can be drawn, which may reveal interesting phenomena and starting points for future research. The important question about possible differences between Finnish and Finland-Swedish learners' productions, which cannot be addressed in the scope of this study, is what causes the pronunciation differences between these two groups that live in the same country and go through the same educational system. Essentially, this study is a descriptive phonetic study, although it has its pedagogical implications. In order to solve the research problem, the following research questions (RQs) will be answered:

1. Which vowels are produced with a significantly non-native quality?
2. Do the learners produce English vowel reduction accurately?
3. Do the learners produce the duration differences between long vowels preceding fortis and lenis consonants?
4. Do Finnish and Finland-Swedish learners of English produce English vowels differently?

All results will be presented in relation to hypotheses that are based on the theoretical background and previous research presented in the previous section. Both Best's (1995) and Flege's (1995) models of L2 pronunciation learning predict this kind of results, although H5 and H6 may be outside these models. H5 concerns a special prosodic feature of English rather than a segment and H6 concerns a contrast between the vowel systems of English and the L1s, namely that Finnish and Finland-Swedish use duration only to distinguish long-short vowel pairs. The hypotheses are as follows:

1. English vowels that are close or identical to L1 vowels will assimilate into the L1 phonetic categories.
2. English vowels that are not close to any L1 vowels will not be assimilated. Instead, new phonetic categories will be established, and the vowel is produced at least somewhat natively.
3. Both groups will use rounding exaggeratedly.
4. Both groups will use full vowels instead of reduced vowels in unstressed syllables.
5. Neither group will produce the durational differences between vowels preceding voiceless and voiced consonants accurately.
6. Both groups will differentiate long and short English phonemes with duration, not quality.

Of these hypotheses, the first two are the most fundamental. Both Flege (1995: 239) and Best (1995: 194) argue that L2 sounds either assimilate into an L1 category or fall in between L1 categories in the phonetic space, and the similarity between the L1 sound and the L2 sound dictates which process takes place. Hypotheses 3 and 4 are expected to result from assimilating L2 sounds into L1 categories. First, the rounded sounds of English are often close enough to Finnish and Finland-Swedish sounds but less rounded, which is the reason for H3. The English reduced vowels /ə ɪ/ are also quite similar to full vowels /i ø/ of Finnish and Finland-Swedish, which again should result into assimilation to L1 categories.

The hypotheses ended up being very similar for the two L1s because of the similarity between Finnish and Finland-Swedish vowels. Although the languages themselves are fundamentally different and belong into different language families, Finnish has influenced Finland-Swedish over time to the extent that both languages have quite similar vowel inventories (Reuter 1971: 240). It is, nevertheless, possible that there are some differences because of the large morphological and phonotactic differences between the languages. Either way, it is very interesting to see whether two phonologically similar but otherwise very different L1s lead into similar L2 pronunciation or not, since L2 pronunciation has not been studied previously in this kind of setting. Also, the English pronunciation of Finland-Swedish learners has not been studied before, and this study provides preliminary information on that as well.

3.2 Participants

The RQs will be answered by analyzing recorded speech from both Finnish (n = 5) and Finland-Swedish (n = 4) ninth-graders. All participants were between 14 and 15 years of age, and it is safe to assume that all had reached puberty. However, the voice change that comes with puberty was not complete for all participants, which can be seen as a shortcoming; for example, some male participants' F_0 was closer to the reference value of adult women. Additionally, all participants had to have an 8 or higher as their latest grade in English (8 meaning "good" on a scale of 4–10), and their target variety varied between British and American English. Both male and female participants were recorded and analyzed. However, the results are listed separately for men and women, because men and women have structurally different speech organs. Men have longer vocal cords as well vocal tracts, which is why the fundamental frequency and formant frequencies, respectively, are lower in male speech. Especially the latter makes it largely irrelevant to compare male and female formant frequencies. The participants' formant values are referred to male and female reference values accordingly. It was in the author's consideration to choose whether the speaker was compared to General American or General British reference values; this decision was ultimately made by listening to the speech sample and determining which variety the speech resembled most. However, many participants did not clearly favor one variety over another, which is why their values were often compared to both varieties.

The Finnish participants (three male, two female) were pupils of a school in Central Finland. All Finnish participants spoke only Finnish as their native language and home language; participants with any other language background were dismissed. The Finland-Swedish participants (two male, two female) were pupils of a school in

Uusimaa. The Finland-Swedish participants were chosen from Uusimaa because they were expected to speak a standard variety of Finland-Swedish. In comparison, Finland-Swedish varieties from the Western coast of Finland are at least somewhat intelligible to other speakers of Swedish but substantially different from the standard variety of Finland-Swedish that is spoken in Uusimaa. Many Finland-Swedish participants spoke both Finnish and Finland-Swedish as their native languages, because Finland-Swedish people from Uusimaa are very often bilingual to some extent. This is why it was impossible to limit the Finland-Swedish participants to monolingual speakers, as was done with Finnish-speaking participants. However, only participants who spoke Finland-Swedish as their home language could participate in the study, others were dismissed.

All data was gathered with the permission of the participants themselves, as well as their parents and the schools they went to. All parties involved were informed about the use of the participants' personal information with a privacy notice. The data collection was organized in collaboration with project "Intelligibility, comprehensibility and accentedness of English spoken by Finns" (ICASEF, Academy of Finland grant number 315980), for which the author also collected data. Because the participants are recognizable through their voices on the recordings, all data was stored securely and handled with discretion, as per GDPR regulations and the privacy policy of the University of Jyväskylä.

3.3 Data and procedure

In the recording procedure, the participants were given a list of stimulus words and passages (hereafter simply "test") that they were asked to read aloud. The words were presented one at a time from a computer screen. This was to help the participants concentrate on the task at hand and also to reduce unwanted noise from handling a paper, flash cards etc. The participants were told to pronounce the words with a clear but natural voice and articulation. It was also emphasized that their pronunciation was not going to be evaluated, but rather analyzed for descriptive purposes, to prevent anxiety.

The recording took place in September 2019 for the Finnish participants and October 2019 for the Finland-Swedish participants. The speech was recorded digitally with a Røde NT-USB condenser microphone which was set at approximately 15cm away from the speaker, which is the reference distance given in the product manual. The microphone has an integrated 48kHz/16bit audio interface and a frequency range of 20–20000Hz. An acoustically transparent pop filter was fitted to the microphone to prevent loud transients (such as plosives) from overloading the microphone capsule. The sound was recorded in Audacity 2.3.2 (The Audacity Team 2019) at a sample rate of 44.1kHz/16bit. No noise reduction was made to ensure that the sounds were not altered. The sound was then cropped and exported into lossless .wav files at the same sample rate, which were the files used in the analysis. The recording equipment itself produced nearly no noise, but some ambient noise from the air conditioning and a construction site next to the building could not be prevented. The empty classrooms where the material was recorded were also somewhat reverberant, which is why some foam rubber was placed behind the microphone to reduce reverb in the recording. In the end, neither ambient noise or reverberations caused difficulties in the analysis due to their relative weakness when compared to the speech.

Švec & Granqvist (2010) have proposed a set of guidelines for selecting microphones for speech analysis. According to them, a microphone should have a) a dynamic range and a frequency range that exceed those of human speech b) a flat frequency response c) an omnidirectional polar pattern. The microphone used in this study only fulfills the first requirement, while there is a 7dB boost at 5500Hz and the microphone has a cardioid (directional) polar pattern. The latter of these lends itself to the proximity effect, where the low frequencies are boosted when the sound source is close to the microphone and suppressed if far from the source (Švec & Granqvist: 2010). In theory, this could compromise the reliability of the analysis. However, the uneven frequency response cannot shift the formants; it only makes them seem stronger than they are. The proximity effect was probably averted by adhering to the reference distance of the microphone given in the product manual. When also considering that the methods of analysis in the present study are basic and previous studies have been successful with far inferior recording equipment (see e.g. Gonzales 2004), I am positive that my results are valid.

There were three sections in the test which were dedicated to RQs 1–3 respectively. To avoid problems with creaky phonation, the stimuli were designed so that the target vowels were most often utterance-medial. The first section included the short and long vowels of English (RQ1). The stimulus words were initially as follows: *heed, hid, head, had, hard, hod, hoard, hood, who'd, herd* and *hud.* These were chosen because of two reasons. First, these words have all GB vowels in an identical phonetic environment [hVd]. They also do not include nonwords, which would have caused problems with non-native speakers. Second, the same set of stimuli has been used in numerous other studies (Wells 1962, Deterding 1990, Hawkins & Midgley 2005, Immonen & Peltola 2018). Most importantly, the same set of stimuli was used in the study of GB vowels to which my results were to be compared (i.e. Deterding 1990). In addition, a slightly modified set (*hawed* instead of *hoard*) was used in Hillenbrand et al. (1995), which provided GA reference values for the present study.

One shortcoming with these words is that some of them are probably not familiar to an intermediate L2 speaker, which is why the author provided rhyming words, or if not sufficient, an example of the word's pronunciation to the participant if they did not know the pronunciation of the word. This was not seen as a problem in reliability, because it was deemed highly unlikely that a L2 learner who normally speaks with an accent would immediately shift into native-like pronunciation when provided with an example. Also, the author only gave instructions on pronunciation only if the participant hesitated for a long time or mispronounced the word completely. In addition, the word *hood* proved to be problematic for gathering data from the vowel /ʊ/ because it is a very common mispronunciation to pronounce it with the vowel /uː/. This may be the result of two things: first, the spelling suggests that the vowel is long, and second, the word has been adopted into Finnish as a direct loan *huudi* /huːdi/, which also reinforces the mispronunciation. This is why the word *put* was added to the list of stimulus words before recording the Finland-Swedish participants. Also, the word *hawed* was added to the stimuli to get a non-rhotic pronunciation of the vowel /ɔː/. This was due to the fact that many Finnish participants pronounced *hoard* with clear rhoticity—even though they had previously stated that their pronunciation is tilted towards British English. The word *hawed* is the only one in the test that is subject to the cot-caught merger (see 2.3.2).

The second part of the test was dedicated to RQ2, and it included words and phrases that include the reduced vowels /ə/ and /ɪ/ in unstressed syllables. Having the target vowels utterance-medially was especially important in this section, because unstressed and reduced vowels are exceptionally susceptible to creaky phonation when utterance-initial or utterance-final. An example of such a phrase would be *get a grip*. The underlined vowels are reduced in normal native speech. The neutral vowel /ə/ was only included in this section because it cannot be studied in a similar environment as the other vowels as it can only occur in unstressed syllables. In addition, as the quality of /ə/ varies greatly depending on its immediate phonetic environment, it is not very reasonable to study its exact formant frequencies; it is only worthwhile to know if it is produced near to a full vowel (i.e. unreduced) or not (i.e. reduced).

The third and last part of the test had minimal pairs that differ only by their final consonant, such as *beat* and *bead*, which provided data for RQ3. The words were presented in randomized order. A native speaker would produce these words with different vowel duration. The stimuli were designed so that the vowel would be in between two obstruents (see example above) in order to ensure accurate measurement of duration. Rhoticity caused some problems in the stimuli for RQ3, because words such as *board* were pronounced rhotically (i.e. /bɔrd/). Consequently, *board* and *bought* were not minimal pairs and their vowel durations could not be compared. This is why some additional word pairs were added to the stimuli for Finland-Swedish participants, such as *seat/seed*.

The pronunciation test proved to be successful in for gathering speech samples for a study like this. Because the test was short and it included mostly easy-to-pronounce and common words, there was no major problems in the recording and the data it produced was easy to analyze. However, if I were to conduct this study again, I would include a short training session with the stimulus words before the recording to ensure that the participants know how to pronounce the words, such as in Hillenbrand et al (1995). The same pronunciation test could be used in a larger-scale study to provide quantitative data of pronunciation; the small sample size of the present study makes it impossible to make generalizations out of the results. More participants would have also enabled me to dismiss all participants whose voice change is still in progress from the analysis. Choosing ninth-graders as my participants involved taking a conscious risk: it was evident from the start the participants would be at different stages of maturing. Adult speakers, for example, would have been a more homogenous group to study.

