English Music in Spare Time and English Pronunciation
– A Quantitative Study on 9th Graders in Finland

Vilma Aaltonen

MA Thesis

English, Language Learning and Teaching Path

School of Languages and Translation Studies

Faculty of Humanities

University of Turku

September 2019

UNIVERSITY OF TURKU

School of Languages and Translation Studies / Faculty of Humanities

AALTONEN, VILMA: English Music in Spare Time and English Pronunciation
                    – A Quantitative Study on 9th Graders in Finland

----

Today, music is a familiar component of people's daily sound environment, and a significant amount of popular music contains English vocals. It has been established by various studies that music can aid language learning, but the studies have mainly been concerned with the effects of purposefully using music to boost learning. The aim of this study, however, was to investigate whether it is possible to acquire English pronunciation through listening to English music (= music with English vocals) and singing in English during one's spare time without the intention of learning. The study examined 44 lower secondary school English students in Southwest Finland via a questionnaire and a picture description task (PDT). The PDT produced speech samples that were evaluated by a jury of nine native speakers of English based on three pronunciation criteria: accuracy, fluency, and intelligibility. The results of the study suggested that the English music people consume in their spare time could actually be beneficial for the development of their English pronunciation both directly and indirectly. The direct factors included natural language input and versatile output opportunities, whereas the indirect factors had to do with music's capability to influence the learner's motivation and affect, such as mood and attitudes. Especially singing in English seemed to be linked to better English pronunciation. The aim of this study is to inform people about the subconscious benefits of the day-to-day English music consumption as well as to encourage English learners to engage in English language music if they wish to improve their English pronunciation.

Key words: English language, pronunciation learning, accuracy, fluency, intelligibility, English music, listening to music, singing, second language learning

Table of Contents

Appendices

Appendix 1: Questionnaire in Finnish

Appendix 2: Questionnaire English translation

Appendix 3: Picture description task

Appendix 4: Native-speaker evaluation

Appendix 5: Artists and bands mentioned by participants

Appendix 6: Finnish summary

List of tables

List of figures

List of abbreviations

L1 = first language

L2 = second language

MT = Motor Theory

PAM = Perceptual Assimilation Model

PDT = Picture description task

RP = Received Pronunciation

SLA = second language acquisition

TL = target language

1 Introduction

We live in a day and age where music is a familiar component of our daily sound environment. According to the latest IFPI Music Consumer Insight Report (2018), people are now listening to music at all points of the day and spend two and a half hours listening to music on an average day. Music has been found to produce a stress relieving effect by raising the brain's serotonin levels, which makes the mind more receptive for new information (Ortis 2008, 216). Not only do people listen to music, but they often also sing along as well as sing the songs that are stuck in their heads. A significant amount of the music people consume today contains English language vocals, and English pop and English hip hop were the two most popular genres among 13–18-year-olds in Finland last year (IFPI and Teosto 2018). All this could potentially have hidden practical implications.

Pronunciation is an important skill because non-native speakers of English need to be understood as well as regarded as an equal interlocutor in a communication situation by other speakers of English. Pennington (1996, 220) has established three criteria that contribute to so called good pronunciation: accuracy, fluency, and intelligibility. Regardless of which criterion one considers the most important one, pronunciation instruction altogether often remains under-represented in the classroom (Gilbert 2010, 2; Lord 2009, 364). When learning to pronounce a new language, a sufficient amount of natural language input and output are required. In other words, the learner needs to both hear the language and to produce it (Lintunen 2014, 175). Listening to non-native sounds has been shown to alter people's non-native language production (Peltola et al. 2017), and several studies have found a link between singing and pronunciation skills (Alisaari and Heikkola 2017; Chen 2016; Christiner and Reiterer 2013). Nevertheless, previous studies have mainly been interested in the effects of listening and singing on pronunciation learning when pronunciation has intentionally been trained through listening and singing exercises, while the effects of the natural language included in the music people consume in their spare time have remained undiscovered. Thus, a question arises: could the language wrapped up in the music that flows into people's ears and escapes their mouths on a daily basis also have a lasting impact? Could English music function as the necessary language input and output for pronunciation acquisition to take place? Could people learn English pronunciation through English music even without the intention of doing so? This is what the present study is set to discover.

Before establishing the research questions, it is worth taking a look at two of the key concepts of the study. In this context, the term *spare time* is used to refer to the time when one

is not at school or at work. The distinction is not always easy to make, and if one was allowed to listen to music while at school or work, it would also count as spare time listening as long as it was voluntary. Another reoccurring term, *English music*, is used to refer to all music with English language vocals regardless of its country of origin and whether the singer is a native or a non-native speaker of English. This decision was made based on the researcher's careful evaluation, according to which the English pronunciation heard in popular music is largely good enough to be considered a suitable model for English learners.

The present study intends to find answers to the following three questions:

1. How could listening to English music and singing in English explain the variation in the native-speaker evaluations of the English learners' pronunciation?
2. To what extent do accuracy, fluency, and intelligibility correlate with listening to English music and singing in English?
3. Which indirect factors could make English music a beneficial means for learning pronunciation?

The thesis begins by covering the theoretical framework of the study in sections 2–4. Section 2 introduces the key concept of pronunciation, explains how speech production and speech perception work, as well as elaborates on how pronunciation is learned both by L1 and L2 learners. It then discusses the intricate notion that is "good pronunciation", after which it sheds some light on how pronunciation can be taught. Section 3 touches upon some of the psycholinguistic factors of language learning that could indirectly influence the pronunciation learning process, namely motivation and affective factors, such as the emotions, the attitudes, and the mood of the language learner. It also briefly ponders the meaning of (musical) idols and role models for adolescents, and how they might contribute to English pronunciation learning. Section 4 comments on the relationship between music and language. First, it provides an overview of people's music consumption habits, after which it discusses whether English music could function as a means of English pronunciation learning. Section 5 begins the empirical part of the thesis by introducing the present study: the research questions and the participants are presented, and the methods of data collection and analysis are explained. Section 6 provides the results of the study, the meaning and implications of which are discussed in Section 7. The limitations of the study are also addressed in this section. Finally, Section 8 concludes the findings of the study, and suggestions are given for future research.

2 Pronunciation

This section, divided into five sub-sections, will cover the most central concept of the present study: pronunciation. Section 2.1 begins by giving definitions and by introducing different components of pronunciation. It shall also address some of the differences between English and Finnish pronunciation. After this, Section 2.2 will discuss the processes of speech production and perception, the knowledge of which is crucial for the understanding the present study. Section 2.3 will elaborate on how both first language (L1) and second language (L2) pronunciation are learned, as well as make some comparisons between the two. The intricate notion of "good pronunciation" shall be addressed in Section 2.4, along with the three individual criteria that can be used for evaluating non-native pronunciation: accuracy, fluency, and intelligibility. At the end, Section 2.5 will introduce a few of the methods in which pronunciation is and can be taught and shall consider whether English music could provide the language learner with similar tools.

2.1 Pronunciation defined

According to Derwing and Munro (2015, 5), *pronunciation* refers to all aspects of the oral production of language. This definition is unnecessarily vast as it would cover some aspects of oral proficiency that are not of interest to the present study, such as non-verbal communication and vocabulary. When searching for an alternative definition, however, it becomes clear that pronunciation is a term people are expected to know the meaning of and no definition is usually provided. Therefore, a new definition was created for the purpose of this study. Here, it is used to refer to the way in which words and strings of words are pronounced.

Fundamentally, pronunciation can be divided into two levels based on the extent of inspection. These levels are called the *segmental* and the *suprasegmental level* (Lintunen 2014, 165). The segmental level of pronunciation is concerned with the smallest components of language that affect meaning: *phonemes* (ibid.). For example, if the first phoneme of the word "hat", /h/, is replaced with the phoneme /k/, the meaning quickly changes from a head covering object into a small carnivorous mammal with sharp claws, "cat". The suprasegmental level of pronunciation, on the other hand, reaches beyond the segmental level and concentrates on vaster units of language. The term *prosody* refers to those vaster units of speech that carry across whole utterances, such as *rhythm, stress,* and *intonation* (Derwing and Munro 2015, 5). These units are important in natural speech and will be discussed further in the following chapter.

3

In the English language, the rhythm is created by the alternation of strong and weak, or stressed and unstressed, syllables (Pennington 1996, 135). The time interval between the stressed syllables tends to be fairly regular, which creates the rhythmic basis of most varieties of English (ibid.). The words carrying stress are normally lexical or content words, as in the utterance "Let's *go* for a *walk*". The unstressed words or syllables are compressed in order to fit into the average interval, which means that they are often pronounced faster and less clearly than the stressed words or syllables. This can pose a challenge for an L2 learner of English, both in hearing the language and producing it. For example, in the Finnish language the primary word stress is fixed to the word-initial syllable (Suomi, Toivanen, and Ylitalo 2008, 75), which means that Finns are accustomed to interpreting stressed syllables as beginnings of words and also to producing stress on the first syllable of a word. This habit, however, does not lend itself to spoken English too well, easily resulting in problems in comprehension. Another unit of pronunciation that is likely to cause problems for the L2 learners of English is intonation, which refers to the raising or lowering of the pitch of the voice by either tightening or loosening the vocal chords during speech (Pennington 1996, 148–149). Within the framework of this study, nonetheless, intonation will not be of great interest because it is not one of the pronunciation features one is likely to acquire through English music. This is due to songs following certain melodies and leaving little room for the melody of the language.

2.2 Speech production and perception

The production of spoken language requires careful control of multiple muscles, and perceiving speech is a complex cognitive process. Before moving on to how pronunciation is learned, it is important to understand how these processes work. As explained by Pennington (1996, 20–22), each speech sound is air that is being modified by the movement of the vocal organs. One of the most central of these organs is the *larynx*, which contains two tiny muscles that can create sound by vibrating in the throat. The space between these two muscles is called the *glottis*, the shape and size of which affect the quality of the sound that is being produced. As an example, when a vowel or a voiced consonant is being produced, for example the English /ɑ/ or /v/, the glottal opening is so narrow that it causes the larynx to vibrate. Again, when a voiceless consonant, such as the English /f/, is being produced, the glottal opening is wide letting the air run through freely. What is more, producing a word requires not only controlling the flow of air through the glottis, but also fine control of the muscles of the tongue, lips, and various regions of the mouth.

