Foreign Accents in Song and Speech

by

Marly Mageau

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs in partial fulfillment of the requirements for the degree of

Master of Cognitive Science

in

Cognitive Science

Carleton University

Ottawa, Ontario

Abstract

Is it more difficult to detect an accent when someone is singing than when they are speaking? Previous work on accents in song has focused on professional singers, who modify their accents when they perform music that is associated with a particular regional accent (Trudgill, 1983, Simpson, 1999, O'Hanlon, 2006, Gibson, 2010). We ask whether there is something about singing per se that causes a shift in accent. In order to answer this question, our study differs from previous work in two important ways: 1) We do not make use of professional singers in our study, and 2) the music in our study is not culturally associated with a particular country.

We recorded twelve speakers: six native speakers of English, and six second-language speakers. They were asked to sing "Twinkle Twinkle Little Star" and read a passage from "Goldilocks". Forty native English listeners had more difficulty detecting a foreign accent in the singing conditions and rated speakers as having less of a foreign accent in song compared to speech. These results suggest that it is more difficult to detect a foreign accent in song compared to speech even when the singers are not influenced by an accent associated with a particular genre. An analysis of the recordings showed that vowel duration is generally longer in song and that the pitch track changes in song. However, there is no clear and consistent difference in how vowels are pronounced. Based on these findings, we argue that accents are more difficult to detect in song than in speech because the rhythm and pitch of song mask important prosodic markers of accent.

Acknowledgements

There are several people I would like to thank for making this thesis possible. First, I would like to thank my thesis supervisor, Dr. Ida Toivonen, for her constant support and guidance during the preparation, execution and writing of this thesis. I am grateful for the compassion she showed and countless hours she put into shaping my Master's experience as a productive and positive one. I would also like to thank my thesis committee members, Dr. Beth Macleod and Dr. Lev Blumenfeld, for their time, contribution and insights, without which this thesis would not be what it is. Also, I would also like to thank my Thesis Defense internal/external examiner, Dr. Anna Hoefnagels, who allowed me to understand this research through an outside, musical perspective. In addition, I was fortunate to work with Dr. Masako Hirotani in both my undergraduate and Master's degrees, as well as her teaching assistant. I would like to thank her for her time and for sharing her expertise in working with participants.

I would like to thank my friends and colleagues for being there to talk, bounce ideas off of and those who were there when I needed some downtime. Specifically, Katherine Baker, Alex Tait, Emily Wang, Kyle Fazackerley and the ICS graduate students. A special thank you goes out to Can Mekik for all of his time and effort in preparing presentations, running participants, and the countless other contributions he made to this thesis despite the pile of work on his desk for his own. An especially special thank you goes out to my friend, partner in crime and fellow nerd Roxana Barbu for all of her editing skills, conversations, guidance and for showing up whenever I needed her (whether I was willing to admit it or not). She directed me to the program and made my experiences unforgettable.

I would also like to thank Dimitri Helis, who has been there every step of the way, from the first day of my undergraduate degree until the end of my Master's degree. He has shared not only my triumphs and successes but also my disappointments and frustrations. He is my greatest friend and has been an invaluable support through this whole process.

Finally, I would like to thank my parents, Mike and Malodie Mageau and my sister, Mallissa Pomerleau for putting up with my roller coaster mood swings and being there whenever I needed them. They have taught and proven to me that hard work and perseverance pays off, especially if it's for something you really want. Their unconditional love, support and endless conversations allowed me to complete my Master's degree. I could not have accomplished this thesis, or anything for that matter, without them; therefore, I dedicate this thesis to them!

Table of Contents

List of Tables

List of Figures

List of Appendices

Chapter 1

Introduction

Professional singers often shift their accents when singing. It has been argued that these singers strive for a particular accent (often American) for cultural reasons (Trudgill, 1983; Simpson, 1999; Morrissey, 2008; Beal 2009; Gibson, 2010; Gibson, 2011). However, to my knowledge, no previous studies have investigated accent shifts in untrained singers. Is there something about singing per se that causes a shift in accent? I specifically address the following questions:

1) Is it more difficult to detect a foreign accent in song than in speech?
2) Does singing per se change the way vowels are pronounced?
3) Does prosody make it more difficult to detect a foreign accent in speech?

In the next three sections, I will provide an overview of previous research on accents in music and foreign accents in speech and song. I will start by a discussion of Hagen, Kerkhoff and Gussenhoven (2011) which is the only previous study of foreign accents in song that I am aware of. I then review the literature on the accents of professional singers who are native speakers of English. It is common for such singers to sing in an accent that is different from their native accent, the accent that they use when they talk. The researchers who have explored professional singers' accents have almost unanimously concluded that the observable accent shifts are due to social and cultural factors (for example, professional singers shift their accent to sound similar to an artist they want to be associated with). Because of this conclusion, the research on accent shifts in professional singers is not directly pertinent to the research question I pursued here. However, I will review this literature anyway because it encompasses the majority of research on accents in song in general. Also, I do agree that the accent shift often occurs for cultural reasons, but only under specific circumstances which are described in the literature reviewed here, and the current study specifically tries to avoid those circumstances. The section on accents in music also includes a brief summary of how singing affects the pronunciation of vowels. I complete the literature review with a discussion of some of the main findings in the research on foreign accents in speech.

1.1 Background Information

1.1.1 Foreign accents in music

Hagen, Kerkhoff and Gussenhoven (2011) conducted a study on foreign accent in song compared to speech. They investigated whether it is more difficult to detect a foreign accent due to the following things:

1. the realization of segments being more authentic in song
2. listeners judging accentedness less harshly in song compared to speech
3. the fundamental frequency (F0) and duration imposed by the music

In order to investigate these three hypotheses, Hagen and colleagues recruited eleven Dutch students with English as a second language (L2) and six native English speakers. The researchers recorded the speakers reading and singing 11 passages; therefore, there were a total of 34 experimental stimuli. The researchers also manipulated the reading recordings in three different ways. In the first, the durations of all utterances of the non-native speakers were matched with the duration of a native English speaker at the segmental level. The first manipulation intended to explore whether duration plays a role in identifying an accent. In the second pitch was monotonized. The second manipulation intended to explore the role pitch plays in identifying an accent. In the third, the duration was manipulated like in the first and the pitch was also monotonized. The third manipulation intended to explore the role pitch and duration together play in identifying an accent. Therefore, there were 44 manipulated stimuli. The stimuli was separated into two separate tests, spoken and sung. The researchers then recruited twenty listeners who were native speakers of English, where ten were exposed to the spoken test and ten were exposed to the sung test. The listeners were asked to rate how native the speaker's accents sounded on a scale from 1 to 7.

The researchers found that the listeners rated non-native speakers as having less of a perceived accent in song compared to speech. The researchers also found that listeners rated the recordings where the duration was manipulated to be similar to an English speaker as having less of a foreign accent. They also found that the listeners rated the pitch monotonized recordings as having less of a foreign accent compared to the original recordings. The researchers concluded that these differences were due to the unavailability of F0 and duration cues to signal a foreign accent in song.

Hagen, Kerkhoff and Gussenhoven (2011) concluded that non-native speakers have less of a perceived foreign accent in song due to the imposition of pitch and duration which masks the signals of a foreign accent. This study is similar to the one conducted in this thesis, but the details of the experimental design differ, as will become evident below. Also, Hagen, Kerkhoff and Gussenhoven had a small sample size whereas I have a larger sample size. Moreover, Hagen and colleagues recruited only Dutch speakers whereas we have recruited individuals with various first languages. Finally, my study differs in that I do not use trained singers.

1.1.2 Accent shifts in professional singers

British singers often shift their accent when they are singing compared to when they are speaking. For example, British singer, Alesha Dixon made the news in the summer of 2015 when she sang *God Save the Queen* at the European Grand Prix in an American accent (see, for example, Rudgard, "Why you put on an American accent when you sing" (2015) in *The Telegraph*; Calderwood, "'Gahaad save our Queen!' Alesha Dixon sets Twitter alight for singing national anthem at British Grand Prix with an American accent (and getting the lyrics wrong)" (2015) in *The Daily Mail*). This is also true for other English dialects, such as Australian and New Zealand (O'Hanlan, 2006; Gibson, 2010; Bell & Gibson, 2011). This shift towards an American accent in song has been noted among lay people and is often discussed in on-line chat forums and discussion groups[1], as well as columns and articles in newspapers such as *The Telegraph* and *The Daily Mail* [2](cited above). Some linguists have also studied accent and music in professional singers with English dialects. Please note that the term dialect is used in reference to regional variation of all aspects of language (lexicon, syntax, pronunciation, etc.), whereas an accent refers specifically to pronunciation (Wardhaugh, 1998).

Early research in this field was conducted by Peter Trudgill (1983). Trudgill (1983) conducted a dialect analysis of accent of British singers in popular music. He claims that the shift in accent was due to "cultural domination" of American music in the 1950s. Cultural domination means that American English music is at the top of the popular music scene; therefore, singers shift to sound more like American singers (Schulze, 2014). The analysis was conducted on

[1] See, for example: "More on Young New Zealand English," http://dialectblog.com/2013/12/16/young-new-zealand-english/ and "Why don't you hear someone's accent in song" http://mentalfloss.com/article/29780/why-dont-you-hear-someones-accent-song.
[2] See also, Gibsone, H., (2016)'s "Talking tactics: Rihanna and the pop stars who change accent" in *The Guardian* and Guo, J., (2016)'s "How Iggy Azalea mastered her 'blaccent'" in *The Washington Post.*

popular British singers, such as Paul McCartney from the Beatles and Mick Jagger from the Rolling Stones. Trudgill proposed that there are major tendencies that British singers follow in order to shift their accent closer to American English. These tendencies are not always followed and there is much variability (Trudgill, 1983; Simpson, 1999), but researchers have nevertheless used these tendencies in order to determine the degree to which singers follow and/or deviate from an American English accent when singing (Simpson, 1999; Morrissey, 2008; Beal, 2009; Gibson, 2010). Paul Simpson (1999) calls this list of tendencies the USA-5 model. The tendencies are as follows:

1. Realization of intervocalic /t/ as the British variants [t] and [ʔ] is not permitted.
2. Words such as *dance*, *last* and *can't* are not permitted to be realized with the vowel /aː/.
3. Non-prevocalic /r/ is generally pronounced.
4. Words such as *life* and *my* are pronounced with a vowel like [aˑ] instead of a diphthong.
5. Words such as *love* and *done* are pronounced with a vowel like [əˑ] instead of a British variant.
6. Words such as *body* or *top* are pronounced with an unrounded vowel [ɑ] instead of the British variant [ɒ].

(Trudgill, 1983; Schulze, 2014)

Trudgill claims that motivation is a key factor in explaining a speaker's accent shift. Motivation to either sound more or less American depends on the singer's genre and cultural domination (Trudgill, 1983; Simpson, 1999; Gibson, 2010; Schulze, 2014). Speakers wish to alter their linguistic behaviour to resemble groups that they wish to identify with (Trudgill, 1983; Bell, 1984).

Even though Trudgill (1983) is not able to account for the inconsistency and variability of a speaker's shift of accent, several sociolinguistic theories have been built on Trudgill's foundational claims. Simpson (1999) built on the USA-5 model by taking context and discourse into account. Specifically, Simpson draws upon the USA-5 model as a general set of characteristics for accent shift and changes in cultural patterns as the main reason why the shift occurs. These changes in cultural patterns may include changes in music preference and cultural domination (Simpson, 1999; Gibson, 2010). A singer who wishes to identify with a specific genre will alter accent (Stone et. al, 1999; Simpson, 1999; Beal, 2009; Gibson, 2010; Gibson, 2011; Schulze, 2014). For example, punk singers will show more variability in the USA-5 model (i.e. less consistent shifts towards the American pronunciation) as they intend to sound more "grungy" than other singers, whereas pop singers show less variability (Simpson, 1999).

The majority of the research on accent in music focuses on how singers modify their accents to sound more American. However, Morrissey (2008) analyzes singers who stray away from an American accent model and maintain their British accent or alter their accent to sound more British (for example Pink Floyd's singer Syd Barrett). Morrissey (2008) found that when looking at accent shifts in music we must look at reference style. Specifically, singers shift their accent either to an unmarked or marked style. An unmarked style is when a singer is projecting a style where the genre and the audience are the same. A marked style is when a singer is projecting a style that deviates from the genre and reference as it is marked with regional meaning (Morrissey, 2008). Morrissey claims that projecting a marked or unmarked style happens unconsciously.

Beal's (2009) analysis of Alex Turner, the singer of the indie music band the Arctic Monkeys, also provides evidence for the claim that professional singers shift their accent either to sound closer to or deviate from their intended audience. Beal argues that indie music singers wish to preserve their accent and authenticity. Beal (2009) found that Alex Turner uses lexical items and linguistic features from his native town Sheffield. Incorporating these linguistic features in song provides evidence that the Arctic Monkeys are preserving their native accent and style (Beal, 2009).

In addition to research on cultural and sociolinguistic aspects of accents in song, a few studies have examined the effect of singing per se on pronunciation. For example, Gibson (2010, 2011) points out that vowels are generally longer in singing than in speaking. This is due to the rhythmic nature of song; vowels need to be lengthened in order to follow the beat. It has also been found that mean pitch (F0) is higher when individuals are singing than when they are speaking (Stone et al., 1999; Clermont, 2002; Gibson, 2010; Gibson, 2011).

Also, formants (F1 and F2) shift in singing. Specifically, F1 is higher in singing compared to reading (Sundberg, 1969; Sundberg & Skoog, 1997; Stone et al., 1999; Clermont, 2002; Gibson, 2010; Gibson, 2011). F2 frequencies are lower for front vowels in singing compared to reading (Sundberg, 1969; Stone et al., 1999; Clermont, 2002; Gibson, 2010; Gibson, 2011). A higher F1 and lower F2 indicate that singers open their jaw more in song compared to speech. However, it is important to note that F1 and F2 differences of speech and song vary across individuals, different genres and by pitch (Sundberg & Skoog, 1997; Stone et al., 1999; Armstrong, 2004; Coupland, 2011; Bicknell, 2015; Eberhardt & Freeman, 2015). These differences are due to different singing techniques. Also, previous studies have focused on professional singers, who have experience and are trained singers. The difference of F1 and F2 in song compared to speech found in these studies may be due specifically to singers being trained to manipulate their jaw while singing.

The studies discussed in this section have in common with the present research that they focus on accent shifts in singing. However, their focus is different as they explore accents of professional singers, whereas the present study intends to investigate whether untrained singers alter their foreign accent in song. The main consensus is that the major factors contributing to accent shifts in song are social and cultural. In order to explore whether the accent shift is about music per se, or about music training, my study focuses on untrained singers and tries to steer clear of any musical genre that may be associated with a particular region or artist. Using untrained singers and a familiar musical genre will explore whether it is still more difficult to detect an accent when someone is singing compared to reading, which suggests that difficulty with accent detection is due to the complexity of music itself rather than musical training. The next section provides an overview of research on foreign accents in speech, as there is no work on foreign accents in song.

1.1.3 Foreign Accent in Speech

Foreign accent is when the linguistic parameters of an individual's first language (L1) clash with the linguistic parameters of his/her L2. A perceived foreign accent is when a native speaker judges a foreign accent based off of their deviation from the phonetic form (Hagen et al., 2011). This project aims to investigate the vowel and prosody features that influence foreign accent production and perception in music.