3.4 Analysis

The recordings were analyzed in Praat (Boersma & Weenink 2019), which is a free software program intended for acoustic analysis of speech. The program is also widely used by phoneticians. Each speaker was analyzed separately. From Sections 1 and 2 of the pronunciation test, the formant values were taken from a point that was closest to the middle point of the *pure* section of the vowel where the formants were clearly visible. An example of this can be seen in Figure 7. In some words, it was clearly visible that the vowel started to glide towards the place of articulation of the following consonant /d/ before closure, and this glide was left out when approximating the middle point of the vowel. After finding an appropriate place to take the measurements, the formants were extracted by using Praat's standard Burg algorithm, which was set to give F_1–F_5 from the selected point in the sound. The maximum formant value was initially set to 5000Hz for male participants and 5500Hz for female participants. The maximum formant value determines the highest possible value for the last formant (in this case, F_5) and if it is too high or too low, the algorithm will not give correct values for the formants. With one male participant, the average fundamental frequency was found more similar the female participants (close to 200Hz) instead of other male participants (approximately 120Hz) which is why the maximum formant value was set to 5500Hz for his analysis. A typical average F_0 is

120Hz for adult men and 210Hz for adult women, although there are numerous other factors at play besides gender (Traunmüller & Eriksson: 1995).

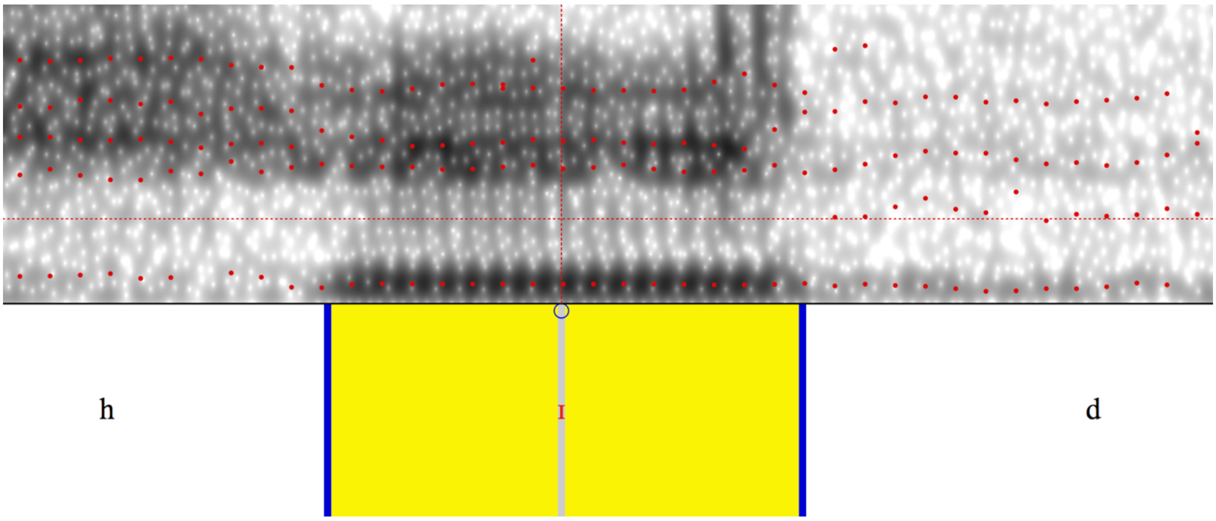

Figure 7. An example of extracting the formants from the vowel /ɪ/ in the word *hid*.

The formant values were listed in Excel, where ERB rates (see 2.1.2) and the effective second formant (F_2') were calculated for each vowel. The formula used for the ERB conversion was the one by Glasberg & Moore (1990):

$$\mathrm{ERB_N\ number} = 21.4\log_{10}(0.00437F + 1)$$

The formulae used for calculating F_2' were the ones by Bladon & Fant (1978). F_1–F_4 represent the formant values in Hz. B_2 was set at 67Hz as authors suggested (ibid.).

$$F_2{}' = \frac{F_2 + c^2(F_3F_4)^{1/2}}{1 + c^2}$$

$$c = K(f)\frac{A_{34}}{A_2}$$

$$\frac{A_{34}}{A_2} = \frac{B_2F_2\left(1 - F_1{}^2/F_2{}^2\right)\left(1 - F_2{}^2/F_3{}^2\right)\left(1 - F_3{}^2/F_4{}^2\right)}{(F_4 - F_3)^2\left(\frac{F_3F_4}{F_2{}^2} - 1\right)}$$

$$K(f) = 12\left(\frac{F_2}{1400}\right)$$

The values gathered from Section 1 were compared to each other and to General British or General American reference values from the literature depending on which variety the pronunciation resembled the most (Cruttenden 2014: 104, Hillenbrand et al 1995). Unfortunately, Cruttenden's (2014) book did not provide F_2' values for GB English nor anything to calculate them with, so it was impossible to compare my F_2' results into General British reference values, which is why the vowels are compared in the more traditional yet fundamentally inaccurate F_1/F_2 space (see 2.1.2 for details).

Both the reference values and the measured values were converted into ERB by the author using the formula presented above. ERB rates were normalized with Kuronen & Kautonen's (2018) method. The initial results were very varied because the participants were in varying stages of developing an adult voice. Therefore, the only way to compare the results to each other was through normalization. As a disclaimer, it must be noted that the method is still experimental and not widely used in phonetic research. Although experimental, the method is also innovative in its simplicity. Testing this new method was a secondary objective for the present study. This method uses the vowel /iː/ as a fixed reference point with an imaginary value and other vowels are placed into a F_1/F_2 ERB chart based on their distance from /iː/. This logic behind the chosen method is similar to Lobanov's (1971) acclaimed normalization method, which used a calculated center point of the the person's vowel constellation as the reference point rather than an actual vowel. Both methods shift the focus from formant frequencies to vowels' positions in relation to each other. This approach can be connected to Liljencrants & Lindblom's (1972) **vowel dispersion theory**. According to this, vowels are distributed at equidistant points in the vowel space. Disner (1984) found that an overwhelming majority of the world's languages seem to follow the theory. This supports my choice to concentrate on vowels' relative distances when the absolute formant frequencies proved to be incomparable. The Lobanov (1971) normalization method has been found the best in a thorough review of normalization methods (Adank, Smits & van Hout: 2004) and his choice of reference point seems more logical than using /iː/. However, data normalized with the Lobanov method cannot be used with psychoacoustic scales such as the ERB scale, which is the main reason why I chose Kuronen & Kautonen's (2018) more recent method instead. A normalization method that combines the accuracy of Lobanov's (1971) method and the possibility to use psychoacoustic scales would be ideal, but it has not yet been invented. The obvious drawback of using /iː/ as the reference point leaves /iː/ outside of the comparison. If the formant values for /iː/ would be radically different between the languages compared, the normalization would hide the difference, which would also compromise the results of the analysis. For example, if the participants' F_1 values for /iː/ would be significantly higher than those of the target language (i.e. the vowel would be more open), it would make all other vowels look more close than they actually are. However, as a typical /iː/ is quite similar between all languages in the study, this was not seen as a major concern.

The reference values chosen for this study were analyzed from words that are similar to the stimuli of the present study, and that were pronounced in citation form, which also matches the setting of this study. In connected speech, vowel productions tend to be less peripheral (Cruttenden 2014: 105). In Section 2, the formant values of the target vowels (i.e. the vowels that would be pronounced with reduced vowel quality in natural native speech) were compared to the formant values of the same speaker's productions of English full vowels to see if the speaker pronounced them as reduced.

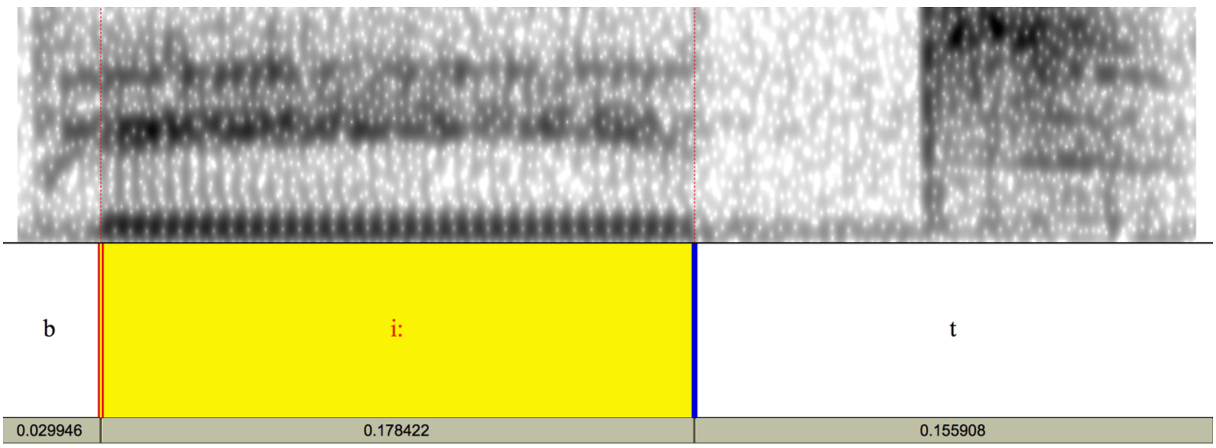

Figure 8. Measuring the duration of the vowel /iː/ in the word *beat*.

From Section 3, only the durations of the vowels were measured. The duration was measured by determining the onset and end of unobstructed voiced phonation (i.e. the vowel) by looking at the spectrogram and measuring the duration of the sound between the two points. An example of this can be seen in Figure 8. This left the release phase of the preceding consonant and the closure of the following consonant out of the vowel duration. After that, the relative durational differences between pre-voiceless and pre-voiced vowels were calculated and compared to Wiik's (1965) results.

4 RESULTS

The results of this study are presented in the following section. Overall, the vowel pronunciation of all participants resembled the target language more than their native language. All results are presented in a normalized ERB F_1/F_2 space. The reason for the normalization is twofold. First, many participants were so young that their formant frequencies were incomparable with the reference values from previous studies. For example, the author got F_2 values as high as 2800Hz for /iː/ from both male and female productions, which is 600Hz (approx. 2 ERB) higher than adult reference values. Immonen & Peltola (2018), who studied the English vowels of slightly younger Finnish children, found similar results. On the other hand, some male participants, whose voice had already changed into an adult man's voice, yielded formant frequencies that were comparable with adult values. Because of this, my results were incomparable to native values and to each other before normalization, which seemed to work well for my data. It neutralized the structural and biological differences between speakers, but it also preserved the differences caused by different linguistic background. This can be seen especially well in the reduced vowels: the large differences are between L1s, not between genders. Results concerning RQ1 and RQ2 (qualities of full and reduced vowels respectively) are presented in Sections 4.1 and 4.2, while results for RQ3 (durations for vowels preceding voiced and voiceless consonants) are presented in Section 4.3. Last, the differences between Finnish and Finland-Swedish participants' productions (RQ4) are covered in Section 4.4.

Because there were no F_2' values for GB vowels, the effective second formant was left out of the comparison altogether. However, some observations about the difference between F_2 and F_2' in the present study can be made. It seems that front vowels are the only ones that are significantly affected by higher formants. As can be observed from Tables 5–12, there is hardly any difference between F_2 and F_2' apart from vowels /iː ɪ e/. The explanation for this is that front vowels have a high F_2 and F_3/F_4 are quite stable across different vowels. Therefore, only front vowels have F_2–F_4 so close together that they subject themselves to formant integration (i.e. the formants fit inside 3.5 Bark), which then drags F_2' higher.

4.1 Finnish speakers' vowel quality

Overall, Finnish male participants' (hereafter FIM) English vowel constellation (Fig. 9) bears clear resemblance to GB. However, some long-short vowel pairs are nearer to each other than in the target language, as was expected.