Speech production is never simply the production of individual sounds or words after one another in an isolated fashion. In the English language, words tend to "run together", as expressed by Celce-Murcia, Brinton, and Goodwin (2010, 163). This is referred to as *connected speech* (ibid.). There are several phenomena that regularly occur in connected speech that need to be considered when striving for natural-sounding and fluent production. When one is pronouncing clusters of words, the environment in which a single sound exist plays an important role; sounds might be modified, something may be added or even completely deleted depending on the neighbouring sounds. To provide a concrete example, a few the phenomena shall be introduced briefly. The phenomena of connected speech include *linking*, *assimilation*, and *deletion*, all of which contain multiple sub-categories (see e.g. Gómez González and Sánchez Roura 2016, 234–58; Celce-Murcia, Brinton, and Goodwin 2010, 163–75). Linking refers to "the smooth connection of sounds" (Celce-Murcia, Brinton, and Goodwin 2010, 165) and includes, for instance, *resyllabification* and *linking r.* Resyllabification occurs when a word terminates in a consonant cluster and is followed by a word starting with a vowel. The final consonant is then pronounced as a part of the following syllable, as in "left arm", /lɛf ˈtɑ:m/ (ibid.). Linking r, on the other hand, appears between two words when the first word ends and the second starts with a vowel sound (Gómez González and Sánchez Roura 2016, 250). For example, although the word "door" is pronounced /ˈdɔ:/ in the standard British pronunciation, the /r/ sound is present once the word appears in the utterance "the door is open", /ðə ˈdɔ: **r**ɪz ˈəʊpən/. Assimilation, again, refers to the process where a sound takes on some characteristics of a neighbouring sound (Celce-Murcia, Brinton, and Goodwin 2010, 167) meaning that the articulation of a phoneme might vary depending on where it occurs. For instance, in *progressive assimilation*, the preceding sound affects the following sound, something that commonly occurs in the regular English plural: "backs" is pronounced /bæk**s**/, while "bags" is pronounced /bæg**z**/; the voiceless /k/ and voiced /g/ condition the pronunciation of the -s ending (Celce-Murcia, Brinton, and Goodwin 2010, 168). Deletion refers to "the disappearance of a sound" (Celce-Murcia, Brinton, and Goodwin 2010, 164) and includes, for example, *syncope*, which refers to the loss of an unstressed medial vowel in certain contexts, such as in the words "chocolate", "camera", and "interesting". These words are often pronounced /ˈtʃɒklɪt/, /ˈkæmrə/, and /ˈɪntrɪstɪŋ/. These, as well as many more, naturally occurring pronunciation phenomena can be a stumbling block for non-native speakers of English, especially if not enough attention is paid on them in formal English teaching.

As for speech perception, a distinction needs to be made between *hearing* and *perceiving*. Whereas hearing is physical and largely passive, perception is the cognitive process

applied to the result of hearing (Tatham and Morton 2011, 130). One way of dividing speech perception theories is into two opposing groups: auditory- and motor-oriented theories. In auditory-oriented theories, the acoustic signal itself – the sounds that are being produced – is the object of perception, whereas in the motor-oriented theories an auditory signal only functions as the carrier of the important information: the motoric gestures that were used to produce the sounds (Scott 2017, 26). Because pronunciation is a highly motoric skill, the present study is particularly interested in the latter, and introduced next will be one of the most famous of the motor-oriented theories: The Motor Theory.

Coined by Liberman et al. in 1967 (Tatham and Morton 2011, 152), The *Motor Theory* (MT) it is still relevant today. MT is an active theory of speech perception, where the listeners decode speech based on what motor control they think was involved in the production of a particular sound, instead of perceiving speech sounds as sounds per se (ibid.). The acoustic signal reflects the gestural information behind it (Liberman, 1985). In other words, while an individual listens to speech, they subconsciously register the movements of the muscles that were used to produce the speech sounds. Several studies have found evidence that the speech-related motor areas in the human brain are indeed activated during speech perception. For instance, it has been found that depending on whether the listeners hear labiodental fricatives, such as /v/ as in "voice", or lingua-palatal fricatives, such as /ʃ/ as in "shoe", different motor areas of the brain are activated (Fadiga et al. 2002). Watkins, Strafella, and Paus (2003) found that listening to speech enhanced the size of motor-evoked potential in the lip muscles, which provided evidence that auditory speech perception facilitates the excitability of the motor system that is involved in speech production. The results of these studies suggest the existence of a sub-category of the mirror-neuron system, an *echo-neuron system*, that is, a system in humans that resonates motorically when a person listens to verbal material (Rizzolatti and Craighero 2004, 186).

More recently, Peltola et al. (2017) studied whether mere exposure to auditory stimulation could result in changes in non-native language production. They conducted two types of training on two groups of native Finnish speakers of the age 18–32 who had not studied any languages at the university level. Without any prior screening, the subjects were divided into a "Passive listeners" and a "Listen and repeat" group that were balanced in relation to age and gender. The "Passive listeners" were instructed to simply listen to the stimuli, a pseudoword pair /ty:ti/ - /tʉ:ti/, whereas the "Listen and repeat" group were instructed to repeat the same stimuli according to the model provided. The training took place on two days. While the control vowel /y/ is present in the Finnish language, the non-native target vowel /ʉ/ is not, and its

proximity with the Finnish vowel sounds /y/ and /u/ makes it theoretically extremely challenging to learn. The result of the study was that the groups did not differ in their production patterns and showed identical changes as a function of practice. The articulation improved both when it was practised by repetition and when no production exercises took place. This suggests that even mere listening to a language without productive exercises could be beneficial for learning pronunciation.

2.3 Pronunciation learning

Whenever language learning in concerned, it is worth mentioning that a distinction is often made between *second language learning* and *second language acquisition* (SLA) so that language learning is used to refer to the conscious and active learning process, such as the one typically taking place at school, while language acquisition refers to the type of subconscious learning through which a child acquires their first language (Pietilä and Lintunen 2014, 12–13). Nonetheless, SLA can be used as an umbrella term that covers both language acquisition and learning (ibid.), and in the context of the present study, the terms will be used interchangeably; a difference will be made between implicit and explicit learning instead. *Second language* (L2) refers to a language that is known in addition to one's first language, but it needs not be the language that was chronologically learned second. *First language* (L1), on the other hand, means the person's native language or mother tongue (Pietilä and Lintunen 2014, 14–15).

Fundamentally, two components are required for both L1 and L2 pronunciation learning to take place: language *input* and *output*. In other words, it is crucial for the learner to hear a sufficient amount of the language they are learning and to also produce it. Input refers to all the language material surrounding the learner from which it is possible to acquire new information about the language (Pietilä and Lintunen 2014, 16), whereas output refers to the language the learners themselves produce (Swain 1995, 125). Because pronunciation differs from other language proficiency areas by being a highly motoric skill, Lintunen (2014, 175) highlights the importance of output by saying that pronunciation is basically learned by pronouncing, as long as the cognitive side is also kept in mind. Producing the language oneself also pushes the learner to process the language more deeply than input (Swain 1995, 126). The term *comprehensible input* is included in one of Krashen's most famous L2 learning theories, the *Input Hypothesis* (1977), and it refers to language that is slightly above the learner's current proficiency level yet still understandable. For instance, one might come across a strange word or phrase while watching TV or listening to music but will be able to figure out its meaning with the help of the context. According to Krashen, comprehensible input is the single most important source of L2

learning. The Input Hypothesis is based on the principle that L2 acquisition resembles L1 acquisition in numerous ways, and Krashen argues that the mechanisms of L2 learning are essentially very similar to those of L1 learning. The L1 mechanisms will be discussed briefly in the following paragraph.

Children learn their first language by listening to L1 input and by producing L1 output. In fact, the learning process begins before the child is even born. The human auditory system enables response to sound by the gestational age of approximately 25–28 weeks, despite the functional maturation still being in progress (Litovsky 2015). A study by Moon, Lagercrantz, and Kuhl (2013) measured differences in newborn babies' responses to familiar and unfamiliar vowels, and the results suggested that the auditory input to which unborn babies are exposed to in the womb shapes their perception of their native language at a phonetic level. Similar results have been found in Finland by Partanen et al. (2013), who exposed fetuses to pseudowords. The fetuses that had been exposed to the stimuli showed enhanced brain activity when presented with the same stimuli after birth. This suggests that by the time a baby is born, it has already developed an idea of the sound patterns of their first language through listening. Also later on, children acquire their L1 by extracting information from the speech stream (de Carvalho et al. 2018, 18). In practice this occurs through observing and listening to their caretakers, and by repeating what they hear. It is a top-down process, whereby knowledge of the language is built little by little from the available cues (Lintunen 2014, 183).

When it comes to L2 learning, it cannot directly be assumed that adolescents and adults would function the same way as young children learning their L1, and in the course of SLA research, the trends have been alternating between emphasizing the similarities and dissimilarities of L1 and L2 acquisition. According to Dörnyei (2009, 21), the main differences lie mostly in success rate and the role of motivation; under normal circumstances L1 acquisition is always successful, whereas in L2 learning the ultimate attainment, or the final stage of learning, varies enormously among learners. As for motivation, it does not play a role in L1 learning like it does in L2 learning. Dörnyei (ibid.) notes that although the L1 learning mechanisms are less powerful in the L2 learner, it is still possible to access them at least partially. Children are sensitive to acquiring their first language, and a *sensitive period*, or sensitive periods, also exist for L2 learning. A sensitive period is a time of heightened sensitivity to the stimuli to trigger learning, for example the sensitivity to subtle sound distinctions in early childhood (Long 2013, 4). However, the human brain is very plastic throughout life. *Plasticity* refers to the ability of the brain's nervous system to be shaped in response to people's experiences (Møller 2009, 27). For example, Barbeau et al. (2017) found

that learning-induced plasticity in the adult brain occurred as early as 12 weeks into an L2 training period.

Before it is possible for the language learner to produce the target language (TL), the language they are learning, they need to listen to it first. Listening to the TL improves the learner's ability to distinguish its sounds and gives birth to new memory traces, or engrams, in the learner's brain. According to the *Perceptual Assimilation Model* (PAM) by Best (1995), people assimilate non-native phones to those native phoneme categories that include the most similar phones. In other words, before learning more about the sounds of a new language, people tend to listen to them through their native language. If a certain sound does not exist in their L1, it is then interpreted as something that does. For example, the aspiration of plosive sounds is not typical of the Finnish language, so the English word "tap", [tʰæp], easily becomes [tæp] when interpreted by a Finnish speaker. Because the sound has not been heard accurately, it cannot be produced accurately either and, as a result, the word could turn into something closer to "dap" once produced by a native Finnish speaker. Consequently, it is crucial for an L2 learner to learn to listen to the TL correctly, only after which can they learn to produce it themselves. For the listening ability to develop, the learner has to receive plenty of L2 input, typically in the form of L2 speech. An L2 learner with access to large amounts of natural L2 input would most likely achieve a more advanced L2 pronunciation than a learner with less input. What language learners may not realize is that spoken language is not the only possible source of natural input: vocal music, for example, could also provide the learner with vast amounts comprehensible input – implicitly.

According to Ellis (1994, 1), language can be learned in two ways: implicitly and explicitly (Ellis 1994, 1). *Explicit learning* refers to the learner's conscious and deliberate attempt to learn a new skill (Dörnyei 2009, 136) and it is typical of classroom instruction. For example, when a teacher introduces the pronunciation of a new English word to the class, the students aim to learn it. Explicit instruction helps the students notice certain features in the TL input, and also make comparisons between the input and their own output (Ellis 1994, 106). *Implicit learning*, on the other hand, takes place naturally and without conscious attempts from the person to learn something. In implicit learning, certain regularities in the language patterns are unconsciously picked up in environmental or contextual stimulus, resembling the way children are hypothesized to learn their first language. For instance, as one listens to music in English, English words keep coming up, which could potentially result in the implicit learning of pronunciation patterns. Information acquired implicitly becomes unconscious and automatized *procedural knowledge*, whereas classroom teaching would normally produce

declarative knowledge that the learner will be able to explain verbally (Sundman 2014, 121). *Automatization* refers to "the absence of attentional control in the execution of an activity" (Segalowitz and Hulstijn 2005, 371), and in this case it could mean the correct pronunciation of a word without having to think about how it is pronounced.