The vowel inventory of individuals' L1 influences their L2 pronunciation, and improves as individuals use their L2 more often (Flege et al., 1997; Flege 1981). The closer the individuals' L1 vowel inventory is to their L2 vowel inventory, the stronger the foreign accent compared to if their L1 were further away from their L2. This is due to speakers' requiring to add additional phonetic features from their L2 to their L1 rather than creating new phonetic categories. This leads to a stronger perceived foreign accent as their L1 is competing with their L2. For example, a perceptual and production experiment found that L1 German speakers were less successful at producing and perceiving English vowels compared to similarly matched L1 Spanish speakers (Flege et al., 1997). This illustrates that speakers' L1 vowel inventory influences the way they perceive and produce their L2 (Flege et al., 1999).

Prosodic features are important in foreign accent perception. Perceptual experiments have found that prosodic features play a key role in determining the intelligibility of speakers' foreign accent (Anderson-Hsieh et al., 1992; Munro & Derwing, 1995; Munro & Derwing, 1999). For example, a study conducted on non-native speakers' accentedness found a correlation between intelligibility and prosodic variables. The researchers found that native speakers rated non-native speakers as being less intelligible when prosody was poor even if the speaker had good pronunciation. This suggests that poor prosody impacts intelligibility more than aspects of pronunciation (Anderson-Hsieh et al., 1992).

Also, speech synthesis manipulations provide evidence that prosody plays a role in identifying individuals' foreign accent (Boula de Mareuil & Vieru-Dimulescu, 2006; Vassiere & Boula De Mareuil, 2007; Munro, 2008). For example in an experiment that transplanted Spanish prosody on Italian phrases and Italian prosody on Spanish phrases, listeners had more difficulty determining the language of the speaker in the manipulated phrases compared to regular prosodic phrases (Vassiere & Boula De Mareuil, 2007). In another experiment, recordings of non-native speakers were played backwards and native speakers were asked to judge whether the speaker was a native English speaker. The reversed recordings intended to eliminate the phonetic pronunciation and leave the prosodic parameters. The results found that listeners were able to accurately judge whether the speaker was native and non-native (Munro, 2008).

This section addresses the third research question that is being investigated, whether prosody makes it more difficult to detect a foreign accent. Perhaps it is more difficult to detect a foreign accent in song, as prosodic features are leveled out in song (Bicknell, 2015).

1.2 Preliminary research

This thesis was inspired by the preliminary project and analysis carried out by Mekik and Boucher (2015) and Mageau, Mekik and Wang (2015). This project included ten speakers, four native English and six non-native English. Speakers were recorded singing *Twinkle, Twinkle, Little Star* accompanied and not accompanied by piano music and reading *Goldilocks and the Three Bears*. Eleven listeners were recruited to judge the language of the speaker in the recording. It was found that listeners were less accurate at judging a speaker's accent when singing with piano accompaniment (29%) than when reading (69%). This pilot study suggests that it is in fact more difficult to detect a foreign accent in song compared to speech.

1.3 *Goals of the present thesis*

Considering the results from the pilot study together with the literature on foreign accents (section 1.1.3), I hypothesize that it is more difficult to detect a foreign accent in song than in speech. I further hypothesize that this is due (at least in part) to duration and pitch being imposed by the music, eliminating cues to detecting an accent.

I conducted three studies to tackle my research questions. In Study 1 native and non-native Canadian English (CE) speakers sang a "familiar-genre" song, *Twinkle Twinkle Little Star,* and read a short excerpt from *Goldilocks and the Three Bears.* Native English listeners were recruited in order to judge if the voice belonged to a native English speaker or not. Results from a pilot study suggested that it is more difficult to detect a foreign accent in song compared to speech, and this result was replicated in Study 1.

In Study 2 the segmental (F1 and F2) and suprasegmental (pitch and duration) features of the vowels were analyzed in order to investigate the phonetic differences between native and non-native singers in song and speech which may have given rise to the results found in Study 1. Study 3 used speech manipulation in order to isolate the pitch and prosodic features from native and non-native English speakers. Native English listeners were recruited in order to judge if the voice belonged to a native English speaker or not. Study 3 was intended to investigate the importance that prosody plays in foreign accent perception.

This study aims to determine whether it is more difficult to detect an accent in song than in speech also when speakers have no musical training, as a consequence of music itself rather than musical training.

Chapter 2

Methods

The three studies conducted in order to address the current research questions made use of native and non-native speakers singing and reading. This section presents the specifics of the speakers, materials and procedure for the recordings. The recordings used in the studies are described in Chapter 3, 4 and 5.

2.1 Speakers

Twelve female speakers were recruited from the Ottawa area. The speakers ranged from 20 to 44 years of age and were compensated for their participation with $5 Tim Horton's gift cards. Six of the speakers were native Canadian English and six were non-native Canadian English speakers who differed in their degrees of foreign accents.

None of the speakers had any previous musical training. Musical training has not been found to affect a speaker's degree of foreign accent when speaking (Flege et al, 1995); however, musical training has been shown to affect the ability to mimic unfamiliar speech sounds (Piske et al. 2001). Therefore, it is important that no speaker has musical training, in case they would attempt to use their musical knowledge to mimic unfamiliar sounds of their L2. Also, trained singers might manipulate their phonetic parameters when they sing as they might have been trained to do so when singing a specific genre (Sundberg & Skoog, 1997; Stone et al., 1999; Armstrong, 2004; Gibson, 2010; Gibson, 2011; Coupland, 2011; Bicknell, 2015; Eberhardt & Freeman, 2015). In what follows, the speakers will be identified as either a native or a non-native speaker of English.

2.1.1 Native

The native speakers had been exposed to English since birth and had little exposure to a second language.

2.1.2 Non-native

The non-native speakers had not been immersed in an English dominant environment until adulthood. However, all speakers had some English exposure and/or training while growing up.

The non-native speakers' first languages included: Farsi (length of L2 immersion = 2 years), French (length of L2 immersion = 2 years), Romanian (length of L2 immersion = 10 years), Spanish (length of L2 immersion = 1 year), Tamil (length of L2 immersion = 15 years). The fact that the non-native speakers' first languages and length of exposure differ are not confounding variables as language transfer or a measure of L1 influence will not be addressed in this project. As was previously mentioned, individuals' L1 plays a role when producing their L2; therefore, it might seem that the speakers' L1 influence the results of our study. However, we are not interested in language specific differences, we are simply looking at the distinction between foreign and native accent. Also, individual differences will be measured and reported when analyzing vocalic properties.

2.2 Materials

The materials consisted of the first two verses of the song *Twinkle, Twinkle, Little Star* (Appendix A)[3], and the first paragraph from the story *Goldilocks and the Three Bears* (Appendix

[3] The song *Twinkle, Twinkle Little Star* and the first paragraph of *Goldilocks and the Three Bears* was retrieved from internet sources.

B). *Twinkle, Twinkle, Little Star* was used because of its familiarity and it does not appear be associated with a specific dialect, geographical region or singer. It is important that the music excerpt is not associated with any geographically specified musical genre likely to influence the singer's linguistic style or motivate the singer to sing like a celebrity of the specific genre (Simpson, 1999; Beal, 2009; Gibson, 2010; Gibson 2011). *Goldilocks and the Three Bears* was chosen as it is not associated with any author in particular, eliminating the possible influence of a specific author.

2.3 Procedure

The speakers were first asked to answer four language background questions:

1. What is your first language?
2. How old were you when you first learned English?
3. How long have you been in an English environment?
4. What other language(s) do you speak, if any?

The language background questions provide insight into the speakers' linguistic profiles. The speakers were recorded in a quiet room.

There are four conditions in which speakers were recorded twice. The conditions are as follows:

i) singing *Twinkle, Twinkle, Little Star* with piano accompaniment

ii) singing *Twinkle, Twinkle, Little St*ar without piano accompaniment

iii) reading *Twinkle, Twinkle, Little Star*

iv) reading *Goldilocks and the Three Bears*

The piano accompaniment was very basic and consisted of just the melody without harmonies. Also, before the actual recordings began, the speakers were given practice trials in order to familiarize themselves with the material. All of the speakers were familiar with *Twinkle, Twinkle, Little Star,* which was important as the researchers did not want to sing the song to the speaker to avoid influencing them. Each condition was recorded twice. We[4] collectively chose the best recording for each condition.

[4] Can Mekik and Emily Wang assisted with the twelve recordings. Can Mekik also coded the PsychoPy experiment.

Chapter 3

Study 1

The purpose of Study 1 was to determine if listeners are able to correctly judge whether a speaker has a native English accent or a foreign English accent when speaking or singing. Listeners were exposed to the recordings of *Twinkle, Twinkle, Little Star* and *Goldilocks and the Three Bears*. They were asked the following question: "Was Canadian English the first language of the speaker or singer?" Press y for "yes" or n for "no".[5] If they answered "no," they were then asked to judge how native-like the speakers sounded. Based on Hagen et al. (2011) as well as the pilot study (Hagen, Kerkhoff & Gussenhoven, 2011; Mekik & Boucher, 2015) it was hypothesized that it is more difficult to detect a foreign accent in song than in speech.

3.1 Listeners

Forty native Canadian English speaking listeners were recruited from the Ottawa area. The listeners were paid for their participation. They were asked to complete a language background questionnaire (Appendix C) in order to gain insight into their linguistic profiles.

3.2 Materials

The materials consisted of a total of 48 recorded excerpts from *Twinkle, Twinkle, Little Star* and *Goldilocks and the Three Bears*, which were sung or read by twelve speakers (native = 6; non-native = 6). As previously mentioned, the four conditions were as follows: *Twinkle, Twinkle, Little Star* sung twice (once with piano accompaniment and once without piano accompaniment) and read once, and the beginning of *Goldilocks and the Three Bears*.

---

[5] It could be argued that it would have been preferable to simply ask about "English" instead of "Canadian English" We will return to this issue in Chapter 6.

The recordings were divided into four different blocks. There were 12 recordings in each block; all blocks contained an equal amount of each condition (i.e. three recordings of a speaker singing *Twinkle, Twinkle, Little Star* with piano accompaniment, three recordings of a speaker singing *Twinkle, Twinkle, Little Star* without piano accompaniment, three recordings of a speaker reading *Twinkle, Twinkle, Little Star*, and three recordings of a speaker reading the first paragraph of *Goldilocks and the Three Bears)*. In each block there was one recording per speaker, and the conditions were randomized. The design is illustrated in the table presented in Appendix D. It was important that each listener did not hear a speaker more than once. I wanted to avoid the possibility of a listener recognizing the voice of a specific speaker. This could have altered their response if for example they heard a speaker read first and sing in a later recording or vice versa. The decision they made the first time might influence their second decision (Schacter & Church, 1992; Church & Schacter, 1994). This could have altered the listeners' judgements and led to misleading information.

3.3 Procedure

The experiment was administered to the listeners in a quiet room. Listeners were seated in front of a laptop and instructed to wear headphones. The listeners heard one of the four blocks, which contained twelve sound files. After each sound file, the listeners were asked to judge whether the speaker was a native Canadian English speaker or not. If listeners responded (correctly or incorrectly) that the speaker was not a native Canadian English speaker, they were asked to rate how close the speaker sounded to Canadian English.

The listeners rated the speakers' accents by using a 9-point Likert scale, which has been used in previous studies examining the degree of foreign accent, due to its sensitivity to small phonetic differences (Anderson-Hsieh et al., 1992; Flege et al, 1992; Munro & Derwing, 1995;

Piske et al., 2001). Also, previous research has found that an interval scale can provide indicators of accentedness and a good measurement of perceived foreign accent (Southwood & Flege, 1999). Listeners were instructed that a 1 on the scale corresponded to "closer to native Canadian English" and a 9 corresponded to "less close to native Canadian English" (Magen, 1998).

The material was presented through a software called PsychoPy (Peirce JW, 2009), where the listeners' raw responses were stored in a format compatible with a Microsoft Excel worksheet. The responses were categorized according to whether the listener correctly identified if the speaker were a native or a non-native speaker. If listeners responded that the speaker is not a native Canadian English speaker, their score on the 9-point Likert scale was recorded.

3.4 Results

Forty listeners were recruited and asked to judge 12 recordings (i.e. one of the four blocks). In total there were 480 responses (96 responses in each condition). In total the listeners responded to 240 recordings of speakers singing and 240 with speakers reading. Half of these responses were of native English speakers and the other half of non-native English speakers. We hypothesize that the listeners will have more difficulty detecting a foreign accent in the singing conditions compared to the reading conditions. Specifically, we hypothesize that listeners will have more difficulty when the speaker is singing *Twinkle, Twinkle, Little Star* with musical accompaniment, followed by the speaker singing *Twinkle, Twinkle, Little Star* with no musical accompaniment, reading *Twinkle, Twinkle, Little Star* and the least difficulty with when the speaker is reading *Goldilocks and the Three Bears.* We hypothesize that the singing conditions would be different due to the added complexity that the accompaniment provides. We hypothesize that the reading conditions would be different due to the poetic rhythm of *Twinkle,*

Twinkle, Little Star. We also hypothesize that listeners will rate non-native speakers as having less of a foreign accent in song compared to speech.

There were a total of 480 guesses in Study 1. Of these guesses there were 70 incorrect responses (15%) and 410 correct responses (85%).

Figure 1. *Incorrect Responses by Condition*

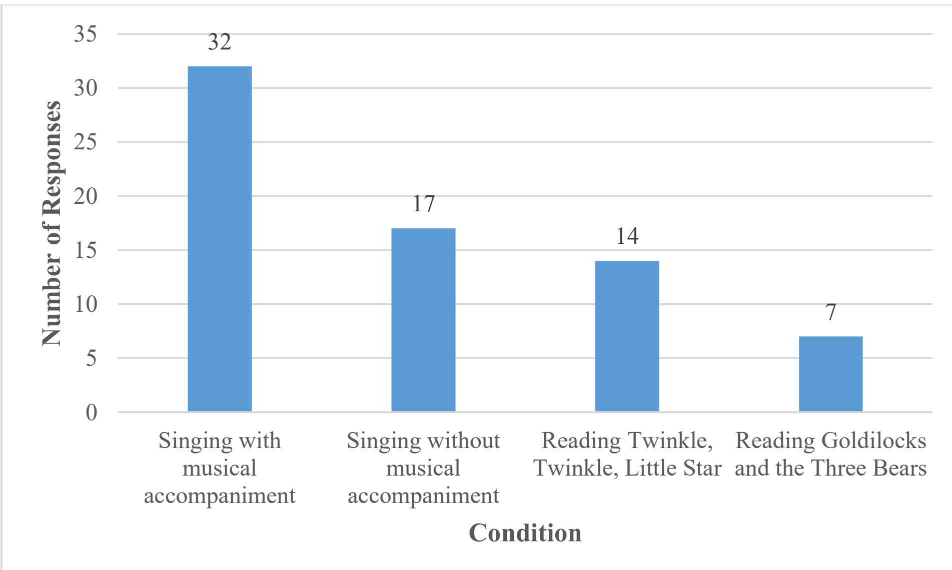


Figure 1 illustrates that of the incorrect responses (n = 70) majority of the responses are in the singing conditions (n= 49). Specifically, 70% of the incorrect responses are in the singing conditions and the remaining 30% are in the reading conditions. This suggests that a listener who has difficulty detecting a foreign accent has greater difficulty when the speakers are singing compared to when they are reading. Figure 1 also illustrates that listeners have more difficulty detecting a foreign accent when the speakers are singing with musical accompaniment (46%), followed by speakers singing without musical accompaniment (24%), reading *Twinkle, Twinkle, Little Star* (20%) and reading *Goldilocks and the Three Bears* (10%). These descriptive results suggest that our hypothesis was correct as the participants made more errors detecting a foreign accent in song with musical accompaniment than song without musical accompaniment than reading *Twinkle, Twinkle, Little Star* and finally reading *Goldilocks and the Three Bears.* The results illustrated in Figure 1 replicate the pilot study that was conducted in the summer of 2015 which was also looking at foreign accents in song compared to speech. This suggests that a listener has more difficulty detecting a foreign accent when speakers are singing compared to when they are reading.