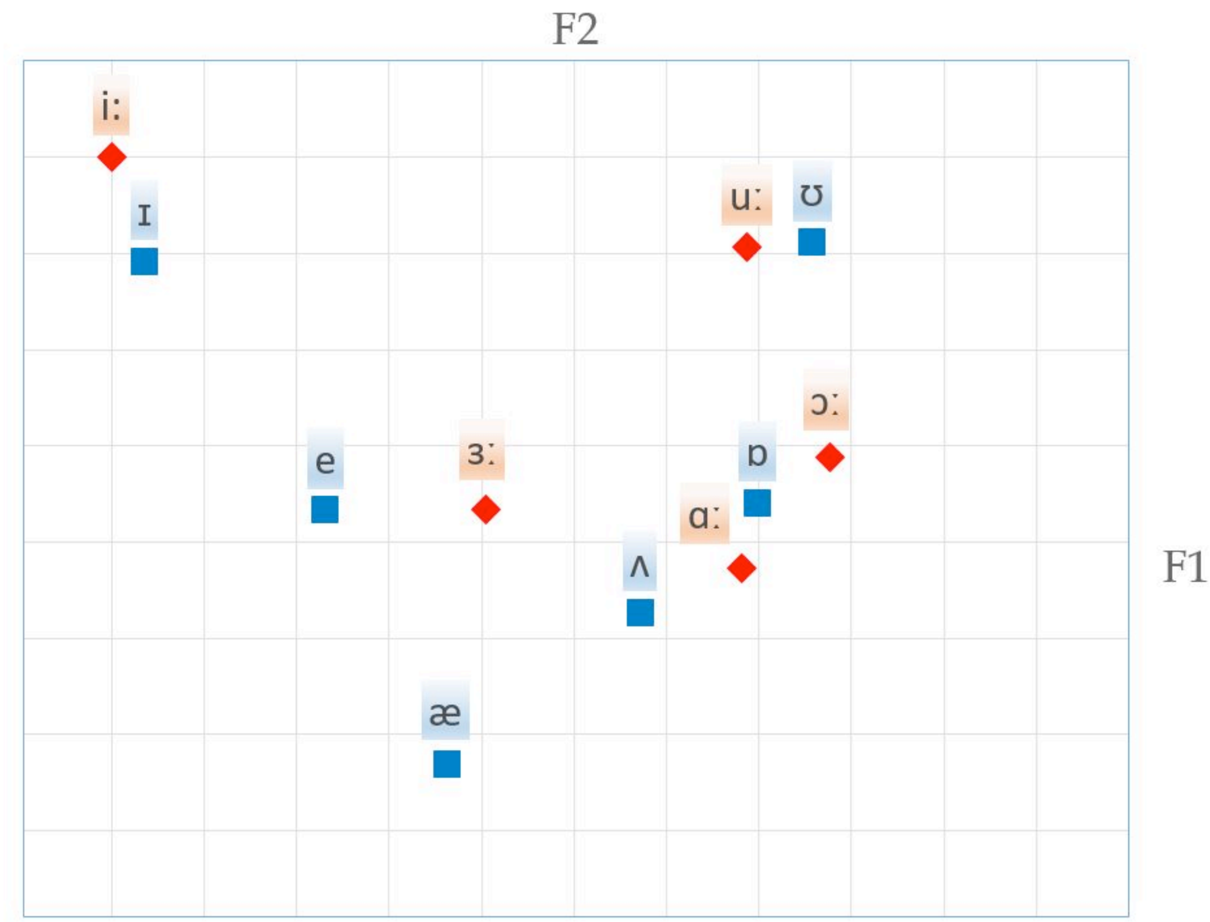

Figure 9. English vowels of Finnish male speakers (normalized ERB). The grid spacing is 1.06ERB.

For example, /ɪ/ is about twice as far from /iː/ in GB when compared to FIM in both vowel height (F_1) and vowel advancedness (F_2). In addition, whereas the difference is over 1.06ERB (i.e. the just-noticeable-difference, JND) in both height and and advancedness in GB and GA, the FIM vowels are noticeably different only in vowel height. In addition, the vowel pairs /uː ʊ/ and /ɔː ɒ/ are both perceptually similar in quality, which is not the case in either varieties of the target language. The first pair is located somewhat in between GA and Finnish /uː/, being less than 1.06ERB away from both. In this case, it is likely that the English /uː/ has assimilated to the L1 vowel category. The stimulus for /ʊ/ was somewhat misleading (see 3.3) for the Finnish participants, so I can only speculate whether the similarity of /uː/ and /ʊ/ is caused by same-category assimilation or the fact that the participants thought that the stimulus word was to be pronounced with /uː/. In addition, it was unexpected that that /ʊ/ was slightly more peripheral than /uː/, although not noticeably so. It is, after all, a linguistic universal that long allophones of the same vowel are more peripheral than short ones. When it comes to the vowels /ɔː ɒ/, their similarity is most probably caused by same-category assimilation. The short /ɒ/ was within 1.06ERB of its GB counterpart whereas the long /ɔː/ was not pronounced similarly to GB or GA. The vowels /ɑː ʌ/ were within 1.06ERB in both GB and FIM, but the vowel pair was noticeably more advanced and slightly more open in FIM than GB. Last, /e æ ɜː/ were all within 1.06ERB of GB values. Other noteworthy features are the overlap between the vowel pairs /ɔː ɒ/ and /ɑː ʌ/ as well as the fact that there was no sign of the cot-caught merger (see 2.3.2) in any of the FIM participants. The a's and o's as a whole are somewhat different to both target varieties. The single feature that stands out the most is the qualitative similarity of long-short vowel pairs.

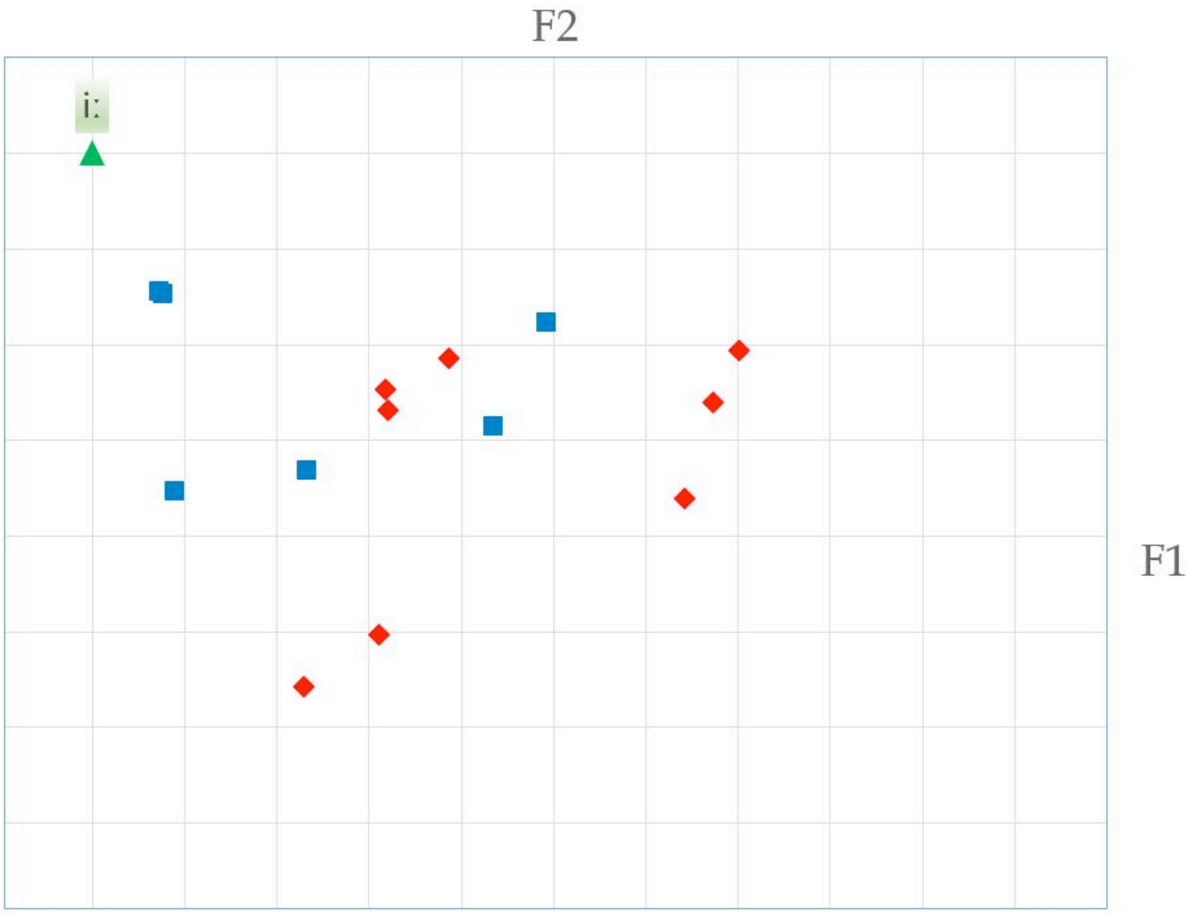

Figure 10. English reduced vowels of Finnish male speakers (normalized ERB). Blue markers are vowels that can be pronounced as either /ɪ/ or /ə/, red markers are vowels that can only be pronounced as /ə/. The grid spacing is 1.06ERB.

The reduced vowels of FIM speakers (Fig. 10) do not seem very target-like at first glance, as almost full range of both F_1 and F_2 is in use. Some vowels are clearly pronounced as /æ/, /e/ or even further front than the average /e/. There are probably numerous factors behind this. First, the student is probably not certain on the pronunciation of the word. Second, the orthography points towards /e/ or /æ/ — especially for a Finnish learner, who is accustomed to transparent orthography through their L1 (Suomi, Toivanen & Ylitalo 2008: 37). Last, the learner is probably not familiar with the concept of vowel reduction. For example, the vowel /ɪ/ is not very different to Finnish /e/, which is why pronouncing *example* as /egˈzɑːmpl̩/ or /egˈzæmpl̩/ does not sound or feel "wrong" to a Finnish speaker. However, there is an explanation for some seemingly unreduced pronunciations. It is widely known that the realizations of /ə/ are very varied (see Cruttenden 2014: 138). For example, the cluster of three back vowels were all preceding the bilabial approximant /w/, which unavoidably introduces roundedness into the vowel, which in turn leads into lower F_2. This, along with the author's auditory evaluation of the speech samples, suggests that FIM vowel reduction is not target-like, as there are numerous unreduced productions.

Table 5. The average formant values of English vowels produced by Finnish male speakers (Hz/ERB).

FI M S	ɪ	e	æ	ɒ	ʊ	ʌ
F1 (Hz)	358	559	836	552	345	661
F2	2368	1851	1561	991	912	1181
F2′	2734	1938	1601	997	916	1191
F3	2956	2720	2592	2742	2421	2617
F4	3908	3927	3764	3607	3513	3689
F1 (ERB)	8.76	11.49	14.29	11.41	8.54	12.62
F2	22.58	20.51	19.12	15.55	14.93	16.90
F2′	23.80	20.89	19.32	15.60	14.97	16.96
F3	24.47	23.76	23.34	23.83	22.76	23.43
F4	26.90	26.95	26.57	26.20	25.97	26.40

Table 6. The average formant values of English long vowels produced by Finnish male speakers (Hz/ERB).

FI M L	iː	ɑː	ɔː	uː	ɜː
F1 (Hz)	290	616	511	349	559
F2	2477	1015	886	1007	1477
F2′	3184	1018	887	1018	1506
F3	3185	2575	2672	2358	2533
F4	3937	3622	3702	3488	3686
F1 (ERB)	7.61	12.14	10.91	8.60	11.49
F2	22.96	15.73	14.72	15.68	18.67
F2′	25.11	15.76	14.73	15.75	18.82
F3	25.12	23.29	23.60	22.54	23.15

F4	26.97	26.24	26.43	25.91	26.39

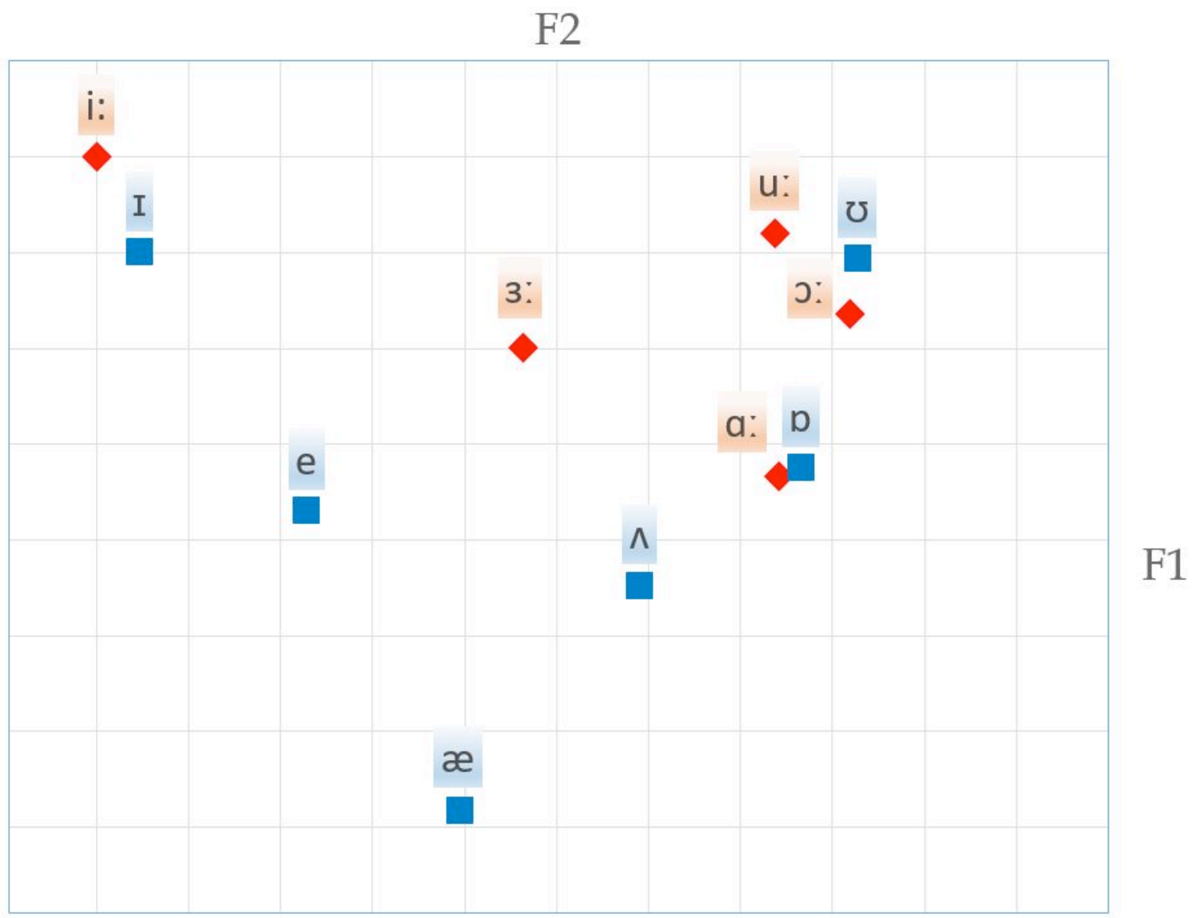

Figure 11. English vowels of Finnish female speakers (normalized ERB). The grid spacing is 1.06ERB.