Learning something implicitly does not imply that attention would not be paid to the language input at all. Usually, attention needs to be paid to the stimulus for learning to take place (Dörnyei 2009, 138). If attention is paid to the input, the input becomes *intake*, which refers to the language input that is being utilized by the learner (Pietilä and Lintunen 2014, 17). Schmidt presented the *Noticing Hypothesis* (1990) that was concerned with how consciousness is related to the acquisition and development of language skills, including discourse abilities. He defined *noticing* as "the basic sense in which we commonly say that we are aware of something" (Schmidt 1990, 132). His argument was that subconscious language learning is impossible for conscious noticing is a necessary condition for language input to become intake (Schmidt 1990, 149). That is to say, if one does not notice a certain feature of the language, one will not learn it either. For example, listening to English music is unlikely to result in language learning if no attention is paid to the language of the vocals. In support of Schmidt's hypothesis, it has been shown that sheer frequency of input does not automatically lead to language acquisition (Carroll 1999).

2.4 "Good pronunciation"

An English learner may wish to acquire a good pronunciation of English. The concept of "good pronunciation", nonetheless, is somewhat complicated. The questions that arise include how pronunciation should be evaluated, as well as whether one should aim at a native-like accent or settle for a non-native pronunciation model. To ease the evaluation, Pennington (1996, 220) established three criteria that together contribute to good pronunciation. These are *intelligibility, accuracy,* and *fluency.*

Celce-Murcia, Brinton, and Goodwin (1996, 8) note that only few non-native speakers will ever reach a native-like level, so simply enabling the learners to rise above the threshold that their pronunciation does not interfere with their ability to communicate might be a more realistic goal. This is also the idea behind the *intelligibility principle*, according to which it is enough for the language learner to achieve understandable speech irrespective of its native-likeness (Levis 2005, 370). Intelligibility refers to the degree of match between the speaker's intended message and the listener's comprehension (Derwing and Munro 2015, 5). Intelligibility is hearer-based, meaning that the listener functions as an important variable in

whether the speaker's pronunciation is understood (Levis 2006, 259). For example, two Finnish speakers of English might be able to understand each other's accents perfectly while a native speaker of English has no idea what they are talking about. According to Pennington (1996, 220), intelligibility is the most important criterion since a sufficient amount of mutual intelligibility between the speakers is required for successful communication to take place.

A competing orientation, the *nativeness principle,* focuses on the accuracy of pronunciation (Levis 2005, 370). Accuracy was defined by Skehan (1996, 23) in the terms of how well the language is pronounced in relation to its rule system. Basically, it is a question of whether the speaker pronounces words correctly. The goal of the nativeness principle is for the learner to develop such advanced and accurate non-native speech that it is essentially indistinguishable from that of a native speaker. However, many non-native speakers have a noticeable accent that does not interfere with their ability to communicate, and an utterance can be heavily accented yet fully intelligible (Derwing and Munro 2015, 5). The pronunciation teaching techniques that simply focused on the accuracy of pronunciation were mainly abandoned in the late 1970s following the take-over of the *Communicative Approach*, according to which the primary purpose of language is communication (Celce-Murcia, Brinton, and Goodwin 1996, 10). Today, as well, using language to communicate is usually the central view in classroom pronunciation instruction.

The third criterion of good pronunciation is fluency that can be defined as "[t]he flow, fluidity, or smoothness of speech" (Derwing 2018, 321). "Fluent speech" is sometimes used as a synonym for "good speech" (Pietilä and Lintunen 2014, 22) for it is a question of how easily the speech flows without too many hesitations, pauses, and self-corrections (Derwing and Munro 2015, 4). In natural, fluent speech, sounds and words blend into one another (Lintunen 2014, 166) and are not always pronounced clearly and accurately. Therefore, practicing vaster entities is beneficial when it comes to practicing pronunciation for they have been found to have the biggest impact on the overall comprehensibility of a speaker (Munro and Derwing 1995). For instance, instead of repeating the words "hat" and "cat" independently, one should come up with a sentence that includes both words and repeat the whole sentence: "There's a cat in the hat".

Still, according to Tergujeff (2013), the formal instruction of English in Finland often tends to focus on the segmental level while little explicit instruction on the suprasegmental, or prosodic, features of speech is offered. This does not fully follow the communicative recommendations for pronunciation instruction. Generally speaking, it is not uncommon for pronunciation to be under-represented inside the English classroom (Gilbert 2010, 2). Lord

(2009, 364) addresses the issue as follows: "Many [language instructors] assume that students will eventually learn the pronunciation on their own". In the Finnish education system, for example, vocabulary and grammar often receive more attention than pronunciation. According to Lord (ibid.), if the pronunciation of a language learner does not match the native form, they will immediately be recognized as foreigners or "outsiders". This is not necessarily a problem, but in some cases listeners may evaluate the non-native speaker negatively because of their heavily accented speech (Derwing and Munro 2015, 2).

Some native accents can also be viewed more positively than others (Lintunen 2014, 168), and non-native speakers of English should be aware of the attitudes their accents may provoke. Especially native speakers are extremely sensitive to accents different from their own (de Jong 2018, 10), and there being so much variability in native English, deciding which accent norm is the most appropriate for a learner is complicated. When it comes to choosing a pronunciation model, the learner should bear in mind that every existing accent, even the ones considered standard, can evoke prejudice in other speakers of the language. Someone who has been taught *Received Pronunciation* (RP) of British English, also known as *the Queen's English*, could come across as "posh" to a native speaker even if it was never their intention. Also, if an English learner was to choose a singer's accent as their pronunciation model, they could sometimes end up with a very simplified accent. For example, British artists often tend to sing in a more "neutral" or American-like accent so that even if they normally pronounced the word "can't" as /kɑːnt/, they might have chosen to pronounce it as /kæːnt/ when singing. Even so, this type of modified accent would be a rather safe choice for an L2 learner. In the end, it depends on the language learners' personal motives and goals which accent and pronunciation learning principle they find the most suitable.

2.5 Pronunciation teaching methods

Celce-Murcia, Brinton, and Goodwin (1996, 8–9) provided a comprehensive list of techniques that have traditionally been used in pronunciation teaching, and introduced here are three of them that could have implications for the present study. These include *tongue twisters, minimal pairs*, and *listen and repeat exercises*. Tongue twisters are challenging combinations of sounds and words that were originally used for native speakers' speech correction but are useful for L2 learners as well (ibid.). "Sheila sells sea shells on the sea shore" is one of the most classic tongue twisters, the challenge being not to confuse the /ʃ/ and /s/ sounds in fast repetition. Similar wordplay is often found in song lyrics, for instance try repeating "And now a piece of me is a piece of the beach and it falls just where it needs to be and rests peacefully" (Sheeran

and Gosling, 2011). Also, Eminem's "Without Me" (Mathers et al. 2002) is basically four minutes and 50 seconds of consecutive tongue-twisters. These are only two examples. Minimal pairs and contextual minimal pairs, on the other hand, are pairs of words that differ only in a single sound, for instance "ship" and "sheep" (/ʃɪp/ and /ʃiːp/), where the vowel is different. In minimal pair drill exercises two similar sounding words are introduced to learners either independently or in a context in order to help the learners distinguish between the sound contrasts in the TL through listening discrimination and highly controlled spoken practice (Celce-Murcia, Brinton, and Goodwin 1996, 8–9). Presenting the words in an established context facilitates sound discrimination: for example, imagine having to decide which of the two words is more likely to be present in the following sentence: "The ship/sheep was sailing on an open sea." Also songs tend to present words in a context. In listen and repeat exercises learners listen to a model and try to repeat it as well as they can. These types of exercises have been found very effective (Macdonald, Yule, and Powers 1994), and again, it is not difficult to see why music could be thought to function as a listen and repeat exercise. When a language learner listens to a song, possibly singing along and perhaps even singing it later without the original music playing at the same time, they listen and repeat. Tergujeff (2012) found listen and repeat exercises to be the most commonly used pronunciation task during her observations on English pronunciation teaching in Finland.

In English teaching, the teachers' pronunciation acts as the model that the learners imitate, which is a major responsibility. At the University of Turku, a minimum grade of 3 is required for future English teachers in the mandatory English pronunciation course. Considering that the course is graded from 1 to 5 (1 being the lowest grade and 5 being the highest), 3 implies average knowledge. Since average pronunciation skills are enough for one to become a teacher, the case could sometimes be that a teacher does not feel fully comfortable with their own English pronunciation. This might result in pronunciation being over-shadowed by other proficiency areas in their teaching.

What is also worth mentioning about English teaching in Finland is that the written forms of English words are typically present from the beginning. Unlike English, Finnish has a transparent writing system, indicating a strong correspondence between letters and phonemes, and it was found by Peltola et al. (2015) that visual, orthographical cues altered the pronunciation of a word towards the visual cue. The fact that English words are presented in their orthographical form could confuse the learners and cause them to pronounce English following the same habit in which they pronounce Finnish: the way it is written. Providing the

learners with natural, phonemic models of the TL before being introduced to the written form could be realized by, for example, incorporating music into the teaching.

Heikkola (2018) investigated how three different musical teaching methods affect pronunciation with the help of another researcher, Jenni Alisaari. The methods were singing, listening to songs, and repeating song lyrics rhythmically. The 67 test subjects were university students aged 18–33 who studied Finnish as a second language. The students were of two proficiency levels, beginners (I) and advanced beginners (IIA). All in all, there were six different groups: singing group I and IIA, listening group I and IIA, and rhythm group I and IIA. The same songs were used for teaching all the groups, but the teaching method was different for each one. The songs were mainly Finnish children's songs but also included a couple of pop songs. The singing groups practised pronunciation by listening to one researcher sing a verse after which they sang the same verse with the other researcher. Pop songs were listened to from a tape and were sung with the researchers to the original music. The listening groups listened to the songs multiple times on various days. The rhythm groups recited the song lyrics verse by verse after listening to the researcher read the verses and then read through the lyrics one more time.

To evaluate the pronunciation, Heikkola used native Finnish evaluators. The results showed that out of the three teaching methods used in the study rhythmical repetition of song lyrics was the most effective manner to develop a more native-like accent; a significant improvement was only found here, although all of the groups came across slightly less foreign in the posttest compared to the pretest. This was believed to be due to the fact that rhythmical repetition was closest to natural speech production. A surprising finding was that among the IIA listening group the pronunciation actually seemed to have worsened, which according to Heikkola could be explained by listening to music being the most passive teaching method used in the study. The 2018 result differs from another study by Alisaari and Heikkola (2017), where singing was found the most effective way to improve one's L2 pronunciation.

Section 2 has covered one of the fundamental concepts of the present study: pronunciation. It has explained what pronunciation is and what it consists of, how speech is produced and perceived, how pronunciation is learned and taught, and whether "good pronunciation" exists. A few previous studies related to the study at hand were also introduced. The aim has been to lay a well-founded basis for the hypotheses that shall be presented in Section 5 by shedding light on why listening to English music and singing in English could be beneficial activities when it comes to learning English pronunciation. Discussed next are a few factors of a more indirect nature through which music might influence pronunciation learning.

3 Psycholinguistic perspectives

In this section, the two tightly intertwined concepts, motivation and affect, shall be discussed in the light of pronunciation learning and language learning in general. The section is accordingly divided into two sub-sections. First, Section 3.1 shall address the learner's motivation to acquire a certain language, which is generally considered a strong predictor of L2 achievement. The section will consider integrative and instrumental motivation and how they affect learning, as well as ponder how the concept of ideal self and English-speaking (or English-singing) musical idols are related to language-learning motivation especially in adolescents. Second, Section 3.2 shall discuss affective factors that are linked to L2 learning. Affective factors, such as the learner's mood and attitudes, alter people's receptivity to take in new information. They can also have an impact motivation. Music, for its part, has been found to influence affect. At the end of the section, the difficulty of studying motivation will also be addressed briefly.