A proportion test was also used in order to determine whether the descriptive statistics were significant. According to the results from the proportion test, there was a statistically significant increase of incorrect response for the singing conditions compared to the reading conditions (p <0.01). These results suggest that I was able to confirm my hypothesis that it is more difficult to detect a foreign accent in singing compared to speaking.

I used R (R Core Team, 2015) and the lme4 package (Bates, Maechler, Bolker and Walker, 2015) to perform a linear mixed effects analysis to illustrate the relationship between accuracy and the interaction of nativity and condition. As a fixed effect, I entered the interaction of nativity and condition into the model. As random effects, I entered participants and speakers as these variables could play a role in accuracy. P-values were obtained using a likelihood ratio test of the full model with and without the interested effect (i.e. the interaction of native/non-native and condition). The likelihood ratio test found that there was a significant difference between the two models which indicates that accuracy is significantly different in each condition between native and non-native speakers ($\chi 2(7)$= 4.57, p <0.0001), where the difference in condition is about 0.11 +/- 0.2 (standard error).

Interestingly, the actual recordings from the two singing conditions for each speaker were the same. The recordings were of the singers singing *Twinkle, Twinkle, Little Star* with piano accompaniment. The recordings presented to the listeners were one with musical accompaniments and one without. This suggests that the musical accompaniment adds additional complexity to the recording, making it more difficult to detect a foreign accent. However, recall that the accompaniment was very basic, consisting of simply playing the bare melody on a piano. Also, it is interesting that the reading *Twinkle, Twinkle, Little Star* condition (20%) is similar to the singing *Twinkle, Twinkle, Little Star* without accompaniment condition. I hypothesize that this similarity is due to the rhythmic nature of *Twinkle, Twinkle, Little Star* which affects duration, and possibly also pitch, even when being read rather than sung.

Figure 2. *Listener's Judgment of Non-Native Accent*

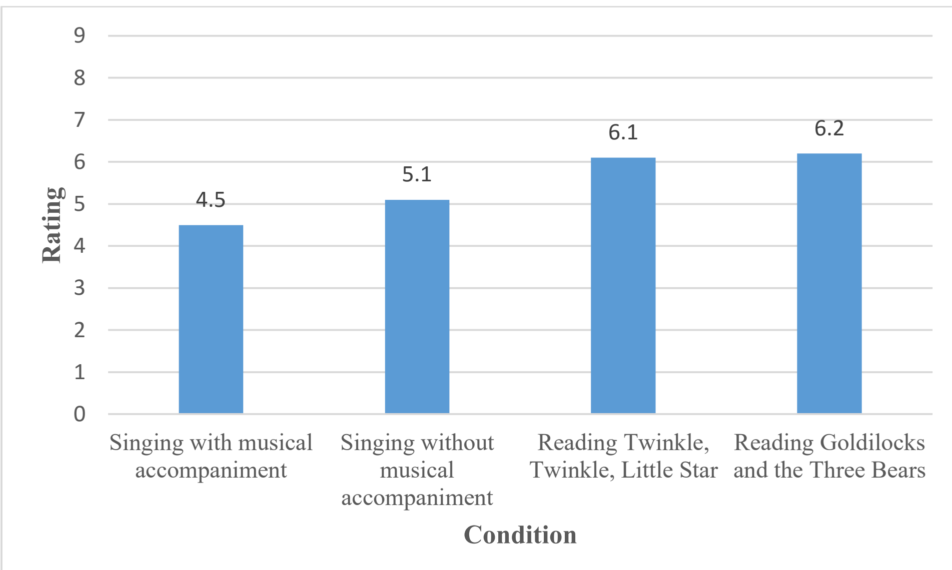


Table 1. *T-values and p-values comparing scores per condition*

Comparing Conditions	t-values	p-values	significance[6]
Singing vs. Reading	3.83	0.0002	***
Singing with musical accompaniment vs. Singing without accompaniment	1.34	0.18	
Reading *Twinkle, Twinkle* vs. Reading *Goldilocks*	0.13	0.89	
Singing with musical accompaniment vs. Reading *Goldilocks*	3.38	0.001	**
Singing with musical accompaniment vs. Reading *Twinkle, Twinkle*	3.33	0.001	**
Singing without musical accompaniment vs. Reading *Goldilocks*	2.18	0.03	*
Singing without musical accompaniment vs. Reading *Twinkle, Twinkle*	2.1	0.03	*

Figure 2 illustrates that native English listeners rate non-native speakers as having less of a foreign accent in the singing conditions compared to the reading conditions.

I used R (R Core Team, 2015) and the lme4 package (Bates, Maechler, Bolker and Walker, 2015) to perform a linear mixed effects analysis to illustrate the relationship between rating and speech or song. As a fixed effect, I entered the speech/song variable into the model. As random effects, I entered participants, speakers and conditions as these variables could play a role in the rating. P-values were obtained using a likelihood ratio test of the full model with and without the interested effect (i.e. the difference in speech and song). The likelihood ratio test found that there was a significant difference between the two models which indicates that speech

[6] * illustrates significant p-values < 0.05; ** illustrates significant p-values < 0.01; *** illustrates significant p-values < 0.001

and song have an effect in ratings ($\chi 2(1)= 5.39$, $p= 0.02$), where the rating in song is lower by about 1.1 +/- 0.4 (standard error). [7]

A Welch Two Sample t-test was used in order to determine whether the ratings between the singing and reading conditions were significantly different. A Welch t-test was used because the responses in each condition were of unequal variances due to having a different number of responses in each condition. It was found that these differences were significant ($T(208.94) = 3.83$, $p < 0.01$). There was no significant differences between the two singing conditions ($T(95.41)= 1.34$, $p= 0.18$). There was also no significant differences between the two reading conditions ($T(108.91)= 0.13$, $p= 0.89$). There was a significant difference between singing *Twinkle, Twinkle, Little Star* with musical accompaniment and reading *Goldilocks and the Three Bears* ($T(97.29)= 3.38$, $p < 0.01$). There were also significant differences between singing *Twinkle, Twinkle, Little Star* with musical accompaniment and reading *Twinkle, Twinkle, Little Star*; singing *Twinkle, Twinkle, Little Star* without musical accompaniment and reading *Goldilocks and the Three Bears*; singing *Twinkle, Twinkle, Little Star* without musical accompaniment and reading *Twinkle, Twinkle, Little Star.* This suggests that there is a clear difference between both singing conditions and both reading conditions. The significant differences are illustrated in Table 1 in order to easily identify the differences between the scores in the conditions.

These results suggest that listeners perceive non-native speakers as having less of a foreign accent when they are singing compared to when they are reading. Similarly to the results

---

[7] When conducting mixed-effects models in R p-values are not computed; therefore, we used a likelihood ratio test. We also used formula 2 * (1-pt(abs(-2.943), 265-2)) included in *Analyzing Linguistic Data: A Practical Introduction to Statistics using R*, Baayen (2008, p. 248) which computes the upper bound of the degrees of freedom and calculates an estimated p-value using the t-statistic. The result also suggests that there is a significant difference of rating between speech and song ($t(1)= 2.94$, $p < 0.01$).

found in the study conducted by Hagen and colleagues (2011), my results suggest that there is less of a perceived foreign accent when non-native speakers are singing compared to when they are reading.

The results from Study 1 suggest that a native listener has more difficulty determining whether an accent is present in the singing condition with musical accompaniment most often and in reading regular prose the least often. The results from Study 1 also suggest that there is less of a perceived foreign accent in song compared to speech. The following two studies explore the reasons behind the results of Study 1. Study 2 compares the properties of vowels in speech and song. Study 3 analyzes the prosody of song and speech.

3.5 Summary of Study 1

The aim of Study 1 was to determine whether it is more difficult to detect an accent in song than in speech. Also, Study 1 intended to investigate whether listeners perceived less of a foreign accent in song compared to speech. The results presented in section 3.4 suggest that it is indeed more difficult to detect a foreign accent in song as 46% of the incorrect responses were in the *singing with musical accompaniment* condition. The results also suggest that listeners have less of a perceived foreign accent in song compared to speech which is evident in the lower accent ratings in song.

Chapter 4

Study 2

The results from Study 1 suggest that native listeners have more difficulty detecting an accent in singing compared to reading. Study 2 explores the phonetic properties that could have led to these results. The analysis focuses specifically on vowels.

4.1 Materials

The materials for Study 2 consist of a subset of the recordings discussed in section 2.2. Specifically, Study 2 used the recordings where i) *Twinkle, Twinkle, Little Star* was sung with piano accompaniments and ii) *Twinkle, Twinkle, Little Star* was read. A total of 24 recordings were included in Study 2 (native = 6; non-native = 6).

4.2 Method

I analyzed the acoustic properties of the vowels in the reading and singing recording conditions. The measurements were conducted using PRAAT (Boersma & Weenink, 2001). Each speaker was analyzed individually. All of the vowels were analyzed; however, the only vowels being reported are /ɪ/ and /ʌ/. These are the only vowels being reported as there are more instances of both compared to the other vowels. I also excluded diphthongs, sing splitting the vowel from the diphthong proved to be complex and led to concerns of inaccurate reporting.

Vowel duration was measured from the offset of the pre-vocalic consonant to the onset of the post-vocalic consonant. The frequencies of fundamental frequency (F0) and the first two formants (F1-F2) were measured at the midpoint of the vowel, as is common place in vowel analyses (e.g. Flege et al., 1997).

4.3 Results

4.3.1 Autosegmental features: duration and pitch

I hypothesize that vowels will be significantly longer in the singing condition compared to the reading condition. This is due to the rhythmic nature of song where duration is dictated by the music (Gibson, 2010). Table 2 presents the average means of duration for vowels /ɪ/ and /ʌ/ (in seconds). The mean values illustrate that vowels are always longer in song compared to speech across speakers, both native and non-native (Appendix E (i) – (iv)). There also appears to be greater variation in song compared to speech which is illustrated through the higher range values in song (Appendix F). A Wilcoxon signed-rank test was conducted on individual speakers to determine the statistical significance of the durational differences in song compared to speech. The results found that these differences were significant per speaker with p-values of <0.001.

Table 2. *Average duration values of vowels ɪ and ʌ in speech and song*

	Speech		**Song**	
Native	*ɪ Average*	*ʌ Average*	*ɪ Average*	*ʌ Average*
	0.07	0.09	0.15	0.17
Non-Native	*ɪ Average*	*ʌ Average*	*ɪ Average*	*ʌ Average*
	0.07	0.10	0.18	0.18

The results of the duration values suggest that vowels are always longer in song compared to speech, with no differences between native and non-native speakers.

Previous research has found that the fundamental frequency (F0) is higher in singing compared to speaking (Sundberg & Skoog, 1997; Stone et al. 1999; Gibson, 2010). I expected that pitch will be higher in the singing condition compared to the reading condition in the native and non-native speakers. This would indicate that pitch is higher in song than in speech. The mean pitch values presented in Table 3 suggest that pitch is typically higher in song compared to speech; however, this does not hold at the individual level (for example, S4, S5 and S6 presented in Appendix G). A Wilcoxon-signed rank test was also conducted to determine whether there were statistically significant differences between speech and song. There were no overall patterns as only some native and non-native speakers had significant differences. Where I notice a difference is in the range of the pitch values. Table 3 also illustrates that the range values are overall higher in song compared to speech which suggests that there is more movement occurring in the singing condition compared to the reading condition (Appendix G). This is also illustrated in Appendix E (section (v)-(viii)) where there is more movement in song compared to speech for almost all speakers. These results suggest that pitch is not always necessarily higher in song but it is different. There was no consistent pattern as the difference for some speakers was significant and others was not. I speculate that this is due to some speakers reading and singing in what sounded like a similar pitch. I also speculate that the difference between our results and the results of previous researchers (e.g. Stone, Cleveland & Sundberg, 1999) is due to the use of untrained singers.

Table 3. *Average pitch values in Hz of vowels ɪ and ʌ in speech and song*

	Speech				**Song**			
Native	*ɪ Average*	*ʌ Average*	*ɪ Range*	*ʌ Range*	*ɪ Average*	*ʌ Average*	*ɪ Range*	*ʌ Range*
	188	174	120	61	210	214	129	116
Non-Native	*ɪ Average*	*ʌ Average*	*ɪ Range*	*ʌ Range*	*ɪ Average*	*ʌ Average*	*ɪ Range*	*ʌ Range*
	206	187	113	83	209	223	134	136

4.3.2 Segmental features: formant shift

Sundberg (1969), Stone et al. (1999), Clermont (2002), Gibson (2010) and Gibson (2011) all find that F1 is higher in song compared to speech; therefore, we expect to find that native and non-native speakers all have higher F1 in song. This would indicate that vowels are realized as being lower in singing than in reading (Gibson, 2010). Based on the average values presented in Table 4 there is no apparent trend or generalization (Appendix H). This is also presented in the graphs in Appendix E ((ix)-(xii)) which illustrates that F1 values do not always increase and or decrease in song compared to speech. There is also no difference between native and non-native speakers.

Table 4. *Average F_1 values of vowels ɪ and ʌ in speech and song*

	Speech		**Song**	
Native	*ɪ Average*	*ʌ Average*	*ɪ Average*	*ʌ Average*
	526	704	520	705
Non-Native	*ɪ Average*	*ʌ Average*	*ɪ Average*	*ʌ Average*
	400	717	452	730

Also, based on previous results in the literature, the native and non-native speakers are expected to all have lower F2 frequencies in the singing condition compared to the reading condition. This would suggest that vowels are more back when singing than when reading. Based on the average F2 values presented in Table 5, it appears there is no difference between speech and song in native or non-native speakers. Also, the mean, median and range values for individual speakers indicate no clear pattern in song compared to speech (Appendix I). Similarly to F1 frequencies, there does not appear to be a difference between native and non-native speakers (Appendix E (xiii)-(xvi)).

Table 5. *Average F_2 values of vowels ɪ and ʌ in speech and song*

	Speech		**Song**	
Native	*ɪ Average*	*ʌ Average*	*ɪ Average*	*ʌ Average*
	1969	1503	2122	1571
Non-Native	*ɪ Average*	*ʌ Average*	*ɪ Average*	*ʌ Average*
	1762	1388	1915	1522

Wilcoxon signed-rank tests were conducted in order to determine whether there were any statistically significant differences of F1 and F2 between speech and song per speaker. Significant differences could have suggested that there is an underlying pattern in regard to formant frequencies which was not evident in the descriptive statistics. However, neither set of tests revealed any clear generalizations which suggests that formant differences do not play a clear and significant role in detecting an accent as there are no general patterns.

4.4 Summary of Study 2

The original hypothesis was that there would be autosegmental and segmental differences between the reading and singing conditions based on previous research. However, current results suggest that there are only autosegmental differences between speech and song as there were no general patterns of the formant frequency analyses. The differences in formant frequencies between song and speech found in previous studies could be due to musical training whereas in this study the speakers had no musical training. Also, there are no differences between native and non-native speakers.

I speculate that the results suggest that the difficulty to detect an accent in song is due to duration and pitch differences in song compared to speech. As duration and pitch are key variables of prosody, I argue that these results reflect the importance that prosody plays in detecting a foreign accent. The prosodic cues are being masked by the imposed duration and pitch of the song, making the listeners perceive less of foreign accent, if any accent at all (as suggested by the results of Study 1).