Finnish female participants' (hereafter FIF) vowel constellation (Fig. 11) is also quite similar to GB but there are more signs of American influence. The same phenomenon as with FIM can be observed in long-short vowel pairs, although to a lesser extent: the vowel pairs /iː ɪ/ and /uː ʊ/ are barely distinguishable whereas /ɔː ɒ/ and /ɑː ʌ/ are clearly over the JND limit. Although /ɪ/ is within 1.06ERB from GA and GB, it is more important to look at its distance from /iː/ than its position. The vowel /uː/ is similar to GA, as was the case with FIM. However, /ʊ/ is significantly more peripheral than it should be, being even more rounded and/or back than /uː/ and quite different from GB and GA. The mid vowel /ɜː/ is pronounced similarly to GA, which is most likely caused by rhoticity. However, there is reservations about the stimulus *hood*, as explained in the previous paragraph. The front series /ɪ e æ/ are pronounced by similarly by FIM, FIF and GB speakers. The vowels /ɔː ɒ/ are both produced closer than GB and GA. The same applies to /ɑː/ and /ʌ/. As a result, the o's are too near to the u's: the vowels /ɔː/ and /uː/ are nearly indistinguishable. In addition, /ɑː/ and /ɒ/ are practically identical, which is an American feature. The reason for this is not clear: if the u's would also be questionably close, it would signify that the analysis of back vowels is somehow distorted. One factor that affects this the small sample size of the study. In addition, the close /ɔː/ could be explained by rhoticity, because in Finland-Swedish participants' results, /ɔː/ was closer in the word *hoard* than in the word *hawed*. The cot-caught merger could also explain the difference between the aforementioned vowels, but there was only one participant in the entire study (a Finland-Swedish female) that had the merger. What also argues against the cot-caught merger is that it was indeed the a's that moved towards the o's and not the other way around.

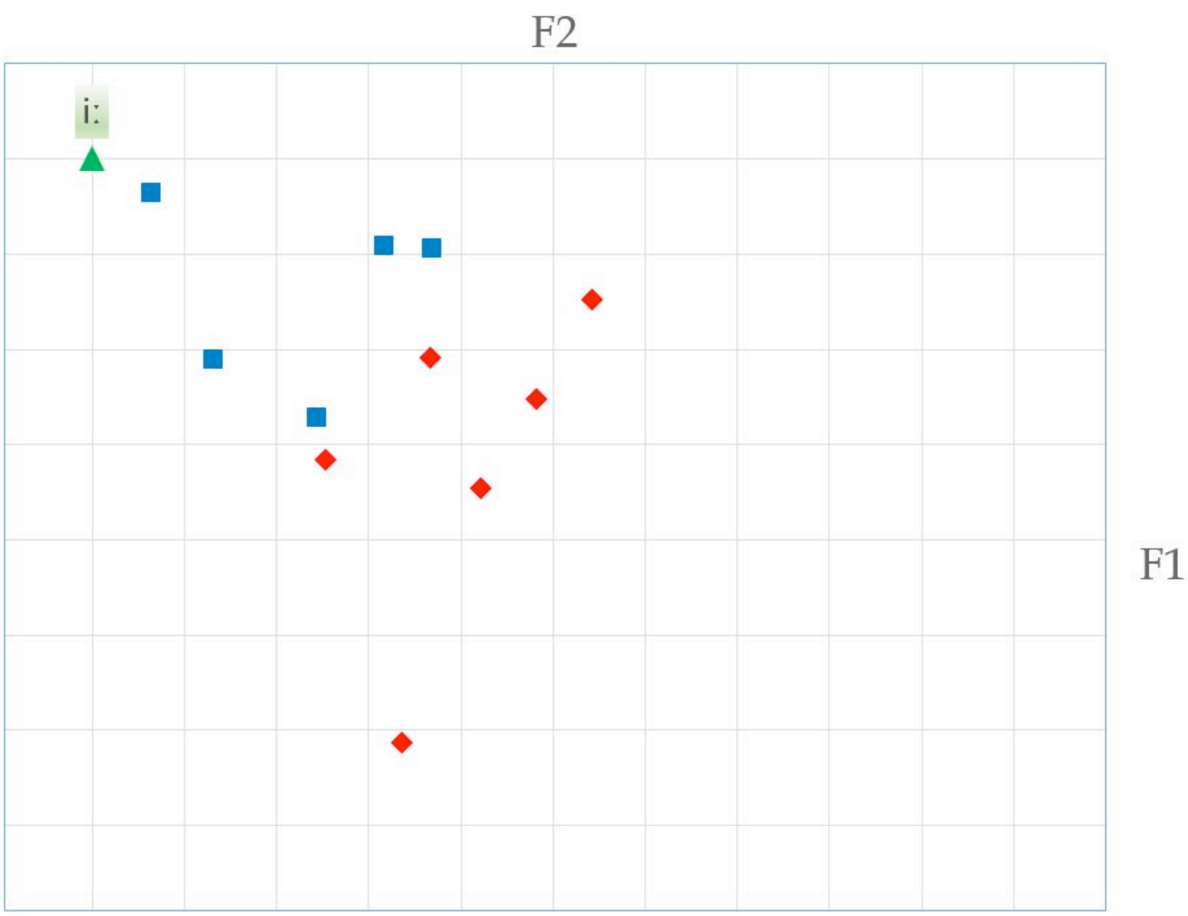

Figure 12. English reduced vowels of Finnish female speakers (normalized ERB). Blue markers are vowels that can be pronounced as either /ɪ/ or /ə/, red markers are vowels that can only be pronounced as /ə/. The grid spacing is 1.06ERB.

The reduced vowels of FIF speakers (Fig. 12) seem to be generally less spread out than those of FIM speakers; the sounds are less varied especially F_2-wise. However, the extremes are even more extreme, as there is a pure /æ/ and /i/ in the results. These two vowels were gathered from the stimuli *get <u>a</u> grip* and *<u>re</u>spect* respectively. There is also clearer division between /ɪ/ and /ə/ than in FIM results.

Table 7. The average formant frequencies of English short vowels produced by Finnish female speakers (Hz/ERB).

	ɪ	e	æ	ɒ	ʊ	ʌ
F1 (Hz)	414	645	1020	602	419	727
F2	2686	2144	1733	1056	969	1340
F2'	3112	2256	1743	1057	970	1343
F3	3213	3006	2867	2721	2431	2482
F4	4180	4295	4427	4028	3891	3974
F1 (ERB)	9.60	12.45	15.77	11.98	9.66	13.28
F2	23.65	21.74	19.97	16.04	15.38	17.89
F2'	24.92	22.16	20.02	16.04	15.39	17.91
F3	25.19	24.62	24.21	23.76	22.80	22.97
F4	27.49	27.73	28.00	27.17	26.86	27.05

Table 8. The average formant frequencies of English long vowels produced by Finnish female speakers (Hz/ERB).

	iː	ɑː	ɔː	uː	ɜː
F1 (Hz)	345	611	463	400	491
F2	2847	1092	980	1098	1584
F2'	3411	1093	981	1103	1599
F3	3394	2534	2757	2548	2183
F4	4294	3905	3855	3817	3626
F1 (ERB)	8.54	12.08	10.28	9.39	10.65

F2	24.15	16.29	15.47	16.33	19.24
F2'	25.71	16.30	15.48	16.37	19.31
F3	25.67	23.15	23.87	23.20	21.89
F4	27.73	26.90	26.78	26.70	26.25

4.2 Finland-Swedish speakers' vowel quality

All in all, the Finland-Swedish male (hereafter FSM) vowel constellation (Fig. 13) is somewhat more compressed than that of Finnish participants and not as clearly British. Contrary to Finnish speakers, all long-short vowel pairs seem to be over 1.06ERB from each other.

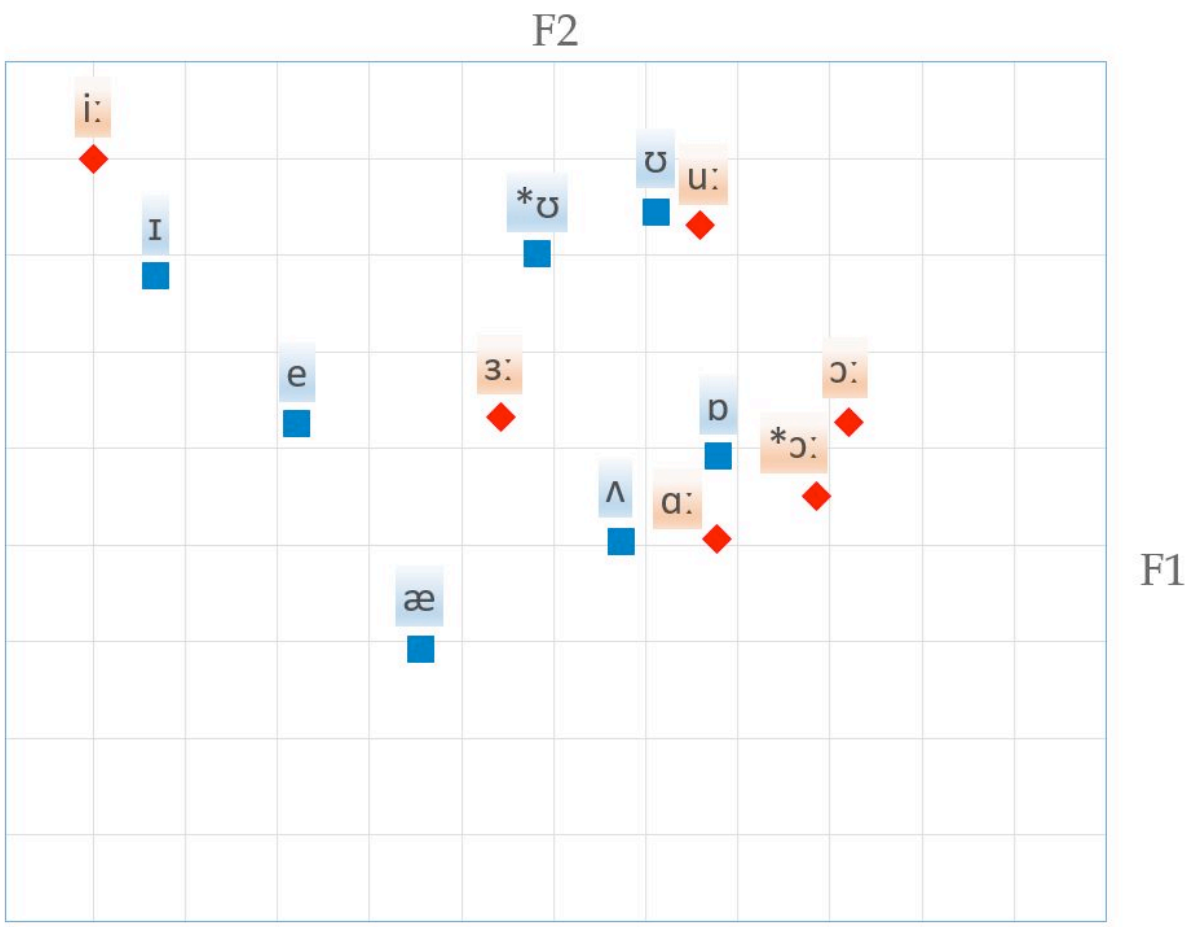

Figure 13. English vowels of Finland-Swedish male speakers (normalized ERB). The grid spacing is 1.06ERB.