3.1 Motivation

Motivation is a dynamic, multifaceted construct consisting of various components. In a simplified manner, it could be described as the driving force behind the actions people take, how much time they are ready to invest in the actions, and how diligently they engage in them (Dörnyei and Skehan 2003, 614). That said, motivation is also a strong predictor of L2 learning achievement, which was demonstrated by Masgoret and Gardner (2003), who conducted a meta-analysis of empirical studies involving 75 independent samples and a total of 10,489 L2 language learners. The results suggested that out of five attitudinal-motivational variables, all of which have been found to have a positive relationship with L2 achievement, motivation was the strongest single predictor of achievement. Other variables included attitudes towards the learning situation and integrative orientation, both of which shall be commented on later in this section. Directly related to pronunciation learning, Moyer (1999) and Bongaerts et al. (1997) found that motivation was the most important factor in explaining why the pronunciation of some non-native speakers of English resembled that of native-speakers more than others'. Also according to Pietilä (2014, 49), motivation is considered one of the most important factors in causing individual differences in L2 learning. Other factors include personality, age, learning style, and learning strategies (Pietilä 2014, 45–66).

There are multiple ways of categorizing motivation. Chosen for the purposes of the present study is the one used in the *Socio-Educational Model of SLA*, developed by Gardner

(1985), which divided L2 learning motivation into two types. The model makes a clear distinction between *instrumental* and *integrative motivation*, and this distinction is illustrative enough for addressing motivation in the context of this study. Instrumental motivation refers to the learner's interest to learn the language in order to achieve a practical goal (Pietilä 2014, 50). Studying English vocabulary merely to pass a test would be a fine example of instrumental motivation. In this case, the learner has no inner aspirations to master the language. Integrative motivation, for its part, is characterized by the learner's sincere and personal interest in the TL and the TL culture. It may involve the desire and goal to identify with the speakers of the language (ibid.) or the desire to "be one of them". The process whereby members of one cultural group adopt perspectives and behaviour patterns from another group is called *acculturation* (Singleton 2014, 91). Schumann (1986, 379–392) proposed an *Acculturation Model,* according to which the extent and quality of contact between the language learner and the TL and TL culture predicts the success of L2 acquisition. Also according to Mishan (2005, 27), integratively motivated learners have greater success in learning a new language than instrumentally motivated learners.

An English learner might have an English-speaking musician as an idol or role model. In addition to listening to their music, many also follow their idols through social media. According to Raviv et al. (1996, 632), popular musicians serve as idols or role models that can influence adolescents through social learning. The idolization of (pop) singers by adolescents is based mostly on two components, *modelling* and *worship*. In the context of the present study, the concept of modelling is of special interest, because it refers to the desire to be like the idol, which might involve imitation by dressing up the same way, copying the idol's hairstyle, activities, speech, or any other social behavioral patterns (ibid.). The pronunciation of English by the idol is one of the features that may be copied, intentionally or unintentionally. Worship, for its part, refers to an unusually intense admiration and respect of an idol, which can be expressed in behaviours such as actively collecting information and artefacts related to the idol. Some "worshippers" may even try to meet the idol personally (ibid.). Dreaming about meeting an English-speaking idol could motivate the person to practise English, for example in the form of imaginary conversations, in case this encounter was to happen. The learner might also want to be able to dive deeper into the musician's English-speaking world, which could increase the learner's overall motivation to learn English. In a sense, music could act as a gateway to personalized learning content that is relevant and meaningful, and having this type of learning content is a powerful strategy for promoting inner, or integrative, motivation (Chambers, 1999).

The concept of *ideal self* was introduced by Dörnyei (2009, 217–218) along his three-component *L2 Motivational Self System*, a model centered around people's visions of their future selves. People have a psychological desire to reduce the discrepancy between their actual selves and their ideal selves (ibid.). The ideal self may possess features that the learner sees in other people, either in someone they know personally or perhaps in the English-speaking role model. For example, if the English learner's ideal self is a fluent user or a native speaker of English, this could function as a powerful motivator for learning English. Another component of the L2 Motivational Self System is the *ought-to self*, concerned with those attributes that the learner believes they should possess in order to meet expectations, yet this is not of interest to the present study. In the field of developmental psychology, there is an increasing body of support for the developmental importance of music to adolescents. Music is a resource through which young people explore possible selves, rehearse social roles, and envision future orientations by observing their favourite musicians (Miranda, 2012). The final component of Dörnyei's model is *L2 learning experience*, which includes the learning experience and the learning environment. A language learner always exists in a learning environment. Learning experience and environment function as independent external variables that influence the learner-internal variables, such as motivation, development, and performance (Ushioda 2015, 47). Positive learning experiences and engaging with the TL can result in images on increasingly sophisticated language use and cross-cultural encounters, and therefore the enhancement of the ideal L2 self (ibid.). Positive learning experiences and success in language learning can also improve the learner's attitudes towards learning that language (Singleton 2014, 92).

3.2 Affect

In the SLA context, *affect* can be defined as "aspects of emotion, feeling, mood or attitude which condition behavior" (Arnold and Brown 1999, 1). Language learners' attitudes are closely related to the learners' motivation, and they also influence the degree of L2 pronunciation proficiency that is achieved (Celce-Murcia et al. 2010, 35). The term *attitude* refers to "a psychological tendency that is expressed by evaluating a particular entity with some degree of favor or disfavor" (Eagly and Chaiken 1993, 1). For instance, every English learner evaluates English in some way; they may think it is a fascinating language whose speakers are nice, or perhaps that English is a complete waste of time and the speakers are annoying. The attitudes of the language learner towards the TL are highly determined by the learner's beliefs about that language, its speakers, and culture, and they feed into the processes related to

becoming proficient in a particular language (Singleton 2014, 92). Especially for adolescents, musical preference can act as a "badge of identity" that also predicts other aspects of lifestyle and attitude (North and Hargreaves 1999).

According to Dörnyei (2009, 184), the L2 learning process is highly emotionally loaded, and empirical findings have demonstrated that positive mood enhances learning (Lee and Sternthal 1999). This has to do with *affective filters*, a term coined by Krashen (1982) that refers to any psychological obstacle that prevents the learner from realizing their full potential in the learning situation. The lower the filter, the more receptive the person is for learning. Affective filters are part of Krashen's *Affective Filter Hypothesis*, according to which a language learner's mental state affects learning either by facilitating or hindering it (ibid). Affective filters, such as anxiousness, can prevent one from communicating in the L2 by making them reluctant to do so (Szyszka 2011, 295). This will obstruct them from improving their oral communication skills, and therefore it would be important to reduce anxiety in a pronunciation learning situation. Being motivated, on the other hand, lowers the affective filter.

Emotions are also important motivators (MacIntyre, Mackinnon, and Clément 2009, 47). Aguirre, Bustinza, and Garvich (2016) studied how the use of English songs could help motivate L2 learners of English in Peru. They examined 3rd grade students through observation and a questionnaire. Whereas the questionnaire results did not reveal a positive change between the classes with and without music, the observations showed that the environment in the classroom became more positive and less restless when a song was playing in the background. The learners also paid more attention and showed greater desire to participate. Sigurðardóttir (2011) noticed a similar effect as she incorporated music into her classroom teaching in order to make her classes more interesting; the students became more actively engaged in the teaching materials. These results together suggest that by affecting the learners' emotional state in a positive way, music could help motivate language learners as well as make them more receptive to teaching and open for learning. In this way, music could also help build a less stressful learning environment. In addition, it has been found that especially music played at 60 beats per minute produces a state of relaxation in both children and adults, because this tempo resembles the ideal resting heart rate of humans (Chalmers, Olson, and Zurkowski 1999, 44).

Because of its multifaceted nature, studying motivation and its related concepts comes with a challenge. Not being directly observable, motivation is notoriously difficult to measure objectively. In the context of second language acquisition research, motivation has generally been studied by collecting self-report data from L2 learners (Ushioda and Dörnyei 2012, 397). The information gathered via self-reporting, however, is not always entirely reliable. According

to Pietilä, evaluating one's own motivation can be challenging and, also, the answers might be embellished (2014, 49).

This section has touched upon the cognitive side of language learning by discussing motivation and affect, both of which could potentially be positively influenced by English music, as was suggested by the studies that were introduced. The section has also addressed the concept of idols since it is very often connected to music. The purpose of this section has been to demonstrate how English music could function as an indirect factor impacting English learners' pronunciation learning, and in this way create further basis for the hypotheses that are to be presented in section 5. Before that, however, Section 4 shall discuss the relationship between music and language.

4 Music and language

In this section, music is considered from the perspective of language learning. Section 4.1 begins by providing an overview of people's contemporary music consumption habits. Section 4.2 shall then comment on why music could be thought of in more than solely cultural terms. This is followed by a description of what happens in the human brain when people hear music. A previous study conducted on the effects of integrating music into pronunciation teaching will also be introduced in this section. Finally, the relationship between musicality and language aptitude shall be addressed briefly.

4.1 Music consumption of today

IFPI and Teosto (2017) interviewed 1.019 Finns aged 16–65 about their music listening habits. IFPI (the International Federation of the Phonographic Industry) is an organization that represents the interests of the music recording industry worldwide, whereas Teosto is a Finnish performance rights organization that collects royalties on behalf of composers and songwriters. Based on the interviews by these two organizations, the average time of focused listening was 15 minutes per day, whereas the time people reported listening to music on the background was one hour per day. In a week, this makes up for 1 hour and 45 minutes of focused listening, and 7 hours of background music. Statistics by age group were not provided in the report, and this was also mentioned as a limitation of the study. In order to gain a reliable picture, the results should be analyzed both by age group and as a whole. However, a similar study (Delmonte 2017) conducted in the United States on respondents of the age 16 and older indicated that, on average, younger people tend to listen to more music than older people. In this study, the distribution of focused and unfocused listening was not specified.

Globally, people spend 2 hours and 30 minutes per day, and 17 hours and 48 minutes per week listening to music, as stated by the latest IFPI Music Consumer Insight Report (IFPI 2018). This study was carried out on a demographically representative sample of 16–64-year-olds in twenty countries. The report also revealed that people are now listening to music at all points of the day, which demonstrates the importance and value music holds in people's lives. Music was listened to while commuting, driving, working or studying, relaxing, cooking and cleaning, exercising, as well as when going to sleep. Younger music consumers were more likely to listen to music during any activity than older consumers. In Finland, car and home were the two most popular places for listening to music and, for people under the age of 25, listening to music on the go seemed to be almost a standard practice (IFPI and Teosto 2018).

As for the language and type of the music, according to the most recent IFPI and Teosto report (2018), the music Finns had listened to the most during the past week was Finnish pop (67%), yet English pop had been listened to by the almost same percentage (65%). It was also found that 16–25-year-olds are especially interested in artists they can relate to, which describes the personal significance music has to young people.

4.2 Music and pronunciation learning

According to Wolfe (2001, 160–161), people are accustomed to thinking of music only in cultural or artistic terms, yet it has been found to be a highly complex neural activity. She notes that the mental mechanisms present at music processing are deeply entwined with other basic functions of the brain – including emotion, memory, and language. Singing and speaking are, after all, sustained by the same speech generation process (Christiner and Reiterer 2013), and similar to language, music is a way of communication. It has even been called the universal language of mankind by some, because even if people do not share a spoken language, they are often able to connect through songs and melodies that are familiar to both. In addition to the similarities in processing music and language, music possesses other elements that could be beneficial for language learning.