Chapter 5

Study 3

In Study 2, the autosegmental and segmental differences were illustrated in native and non-native singers in song and speech. The results suggest that there are autosegmental differences (i.e. duration and pitch) between song and speech, whereas there are no consistent segmental differences (i.e. F1 and F2) between the two conditions.[8] These results suggest that prosody plays a role in detecting an accent, and singing masks prosodic cues that listeners use to detect a foreign accent. Study 3 is an experiment that explores the role of prosody in accent detection.

5.1 Materials

The materials for Study 3 consisted of the reading conditions previously outlined in Chapter 2 (section 2). The conditions are as follows:

i) reading *Twinkle, Twinkle, Little Star*

ii) reading *Goldilocks and the Three Bears*

5.2 Method

The recordings were run through a pass band filter in PRAAT with the band from 0Hz to 300Hz. A new, manipulated, sound file was created where the signal was cut off above 300Hz. The aim of running a pass band filter was to extract the actual pronunciation of the recordings and keep only the intonation. After the recordings were manipulated, they sounded muffled. As prosodic cues are important in detecting a foreign accent (refer to section 1.1.3), we expected

[8] This was found only in vowels as we did not analyze the consonantal differences.

that native listeners to use the prosodic cues and accurately detect whether the speaker was native or non-native.

There were a total of 24 recordings used for the experiment, which were recorded by twelve speakers (native = 6, non-native = 6). The recordings were then divided into the four different blocks (Appendix D, excluding the two singing conditions).

5.3 Listeners

Forty native English listeners were recruited from the Ottawa area.[9] The listeners were paid for their participation. The speakers were asked to complete a background questionnaire (Appendix C) which provided insight into their linguistic background.

5.4 Procedure

The experiment was administered in a quiet room. Listeners were seated in front of a laptop and were instructed to wear headphones. The listeners were presented with one of four blocks and heard six of the manipulated recordings (native = 3, non-native = 3). The manipulated recordings were presented to the speakers before the recordings from Study 1. After each sound file, the listeners were asked to judge whether the speaker was a native Canadian English speaker or not. If the listener responded that the speaker was not Canadian English, they were asked to rate how close the speaker sounded to Canadian English. The listeners used a 9 point Likert scale to respond, where a 1 on the scale corresponded to "closer to native Canadian English" and a 9 corresponded to "less close to native Canadian English."

Similarly to Study 1, the material was presented through PsychoPy, where the listeners' raw responses were stored in a format compatible with a Microsoft Excel worksheet. The

[9] The listeners used for Study 3 were also used for Study 1.

responses were categorized into whether the listener correctly identified if the speaker was a native speaker of Canadian English. If the listener responded that the speaker was not a native speaker, their score on the 9-point Likert scale was stored.

5.5 Results

Forty listeners were recruited and asked to judge six recordings; therefore, there was a total of 240 responses (native = 120, non-native = 120). Anderson-Hsieh and colleagues (1992) found that prosody is more important than pronunciation when detecting a foreign accent; therefore, we hypothesized that listeners would be able to correctly identify whether the speaker was native or non-native. We predicted that the listeners would use the prosodic cues available to them to make their judgements.

Figure 3. *Listeners' Incorrect and Correct Responses*

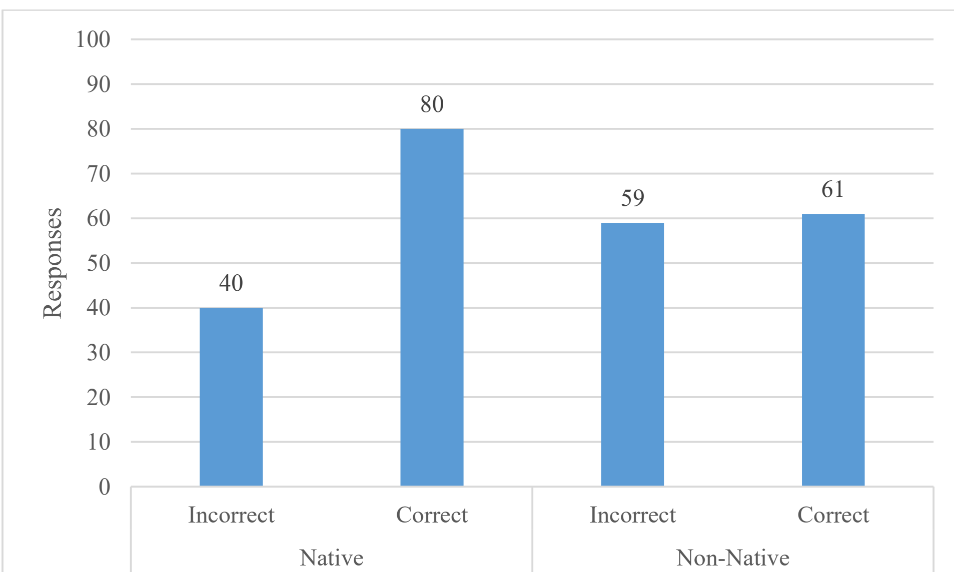


Figure 3 illustrates that listeners correctly judged 141 (59%) of the recordings and incorrectly judged 99 (41%) of the recordings. Listeners judged non-native speakers (*n* = 59) incorrectly more than native speakers (*n* = 40). Listeners were able to correctly judge native speakers (n = 80) more than non-native speakers (n = 61). The descriptive results suggest that listeners are able to detect a native accent more accurately with the use of prosodic cues more so than a non-native accent. Even though listeners were not able to correctly identify non-native speakers' accents, it appears that the responses are moving towards our hypothesis as listeners were able to judge the recordings correctly more than incorrectly overall. This difference was illustrated with a proportion test, which found that there is a significant difference between correct and incorrect responses (p <0.001).

It is important to note that there was no difference between the reading *Twinkle, Twinkle Little Star* and reading *Goldilocks and the Three Bears* conditions. The results between the two conditions are illustrated in Table 6.

Table 6. *Responses by condition and accuracy*

	Native		**Non-Native**	
	Correct	*Incorrect*	*Correct*	*Incorrect*
Twinkle, Twinkle Little Star	39	21	28	32
Goldilocks and the Three Bears	41	19	33	27

Figure 4. *Study 1 Listeners' Incorrect and Correct Responses*

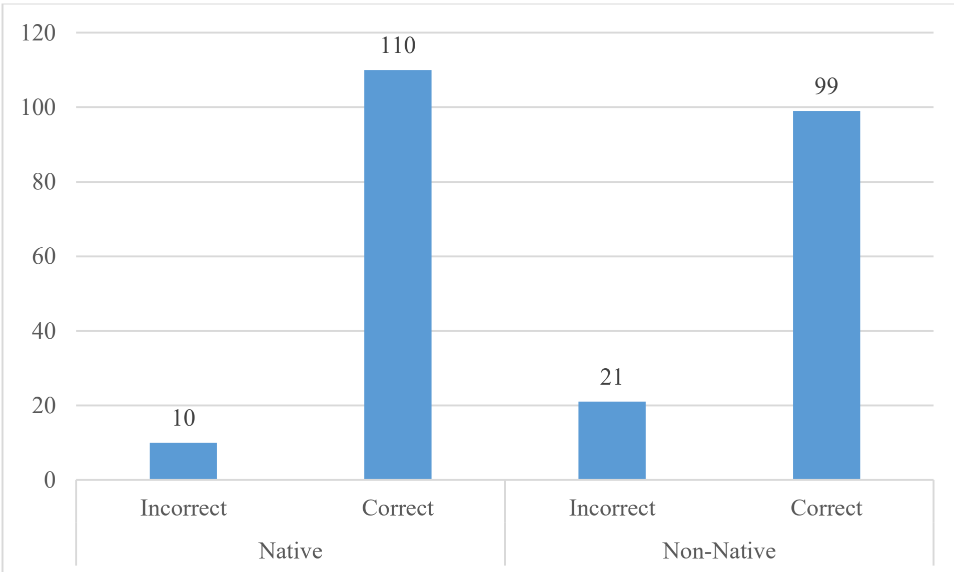


Figure 4 illustrates the listeners' incorrect and correct responses in Study 1 where the recordings included both prosodic and segmental information. In study 1, native listeners rated the speakers correctly (87%) more than incorrectly (13%) in both reading conditions. The difference between the results from Study 1 and Study 2 are presented in Table 7. The differences suggest that native listeners are better able at detecting a non-native speaker when they are presented with both segmental and prosodic information.

Table 7. *Accuracy differences between Study 1 and Study 3*

	Native		**Non-Native**	
	Incorrect	*Correct*	*Incorrect*	*Correct*
Study 1	8%	92%	17%	83%
Study 3	33%	67%	49%	51%

Table 8. *Number of responses for each person by accuracy*

Speaker	Correct	Incorrect
S1	11	9
S2	12	8
S3	11	9
S4	14	6
S5	3	17
S6	10	10
S7	15	5
S8	6	14
S9	13	7
S10	14	6
S11	16	4
S12	16	4

--S1-S6 are non-native speakers; S7-S12 are native speakers

Table 8 was constructed in order to understand whether there was a specific speaker who was judged completely incorrect or correct which could have skewed the data. Table 8 illustrates that listeners were able to correctly detect a native accent in almost all speakers; however, there appears to be an equal distribution in the non-native speakers. As is illustrated, speaker 5 had more incorrect responses compared to all of the other non-native speakers. This could be due to listeners perceiving this speaker as a native speaker.

Table 9. *Listeners' rating of non-native accent*

Native Speaker	Non-Native Speaker
5.1	5.2

Table 9 illustrates that native English listeners' rate native and non-native speakers, either correctly or incorrectly, similarly. I had expected that the listeners might judge the native speakers as having less of an accent compared to the non-native speakers. This would have illustrated that the listeners were close to accurately detecting whether the speaker was native or non-native using only their prosody. However, when the listeners thought the speaker was non-native, either correctly or incorrectly, they rated the native and non-native speakers similarly.

The results presented in this section suggest that native listeners are able to correctly identify native intonation more than non-native intonation. When asked to rate the strength of the speakers' accent, the listeners rated the native and non-native speakers similarly.

5.6 Summary of Study 3

In sum, the results of Study 3 suggest that the listeners are more able to detect a native accent with prosodic cues; however, it is more difficult to detect a non-native accent. We predicted that listeners would be able to accurately detect whether a speaker is native or non-native using the prosodic cues available to them. However, the results suggest that in order to detect a non-native accent some segmental features are required, as native listeners require hearing something that is clearly non-English. This is confirmed when comparing the results from Study 1 to Study 3. There is a clear difference between the two studies and the results suggest that native listeners require both segmental and prosodic cues in order to accurately detect a non-native speaker.

The lack of significant correct responses for non-native speakers could be due to the limitations of the experiment, as it appears that the responses were moving in the direction that we had expected. One limitation could have been the way the pass band filter was used. The settings may have been set incorrectly where 300Hz was too high or too low to capture all of the prosodic cues and eliminate the actual pronunciation of the recordings. Considering that listeners were correct at judging more than they were incorrect, future research could control for the limitations presented in order to find results closer to what was expected.

Chapter 6

Discussion

The studies presented in this thesis suggest that it is more difficult to detect an accent when speakers are singing than when they are reading. This difference is due (at least in part) to duration and pitch being dictated by the music which masks the prosodic cues needed to detect an accent.

The results from Study 1 suggest that it is more difficult to detect an accent in song. Listeners were also asked to rate the non-native speakers accent on a scale from 1 to 9 where 1 corresponded to an accent closer to English and 9 corresponded to an accent further away from English. The results of the rating found that native listeners rated non-native speakers as sounding closer to an English accent when singing and further away when reading. These results suggest that native listeners have less of a perceived foreign accent in song compared to speech.

Study 2 analyzed the vowel properties of the singing *Twinkle, Twinkle, Little Star* accompanied with a piano condition and the reading *Twinkle, Twinkle, Little Star* condition. The results from Study 2 found that there were autosegmental differences between the two conditions. Specifically, duration was always higher in song compared to speech and there was more variation of pitch in singing compared to speaking. The changes that occurred in pitch were not significant overall; however, the graphs in Appendix E illustrate that speakers were attempting to change their pitch while singing. Study 2 also found that there were no consistent segmental differences between the singing and reading conditions. This was illustrated through the mean values presented in Appendix H, Appendix I and Appendix E, which found that there are no general patterns that arise between speech and song in F1 and F2 formant frequencies. It is important to note that the results presented in Study 2 did not find any significant differences between native and non-native speakers.

The final study, Study 3, tested the role that prosody plays in detecting an accent. In order to explore this research question, speech manipulation was used in order to filter out the segmental features of the two reading conditions (i.e. reading *Goldilocks and the Three Bears* and reading *Twinkle, Twinkle, Little Star*). The manipulated recordings were presented to native listeners who were instructed to respond to whether they thought the speaker was a native or non-native English speaker. The results from Study 3 suggest that native listeners are able to correctly identify a native speaker with only the prosodic features available to them; however, they are not successful at identifying non-native speakers. The results suggest that listeners require some segmental features in order to detect an accent.

The results from the three studies presented indicate that not only do cultural factors play a role in accent and music, but so do vowel prosodic features such as pitch and duration. The fact that listeners performed the worst at detecting accents when the song was accompanied by a piano further indicates that music has a generally distracting influence that makes it more difficult to detect accents. Future research should include a condition where speakers are reading with piano accompaniment to explore the role that accompaniment plays on accent detection. Furthermore, the results suggest that it is more difficult to detect an accent when someone is singing due to the prosodic cues being masked by the music which are required to detect an accent.

The results presented differ from previous research conducted on accent and music as we did not find segmental differences in speech and song. These differences could be due to the use of untrained singers which indicates that duration and pitch are key in detecting an accent in song.

The current thesis has implications for Second Language learning as it indicates that it is important to teach both segmental and suprasegmental features in an individual's L2. This may lead to an increase in the L2 comprehensibility and intelligibility of students (Anderson-Hsieh, 1992; Boyd, 2003; Ikeno & Hansen, 2007; Munro & Derwing, 1995; Munro & Derwing, 1999; Derwing & Munro, 2005; Munro, 2008). The results suggest that when learning a second language it is important to not only learn the pronunciation of the language, but also the prosodic features. Learning the prosodic features allow you to be more intelligible in the language which may lead to more opportunities in the language (e.g. career and educational opportunities).

In future research, it might prove fruitful to analyze the vowel properties of trained, professional singers in an experiment similar to Study 2. Having trained singers sing a familiar song, with minimal social cues, may reveal the differences between trained and untrained singers. Knowledge of these features could facilitate native speaker perception of foreign accent, as well as provide information about how to teach second language speech. Also, future research should also replicate the studies presented in different languages to explore the cross linguistic differences that music plays on accent detection.

6.1 Limitations

The studies conducted in this thesis include two main limitations which could be improved upon in future research. For example, the design of Study 1 instructed native listeners to respond to whether they thought the speaker's first language was "Canadian English". This limits the implications of the Study as the results may only apply to Canadian English rather than

English. Listeners could have also heard a speaker and thought their first language was English but not Canadian English, which would have impacted their response. However, considering the results of Study 1 replicated the research conducted by Can Mekik and Catherine Boucher (2015) and Hagen, Kerkhoff and Gussenhoven (2012) on the same topic, we conclude that this limitation does not impact the results significantly. Also, it is unclear whether this matters that much.

Moreover, the results of Study 1 found that native listeners were incorrect at detecting a foreign accent 15% of the time which could indicate that the task was too easy. Future research could include similar "genre-familiar" songs (with minimal social cues) and omit words which clearly demonstrate a speaker's foreign accent. For example, the phonemic inventory for Spanish does not include /ð/ or /θ/ sounds; therefore, words like /nʌθɪŋ/ could be extracted from all recordings as they may have made it easier for listeners to detect a foreign accent.