The vowel /ɪ/ is pronounced quite accurately, being closest to GA. /e/ is also quite near to GA /ɛ/. The vowel /æ/ is within 1.06ERB of the GB vowel, but noticeably closer than in either of the Finnish groups. This seems strange, given that the Finland-Swedish /æ/ is, in turn, more open than its Finnish counterpart. /ɜː/ is approximately in between GA and GB values, being marginally closer to GA. All participants also pronounced it as a rhotic vowel. Judging by the results, *hood* and *who'd* were both pronounced /huːd/, which is reinforced by the lack of durational difference between the vowels. Both are within 1.06ERB from GA /uː/. This was also the case with Finnish speakers, and is most probably caused by misleading orthography. However, the word *put* was pronounced /pʊt/, which in turn gave a valid production of the vowel /ʊ/. This was over 1.06ERB away from /uː/, although pronounced as the Finland-Swedish /ʉ/ rather than either of the target language vowels. /ɑː/ is quite near the GB /ɒ/ and therefore closer than it should be. The o-like quality of /ɑː/ is a feature that is commonly connected to Sweden-Swedish because the equivalent vowel is slightly rounded in Central Swedish (Riad 2014: 35). /ʌ/, on the other hand, is within 1.06ERB from GB. */ɔː/ is just over 1.06ERB away from GB, whereas /ɒ/ is significantly different to GB, being closer. It can also be seen that F_1 is lowered when /ɔː/ is followed by /r/, which resembles the lower F_1 in /ɚ/ when compared to /ɜː/. As a whole, the a's and o's are compressed together and closer when compared to GB, which was also found in Finnish speakers' vowels.

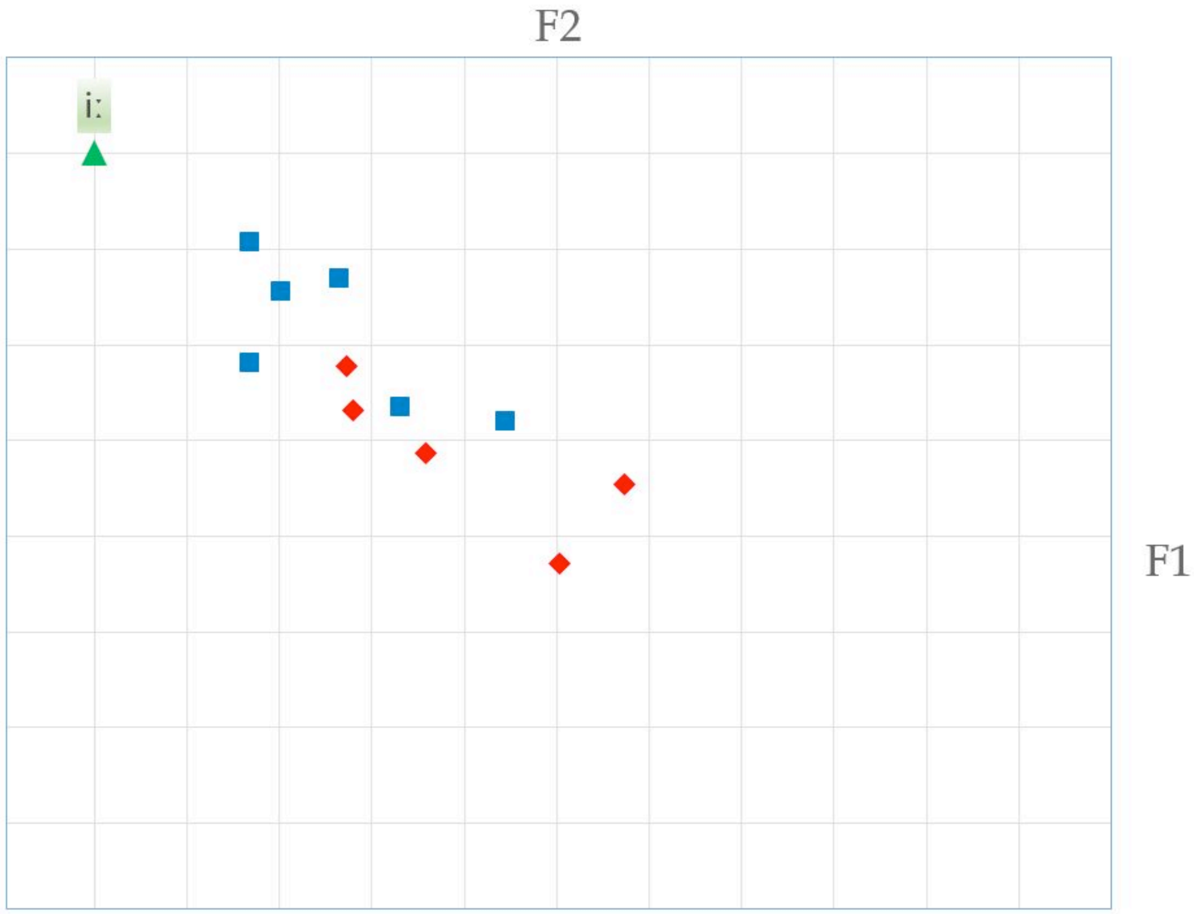

Figure 14. English reduced vowels of Finland-Swedish male speakers (normalized ERB). Blue markers are vowels that can be pronounced as either /ɪ/ or /ə/, red markers are vowels that can only be pronounced as /ə/. The grid spacing is 1.06ERB.

There is a clear division between /ɪ/ and /ə/ in FSM reduced vowels (Fig. 14). Furthermore, it seems that all instances of /ə/ (all red tokens and the two blue ones amidst them) are beautifully scattered around the approximate middle of the vowel constellation. Likewise, /ɪ/ seems to be slightly more central than it is in stressed syllables. All things considered, the reduced vowels of FSM speakers are highly target-like and neatly compressed when compared to either of the Finnish groups. One reason for this could be that Swedish speakers are familiar with the idea of complex stress patterns on both word and sentence level. This might help Finland-Swedish speakers produce correct stress patterns, which go hand in hand with vowel reduction.

Table 9. The average formant frequencies of English short vowels produced by Finland-Swedish male speakers (Hz/ERB).

	ɪ	e	æ	ɒ	*ʊ	ʊ	ʊ avg.	ʌ
F1 (Hz)	403	523	753	552	386	357	371	635
F2	2448	2019	1698	1105	1440	1212	1326	1276
F2'	2758	2258	1774	1115	1474	1218	1346	1299
F3	2946	2875	2773	2675	2703	2551	2627	2845
F4	4032	3905	3847	3689	3921	3873	3897	3776
F1 (ERB)	9.44	11.05	13.54	11.41	9.19	8.73	8.96	12.35
F2	22.86	21.24	19.80	16.38	18.47	17.10	17.81	17.50
F2'	23.87	22.17	20.16	16.45	18.65	17.14	17.93	17.65
F3	24.44	24.23	23.92	23.61	23.70	23.21	23.46	24.14
F4	27.18	26.90	26.77	26.40	26.93	26.82	26.88	26.60

Table 10. The average formant frequencies of English long vowels produced by Finland-Swedish male speakers (Hz/ERB).

	iː	ɑː	ɔː	*ɔː	ɔː avg.	uː	ɜː
F1	321	633	522	590	556	366	517
F2	2662	1108	905	952	928	1136	1517
F2'	3405	1122	909	961	935	1140	1535
F3	3366	2609	2744	2736	2740	2594	2159
F4	4060	3458	3604	3481	3543	3877	3448
F1	8.15	12.32	11.04	11.85	11.46	8.87	10.99
F2	23.57	16.40	14.88	15.25	15.06	16.60	18.89
F2'	25.70	16.50	14.90	15.32	15.11	16.62	18.98
F3	25.60	23.40	23.83	23.81	23.82	23.35	21.80
F4	27.24	25.83	26.19	25.89	26.04	26.83	25.81

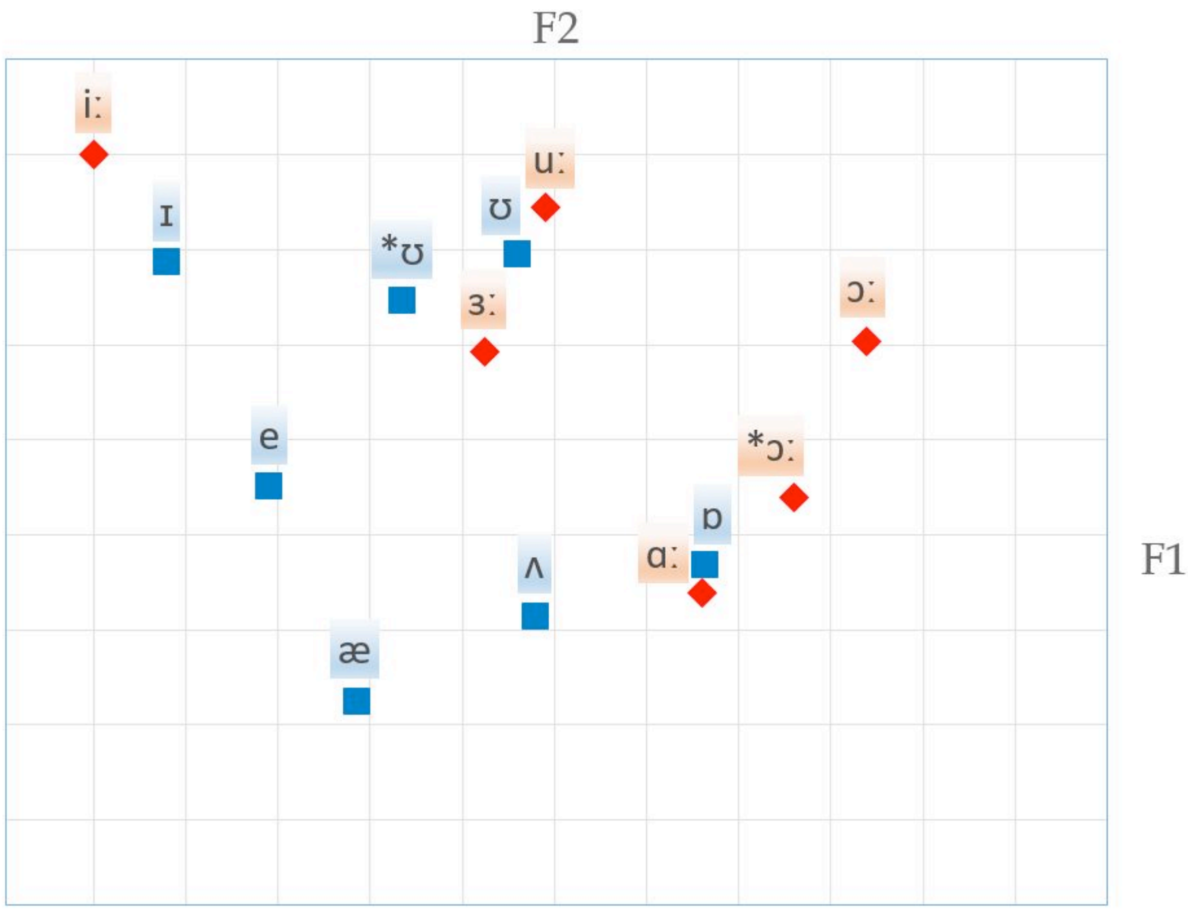

Figure 15. English vowels of Finland-Swedish female speakers (normalized ERB). The grid spacing is 1.06ERB.

Finland-Swedish female participants' (hereafter FSF) vowel constellation (Fig. 15) bears more resemblance to GB than FSM vowels. The front vowels /ɪ e æ/ are all near to GB vowels, much like both Finnish groups. In addition, the distance between long-short vowel pairs is over the JND limit in all cases. Of all groups in this study, FSF was the only one that pronounced /uː/ as the more front GB vowel. The distance to */ʊ/ is well over the JND limit, but similarly to FSM, the vowel is further front when compared to GB. There is hardly any difference in F_2 between GB /uː/ and /ʊ/, but FSF speakers differentiate these vowels mostly by F_2. The middle vowel /ɜː/ was rhotic and essentially identical to GA /ɚ/. /ʌ/ is very near GB, as is /ɒ/. However, there is once again overlap between a's and o's—a feature that was also found in both Finnish groups. F_2 is quite low in /ɑː/, which is why it essentially identical to /ɒ/. The similarity of these phonemes is an American feature that was also found in FIF results. /ɔː/ is approximately in between GB and GA, being more open than in GB but closer than in GA. This result has probably been affected by the cot-caught merger, since one FSF participant showed clear signs of it. The merger also explains the large difference between the two values for /ɔː/, since the merger does not affect /ɔː/ if it is followed by /r/.