Ortis (2008, 216) explains what happens in the human brain when people hear music. The emotional response to music releases *serotonin*, a neurotransmitter sometimes called "the happy chemical", which has been linked to various emotional and motivational aspects of human behavior, including anxiety and depression (Meneses and Liy-Salmeron 2012). Among other things, it contributes to people's wellbeing by regulating their mood (Ortis 2008, 216). When people hear music, the serotonin levels in the brain rise producing a stress relieving effect, and once this response is involved in a learning situation, the mind is in a receptive state for accepting new information (ibid.). Although people usually listen to music for reasons other than learning, be it pure enjoyment or stress relief, the increasement of serotonin is still making the mind receptive for information, for example, English pronunciation patterns.

The effects of integrating music into the English pronunciation instruction were explored by Chen (2016). The participants, 95 advanced English-learners in Taiwan, were advised to familiarize themselves with six songs chosen for the purposes of the study and were told to pay attention to the meaning of the lyrics, as well as asked to memorize the lyrics so that they would be able to sing the songs on their own. The participants were provided with the lyrics, lists of the vocabulary, as well as the phonetic transcriptions of the words. The data were collected via three questionnaires, four quizzes, and two exams. The results of the final

questionnaire indicated that 97% of the students felt like they had made progress in their pronunciation and felt more familiar with the pronunciation rules of English. A similar result was revealed by the quizzes: by having memorized the lyrics the students were able to pronounce words with more ease. As for the enjoyability of musical exercises, 77% of the students reported having enjoyed singing, whereas as many as 98% had enjoyed listening to music. The participants were also asked to indicate which song was their favourite, and there seemed to be a relatively direct correlation between the song preference and the success in the quiz concerning that particular song. The results suggest that music could be a useful tool for pronunciation learning because it makes the learning experience pleasant and subjectively positive, and helps the learners remember the pronunciation of words.

Memory is an important component in acquiring new information and recalling it later (Pietilä 2014, 47). According to Wolfe (2001, 162), "[t]here is little doubt that when information is embedded in music or rhyme, its recall is enhanced". Surely, most people have been in a situation where they hear a song from years ago and somehow remember the lyrics. Similarly, many people have been in a situation where they have failed to recall crucial information during an exam. However, people might be able to remember a certain language feature, for example a grammatical pattern or the pronunciation of a challenging word, had it been incorporated into a song they know. Ludke, Ferreira, and Overy (2014) also suggested that melody can act as a facilitator for language learning. In addition, music could be beneficial for pronunciation learning also in the sense that song lyrics are likely to contain words that are completely new to the learner.

What cannot go unmentioned when music and language are being discussed is the relationship between *musicality* and language aptitude, a very popular research subject in the 21st century. Musicality or *musical ability* refers to people's ability to make sense of music (Sloboda 1993, 106). The link between musicality and language aptitude has been established by several studies (see e.g. Milovanov 2009; Nardo and Reiterer 2009; Slevc and Miyake 2006). The main finding has been that, compared to their non-musical peers, musically talented people are often more successful at learning foreign languages What is more, musicality has been found to be linked to phonetic language aptitude in particular. For instance, according to Christiner and Reiterer (2013), good singers often exhibit excellent speech imitation abilities, which was thought to stem from their vocal flexibility achieved through the physical training of the vocal organs and the fact that singers are open to new sound combinations throughout adulthood.

Therefore, Christiner and Reiterer (ibid.) suggest that singing training could also be applied to L2 pronunciation teaching. However, as Cockburn (1991, 73) points out, some teachers can be very uncomfortable with singing, and if a teacher was to use singing exercises in the classroom, it would be important for them to come across comfortable. Students can recognize the discomfort of the teacher, which will make them uncomfortable and reduce the effectiveness of the pronunciation exercise. In case the teacher wishes to use music despite not being able to carry out the exercises by singing themselves, using recorded music, especially chosen by the students, could be the solution. It has been found that exposing learners to variable input and many different voices appears to assist the development of accurate cognitive representations of sounds (Iverson, Hazan, and Bannister, 2005), which suggests that listening to different artists may benefit the language learner by providing them with the TL produced by different people. An advanced learner is likely to understand why some words are pronounced differently by different singers, whereas a beginner might get confused. In this sense learning English pronunciation through popular music without any formal instruction would work better for more advanced L2 learners.

This section has presented the music consumption habits of the contemporary people both globally and in Finland, as well as discussed the relationship between music and language. It has introduced a related study, the results of which suggested that listening to English music and singing English songs could be a useful tool for practising pronunciation. Also, the alleged connection between musicality and language aptitude has been addressed. The aim of this section has been to further demonstrate why musical activities in the English language also in one's spare time could be beneficial for pronunciation learning purposes. Section 5 shall now introduce the present study.

5 The present study

This section of the thesis will cover the methodology of the present study, and it is divided into three sub-sections. Section 5.1 will begin by introducing the participants of the study. Section 5.2 shall then move on to the explaining how the data were collected and elaborate on the procedures that were applied in the analysis of the data. Section 5.3 will account for the so-called musical index that was counted for each participant in order to facilitate the analysis of the results. At first, however, the section will be returning to the three research questions that were established in Section 1 and shall present them along with their hypotheses. The present study intends to answer the following questions:

1. How could listening to English music and singing in English explain the variation in the native-speaker evaluations of the English learners' pronunciation?
2. To what extent do accuracy, fluency, and intelligibility correlate with listening to English music and singing in English?
3. Which indirect factors could make English music a beneficial means for learning pronunciation?

Firstly, the study is concerned with how listening to English music and singing in English could explain the variation in the native-speaker evaluations of the English learners' pronunciation. In other words, the question here is how great of an effect these two activities could have on the English learners' pronunciation and why. It was hypothesized that the consumption of English music of one's own choosing in one's spare time could be one of the several factors improving people's English pronunciation, because listening to music and singing, both independently and combined, contain elements of a natural language learning situation (de Carvalho et al. 2018, 18; Moon, Lagercrantz, and Kuhl 2013; Partanen et al. 2013; Krashen 1977) and provide useful tools for learning pronunciation (Alisaari and Heikkola 2017; Chen 2016; Ludke, Ferreira, and Overy 2014; Christiner and Reiterer 2013; Wolfe 2001, 162; Celce-Murcia, Brinton, and Goodwin 1996; Best 1995; Munro and Derwing 1995) This question will be approached by performing the procedures that will be introduced soon in sections 5.1 to 5.3.

Secondly, the study sets out to discover the extent to which accuracy, fluency, and intelligibility correlate with listening to English music and singing in English (see Pennington 1996). This question will be examining the three above-mentioned pronunciation areas separately, and shall be approached purely quantitatively; is there a correlation between the learners' engagement in musical activities in English and how well they scored in the three

individual pronunciation areas? The hypothesis is that the scores obtained from the three individual areas do not correlate equally with the engagement in musical activities, because accuracy, fluency, and intelligibility features are not equally represented in music. For instance, sung language may be fluent yet not precisely accurate, which suggests that fluency features are more likely to be acquired than accuracy features. Nevertheless, it remains to be seen which area reveals the strongest link to the musical activities.

Thirdly, the study is concerned with the indirect factors that could make English music a beneficial means for acquiring English pronunciation. Listening to music and singing have some psycholinguistic bearings that could be advantageous for language learning in a more general sense, and it is hypothesized that English music of one's own choosing could boost the learning of English pronunciation by increasing the learner's motivation (Aguirre, Bustinza, and Garvich 2016; Sigurðardóttir 2011; Dörnyei 2009, 217–218; Mishan 2005, 27) and by altering their emotions towards a more receptive mind state (Ortis 2008, 216; Lee and Sternthal 1999). Also, the learner's attitudes towards English, its speakers, and learning English are hypothesized to improve as a result of English music. Previous studies on language learning motivation have found motivation to be a strong predictor of L2 learning and L2 pronunciation learning achievement (Masgoret and Gardner 2003; Moyer 1999; Bongaerts et al. 1997). Affective factors are believed to impact the learners' receptivity to the TL (Szyszka 2011, 295; Krashen 1982). Positive attitudes, on the other hand, feed into the processes that is related to becoming proficient in an L2 (Singleton 2014, 92; Celce-Murcia et al. 2010, 35). This issue will be approached by the quantitative examination of questions interested in music-related affective factors.

A quantitative design was chosen in order to conduct an analysis on the complex relationship of listening to English music and singing in English, and the pronunciation of the English learners in the most reliable manner possible. Quantitative studies are theory-driven, and their advantages are that they are systematic and controlled and produce generalizable data that are quick to analyze due to their directly quantifiable nature (Magnan and Lafford 2012, 530). This was thought to facilitate the investigation of the phenomenon as well as to enhance the reliability of the study by leaving little room for researcher subjectivity. Quantitative research is sometimes criticized for obscuring individual differences and lacking their possible explanatory factors (ibid.), and this was acknowledged by accompanying the quantitative questions with a few open-ended questions where the participants could elaborate their answers.

5.1 Participants

The sample of the present study consisted of 44 voluntary participants. They were 9th graders in a lower secondary school in Southwest Finland, and were aged 14–16, the vast majority (88.6%) being 15. There were two groups: a regular English class (English A1, n = 24) and a German/Russian language class (English A2, n = 20). The regular English class had received the standard number of English lessons whereas the language class had received less due to their main focus being on German or Russian. The purpose of the study is not, however, to make comparisons between these two groups; they were chosen based on availability. In order to enroll to a language class at the age of 6 or 7, there is an entrance exam that tests the pupils' overall readiness for starting school. Since it is possible that the language aptitude of the language class students has been higher to begin with, an Independent Samples T-Test was performed to compare the overall pronunciation scores of the two groups of students that participated in the study. This was done in order to bring forward any underlying differences there might be between the regular English class and the German/Russian language class. The results of the t-test can be seen in Tables 1 and 2 below.

Table 1 Comparison of the regular and the language class

	n	Mean	Std. Deviation
Regular class pronunciation score	24	113.08	18.58
Language class pronunciation score	20	115.05	15.54

Table 2 Independent Samples T-Test

	Levene's Test for Equality of Variances		T-test for Equality of Means		
	F	Sig.	t	df	Sig. (2-tailed)
Equal variances assumed	.625	.434	-.376	42	.709

There was no significant difference in the pronunciation scores for the regular class (mean = 113.08, st. dev. = 18.58) and the language class (mean = 115, SD = 15.54); t (42) = -0.376, p = 0.709. These results suggest that class does not need to be considered as a variable in the analysis of the results, because the students of the two classes are comparable with each other in the context of the present study despite their different language learning backgrounds.

       Out of the 44 subjects, 26 were male (59.1%), and 18 were female (40.9%). 43 were native speakers of Finnish, yet five of those were also bilingual in Finnish and either Arabic (n

= 2), German (n = 1), Russian (n = 1), or Thai (n = 1). There was also one person whose native language was Chinese, but considering they attend a Finnish school it is expected that they know an adequate amount of Finnish to participate in the study. The bilinguals were included in the study for they could potentially contribute to the generalizability of the results. The purpose of this study is not to examine how English music could have an impact on the pronunciation of Finnish learners of English in particular, but possibly that of the native speakers of other languages as well. Due to the circumstances, nevertheless, the sample was gathered in Finland, which led to the majority of the participants being monolingual native speakers of Finnish. Two people were excluded from the study: one who was a Finnish-English bilingual, and one who did not get a chance to take part in a task due to scheduling problems from the researcher's side. Five people reported having spent over three weeks in an English-speaking country. The mean pronunciation score of those people was higher than that of those who had not spent over three weeks in an English-speaking country (mean = 123.00, SD = 19.20 and mean = 112.82, SD = 16.733) but the difference was not statistically significant (t = 1.262, df = 42, p = 0.214).