Bibliography

Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning*, *42*(4), 529-555.

Armstrong, E. G. (2004). Eminem's construction of authenticity. *Popular Music and Society*, 27(3), 335-355.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press.

Bates, D., Maechler, M., & Bolker, B., Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1-48.

Beal, J. C. (2009). "You're Not from New York City, You're from Rotherham": Dialect and Identity in British Indie Music. *Journal of English Linguistics*. *37*(3), 223

Bell, A. (1984). Style as Audience Design. *Language in Society*, 13(2), 145-204

Bell, A., & Gibson, A. (2011). Staging language: An introduction to the sociolinguistics of performance. *Journal of Sociolinguistics*, 15(5), 555-572.

Boersma, P., & Weenink, D. (2001). Praat, a system for doing phonetics by computer. (Version 5.1.04)[Computer program].Downloadable at http://www.praat.org/

Boyd, S. (2003). Foreign-born teachers in the multilingual classroom in Sweden: The role of attitudes to foreign accent. *International Journal of Bilingual Education and Bilingualism*, *6*(3-4), 283-295.

Calderwood, I. (2015, July 5). 'Gahaad save our Queen!' Alesha Dixon sets Twitter alight for singing national anthem at British Grand Prix with an American accent (and getting the lyrics wrong). *The Daily Mail.* Retrieved from http://www.dailymail.co.uk/, October 23, 2015

Church, B. A., & Schacter, D. L. (1994). Perceptual specificity of auditory priming: implicit memory for voice intonation and fundamental frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(3), 521.

Clermont, F. (2002). Systemic comparison of spoken and sung vowels in formant-frequency space. Proc. 9th Australian SST. 124, 129.

Coupland, N. (2011). Voice, place and genre in popular song performance. *Journal of Sociolinguistics, 15*(5), 573-602.

Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *Tesol Quarterly*, *39*(3), 379-397.

Eberhardt, M., & Freeman, K. (2015). 'First things first, I'm the realest': Linguistic appropriation, white privilege, and the hip-hop persona of Iggy Azalea. *Journal of Sociolinguistics, 19*(3), 303-327.

Flege, J. E. (1981). The phonological basis of foreign accent: A hypothesis. *Tesol Quarterly*, *15*(4), 443-455.

Flege, J. E., & Fletcher, K. L. (1992). Talker and listener effects on degree of perceived foreign accent. *The Journal of the Acoustical Society of America*, *91*(1), 370-389.

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

Flege, J. E., Bohn, O. S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, *25*(4), 437-470.

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

Gibson, A. (2010). Production and perception of vowels in New Zealand popular music (Master's thesis, Auckland University of Technology).

Gibson, A. (2011). Flight of the Conchords: Recontextualizing the voices of popular culture1. *Journal of Sociolinguistics*, *15*(5), 603-626.

Gibsone, H. (2016, February 4). Talking tactics: Rihanna and the pop stars who change accent. *The Guardian.* Retrieved from: http://www.theguardian.com/music/2016/feb/04/talk-that-talk-rihanna-the-cunning-linguist, February 5, 2016

Guo, J. (2016, January 4). How Iggy Azalea mastered her 'blaccent'. *The Washington Post.* Retrieved from: https://www.washingtonpost.com/news/wonk/wp/2016/01/04/how-a-white-australian-rapper-mastered-her-blaccent/, January 6, 2016

Hagen, M., Kerkhoff, J., & Gussenhoven, C. (2011). Singing your accent away, and why it works. In the proceedings of *The 17th International Congress of Phonetic Sciences* (ICPhS XVII), pp. 799-802

Hansen, J. H., & Arslan, L. M. (1995, May). Foreign accent classification using source generator based prosodic features. In *Acoustics, Speech, and Signal Processing*, 1995. ICASSP-95, 1995 International Conference on (Vol. 1, pp. 836-839). IEEE.

Ikeno, A., & Hansen, J. H. (2007). The effect of listener accent background on accent perception and comprehension. *EURASIP Journal on Audio, Speech, and Music Processing*, 2007(3), 4.

Lehiste, I (1970). *Suprasegmentals.* Cambridge, Mass: M.I.T. Press

Mageau, M., Mekik, C., & Wang, E. (2015). *Foreign accents in speech and song*. Presented at the Ottawa Conference for Undergraduate and Masters students conference. University of Ottawa, Ottawa.

Magen, H. S. (1998). The perception of foreign-accented speech. *Journal of Phonetics*, *26*(4), 381-400.

Mareüil, P., & Vieru-Dimulescu, B. (2006). The contribution of prosody to the perception of foreign accent. *Phonetica*, *63*(4), 247-267.

Mekik, C., & Boucher, C., (2015). *Perception of accents in song and speech.* Research project, Carleton University

Morrissey, F., (2008). Liverpool to Louisiana in one lyrical line: Style choice in British rock, pop and folk singing. In Locher, M. & Strässler, J. (eds.) *Standards and Norms in the English Language*: *Contributions to the Sociology of Language* (Vol. 95, pp. 195-218). Berlin: De Gruyter

Munro, M. J. (2008). Foreign accent and speech intelligibility. In J. Edwards & M. Zampini (Eds.), *Phonology and Second Language Acquisition* (Vol. 36, pp. 193-218). Amsterdam: J. Benjamins Pub.

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, *45*(1), 73-97.

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*(supplement 1), 285-310.

O'Hanlon, R. (2006). Australian hip hop: A sociolinguistic investigation. *Australian Journal of Linguistics*, *26*(2), 193-209.

Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, 2, 10. Retrieved from https://github.com/psychopy/psychopy/releases

Peterson, G. E., & Lehiste, I. (1960). Duration of syllable nuclei in English. *The Journal of the Acoustical Society of America, 32*(6), 693-703.

Piske, T., MacKay, I. R., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics*, *29*(2), 191-215.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

Rudgard, O. (2015, July 6). Why you put on an American accent when you sing. *The Telegraph*. Retrieved from http://www.telegraph.co.uk/, October 23, 2015

Schacter, D. L., & Church, B. A. (1992). Auditory priming: implicit and explicit memory for words and voices. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(5), 915.

Schulze, C. (2014). Identity Performance in British Rock and Indie Music: Authenticity, Stylization, and Glocalization. (Master's thesis, Lund University).

Simpson, P. (1999). Language, culture and identity: With (another) look at accents in pop and rock singing. *Multilingua*, *18*(4), 343-367.

Southwood, M. H., & Flege, J. E. (1999). Scaling foreign accent: Direct magnitude estimation versus interval scaling. *Clinical Linguistics & Phonetics, 13*(5), 335-349.

Stone, R. E., Jr., Cleveland, T. F., & Sundberg, J. (1999). Formant frequencies in country singers' speech and singing. *Journal of Voice*, *13*(2), 161-167.

Sundberg, J. (1969). Articulatory differences between spoken and sung vowels in singers. Speech Transmission Laboratory Quarterly Progress and Status Report (STL-QPSR), *10*(1), 033-046.

Sundberg, J., & Skoog, J. (1997). Dependence of jaw opening on pitch and vowel in singers. *Journal of Voice*, *11*(3), 301-306.

Trudgill, Peter. 1983. On dialect: Social and geographical perspectives. Oxford: Blackwell.

Vaissière, J., & De Mareüil, P. B. (2004). Identifying a language or an accent: from segments to prosody. In *Colloque MIDL* (pp. 1-6), 1-4.

Wardhaugh, R. (1998). *An introduction to sociolinguistics* (3rd ed.). Malden, Mass: Blackwell.

Appendices

Appendix A

Twinkle, Twinkle, Little Star

Twinkle, twinkle, little star
How I wonder what you are
Up above the world so high
Like a diamond in the sky
Twinkle, twinkle little star
How I wonder what you are

When the blazing sun is gone
When he nothing shines upon
Then you show your little light
Twinkle, twinkle, all the night
Twinkle, twinkle, little star
How I wonder what you are

Appendix B

Goldilocks and the Three Bears

Once upon a time there were three bears: A father bear, a mother bear and a little bear. They lived all together in a yellow house in the middle of a big forest. One day, Mother Bear prepared a big pot of delicious hot porridge for breakfast. It was too hot to eat, so the bears decided to go for a walk while waiting for the porridge to cool. Near the forest lived a little girl named Goldilocks.

Appendix C

Language proficiency question

List each language you know, and give an approximate rating of your knowledge of each language:

1) _______________

native,   near native, fluent,   almost fluent,   basic

2) _______________

native,   near native, fluent,   almost fluent,   basic

3) _______________

native,   near native, fluent,   almost fluent,   basic

4) _______________

native,   near native, fluent,   almost fluent,   basic

5) _______________

native,   near native, fluent,   almost fluent,   basic

6) _______________

native,   near native, fluent,   almost fluent,   basic

Appendix D

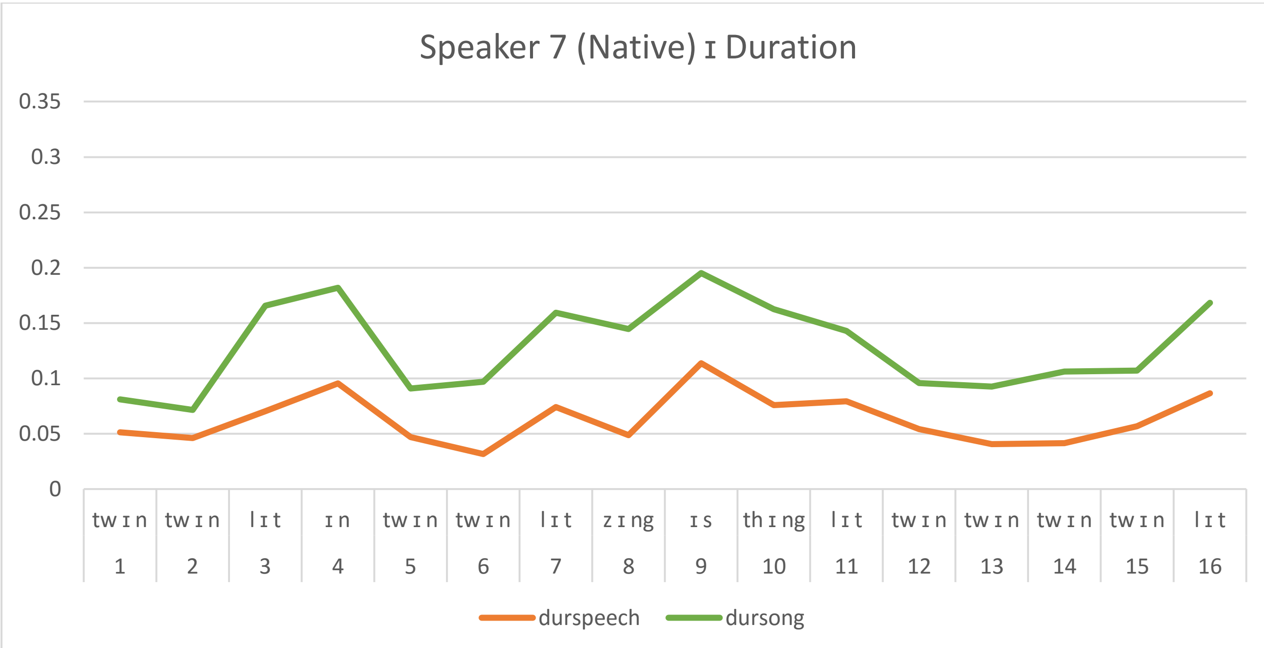
Design of Experiment 1 where the recordings in each block were randomized across listeners.

	Block 1	**Block 2**	**Block 3**	**Block 4**
Speaker 1	Singing with piano accompaniment	Singing without piano accompaniment	Reading *Twinkle, Twinkle*	Reading *Goldilocks*
Speaker 2	Singing without piano accompaniment	Reading *Twinkle, Twinkle*	Reading *Goldilocks*	Singing piano accompaniment
Speaker 3	Reading *Twinkle, Twinkle*	Reading *Goldilocks*	Singing with piano accompaniment	Singing without piano accompaniment
Speaker 4	Reading *Goldilocks*	Singing with piano accompaniment	Singing without piano accompaniment	Reading *Twinkle, Twinkle*
Speaker 5	Singing with piano accompaniment	Singing without piano accompaniment	Reading *Twinkle, Twinkle*	Reading *Goldilocks*
Speaker 6	Singing without piano accompaniment	Reading *Twinkle, Twinkle*	Reading *Goldilocks*	Singing with piano accompaniment
Speaker 7	Reading *Twinkle, Twinkle*	Reading *Goldilocks*	Singing with piano accompaniment	Singing without piano accompaniment
Speaker 8	Reading *Goldilocks*	Singing with piano accompaniment	Singing without piano accompaniment	Reading *Twinkle, Twinkle*
Speaker 9	Singing with piano accompaniment	Singing without piano accompaniment	Reading *Twinkle, Twinkle*	Reading *Goldilocks*
Speaker 10	Singing without piano accompaniment	Reading *Twinkle, Twinkle*	Reading *Goldilocks*	Singing with piano accompaniment
Speaker 11	Reading *Twinkle, Twinkle*	Reading *Goldilocks*	Singing with piano accompaniment	Singing without piano accompaniment
Speaker 12	Reading *Goldilocks*	Singing with piano accompaniment	Singing without piano accompaniment	Reading *Twinkle, Twinkle*

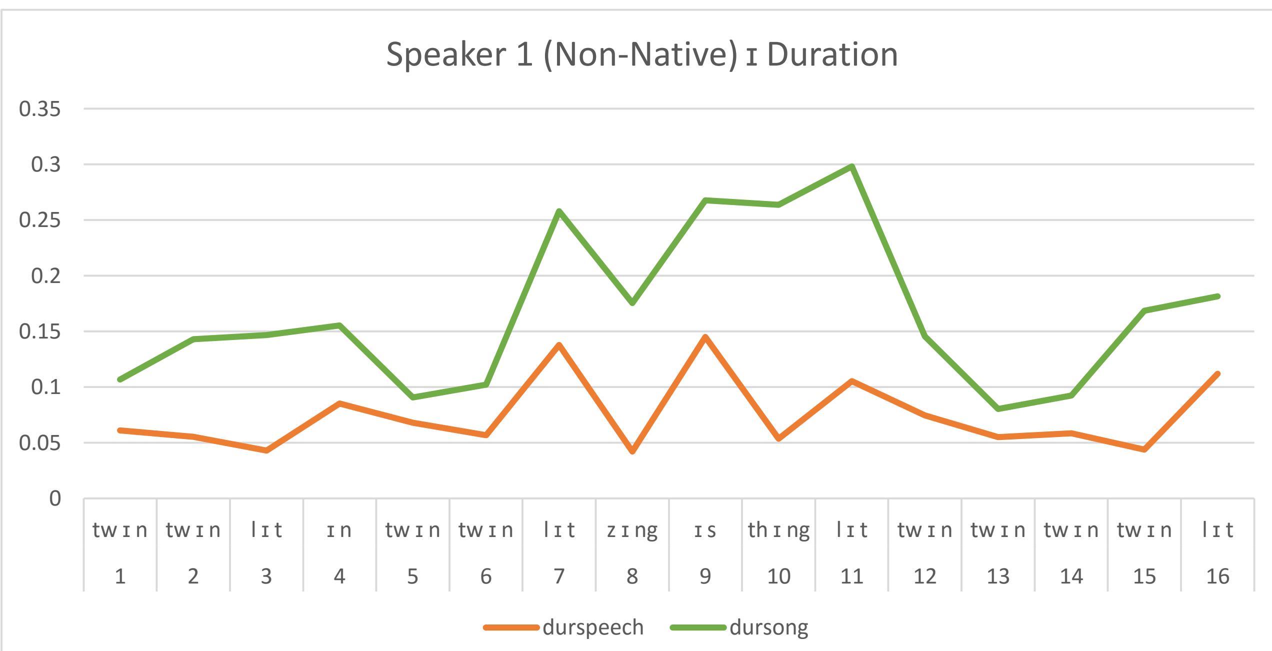
Appendix E

Graphs i) and ii) illustrate the durational differences in speech and song in a native and non-native speaker for vowel /ɪ/.