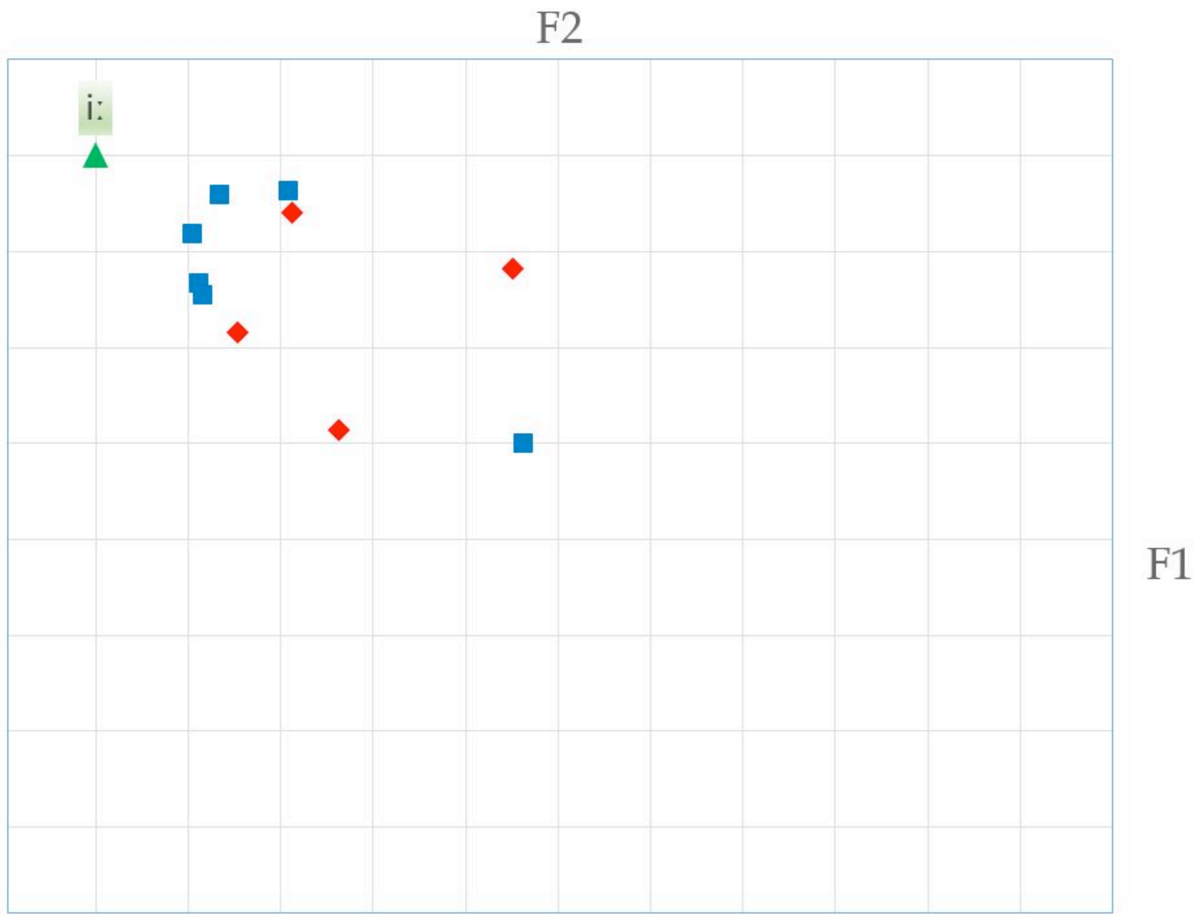

Figure 16. English reduced vowels of Finland-Swedish female speakers (normalized ERB). Blue markers are vowels that can be pronounced as either /ɪ/ or /ə/, red markers are vowels that can only be pronounced as /ə/. The grid spacing is 1.06ERB.

The FSF reduced vowels (Fig. 16) clearly form a group between /ɪ/ and /ɜː/. However, the vowels have generally quite low F_1 and high F_2 when compared to other groups. As can be seen from Figure 16, some of the blue tokens are in the same area as the Finland-Swedish /yː/. The reason for this might be that the FSF /ɜː/, too, is closer and further in the front when compared to other groups. Furthermore, the division between the two vowels is not as clear as with FSM. What is noteworthy is that

Finland-Swedish speakers' reduced vowels are not scattered around the vowel space in the way that Finnish speakers' are.

Table 11. The average formant frequencies of English short vowels produced by Finland-Swedish female speakers (Hz/ERB).

	ɪ	e	æ	ɒ	*ʊ	ʊ	ʊ avg.	ʌ
F1	423	623	876	709	453	416	435	768
F2	2384	2075	1838	1113	1725	1464	1594	1428
F2'	2585	2212	1879	1119	1843	1492	1667	1446
F3	2893	2846	2829	2717	2607	2440	2523	2729
F4	4019	4026	4138	3862	3764	3781	3772	4354
F1	9.72	12.22	14.63	13.10	10.15	9.63	9.89	13.68
F2	22.63	21.46	20.45	16.44	19.93	18.60	19.29	18.40
F2'	23.32	22.00	20.63	16.48	20.48	18.75	19.65	18.50
F3	24.29	24.14	24.09	23.75	23.39	22.83	23.12	23.78
F4	27.15	27.16	27.40	26.80	26.57	26.61	26.59	27.85

Table 12. The average formant frequencies of English long vowels produced by Finland-Swedish female speakers (Hz/ERB).

	iː	ɑː	ɔː	*ɔː	ɔː avg.	uː	ɜː
F1	344	740	488	636	562	382	497
F2	2632	1119	871	974	922	1405	1535
F2'	3319	1120	871	975	923	1434	1555
F3	3316	2392	2440	2693	2566	2569	2170
F4	4131	3683	3697	3717	3707	3851	3755
F1	8.53	13.41	10.61	12.35	11.52	9.12	10.72
F2	23.47	16.48	14.59	15.42	15.01	18.27	18.98
F2'	25.47	16.49	14.59	15.43	15.02	18.43	19.08
F3	25.47	22.66	22.83	23.67	23.26	23.27	21.84
F4	27.39	26.38	26.42	26.46	26.44	26.77	26.55

4.3 Comparison between Finnish, Finland-Swedish and English vowel quality

All in all, it seems that both groups pronounced English very well. It must be remembered that although I was able to hear and measure perceivable differences to native vowels virtually all participants' production, none were so far off that they make the speaker unintelligible. This is in line with Munro & Derwing's (1999) conclusion: a speaker may have a distinct foreign accent but be fully intelligible and easy to follow. All speakers distinguished all English vowels in their pronunciation, although the difference between long and short allophones of the same vowel was chiefly durational, which is acceptable in International English (Cruttenden 2014: 343). Moreover, both groups pronounced front and central vowels accurately, and back vowels were intelligible although their formant frequencies were noticeably different from native values. This is due to the human ability to understand disturbed speech through assimilating close-but-somewhat-off phonemes into the nearest L1 phoneme. However, neither group made a perceivable durational difference between long vowels preceding voiced and voiceless consonants. Most participants had a more British way of pronouncing English vowels, i.e. no diphthongization as in American pronunciation. The reason for this is that there is no diphthongization of monophthongs is not a common feature of Finnish and Finland-Swedish, and the phenomenon can be seen as a marked GA feature in English as well. In addition, marked features of General British, such as loss of rhoticity or pronouncing words such as *can't* with /ɑː/ instead of /æ/, were very scarce.

As for differences between the language groups, Finland-Swedish participants pronounced English slightly more accurately. There are two features in which they excelled: 1) making a qualitative difference between long-short vowel pairs 2) the accuracy in producing vowel reduction. The difference was small but noticeable in the first and quite large in the latter. An interesting point is that although Kuronen's (2000) results show that Finnish vowel constellation is more compact (especially for F_1, which has a 1.24ERB wider range), Finnish participants showed more F_1 variation in their English vowels than Finland-Swedish participants. Last, Finland-Swedish speakers seemed to pronounce English with rhoticity more often than Finnish participants. In fact, I initially intended to compare the participants' pronunciation with General British only, but it was practically impossible to find Finland-Swedish participants that pronounce English with a British accent. Thus, General American had to be included in the present study as well.

Although Finnish and Finland-Swedish are quite different in basically all aspects except segmental, the differences in English pronunciation between the L1s were not large. In addition, it is quite likely that the few differences are caused by segmental differences between the L1s. For example, the reason behind the different pronunciations of the English /uː/ was probably the fact that the English vowel assimilated into /uː/ in Finnish and to /ʉː/in Finland-Swedish. However, Finland-Swedish participants produced vowel reduction significantly better and also had qualitatively different long-short vowel pairs more often than Finnish speakers, which cannot be explained by segmental differences in the L1s. The author believes that Finland-Swedish stress placement and timing, which are substantially closer to English than those of Finnish, are at least partly behind this.

4.4 Comparison between Finnish, Finland-Swedish and English vowel duration

The average durations of pre-voiceless and pre-voiced long vowels can be found in Table 13. The relative difference between the vowels tells how much longer pre-voiced vowels are when compared to pre-voiceless vowels.

Table 13. Average vowel durations of pre-voiceless and pre-voiced vowels and their relative differences.

	Pre-voiceless (ms)	Pre-voiced (ms)	Difference (%)
FI male	207	229	+10.6
FI female	191	228	+19.4
FS male	226	268	+18.6
FS female	205	226	+10.2
FI total	199	229	+15.1
FS total	215	247	+14.9

Male total	217	248	+14.3
Female total	198	227	+14.6

On average, pre-voiced vowels were slightly longer than pre-voiceless vowels. However, the duration difference is significantly smaller than in Wiik's (1965: 116) results with native speakers (pre-voiced vowels 64–100% longer). In addition, the participants of this study—except for one speaker—did not produce duration differences in a systematic manner. The average difference varied between -11.4% and +43.4% across all participants of the study. Both extremes were Finnish female speakers, and interestingly enough, Finnish females also seemed to produce greatest duration differences on average. Nevertheless, the differences across L1s and genders are quite small. These results show that no L1 or gender group produced target-like duration differences or excelled another group. Moreover, Klatt (1976: 1219) has argued that duration differences of about 20% or more can be used as primary perceptional cues. This would mean that on average, no group produced duration differences that are even perceptible. However, it has also been found that duration differences are easier to perceive in vowels than in consonants (Fletcher 2010: 526). Therefore the average difference might be perceptible in FIF and FSM speakers' productions, but only just. According to Fletcher (ibid.), there is a consensus on the JND in duration, which varies between 10–40ms depending on the quality, quantity and the context of the segment. In light of this, it seems that the duration differences are not only significantly smaller than they should, but probably even imperceptible.

In light of this, it seems that neither Finnish nor Finland-Swedish learners cannot use vowel duration to make a difference between words like *beat* and *bead.* This, however, does not mean that they make no difference between the words, but rather use other means for it. For example, most participants pronounced the words with an audible release of the last consonant. Furthermore, many participants pronounced the consonant /t/ as a dental rather than an alveolar consonant, which comes from assimilating the English vowel to the L1 category of /t/ that is pronounced dentally in Finnish and (often, not always) Finland-Swedish. Because of this, the closure from the vowel /iː/ to the consonant is slightly different, but the difference is probably too small to notice on its own. If the participants were to engage in a conversation in a noisy environment and/or use unreleased plosives in their speech, the distinction between minimal pairs as described above would probably be difficult. However, this can hardly be considered as a major problem in terms of intelligibility as long as the distinction is made with other, although non-native, means. Nonetheless, it could hinder the acquisition of native-like speech rhythm, if one aims to pronounce English on a near-native level.

5 DISCUSSION

In this section, I will discuss the results of the present study in relation to the theoretical background, namely Flege's (1995) Speech Learning Model and Best's (1995) Perceptual Assimilation Model, which were presented in Section 2.1.3. The aim is to find whether to see whether the theoretical models can predict the way participants have learned English pronunciation. The research process and its reliability is going to be reviewed as well.

To recap, both SLM and PAM suggest that the more phonetically different an L2 phoneme is from the nearest L1 phonemes, the more accurately it will be perceived and produced. The way in which the participants have learned English pronunciation seems to follow this principle – but not entirely. There was a clear pattern in both groups' productions: front and central vowels were near or spot-on when compared to native formant values, whereas back vowels were almost always somewhat off. All front vowels (i.e. /iː æ e/) were pronounced well. The way that I have normalized the results made it impossible to compare /iː/ to native values but for three reasons, I am positive that it pronounced accurately. First, the L1 values of /iː/ across Finnish, Finland-Swedish and English are not very different to each other, which predicts same-category assimilation and accurate production. Second, the formants of the few speakers whose formant values were comparable to those of adult people were close to native values. Last, I did not hear any strange sounding productions of /iː/ during the analysis. The accurate production of these three sounds was expected, because the English /æ/ and /e/ were, in turn, quite different to L1 values. This enables the speaker to establish new phonetic categories for the sounds, which, in turn, leads to accurate perception and production. The formants of the central vowel /ɜː/ were also close to native values. There was great variation in rhoticity, but rhotic productions were close to American values and non-rhotic were close to British values, which indicates that the pronunciation was accurate.