The anonymity of the participants was carefully preserved throughout the study. The participants provided their first names for this left less room for error when matching the questionnaire answers with the correct speech samples. This helped increase the reliability and was also more convenient for the researcher. Nonetheless, the anonymity of the subjects was maintained by assigning each participant to a speaker number immediately after the data had been collected and by deleting their names from the data file completely. The analysis of the data was anonymous, and so is the presentation of the results, which means that single participants cannot be recognized from the study. The students participated with their parents' permission granted via an online platform Wilma that is designed for communication between the school, the students, and the parents of the students. The participants knew they were taking part in a study that was interested in the possible effects English music could have on the pronunciation of English learners and that there would be a questionnaire as well as an oral task that would be recorded, but were not provided with details on what the analysis of their answers and spoken performance would be based on.

5.2 Data collection

The data of the present study were collected using two different methods that will be introduced in this section. First, the participants filled in a Webropol questionnaire online, after which they provided a short speech sample by the means of a picture description task. The students who

had the time after finishing the questionnaire completed both parts during the same class, while the rest participated in the picture description task on a different day. The speech samples elicited through the picture description task were later analyzed by a group of native speakers of English based on three criteria: the accuracy, fluency, and intelligibility of the learners' English pronunciation. The questionnaire data were combined with the native-speaker evaluation data on Excel, and then transferred as a whole to SPSS for statistical analysis.

5.2.1 Questionnaire

The students answered a questionnaire during their English class, with both the teacher and the researcher present, and were allowed to use their own mobile phones to fill it in. The group-administered questionnaire (Ellis and Barkhuizen 2005, 42) consisted of 71 questions altogether, factual background questions included. Group-administration guaranteed a high return rate, and the researcher was able to provide clarifications whenever needed. Two people who were not present in the class when the questionnaire was conducted filled it in during their spare time. The original questionnaire, followed by its English translation, can be found in Appendices 1 and 2. The questionnaire was piloted by six people and modified according to the feedback.

Background questions were used to collect demographic information about the participants, as well as information about their personal language learning histories, time spent in an English environment, and the age of onset for English learning. This could be used to ensure that the participants were similar enough for comparisons to be made between them and could also help explain any potential outliers. Attitudinal questions (Dörnyei 2010, 5) were used to determine the participants' attitudes, opinions, and beliefs about the English language, its speakers, and learning it. Behavioral questions (ibid.) were used to find out about the participants' use of spare time and their habits and preferences regarding music and the English language. The questionnaire included both closed- and open-ended questions, all of which were in Finnish to minimize misunderstandings by the participants. The advantage of closed-ended questions is that there is no room for rater subjectivity, and that the answers can be easily coded numerically (Dörnyei 2010, 26).

The set of questions that was used for providing answers to the first and the second research question was interested in the time the participants normally spend engaged in musical activities in English. It was also used for counting a musical index for each participant, a procedure that will be explained shortly. The questions included various time frames from which the participants had to choose the one that best corresponded with their time-consuming

habits. They were, for example, asked to indicate whether they listen to English music on "6–7 days a week", "4–5 days a week", "2–3 days a week", "One day a week", "Couple of days a month", "One day a month", or "Less than every month". The participants were also asked to indicate the approximate number of hours they spend on the activities during the days they engage in those activities by choosing between the following options: "Over 8 hours", "7–8 hours", 5–6 hours", 3–4 hours", 1–2 hours, and "Less than an hour". These questions only portray the number of hours spent on the activities on the days the participants do engage in them and it does not reveal a daily average per se. The options were coded and their relationship to the participants' pronunciation scores was investigated using a Pearson Correlation test in order to reveal existing correlations between the participants' engagement in musical activities and their pronunciation scores. The negative correlations in the results are due to the formatting of the questions: while the pronunciation scores grew according to how high the participant's pronunciation was rated, most musical activity points were received from the least time spent on the musical activity.

There were also statements concerned with what happens during the participants' engagement in the musical activities, where they were asked to choose between the options "Never", "Rarely", "Sometimes", "Often", and "Always", which were later given numerical values from 1 to 5. These were not included in the musical index, but were addressed under the first research question by conducting separate Pearson Correlation tests. For instance, the statements included the following: "When I listen to English music I sing along" and "If an English song gets stuck in my head, I sing it even without music". The latter was concerned with whether the participants not only sing along to English songs, but also sing the songs on their own. To be clear, in the context of this study, "singing without music" refers to singing without the original, vocal version of the song playing in the background. Accompanying oneself on a guitar or other instrument, as well as singing to an instrumental version or simply "singing as one goes" would all be regarded as singing without music.

The questions included in the first part of the questionnaire were mainly concerned with affective factors of L2 learning and were used to answer the third research question. They contained a Likert scale from 1 to 6, the options being "Strongly disagree", "Disagree", "Slightly disagree", "Slightly agree", "Agree", and "Strongly agree". The participants were asked to take a stand on the statements, for example "Music has helped me learn English pronunciation", and were reminded that there are no right and wrong answers. These questions were forced-choice questions (Bartram 2007), meaning no neutral middle value ("Neither disagree or agree") was provided. This was done in order to improve the validity of the study;

a fair amount of cognitive work is required for providing a truthful answer, and omitting the middle category prevents the participants who are less motivated from repeatedly picking the easy, undecided option (Dörnyei 2010, 28).

The final set contained assorted questions, both closed and open-ended. Some of the closed-ended questions were accompanied with open-ended follow-up questions in order to achieve a fuller understanding of the yes-no answers. These open-ended questions concerned topics such as the participants' English-speaking idols and the type of music the participants listen to. The participants were also asked to list a minimum of three artists or bands that they listen to whose music is performed in the English language so that it would be possible to map the type of the music that is most popular among the participants, as well as to consider its suitability for acquiring pronunciation.

5.2.2 Picture description task

Once the participants had completed the questionnaire, they took part in a picture description task. A *picture description task* (PDT), or as Brown (2004, 151) calls it, a *picture-cued task*, is a popular and versatile way of eliciting oral language performance. PDT is also part of TOEIC (Test of English for International Communication), which is used for assessing English-language skills needed when working in an English environment (ETS 2019). In the TOEIC version of the PDT, the test-takers are advised to provide as many details about the picture as they can in 45 seconds, however, in the present study the time for describing the picture was limited to 30 seconds so that there would not be exhaustive amounts of data for the evaluators to analyze.

The picture used in the PDT (see Appendix 3) was chosen due to its cartoon-like clarity and the fact that it had a lot going on – the main idea was that anyone could come up with things to say about the picture. A winter theme was chosen because the data were collected in November and December, and this was hoped to inspire the participants. The picture, used in the task with the illustrator Aimée de Jongh's (2008) permission, was found online and printed out in size A4. In addition, there were five warm-up sentences before the actual PDT, also found in Appendix 3. They were Christmas-themed and fabricated by the researcher so that they would contain most of the English phonemes. The PDT was untargeted, meaning the participants were free to talk about anything related to the picture. They were simply asked to describe in full sentences what they thought was happening or what they could see in the picture. They were also provided with examples of how to begin the description to help them get started.

The participants had one minute of planning time prior to the performance, during which they were allowed to look at the picture and familiarize themselves with five warm-up sentences they would read out loud before the PDT. Planning time makes it possible for the speaker to pre-monitor their output (Ellis and Barkhuizen 2005, 40), which in this case was hoped to decrease the nervousness experienced by the participants. What still happened to some participants is that they froze completely because they were nervous, which led to longer pauses and decreased overall fluency. Also, due to the short production time, some only got properly started towards the end of the sample. Despite the planning time, the speech production was largely spontaneous, since one minute is a relatively short time and no notes were taken. Still, the accuracy may have increased slightly because the participants had time to plan their performance (Ellis and Barkhuizen 2005, 22). In the TOEIC version, there is a 30-second preparation time, but there are no additional components to the PDT as there were in the one modified for the purposes of the present study.

The participants performed one by one in the presence of the researcher in a small room at the school. There happened to be a music class next door that clashed with some of the students' performances so some singing and piano playing leaked through onto the recordings, and there is a possibility that this background noise may have distracted some of the students. Nevertheless, the noise did not jeopardize the analysis of the samples since the evaluation was based on the subjective experience of the native speakers instead of being a precise acoustic analysis. Both the warm-up and the task were recorded on a laptop using a high-quality microphone (Røde NT-USB) and the recording software and digital audio editor Audacity. Audio recording is a widely used technique in studying naturally occurring language use, yet its main disadvantage is that the recording situation might induce self-consciousness in a learner's speech making it less natural (Ellis and Barkhuizen 2005, 27). After the task some of the 9th graders commented on it saying it was actually quite difficult to talk in such situation. One said, for example, that they had sounded completely different than normally. One also said they tend to perform much better when playing video games. Many felt like they had to put more effort into their pronunciation than in a natural situation, which may have affected their performance.

The final variation of the speech sample length was 30 to 34 seconds depending on how long it took for the speaker to finish a sentence or to arrive to a point where interrupting them was natural. During that time, some students produced more speech than others, which mostly had to do with their speaking pace and how well they managed to come up with things to say. The audio data were stored on a private laptop belonging to the researcher, and a backup copy

was made and preserved on the researcher's private memory stick. In order for the native evaluators to be able to access the audio files, they were uploaded on YouTube as unlisted videos for the reason that Google Forms does not support audio files. Unlisted videos are only visible to people who have the link (the evaluators) and they do not appear in the user's (the researcher) channel page or anyone's YouTube searches. The image on the videos was a black background with a white text displaying the number of the speaker, for example "Speaker 1". The videos, as well as all the audio files, have since been deleted, and no one can access them any longer. The evaluation of the speech samples will be addressed further in the following sub-section.

5.2.3 Native-speaker evaluation

The 44 speech samples collected through the PDT were evaluated by nine voluntary native speakers of English who were selected via convenient sampling, meaning they were available and willing to participate (Dörnyei 2007, 89–99). The evaluation questionnaire (see Appendix 4) was realized via Google Forms and piloted by two people before being sent to the evaluators to ensure the instructions were clear and that everything worked as planned. Native speakers were chosen as evaluators because even though L2 speakers can usually rate the accentedness of the L2, there is no certainty of how capable they are of evaluating other features of speech (Derwing and Munro 2013, 181). *Accentedness* refers to the listener's perception of how different the speaker's accent is from that of the native speakers (Derwing and Munro 2005, 385) and is mostly related to the accuracy of pronunciation. Since fluency and intelligibility were also in the focus of the present study non-native speakers of English were regarded unsuited for the evaluation task.

The evaluators were given instructions for the task in text form due to it not being possible for the researcher to be physically present in three countries. Six of the evaluators were British (66.7%), two were American (22.2%), and one was Canadian (11.1%), as the purpose was to have more than one native English variety represented. The evaluators were aged 23–54, and none of them had studied linguistics at a university level. Five of the evaluators reported hearing non-native English daily or almost daily (55.6%), one weekly (11.1%), and three monthly (33.3%), meaning none of them was completely unfamiliar with non-native pronunciation. This could have affected the evaluations by making the evaluators more forgiving in their ratings, although it was not specified which non-native accent the evaluators were used to hearing. It was also found by Tominaga (2011, 52) that native speakers of English may be generally softer in their evaluation of non-native English pronunciation than non-native

speakers of English. Figure 1 below displays the distribution of answers regarding how often the evaluators reported hearing non-native English.