i)

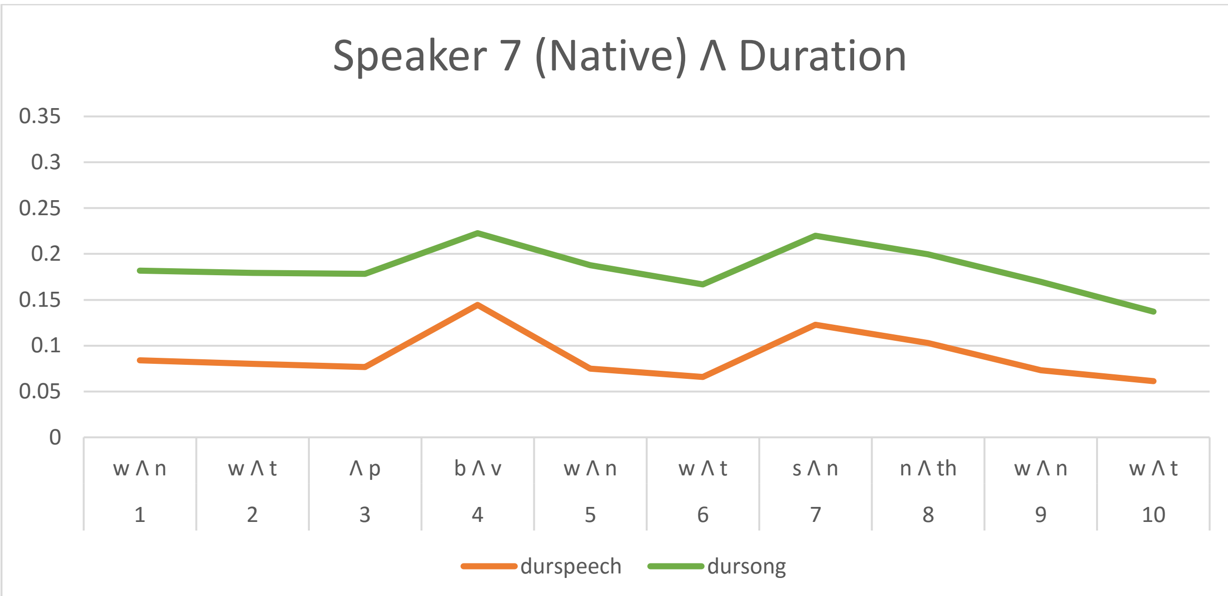


ii)

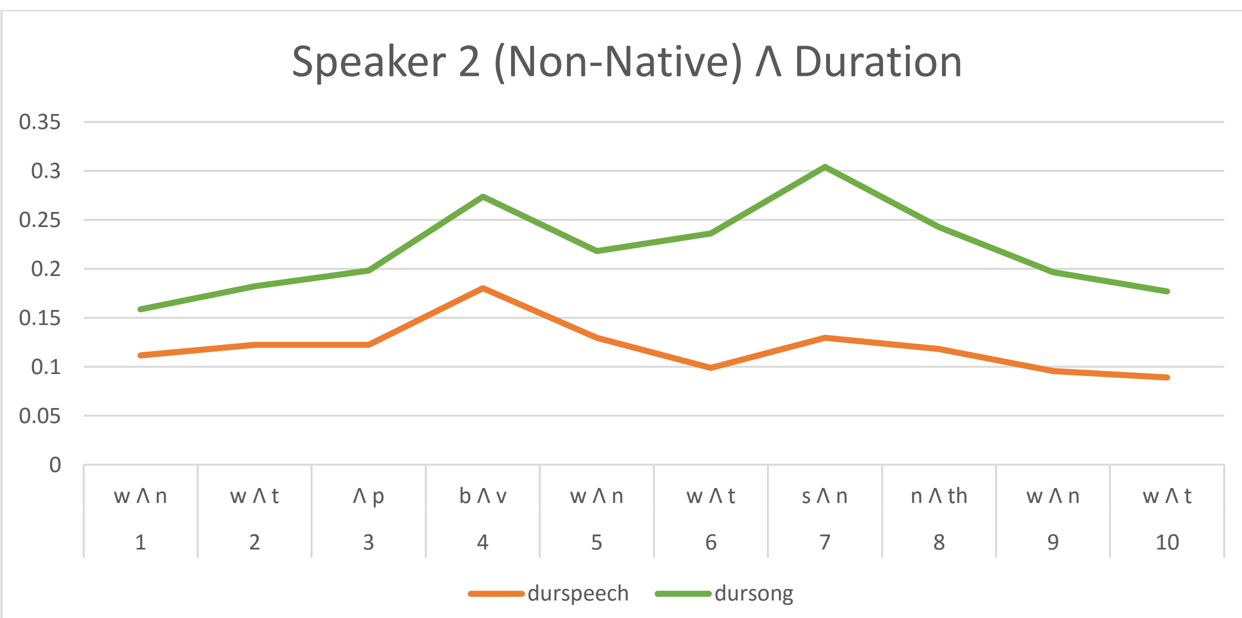


Graphs iii) and iv) illustrate the durational differences in speech and song in a native and non-native speaker for vowel /Λ /.

iii)

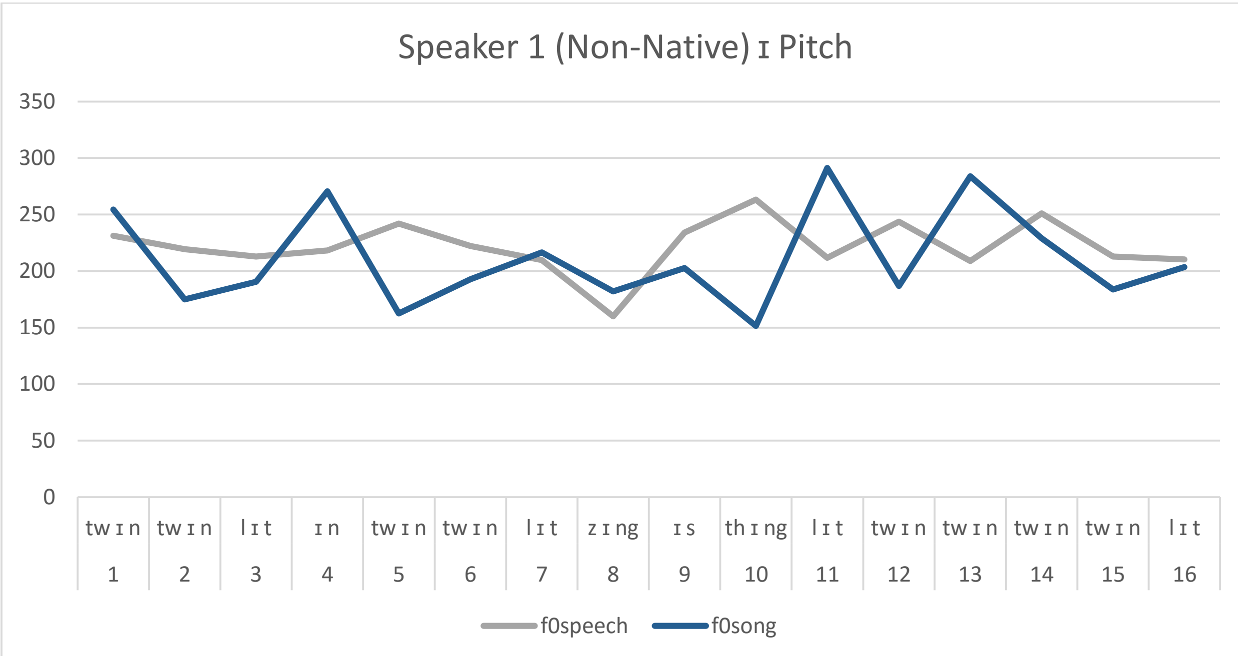


iv)

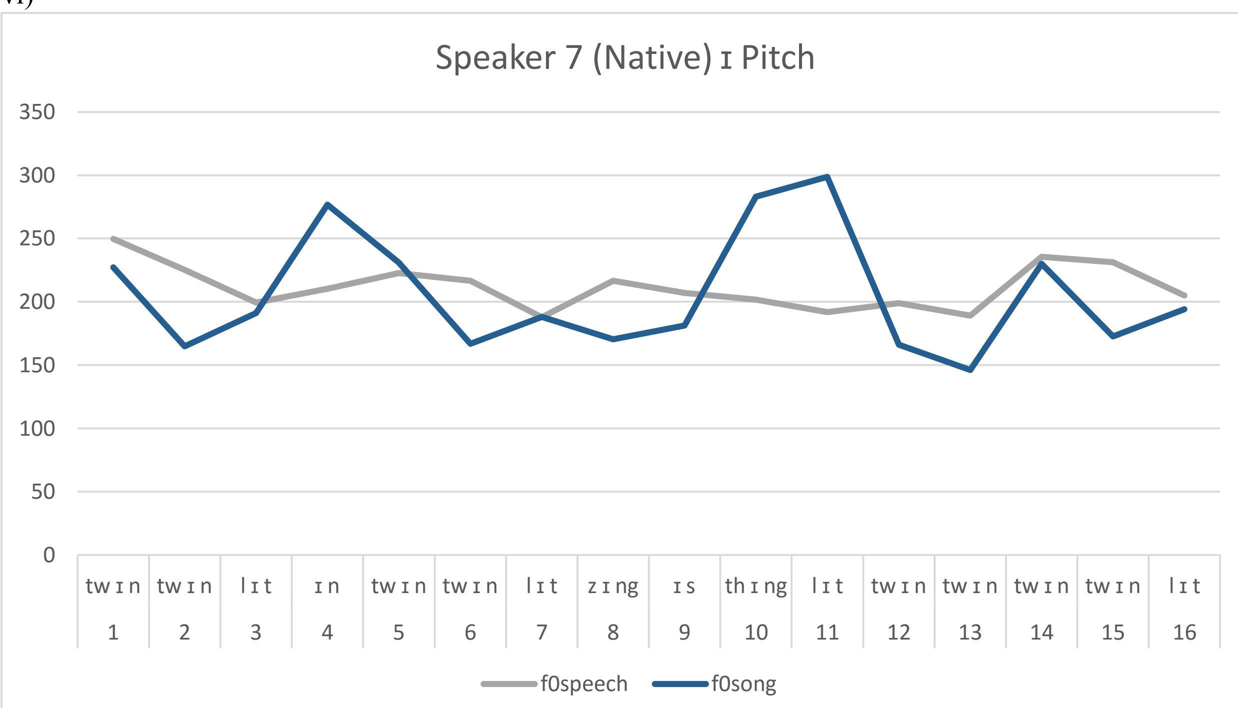


Graphs v) and vi) illustrate the pitch differences between speech and song in a native and non-native speaker for vowel /ɪ/.

v)

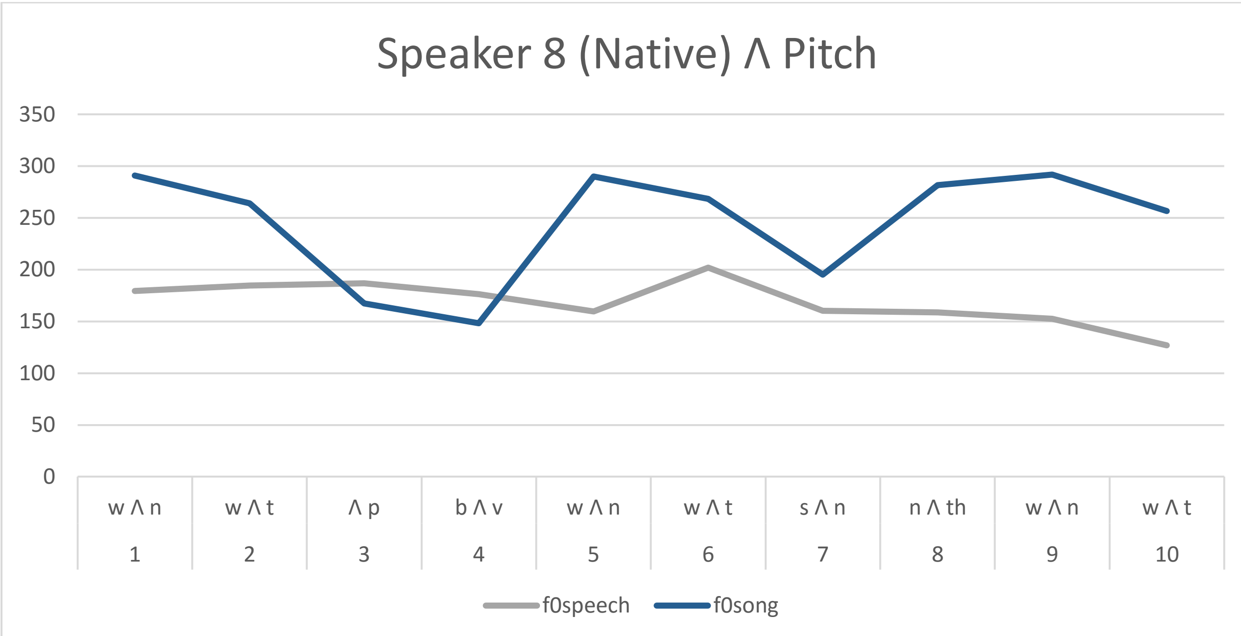


vi)

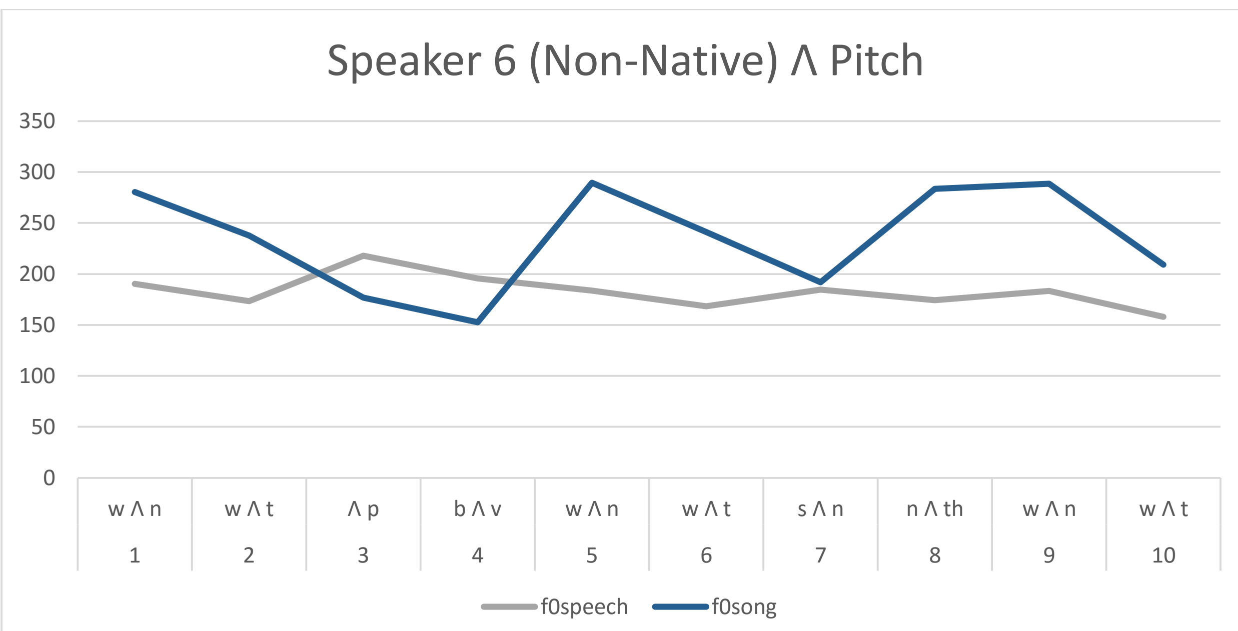


Graphs vii) and viii) illustrate the pitch differences in speech and song in a native and non-native speaker for vowel /Λ /.