However, most back vowels seemed to be somewhat off in all groups. Although /uː/ was produced accurately quite often, the short /ʊ/ was not produced very accurately.

These results are explained by L1 assimilation: The L1 / uː/ of both groups is so similar to the GA vowel that same-category assimilation does not lead into inaccurate pronunciation, whereas assimilating /ʊ/ to L1 categories did. Interestingly, Finnish participants seemed to assimilate it to /u/ and Finland-Swedish participants to /ʉ/. Furthermore, the a's and o's were often significantly more close than those of native speakers. This could be explained by assimilation, as Finnish and Finland-Swedish back vowels are more close than English back vowels. However, the same applies to all other vowels as well: Finnish and Finland-Swedish vowel constellations in their entirety are more compressed F_1-wise than English ones. One cause for this might be that whereas GB and GA reference values were extracted from single words in citation form, Finnish and Finland-Swedish values were extracted from words in a carrier phrase. However, neither context is natural speech, which is why the rather large difference in F_1 range is not likely to be caused by the different contexts. Therefore, the important question is why only back vowels seem to be affected by L1 phonology while front and central vowels are pronounced accurately. This I cannot answer; further studies are needed to find the cause. In addition, the overlap between a's and o's is suspicious, since this overlap is only present in my analysis and not in reference values from the literature. It also cannot be explained by either theory. The author's view is that this might be the result of small sample size and/or faulty analysis. However, if there truly is overlap between the two vowels, it would be a concern for comprehensibility, as Chan & Hall (2019) have found that when the formants of a vowel shift into a more "crowded" area in the vowel space, it reduces comprehensibility and increases accentedness.

One reason for faulty analysis might be that rounded back vowels, such as /o/, have high F_1 and low F_2. Sometimes these are so close together that it is not clear where F_1 ends and where F_2 begins by looking at the spectrogram. They are also subject to formant integration, because they are often less than 3.5 Bark away from each other. In this situation, the researcher can only trust the software because the formants are practically impossible to spot manually by looking at the spectrogram.

Because English has two main varieties, and both have seemingly had an effect on the participants' pronunciation, the present study provides an interesting and new setting as regards theories of L2 pronunciation learning. For example, the e-sounds of GB and GA are quite different, although both fall into the same phonological category of /e/. The GB phoneme is similar to Finnish and Finland-Swedish, whereas the GA phoneme is clearly diphthongized. This predicts that the GB /e/ will be assimilated into the L1 category, whereas GA /ɛ/ gets its own phonetic category. But because there is no need to do both, it is interesting to see which process overrides the other. The results indicate that the tendency to assimilate new phonemes into existing categories is stronger. This is not surprising, because it requires lesser cognitive effort. For example, the participants favoured GB pronunciation in their pronunciation of the aforementioned vowel pair. In turn, Finnish speakers most often pronounced/uː/ as the GA phoneme, because a) the GA vowel is very near to the Finnish vowel, i.e. assimilable b) the GB phoneme is so different from all Finnish vowels that it would require the speaker to establish a new phonetic category.

The influence of orthography on pronunciation was clear in both language groups. The word *hood* was pronounced /huːd/ by almost all participants. This is explained by orthography that suggests a long vowel rather than a short one. This is in line with Jarvis & Pavlenko's (2008: 70) claim that orthography can override a person's ability to hear pronunciation correctly; people living in Finland are exposed to native English and words such as *foot, book* and *hood* through media but nevertheless often pronounce these words with a long vowel. Also the fact that both groups favored the British pronunciation of the word *hot* (i.e. /hɒt/ instead of the American /hɑt/) is at least partly explained by the fact that the British pronunciation reflects the orthography better. The impact of orthography most probably accounts for the choice of rhotic pronunciation over non-rhotic pronunciation, at least partly. In fact, I believe that the majority of Finnish and Finland-Swedish learners make an unconscious decision between using an American or British sound based on two criteria: First, if **the phoneme of only one of the varieties assimilates into an L1 category**, the phoneme of this variety will be chosen. For example, Finnish people systematically chose the GA /uː/ instead of the GB /uː/, because the former is quite readily assimilable to

Finnish /uː/ and the latter would require the establishment of a new phonetic category. However, if **the phonemes of both varieties are assimilable**, orthography dictates the choice. An example of this would be the pronunciation of *hot,* as explained earlier. Both /ɒ/ and /ɑ/ are assimilable to the L1s of the participants of in the present study, and the author's view is that /ɒ/ was chosen because it is more readily associated with the grapheme <o>. This model would not apply if the speaker were skilful and motivated enough to choose one variety and stick to it.

However, this is highly speculative and should be studied more see if it actually predicts phoneme choice. Furthermore, this hypothetic model cannot explain the choice between GB and GA /uː/ for Finland-Swedish speakers. Both vowels are more or less assimilable to Finland-Swedish: The GB vowel is practically identical to /ʉː/ and the GB vowel is not very far from /uː/, although definitely further away than 1.06ERB. The orthography of *hood* clearly suggests /uː/ to a Finland-Swede, since <o> is generally pronounced /uː/ in Finland-Swedish. However, Finland-Swedish girls went for the GB vowel and the boys for the GA vowel. It seems that assimilation to L1 phonetic categories and the impact of orthography are in contradiction with each other in this situation. It would be interesting from a psycholinguistic view to conduct a study on this and see which process seems to affect phoneme choice more.

All in all, I reckon that the research process was reasonably successful. I was faced with two major obstacles during the study, both of which I managed to clear in the end. The first was that the formant frequencies were so wildly different even across speakers of the same gender, which I did not take into account when choosing to study ninth-graders. It would have been easier to study children at around 10 years of age or adults instead, because then the ongoing voice change would not have caused this kind of a problem. The problem was eventually resolved by normalizing the formants with Kuronen & Kautonen's (2018) experimental method. Second, I did not expect to meet so much rhoticity in the speakers' productions, because initially, I only accepted participants that spoke British English in their own view. This was resolved by comparing my results to both General British and General American native values. Both problems could have been avoided by having a large amount of participants to choose from. However, having dozens of participants to choose from was not possible within the limits of the study. Due to the small sample size, the results of the study can also be considered preliminary at best. Further studies will be needed to confirm the results of this study.

Besides increasing the sample size, recording all reference values myself would have increased the reliability of this study greatly. In the present study, there are formant frequencies measured by four different authors (or groups of authors) and with four different recording setups. In addition, I am comparing youngsters' productions to adult reference values because they were the only ones that were available. The advantages would have been numerous: First, differences caused by the setting or age would have been eliminated. Second, examining L2 vowels' assimilation into L1 categories would have been more reliable if I would have been able to compare every speaker's L2 vowels into their own L1 vowels.

When it comes to analysis, there is also a little room for improvement. First, it turned out that spectral change is an important feature in American vowels in particular (Hillenbrand et al. 1995). Spectral change was not taken into account in the analysis of the present study, which can be seen as a shortcoming. In addition, the microphone used in the study did not meet all requirements that Švec & Granqvist (2010) list. However, simple F_1/F_2 formant analysis does not seem to require very sophisticated recording equipment, because I did not get unreasonable results from my analysis. In addition, there are peer-reviewed and published studies that have been made with equipment inferior to the ones used in the present study (e.g. González 2004), which makes this a very minor problem, if a problem at all. Ultimately, the only real problems in this study are related to heterogeneity of my participants and small sample size.

6 CONCLUSIONS AND FURTHER RESEARCH

First, let us go back to the research questions and hypotheses that were presented in Section 3.1. The overall aim of this study was to describe how Finnish and Finland-Swedish intermediate learners pronounce English vowels and especially describe how their pronunciation differs from native English speakers. It seems that Finnish and Finland-Swedish students pronounce English quite well and quite similarly. The largest differences between the two groups are vowel reduction, rhoticity and the qualitative difference between long-short vowel pairs—or lack thereof, as was the case with Finnish participants. The research questions can be answered as follows:

1. Which vowels are produced with a significantly non-native quality?
   a. Finnish speakers pronounce vowels /ɑː ʌ ɔː ɒ ʊ/ with a non-native quality. In addition, long-short vowel pairs are often not qualitatively different.
   b. Finland-Swedish speakers pronounce also pronounce vowels /ɑː ʌ ɔː ɒ ʊ/ with a non-native quality. Contrary to Finnish speakers, their long-short vowel pairs are most often qualitatively different to each other.
2. Do the learners produce English vowel reduction accurately?
   a. Finnish speakers sometimes produce vowel reduction accurately.
   b. Finland-Swedish speakers constantly produce vowel reduction accurately.
3. Do the learners produce the duration differences between long vowels preceding fortis and lenis consonants?
   a. No, neither Finnish nor Finland-Swedish learners produce the duration differences.
4. Do Finnish and Finland-Swedish learners of English produce English vowels differently?
   a. When speaking English, Finland-Swedish speakers have more advanced /uː ʊ/, more qualitative difference between long-short vowel pairs, more rhoticity and more accurate vowel reduction.

The hypotheses also presented in Section 3.1 are answered as follows:

1. English vowels that are close or identical to L1 vowels will assimilate into the L1 phonetic categories.
    a. Mostly true. A good example of this is the vowel /uː/, which was pronounced as the Finnish /uː/ and the Finland-Swedish /ʉː/ by the respective groups. However, some vowels that seem to assimilate into L1 categories are actually in between L1 and L2 values. It might be that the two have assimilated together and "met halfway". This would mean that the L1 sound has changed slightly. After all, all phonetic categories are subject to change over the course of life, even those established during L1 learning (Flege 1995). However, this cannot be confirmed in the limits of this study.
2. English vowels that are not close to any L1 vowels will not be assimilated. Instead, new phonetic categories are established, and the vowel is produced at least somewhat natively.
    a. Again, mostly true. The best example of this is /æ/, which was pronounced similarly to GB by all four groups, probably because it is so different to the equivalent vowel in both L1s. The back vowels of English seemed to get their own phonetic categories, which were, however, somewhat off when compared to native values.
3. Both groups will use rounding exaggeratedly.
    a. Difficult to say, because F_2 correlates with both rounding and advancedness. However, it does not seem that the rounded vowels are too rounded, because there was not much difference in F_2, but rather F_1.
4. Both groups will use full vowels instead of reduced vowels in unstressed syllables.
    a. Not entirely true. Finnish participants sometimes used full vowels instead of reduced vowels, whereas Finland-Swedish participants used vowel reduction in an almost native-like manner.
5. Neither group will produce the durational differences between vowels preceding voiceless and voiced consonants accurately.

   a. True. Not only are the durational differences lesser than what native speakers produce, but they are also nearly imperceptible.
6. Both groups will differentiate long and short English phonemes with duration, not quality.
   a. Not entirely true. Finnish speakers had a difference of 1 ERB and no more in some long-short vowel pairs, such as /iː/ and /ɪ/. On the other hand, the o-sounds were not qualitatively different in Finnish participants' speech samples. In comparison, Finland-Swedish speakers had a difference of 1 ERB or more in all long-short vowel pairs, and the difference was often larger than in Finnish participants' productions.

This study has several achievements to its credit. First, a systematic difference between Finnish and Finland-Swedish ninth-graders' English pronunciation was found. Finland-Swedish participants excelled in practically every aspect of pronunciation that was analyzed in this study. Second, this study is one of the first to research Finland-Swedish speakers' English. As phonetic studies are usually done with adult participants, this study also provides valuable data of adolescent learners' speech that is otherwise scarce. Third, using adolescent speakers as informants forced the author to find and use an experimental method of normalization, which also revealed the absence of an established normalization method that is compatible with psychoacoustic scales such as Bark and ERB. Last, the study has some implications for English teachers in Finland. Although vowel pronunciation of my participants does not seem to threat intelligibility, there are some features that could be given more attention. In the author's view, non-native pronunciation in both groups is caused by not being aware of fundamental differences between English and the participants' L1 rather than pure lack of practice. For example, the fact that there is both a qualitative and a quantitative difference in short-long vowel pairs of English might come as a surprise for the participants of this study despite their good command of English. They might be accustomed to the idea that *beat* is exactly the same as *bit* but with a longer vowel, because that is how the L1 vowel system works. The way to improve this kind of proficient learners' English pronunciation is to help them notice these fundamental differences that they might miss without instruction.