Figure 1 Hearing non-native English
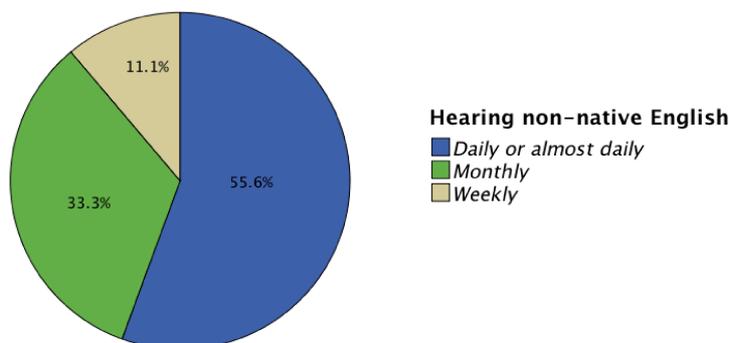
In the questionnaire, the evaluators were first provided with definitions of the three concepts their evaluation would be based on and were asked to familiarize themselves with these definitions. The concepts were pronunciation accuracy, fluency, and intelligibility, and their meanings were made as clear as possible by combining and modifying previously coined definitions. The definitions given to the evaluators of the present study were the following: 1) Accuracy = "How well the language is spoken in relation to its pronunciation rules, i.e. are English sounds and words pronounced correctly", 2) Fluency = "How easily and naturally the speech flows without too many hesitations, pauses, and self-repetitions", and 3) Intelligibility = "How understandable the speaker is, i.e. do they get their message across".

The evaluators were advised to listen to the speech samples in a quiet place using headphones to maintain the quality of the audio, as well as to take a 1–2-minute break after listening to a certain number of samples. The three breaks were marked in the questionnaire in advance. As for the confidentiality of the data, the evaluators were informed that it is strictly prohibited to share the speech samples with anyone else. The samples were rated on a Likert scale from 1 to 6 in the three categories mentioned above, 1 being the lowest and 6 being the highest score. Each Likert scale response option was assigned to a number for scoring purposes so that the highest possible pronunciation score to obtain from the task per sample was 162 and the lowest was 27, whereas the highest single criterion score per sample was 54 and the lowest was 9.

Cronbach's Alpha coefficient (Dörnyei 2010, 94) was used to measure the reliability and the internal consistency of the native speakers' evaluations. The results of this test are shown in Table 3 below.

Table 3 Cronbach's Alpha of the native-speaker evaluations

Evaluation	Cronbach's Alpha	Number of items
Accuracy	.98	44
Fluency	.97	44
Intelligibility	.98	44

Cronbach's alpha for the accuracy items was α = .98, α = .97 for the fluency items, and α = .98 for the intelligibility items. This suggests excellent internal consistency of the evaluations by the nine native speakers of English, which indicates that the questionnaire was reliable.

For additional insight, the evaluators answered three open-ended questions about the task itself and about non-native speakers of English at the end of the evaluation questionnaire. They were asked what they found the most distracting about non-native speakers' pronunciation in general, also outside of the study. One person wrote that they find poor use of grammar more distracting than mispronunciation. A couple of others also mentioned wrongly used grammatical elements. Excessive pausing as well as pausing between words were also found distracting, because they "break up the sentences too much". One British evaluator found not pronouncing what they called "hard consonants" or "speaking in an American format", to be distracting. Hard consonants referred to unvoiced plosive sounds, such as /t/ in the RP pronunciation of the word "winter". The Finnish speakers' tendency to pronounce words as they are spelt was also mentioned. Similarly, another person mentioned "strong accents". One of the answers summed it up rather conveniently: "Just the fact that [it] is not quite the way I would pronounce the words, in the same way that my pronunciation of Finnish words is not right."

All in all, the evaluation task was found rather difficult, and the lack of a "reference point" or an "example reading" was mentioned, which was interesting given that the evaluators were told that the evaluation would be based on their own opinions of the speakers' pronunciation. Also, as expected, the three specific evaluation criteria might have been challenging to grasp even though their definitions were provided at the beginning. One evaluator wrote: "Perhaps a little difficult since I haven't paid that close attention to the evaluation criteria before." On the other hand, some evaluators reported finding the task fairly easy.

5.3 Musical index

A musical index was counted for each participant by combining the scores of four questions regarding how often and how much the participants had reported listening to English music and singing in English on average (questions 45–48 in the questionnaire, see Appendix 1 or 2). This was done in order to include the participants' musical habits into one variable. The lowest possible total score was 4, indicating a very high level of engagement in the musical activities, whereas the highest possible score was 26, meaning that the participant was rarely engaged in the activities. In the sample, the scores ranged from the minimum of 8 to the maximum of 26. The mean value was 15.64, which was rounded up to 16 due to the musical index being represented by round figures.

Based on the mean, the participants were divided into two groups. The first group consisted of those who had scored lower than average (8–15 points, n = 24), which was the group with the higher level of engagement in musical activities. The second group consisted of those who had scored average or higher (16–26 points, n = 20), and this was the group with lower engagement in musical activities. The descriptive statistics of the learners' musical indices are presented below in Table 4.

Table 4 Descriptive statistics of the musical index

	n	Mean	Median	Std. Deviation	Range	Minimum	Maximum
Musical index	44	15.64	15.00	3.82	18.00	8.00	26.00

After the participants were divided into two groups based on their musical index, an *Independent-Samples Mann-Whitney U* test was conducted to determine whether the groups with higher and lower musical indices differed significantly from each other in their pronunciation scores. *Pearson's r* was also used to examine the relationship of the pronunciation scores by each participant and the frequency and amount of musical activities as reported by the participants, that is, their musical index. Pearson's r produces a significance figure that tells whether the correlation between two variables is statistically significant. *Statistical significance* means that a certain result has not been obtained in a particular sample simply by chance and that it is likely to be true in the population (Dörnyei 2007, 210).

Separate indices were also counted for listening and singing in order to compare the two activities to the pronunciation scores individually. The lowest possible index was 2, whereas the highest was 13. Median value was used as the cut-off point for dividing the participants into two groups, first based on their listening habits and second on their singing habits. The median

listening score was 6, and two groups were formed: a group with a lower than the median listening score (n = 15), which was the higher engagement group, and a group with the median or higher listening score (n = 29), which was the group with lower engagement levels. The median singing score was 9, and again two groups were formed. The first group had a singing score lower than the median (n = 17), and the second group had a median or higher than the median singing score (n = 27). The former was the group with a higher and the latter with a lower engagement level.

Section 5 has set the research questions of the present study and presented hypotheses supported by existing knowledge and previous studies. After familiarizing the reader with the participants, this section has introduced in a detailed manner the procedures that were applied in the realization of the study. The section has provided an explanation of the how the questionnaire, the PDT, and the native-speaker evaluation of the speech samples were conducted, as well as elaborated on how the data were analyzed. The results of the study shall be presented in the following section, Section 6.

6 Results

Here, the results of the present study shall be given and analyzed in light of the three research questions presented in the beginning of the previous section, Section 5. This section is divided into four sub-sections. Section 6.1 begins by introducing the descriptive results of the speech sample evaluations conducted by the nine native speakers of English. After this, Section 6.2 shall address the first research question (How could listening to English music and singing in English explain the variation in the native-speaker evaluations of the English learners' pronunciation?) by first presenting figures of the listening and singing habits of the participants, and by providing results for the tests that were run in order to compare the participants' musical habits and pronunciation skills. Section 6.3 moves on to addressing the second research question (To what extent do accuracy, fluency, and intelligibility correlate with listening to English music and singing in English?) by presenting correlations between the accuracy, fluency, and intelligibility of the participants' English pronunciation and the musical activities they reported engaging in. Finally, Section 6.4 provides answers for the third research question (Which indirect factors could make English music a beneficial means for learning pronunciation?) by presenting correlations and percentages concerned with motivational and affective factors, as well as by listing typical situations and places for listening to music.

6.1 Evaluation of the speech samples

The speech samples of the 44 participants were evaluated and rated by nine native speakers of English based on three pronunciation criteria: accuracy, fluency, and intelligibility. The descriptive statistics of the pronunciation scores obtained by the participants are presented in Table 5 below.

Table 5 Descriptive statistics of the pronunciation scores

	Range	Minimum	Maximum	Mean	Std. Deviation	n
Accuracy score	26.00	26.00	52.00	38.16	5.83	44
Fluency score	28.00	25.00	53.00	36.34	6.63	44
Intelligibility score	23.00	29.00	52.00	39.48	5.62	44
Pronunciation score	72.00	82.00	154.00	113.98	17.10	44

As seen in Table 5 above, the accuracy scores obtained by the participants ranged from the minimum of 26 points to the maximum of 52 points (SD = 5.83), the average being 38.16. The

mean fluency score of the participants was 36.34, the minimum score being 25 points and the maximum 53 points (SD = 6.63). The mean intelligibility score in the sample was 39.48, ranging from 29 points to 52 points (SD = 5.62). Finally, the average pronunciation score (= accuracy, fluency, and intelligibility scores combined) of the participants was 113.98, with the lowest score being 82 points and the highest 154 points (SD = 17.10). The lowest possible pronunciation score to obtain from the task was 27 and the highest was 162, while the lowest possible score to receive from a single criterion was 9 and the highest 54. These extremes, nevertheless, were not obtained by any of the participants, and the scores show that the level of English proficiency of the participants was quite high overall. Most variation was found in the fluency score and the least in the intelligibility score.

6.2 First research question

The first research question was concerned with how listening to English music and singing in English in one's spare time could explain the variation in the English learners' pronunciation scores. This question was approached quantitatively by examining the relationship between the participants' everyday engagement in musical activities in English and the pronunciation scores they had obtained from the PDT. Displayed first in Section 6.2.1 are the participants' music listening and singing habits. After this, in Section 6.2.2, the results of the correlation tests conducted between the participants' pronunciation scores and musical activities will be presented. Section 6.2.3 will show comparisons between the pronunciation skills of those participants who were more engaged in musical activities in English and those who were less engaged. At the end, Section 6.2.4 shall address the additional, more specific questions related to listening to English music and singing in English that were not included in the musical index but can also provide important information.

6.2.1 Listening and singing habits

The following figures (Figures 2, 3, 4, and 5) present the frequencies of answers to the four questions concerning the average time the participants reported spending on listening to English music and singing in English in their spare time. The answers to these four questions were also the ones based on which the musical index of each participant was counted. First presented in Figure 2 below is the average time the participants reported listening to English music per week or month.