vii)

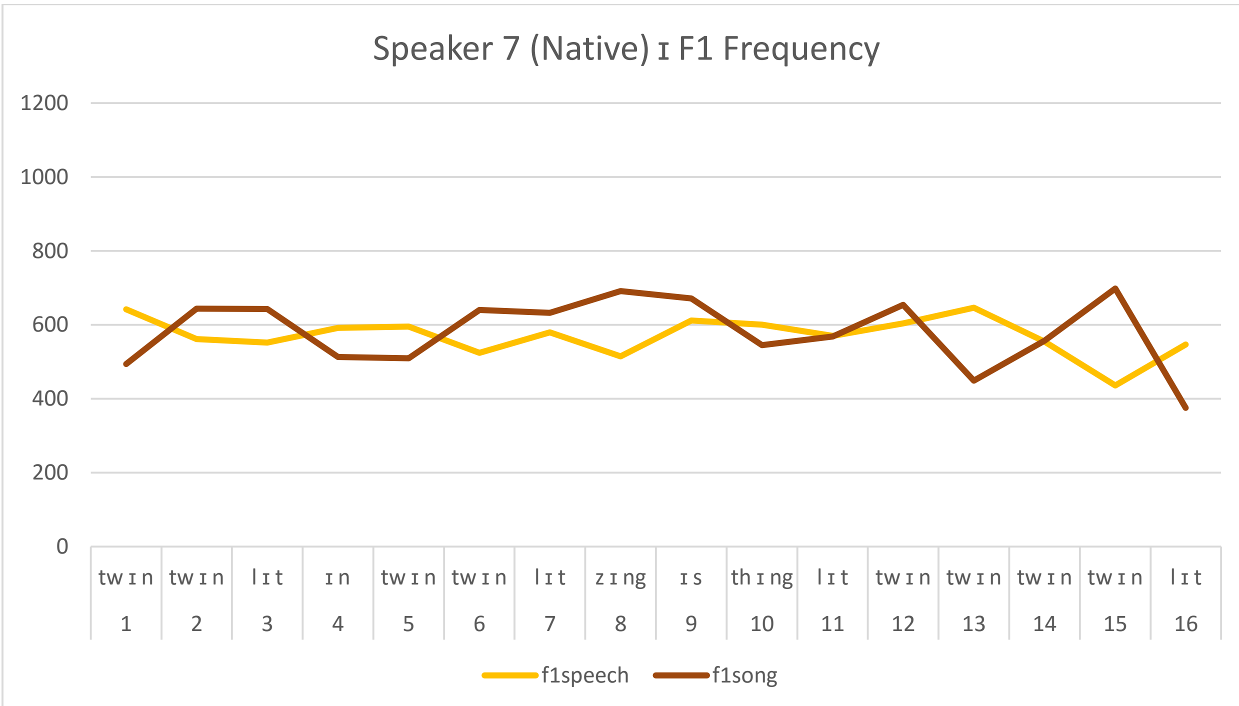


viii)

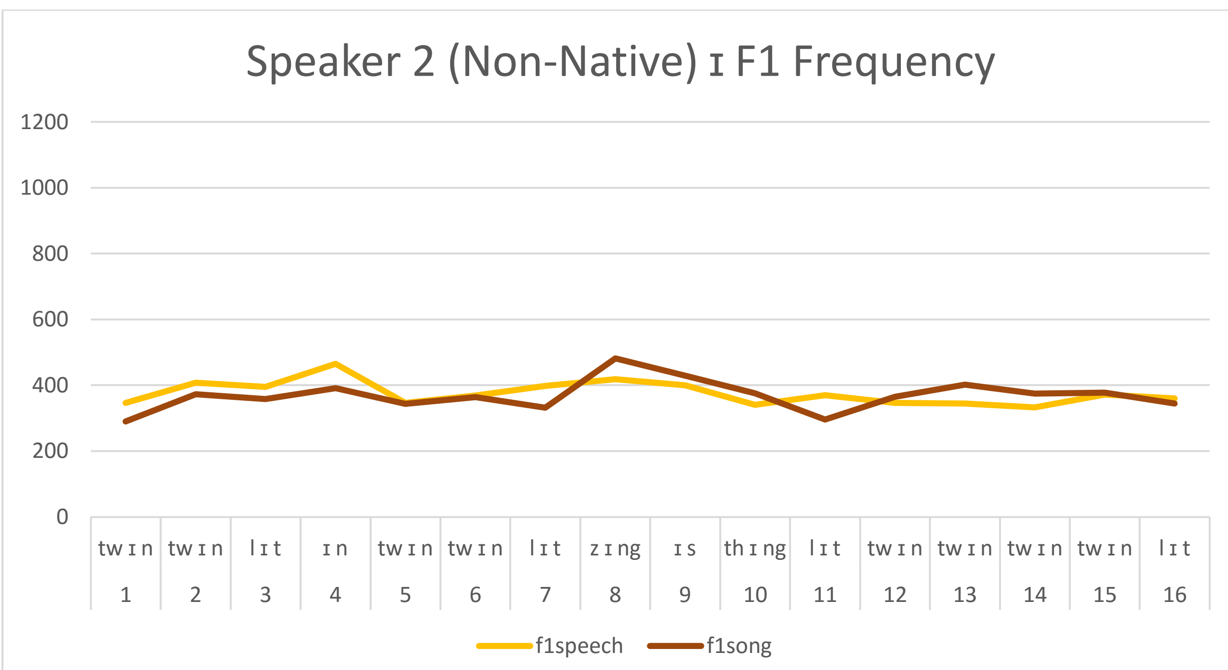


Graphs ix) and x) illustrate the F_1 differences between speech and song in a native and non-native speaker for vowel /ɪ/.

ix)

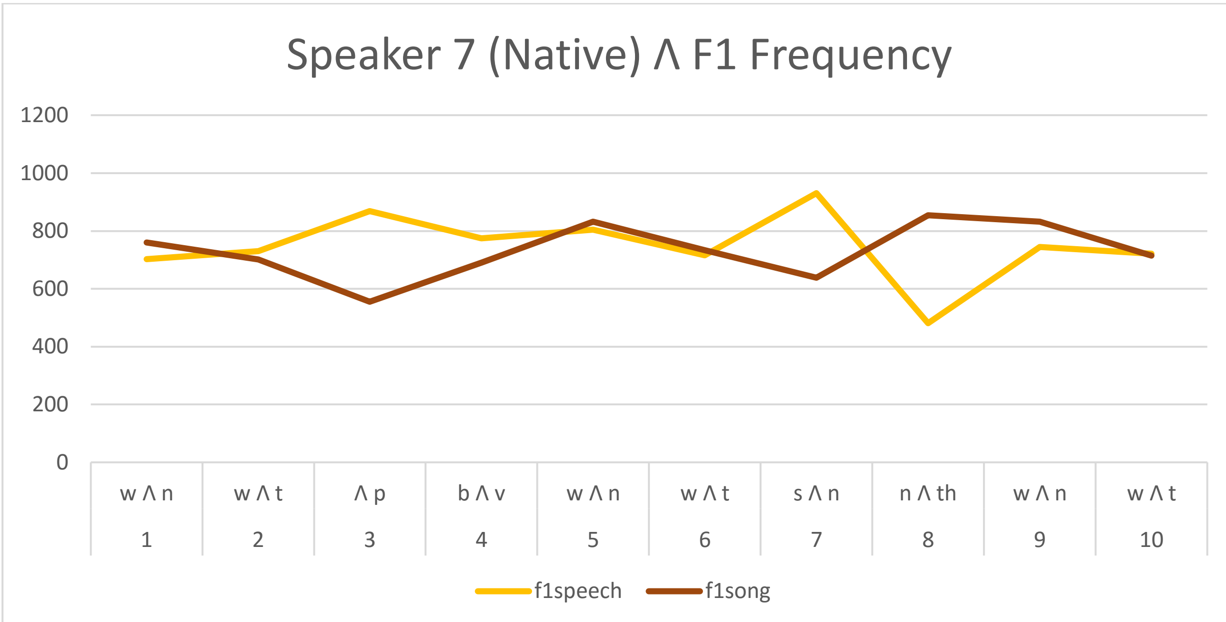


x)

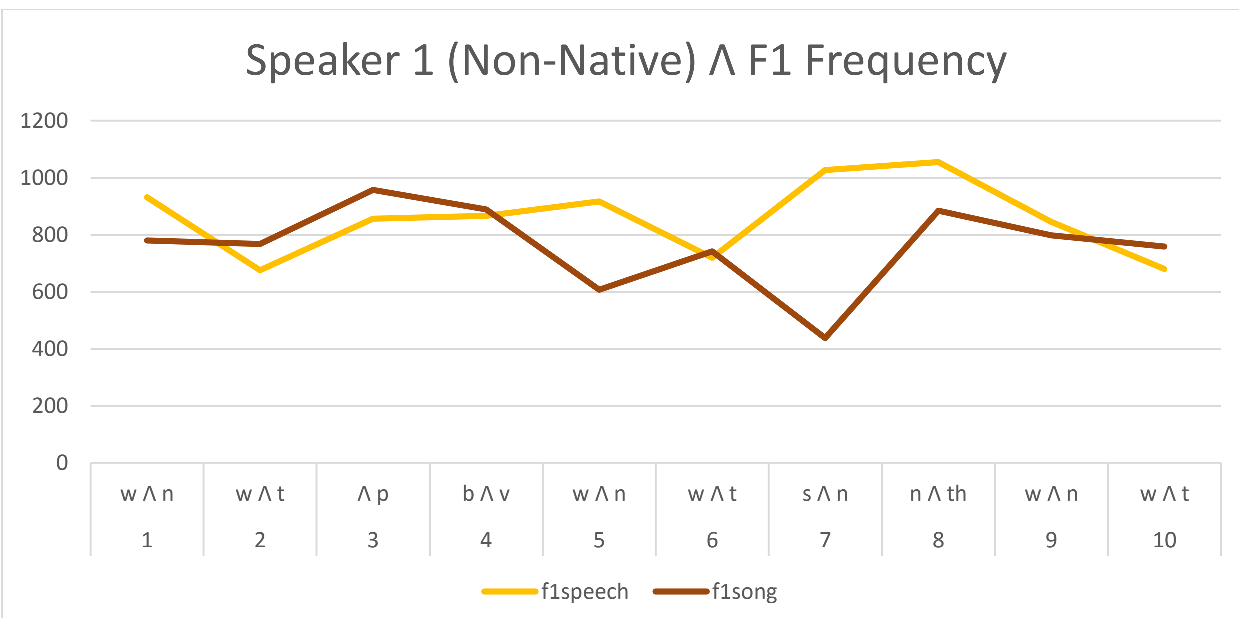


Graphs xi) and xii) illustrate the F_1 differences in speech and song in a native and non-native speaker for vowel /Λ /.

xi)

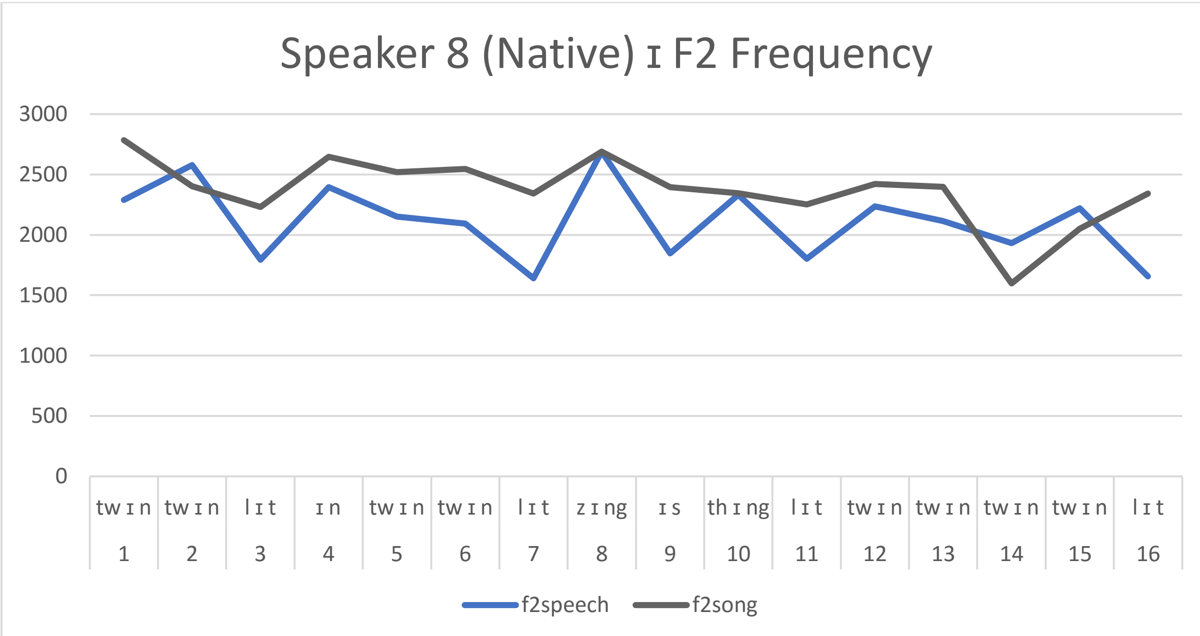


xii)

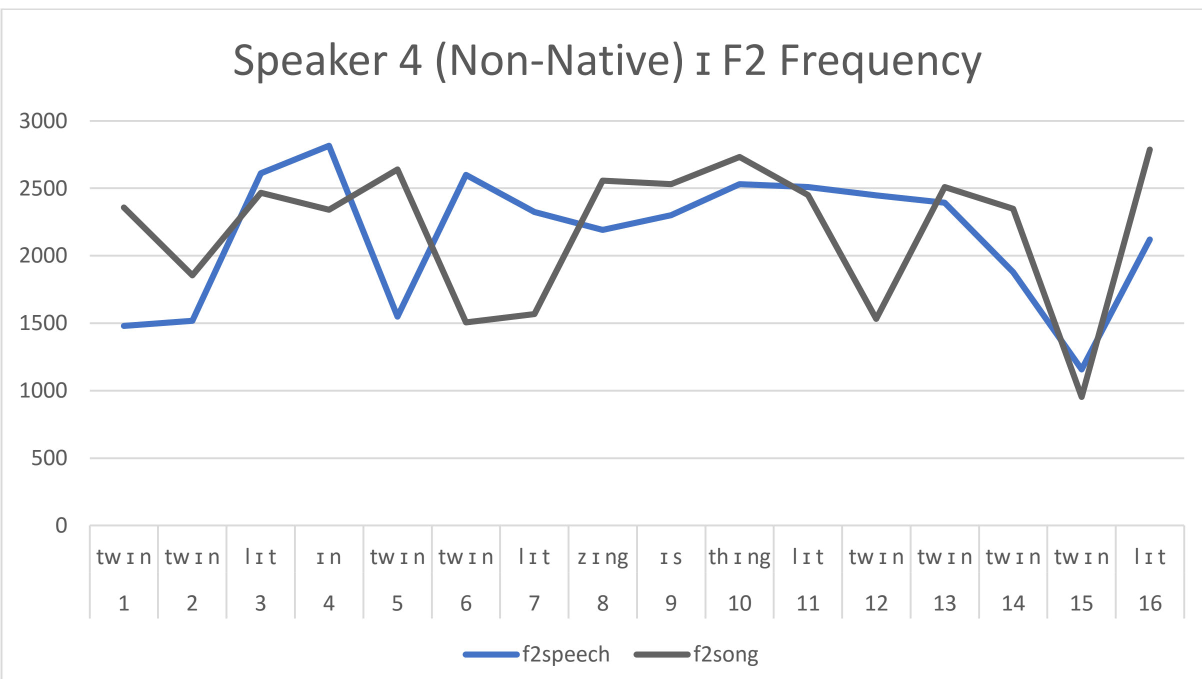


Graphs xiii) and xiv) illustrate the F_2 differences between speech and song in a native and non-native speaker for vowel /ɪ/.