All results should be considered preliminary, and further research is needed to confirm any of my results. The present study has raised a few exceptionally interesting or potentially useful questions that could be answered by further research. The first mission is to simply study and describe the English pronunciation of Finland-Swedish people, since there are virtually no previous studies on this. My preliminary results indicate that Finland-Swedish-accented English pronunciation is quite close to Finnish-accented pronunciation at the segmental level. Therefore, using previous knowledge on Sweden-Swedish learner language would be invalid for designing teaching materials for Finland-Swedish learners, for example. By the same token, using Finnish learner language as the starting point would not be an entirely valid choice, either. This is why more information on Finland-Swedish learner language is needed. In addition, Finnish and Finland-Swedish create an interesting premise for phonetic research just by themselves, because they are segmentally similar but otherwise different languages spoken by socioeconomically similar people, which means that the effects of L1 prosody on L2 learning, for example, can be researched almost in isolation. Furthermore, the overlap between a's and o's, which was found in all four groups, is a puzzling feature. It should be investigated whether this is an actual feature or just an anomaly in my analysis. The indications that Finland-Swedish learners centralize English short vowels more and produce English vowel reduction more accurately are also interesting points that could be studied further. However, the most interesting topic of further research is prosodic differences between Finnish and Finland-Swedish learners of English. It is probable that one would find more significant differences between the two language groups when looking at prosody rather than segments.

REFERENCES

Aaltonen, O. (1985). The effect of relative amplitude levels of F2 and F3 on the categorization of synthetic vowels. *Journal of Phonetics*, 13(1), 1–9.

Adank, P., Smits, R., & van Hout, R. (2004). A comparison of vowel normalization procedures for language variation research. *The Journal of the Acoustical Society of America*, 116(5), 3099–3107.

Boersma, P., & Weenink, D. (2018). PRAAT: Doing phonetics by computer (Version x.x) [Computer program]. Retrieved 1.2.2018 from http://www.praat.org

Best, C. (1995). A direct realist view of crosslanguage speech perception. In W. Strange (ed.), *Speech perception and linguistic experience: theoretical and methodological issues.* Timonium, MD: York Press, 171–204.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M. J. Munro & O.-S. Bohn (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (13–34). Amsterdam: John Benjamins.

Bladon, A., & Fant, G. (1978). A two-formant model and the cardinal vowels. *Speech Transmission Laboratory Quarterly Progress and Status Reports*, 19(1), 1–8.

Chan, K. Y., & Hall, M. D. (2019). The importance of vowel formant frequencies and proximity in vowel space to the perception of foreign accent. *Journal of Phonetics, 77*, 100919.

Cruttenden, A. (2014). *Gimson's pronunciation of English* (8th ed.). London; New York: Routledge.

Davidsen-Nielsen, N., & Harder, P. (2001). Speakers of Scandinavian languages: Danish, Norwegian Swedish. In M. Swan & B. Smith (eds.), *Learner English: A teacher's guide to interference and other problems* (2nd edition). Cambridge: Cambridge University Press, 21–36.

Derwing, T. M. (2003). What do ESL students say about their accents? *Canadian Modern Language Review*, 59(4), 547–566. https://doi.org/10.3138/cmlr.59.4.547

Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research.* Amsterdam, Netherlands; Philadelphia, Pennsylvania: John Benjamins Publishing Company.

Deterding, D. (1990). *Speaker Normalisation for Automatic Speech Recognition* (Doctoral dissertation). Cambridge, UK: University of Cambridge.

Dewaele, J-M., & McCloskey, J. (2015). Attitudes towards foreign accents among adult multilingual language users. *Journal of Multilingual and Multicultural Development*, 36(3), 221–238.

Disner, S. (1984). Insights on vowel spacing. In I. Maddieson (Ed.), *Patterns of sounds*. Cambridge, UK: Cambridge University Press, 136–155.

Fant, G. (1959). *Acoustic Analysis and Synthesis of Speech with Applications to Swedish* [Technical Report].

Fant, G. (1960). *Acoustic theory of speech production: With calculations based on X-ray studies of Russian articulations*. The Hague, Netherlands: Mouton.

Fant, G. (2004). *Speech acoustics and phonetics*. Dordrecht, Netherlands: Kluwer.

Fletcher, H. (1940). Auditory patterns. *Reviews of Modern Physics*, 12, 47–66.

Fletcher, J. (2010). The prosody of speech: Timing and rhythm. In W. J. Hardcastle, J. Laver & F. E. Gibbon (Eds.), *The handbook of phonetic sciences* (2nd ed.) (523–602). Chichester, UK; Malden, MA: Wiley-Blackwell.

Fuertes, J. N., Gottdiener, W. H., Martin, H., Gilbert, T. C., & Giles, H. (2011). A meta-analysis of the effects of speakers' accents on interpersonal evaluations. *European Journal of Social Psychology*, 42, 120–133.

Glasberg, B., & Moore, B. (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing Research,* 47(1–2), 103–138.

González, J. (2004) Formant frequencies and body size of speaker: a weak relationship in adult humans. *Journal of Phonetics,* 32(2), 277–287.

Hawkins, S., & Midgley, J. (2005). Formant frequencies of RP monophthongs in four age groups of speakers. *Journal of the International Phonetic Association,* 35(2), 183–199.

Hillenbrand, J., Getty, L., Clark, M., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America*, 97(5), 3099–3111.

Hunter, G., & Kebede, H. (2012). *Formant frequencies of British English vowels produced by native speakers of Farsi*. Paper presented at Acoustics 2012: 11th Congrès Français d'Acoustique, Nantes, France.

Iivonen, A. (2012). Kielten vokaalit kuuloanalogisessa vokaalikartassa. *Puhe ja kieli,* 32(1), 17–43.

Immonen, K., & Peltola, M. (2018). Finnish children producing English vowels — Studying in an English immersion class affects vowel pronunciation. *Journal of Language Teaching and Research*, 9(1), 27–33.

Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. New York, NY: Routledge.

Johnson, K. (2012). *Acoustic and auditory phonetics* (3rd ed.). Malden, MA: Wiley-Blackwell.

Kautonen, M. (2018). Om hur finskspråkiga och svenskspråkiga lyssnare bedömer och värderar avancerade talares uttal av finlandssvenska. *Nordand: nordisk tidsskrift for andrespråksforskning*, 2(2), 131–150.

Kivistö, A. (2016). The making of rallienglanti: Cross-linguistic influence in Finnish speakers' pronunciation of English (Master's thesis). Jyväskylä: University of Jyväskylä.

Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *The Journal of the Acoustical Society of America*, 59, 1208–1221.

Kuronen, M. (2000). *Vokaluttalets akustik i sverigesvenska, finlandssvenska och finska* (Doctoral dissertation). Jyväskylä: University of Jyväskylä.

Kuronen, M., & Tergujeff, E. (2017). L1 listeners' perception of L2 pronunciation: effect of prosody on accentedness ratings in Swedish. *AFinLA-E: Soveltavan kielitieteen tutkimuksia,* 10, 213–233.

Kuronen, M., & Kautonen, M. (2018). Foneettisten piirteiden ja vieraan aksentin yhteydestä suomen kielessä. *Lähivertailuja*, 28, 207–241. DOI: 10.5128/LV28.06

Labov, W., Ash, S., & Boberg, C. (2006). *The atlas of North American English: Phonetics, phonology, and sound change: A multimedia reference tool*. Berlin, Germany; New York, NY: Mouton de Gruyter.

Lenneberg, E. H. (1967). *Biological foundations of language*. New York, NY: Wiley.

Levis, J. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly*, 39(3), 369–377.

Liljencrants, J., & Lindblom, B. (1972). Numerical simulation of vowel quality systems: The role of perceptual contrasts. *Language*, 48(4), 839–862.

Lobanov, B. (1971). Classification of Russian vowels spoken by different speakers. *The Journal of the Acoustical Society of America*, 49(2), 606–608. DOI: 10.1121/1.1912396

Läromedelssituationen våren 2010 för de finlandssvenska grundskolorna och gymnasierna. Finnish National Board of Education [online]. https://www.oph.fi/download/135472_Laromedelssituationen_varen_2010_for_de_finlandssvenska_grundskolorna_och_gymnasierna.pdf (28 February, 2019)

Moore, B. (2010). Aspects of auditory processing related to speech perception. In: W. Hardcastle, J. Laver & F. Gibbon (Eds.) *The handbook of phonetic sciences* (2nd ed.). Chichester, UK; Malden, MA: Wiley-Blackwell: 454–488.

Morris-Wilson, I. (1999). *Attitudes towards Finnish-accented English* [Doctoral dissertation]. Stirling, UK: University of Stirling.

Morris-Wilson, I. (2003). *English pronunciation exercises for Finns*. Oulu: Oulu University Press.

Morris-Wilson, I. (2004). *English segmental phonetics for Finns*. Oulu: Oulu University Press.

Munro, M. J., & Derwing, T. M. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech,* 38(3), 289–306.

Munro, M. J., & Derwing, T. M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning,* 49, 285–310.

Ng, M., Chen, Y., & Sadaka, J. (2008). Vowel features in Turkish accented English. *International Journal of Speech-Language Pathology*, 10(6), 404–413.

Peltola, M., Lintunen, P., & Tamminen, H. (2014). Advanced English learners benefit from explicit pronunciation teaching: an experiment with vowel duration and quality. *AFinLA-E: Soveltavan kielitieteen tutkimuksia*, 6, 86–98.

Pillai, S., & Salamae, W. (2012). An instrumental analysis of English monophthongs produced by Thai EFL learners. *Pertanika Journal of Social Science and Humanities*, 20(4), 1143–1160.

Reuter, M. (1971). Vokalerna i finlandssvenskan: En instrumentell analys och ett försök till systematisering enligt särdrag. *Studier i nordisk filologi* 58, 240–249.

Riad, T. (2014). The phonology of Swedish. Oxford: Oxford University Press.

Saito, K., Trofimovich, P., & Isaacs, T. (2016). Second language speech production: Investigating linguistic correlates of comprehensibility and accentedness for learners at different ability levels. *Applied Psycholiguistics,* 37(2), 217–240.

Sarmah, P., Verma Gogoi, D., & Wiltshire, C. (2009). Thai English: Rhythm and vowels. *English World-Wide*, 30, 196–217.

Suomi, K. (1976). *English voiceless and voiced stops as produced by native and Finnish speakers*. Jyväskylä: University of Jyväskylä.

Suomi, K. (1990). *Johdatusta puheen akustiikkaan*. Oulu: University of Oulu.

Suomi, K., Toivanen, J., & Ylitalo, R. (2008). *Finnish sound structure: Phonetics, phonology, phonotactics and prosody*. Oulu: University of Oulu.

Švec, J. G., & Granqvist, S. (2010). Guidelines for selecting microphones for human voice production research. *American Journal of Speech-Language Pathology*, 19(4), 356–358.

Tergujeff, E. (2013). *English pronunciation teaching in Finland* (Doctoral dissertation). Jyväskylä Studies in Humanities 207. Jyväskylä: University of Jyväskylä.

The Audacity Team (2019). Audacity (Version 2.3.2) [Computer program]. Retrieved 20.7.2019 from https://www.audacityteam.org/

Traunmüller, H., & Eriksson, A. (1995). The frequency range of the voice fundamental in the speech of male and female adults. Unpublished paper.

Trofimovich, P., & Isaacs, T. (2012). Disentangling accent from comprehensibility. *Bilingualism: Language and Cognition*, 15(4), 905–916.

Wells, J. C. (1962). *Study of the Formants of the Pure Vowels of British English* (MA thesis). London, UK: University of London.

Wiik, K. (1965). *Finnish and English vowels: A comparison with special reference to the learning problems met by native speakers of Finnish learning English* (Doctoral dissertation). Turku: University of Turku.

Zwicker, E. (1961). Subdivision of the audible frequency range into critical bands (Frequenzgruppen). *The Journal of the Acoustical Society of America*, 33(2), 248.

APPENDICES

Appendix 1. A list of stimulus words used. Words marked with an asterisk were only used with Finland-Swedish participants.

1. Short vowels

/ɪ/ - hid

/e/ - head

/ɒ/ - hod

/ʊ/ - hood

/ʊ/ - put*

/æ/ - had

/ʌ/ - hud

2. Long vowels

/iː/ - heed

/uː/ - who'd

/ɜː/ - heard

/ɔː/ - hoard

/ɔː/ - hawed*

/ɑː/ - hard

3. Reduced vowels

Get a grip.

takeaway

seventeen

respect

prevent

example

4. Voiced-voiceless opposition

beat	boot	board
Bert	bard	booed
bird	bead	bought
Bart	seat*	peas*
seed*	peace*	