xiii)

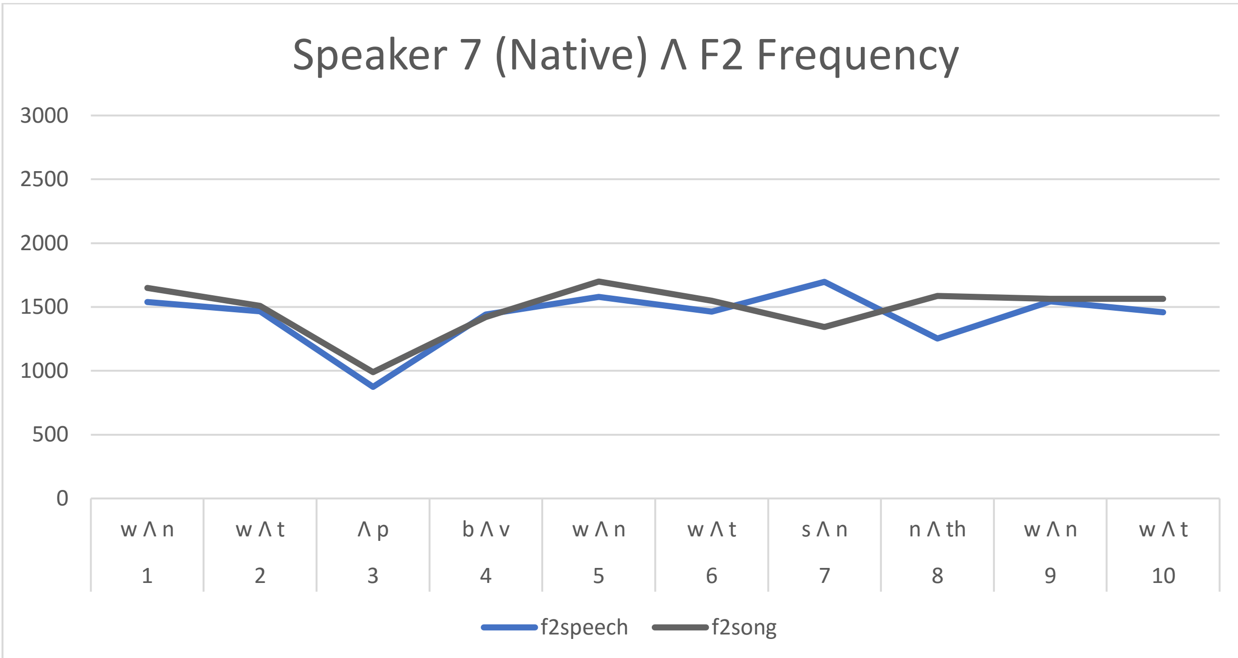


xiv)

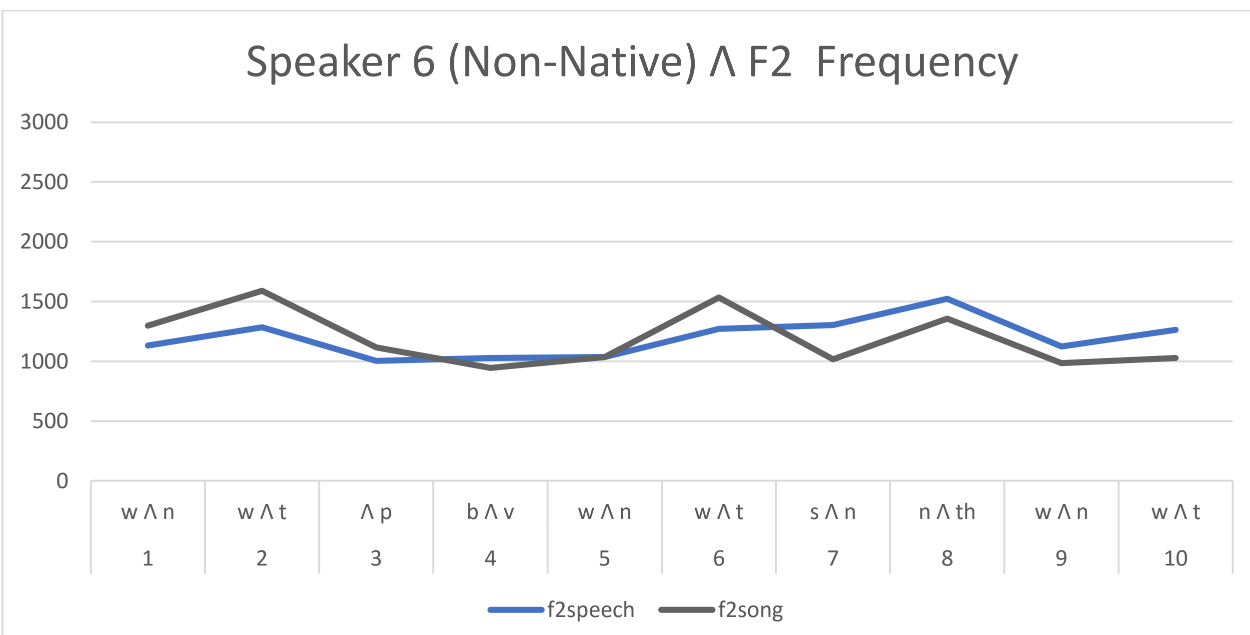


Graphs xv) and xvi) illustrate the durational differences in speech and song in a native and non-native speaker for vowel /Ʌ /.

xv)



xvi)



Appendix F

Vowel I	Duration					
	Mean		Median		Range	
Speaker	Song	Speech	Song	Speech	Song	Speech
S1	0.167284	0.074888	0.150991	0.05973	0.217971	0.102999
S2	0.22248	0.091282	0.199224	0.092251	0.270497	0.067648
S3	0.222551	0.064568	0.184495	0.061789	0.261198	0.061603
S4	0.16458	0.082811	0.135524	0.078717	0.227381	0.087644
S5	0.174618	0.057132	0.167173	0.057618	0.209359	0.02678
S6	0.167284	0.074888	0.132278	0.071414	0.217971	0.102999
S7	0.128858	0.063294	0.124974	0.055388	0.12387	0.081979
S8	0.146376	0.069203	0.107739	0.068555	0.289489	0.077445
S9	0.196259	0.057575	0.136426	0.055589	0.31501	0.038296
S10	0.171531	0.056853	0.13054	0.057718	0.293714	0.040583
S11	0.12609	0.094846	0.094136	0.08034	0.214241	0.11199
S12	0.134507	0.056301	0.101049	0.052144	0.259173	0.062145

-- S1 to S6 are non-native speakers; S7 to S12 are native speakers

Vowel Λ	Duration					
	Mean		Median		Range	
Speaker	Song	Speech	Song	Speech	Song	Speech
S1	0.128163	0.110412	0.108373	0.104526	0.230833	0.108744
S2	0.21887	0.119961	0.208326	0.120524	0.145595	0.091102
S3	0.182612	0.080999	0.17564	0.078461	0.135678	0.089734
S4	0.226589	0.113255	0.221017	0.116859	0.138783	0.099005
S5	0.21871	0.08906	0.207588	0.08049	0.205314	0.100105
S6	0.128163	0.110412	0.191519	0.113613	0.230833	0.108744
S7	0.184279	0.088573	0.180414	0.078458	0.08544	0.083017
S8	0.168578	0.094872	0.1506	0.08873	0.185026	0.136617
S9	0.172751	0.085467	0.163409	0.075317	0.25796	0.14988
S10	0.205207	0.072057	0.183738	0.070344	0.201743	0.055914
S11	0.163765	0.128626	0.155812	0.114424	0.172042	0.126597
S12	0.134884	0.081582	0.128574	0.079051	0.194609	0.074163

Appendix G

Vowel I	Pitch					
	Mean		Median		Range	
Speaker	Song	Speech	Song	Speech	Song	Speech
S1	221.9514	211.0233	197.6619	218.8442	103.1987	139.9055
S2	212.9779	218.1534	222.2679	216.0312	131.5843	100.9459
S3	207.7957	143.2712	198.3253	164.9738	144.6531	142.1379
S4	197.7507	229.1019	190.1955	266.1619	140.1992	149.7696
S5	204.2413	214.3839	196.4644	214.779	147.824	45.66896
S6	211.0233	221.9514	192.0945	188.334	139.9055	103.1987
S7	205.5582	211.8066	189.5785	208.6674	152.5282	62.06677
S8	209.5833	180.1423	195.5236	179.3077	150.7409	76.38291
S9	201.168	185.7909	210.2551	182.8509	96.46646	79.95265
S10	220.6651	181.9001	224.8718	179.7394	85.5703	49.63639
S11	205.6758	181.3757	212.2136	179.9271	78.86965	39.43287
S12	216.31	184.8563	216.1302	182.9403	96.59083	53.02144

Vowel Λ	Pitch					
	Mean		Median		Range	
Speaker	Song	Speech	Song	Speech	Song	Speech
S1	206.3342	164.7088	175.9797	163.347	144.2452	137.376
S2	235.2676	194.1224	232.0341	190.4128	79.25247	54.42034
S3	211.6962	144.1335	224.5933	162.6292	157.4762	110.9819
S4	232.6075	260.9879	260.8533	260.534	147.8576	21.76656
S5	247.4685	196.1208	266.3468	199.1275	143.171	38.51197
S6	206.3342	164.7088	239.4156	183.5622	144.2452	137.376
S7	252.4528	194.6612	264.2718	193.0658	131.849	44.24893
S8	245.3181	158.7306	266.1027	160.0393	143.7001	117.4964
S9	204.3089	182.1209	200.6288	188.2866	132.4438	79.53047
S10	205.8077	172.1379	203.741	171.7917	58.14311	27.00795
S11	195.2019	162.9101	188.2362	162.4646	71.12593	47.20156
S12	181.033	171.5371	181.0699	167.892	165.8017	52.79931

-- S1 to S6 are non-native speakers; S7 to S12 are native speakers

Appendix H

Vowel I	F1					
	Mean		Median		Range	
Speaker	Song	Speech	Song	Speech	Song	Speech
S1	435.9022	299.1083	430.2691	262.9322	506.487	357.275
S2	368.592	375.6557	368.5744	368.839	190.913	132.622
S3	483.3026	497.2543	488.6919	507.4323	292.971	159.299
S4	431.2046	374.7488	418.1457	346.2616	172.866	188.071
S5	561.8383	554.2854	561.4358	556.3958	529.189	216.592
S6	435.9022	299.1083	383.479	367.1574	506.487	357.275
S7	574.0722	570.4998	600.269	574.5466	422.994	210.188
S8	507.1554	477.7887	503.7704	464.8904	398.892	147.541
S9	516.3161	550.6749	515.0143	545.761	271.587	186.725
S10	374.0831	432.8916	374.5227	430.6032	201.522	128.079
S11	579.2665	578.2835	538.2352	578.5444	391.780	330.380
S12	569.3963	547.665	568.9075	539.4472	262.093	173.031

Vowel Λ	F1					
	Mean		Median		Range	
Speaker	Song	Speech	Song	Speech	Song	Speech
S1	762.162	857.1796	773.8894	861.3345	519.5012	379.3917
S2	716.0731	650.0359	673.9812	598.7741	439.4576	332.5341
S3	678.4747	709.1419	691.5586	707.0099	481.0047	274.8268
S4	724.6995	510.0219	731.3913	503.477	179.8566	301.2142
S5	741.4783	723.353	777.1199	747.1825	483.5048	259.0096
S6	762.162	857.1796	595.1553	556.8329	519.5012	379.3917
S7	731.1026	747.3439	724.0957	737.1416	299.1562	449.1985
S8	672.7799	642.3829	649.0469	621.8404	347.0765	152.8575
S9	706.9272	625.137	738.1717	599.7589	278.2368	303.5904
S10	541.7992	615.1165	596.1716	587.5026	599.6334	323.0048
S11	875.5429	901.9074	870.8725	910.7724	192.7113	302.2274
S12	703.224	696.8008	703.3066	722.976	244.9271	289.9351

-- S1 to S6 are non-native speakers; S7 to S12 are native speakers

Appendix I

Vowel I	F2					
	Mean		Median		Range	
Speaker	Song	Speech	Song	Speech	Song	Speech
S1	1642.036	1005.332	1651.247	927.9046	2208.04	1313.839
S2	2498.191	2385.659	2563.081	2420.852	969.3661	742.809
S3	1968.302	2106.48	2336.64	2059.384	1867.371	1285.518
S4	2151.791	2076.883	2403.422	2312.239	1834.62	1897.154
S5	1589.23	1997.605	1939.912	1953.513	1975.744	944.1878
S6	1642.036	1005.332	2370.723	2534.842	2208.04	1313.839
S7	2124.985	1967.935	2260.599	1971.058	1895.892	1242.978
S8	2372.231	2110.106	2396.813	2132.535	1185.681	1046.45
S9	2310.582	1996.883	2415.668	2005.271	1748.599	1123.088
S10	1623.671	1794.265	1852.653	1815.315	1890.309	1250.24
S11	2027.765	2009.853	1969.665	1976.93	858.8912	1299.457
S12	2274.634	1939.793	2308.971	2036.326	857.1541	1210.671

Vowel Λ	F2					
	Mean		Median		Range	
Speaker	Song	Speech	Song	Speech	Song	Speech
S1	1666.714	1489.48	1672.692	1478.946	263.696	389.6065
S2	1256.122	1306.196	1265.179	1345.53	524.0833	454.1499
S3	1650.108	1504.158	1726.49	1489.736	745.0073	575.3761
S4	1365.776	1213.335	1346.203	1220.263	618.4805	366.6453
S5	1532.241	1326.793	1708.024	1325.898	1161.66	524.8151
S6	1666.714	1489.48	1075.913	1197.966	263.6961	389.6065
S7	1487.418	1431.7	1555.689	1465.732	710.5571	824.0409
S8	1589.952	1477.544	1585.215	1483.625	499.9377	376.6257
S9	1695.65	1528.444	1725.056	1548.151	295.8754	426.6699
S10	1421.635	1441.991	1440.542	1401.374	993.0893	443.1225
S11	1633.726	1570.577	1652.061	1557.345	277.3349	312.807
S12	1601.43	1568.311	1590.262	1568.738	235.2842	305.4227

-- S1 to S6 are non-native speakers; S7 to S12 are native speakers