Figure 2 Listening to English music: days per week or month
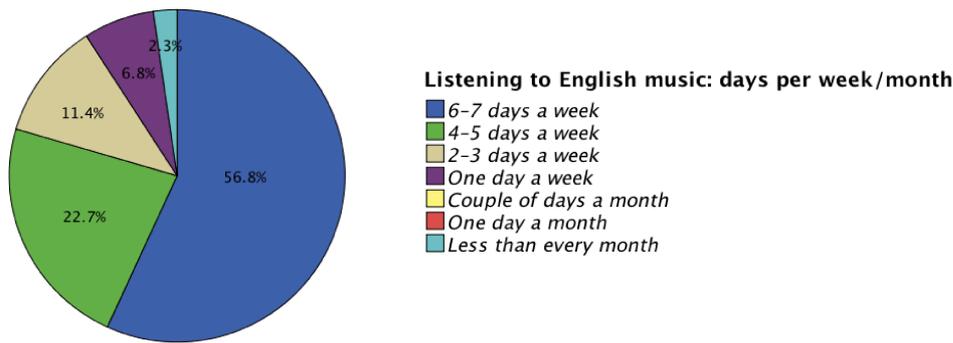
What can be seen in Figure 2 is a fairly uniform listening trend: all of the participants, except for one, reported listening to English music weekly. The majority, 56.8% (n = 25), reported listening to English music on 6–7 days a week, 22.7% (n = 10) reported listening to English music on 4–5 days a week, 11.4% (n = 5) on 2–3 days a week, and 6.8% (n = 3) on one day a week. One participant (2.3%) reported listening to English music less frequently than every month. Within this sample, making a reliable comparison between those who often listen to music and those who rarely do is troublesome for the lack of the latter. Displayed next in Figure 3 is the number of hours the participants reported listening to English music on the days they engage in this activity.

Figure 3 Listening to English music: hours per day
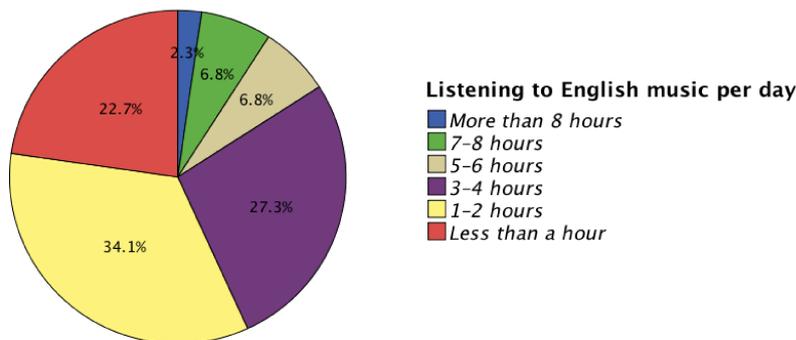
Figure 3 above shows the average hours per day the participants reported spending on listening to English music. The most common time spent on listening to English music per day was 1–2 hours (34.1%, n = 15) and the second most common 3–4 hours (27.3%, n = 12). 6.8% (n = 3) reported listening to English music for 5–6 hours per day, and the same percentage reported listening to it for 7–8 hours. One person (2.3%) reported listening to English music for more than eight hours per day. On the contrary, 22.7% (n = 10) of the participants reported listening to English music for less than an hour a day.

Presented by the following two figures is the time the participants reported spending on singing in English. First, figure 4 below shows the average time they reported singing in English per week or month.

Figure 4 Singing in English: days per week or month
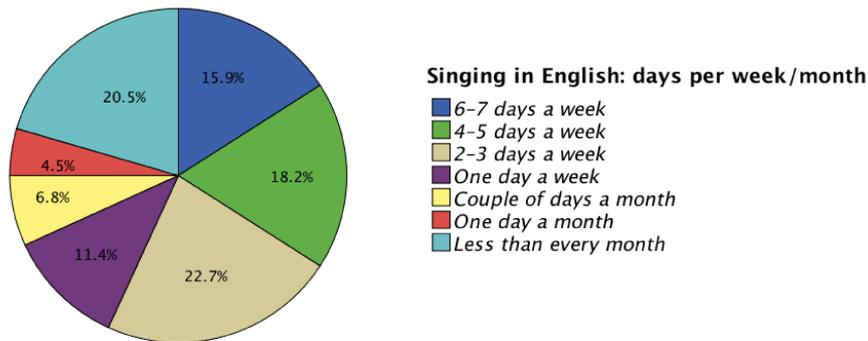
As seen in Figure 4, the distribution of answers was less homogenous for singing in English than it was for listening to English music. 15.9% (n = 7) of the participants reported singing in English almost every day, on 6–7 days a week. 18.2% (n = 8) reported singing in English on 4–5 days a week, 22.7% (n = 10) on 2–3 days a week, and 11.4% (n = 5) on one day a week. A total of 31.8% (n = 14) of the participants had reported singing in English less than weekly, out of which 6.8% (n = 3) sang on a couple of days per month, 4.5% (n = 2) sang on one day per month, and 20.5% (n = 9) less frequently than every month. These numbers enable a comparison to be made between two groups: those who sing in English on multiple days a week and those who sing less. Now, figure 5 below presents the number of hours the participants reported singing in English per day.

Figure 5 Singing in English: hours per day
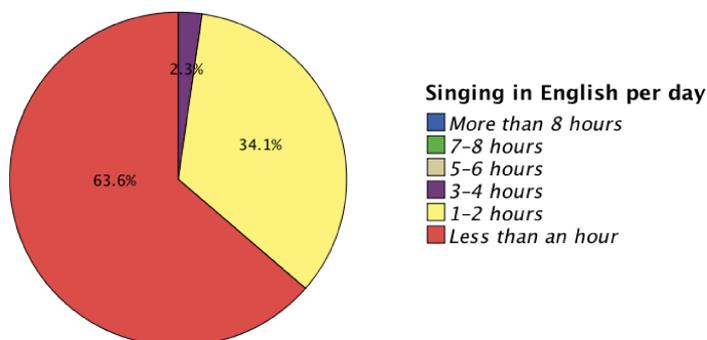
Figure 5 above depicts the number of hours the participants reported singing in English per day. As expected, singing for more than two hours a day was a rare occurrence in the sample; only

40

one participant (2.3%) reported singing in English for 3–4 hours a day. The majority (63.6%, n = 28) reported singing for less than an hour per day. The second largest group (34.1%, n = 15) reported singing for 1–2 hours.

The participants were also asked about the type of music they listen to most frequently. The most popular artists and bands among the participants included Eminem, Travis Scott, Ed Sheeran, Drake, XXXTentacion, Dua Lipa, Shawn Mendez, Lil Peep, Jaden Smith, and Ariana Grande. They all received more than two mentions. These musical preferences seem to be in accordance with the IFPI and Teosto 2018 report according to which English pop and English hip hop were the most popular genres among Finnish adolescents last year. A comprehensive list of the English music preferred by the participants can be found in Appendix 5.

6.2.2 Correlations between musical activities and pronunciation score

The musical index, consisting of the participants' estimations of the time spent on listening to English music and singing in English was compared to their pronunciation scores in search of potential correlations between pronunciation skills and engagement in musical activities in English. The results of a Pearson Correlation test can be seen in Table 6 below.

Table 6 Correlation between pronunciation score and musical index

		Musical index
Pronunciation score	Pearson Correlation	-.379*
	Sig. (2-tailed)	.011
	n	44

*. Correlation is significant at the 0.05 level (2-tailed).

A statistically significant correlation was revealed at the 0.05 level (r = -0.379, p = 0.011), which indicates that while the pronunciation scores grew, the number representing the musical index decreased, indicating an increase in the musical activity. This suggests that the more time was spent on musical activities in the English language, the better the English pronunciation was considered by the native-speaker jury. In order to provide a more comprehensive answer to the first research question, however, it is necessary to also address the two musical activities separately. Demonstrated first in Table 7 is the correlation between the participants' pronunciation score and how much English music they listen to.

Table 7 Correlation between pronunciation score and listening to English music

		Listening to English music
Pronunciation score	Pearson Correlation	-.292
	Sig. (2-tailed)	.054
	n	44

No statistically significant correlation was found between the pronunciation score and listening to English music (r = -0.292, p = 0.054), yet the p-value was approaching significance with only 0.004 difference. Based on this result it cannot be concluded that listening to English music would be linked to better pronunciation, yet a tendency of some sort is likely to exist between the two. This does, nevertheless, require further research. In Table 8 below, the correlation between the pronunciation score and singing in English is presented.

Table 8 Correlation between pronunciation score and singing in English

		Singing in English
Pronunciation score	Pearson Correlation	-.331*
	Sig. (2-tailed)	.028
	n	44

*. Correlation is significant at the 0.05 level (2-tailed).

A statistically significant correlation was found at the 0.05 level (r = -0.331, p = 0.028) between the participants' pronunciation scores and singing in English. This indicates that the pronunciation scores were higher for those participants who had reported spending more time singing in English, which suggests that a link may exist between English pronunciation skills and singing in English.

6.2.3 Pronunciation scores of higher and lower musical activity groups

With the intention of finding answers to the first research question, the participants were also divided into groups based on their musical index, that is, their level of engagement in musical activities in English, after which the pronunciation scores of these groups were compared to one another. In addition, comparisons were made between groups based on a single English music activity, both listening to English music and singing in English. First, the average pronunciation scores of the higher and lower engagement group are presented in Figure 6.

Figure 6 Mean pronunciation score by musical index group
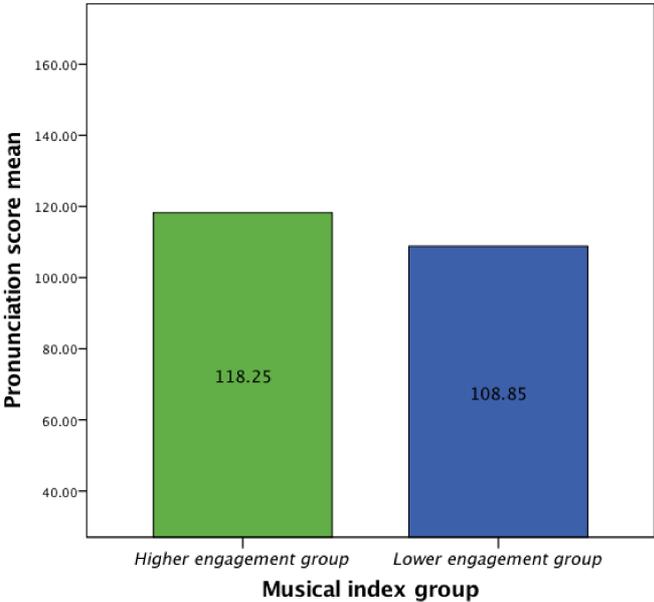
As seen in Figure 6, once the participants were divided into two groups based on their musical index, and the mean pronunciation scores were counted for each group, the mean score of the group with higher engagement in musical activities in English was 118.25, while the mean score of the lower engagement group was 108.85. It seems that, on average, those participants who listened to more English music and sang in English more frequently pronounced English better than those who were less engaged in English music activities.

A Mann-Whitney U test was conducted to compare the pronunciation scores of the higher and lower musical engagement groups statistically. Table 9 below shows the mean ranks obtained from the test, after which Table 10 shall present the significance of the same test.

Table 9 Pronunciation score mean ranks based on musical index group

	Musical index group	n	Mean Rank	Sum of Ranks
Pronunciation score	Higher engagement	24	25.17	604.00
	Lower engagement	20	19.30	386.00
	Total	44		

The mean ranks of the pronunciation scores of the two groups were 25.17 for the group with higher engagement in musical activities in English and 19.30 for the group with lower engagement. This difference in the scores implies that those who were more engaged in English music performed better pronunciation-wise.

43

Table 10 Comparison of pronunciation scores between musical index groups

	Pronunciation score
Mann-Whitney U	176.00
Wilcoxon W	386.00
Z	-1.509
Asymp. Sig. (2-tailed)	.131

Grouping variable: Musical index group.

However, as is evident from Table 10 above, there did not seem to be a statistically significant difference between the higher and the lower musical activity engagement groups after all (U = 176.00, p = 0.131). This means that the null hypothesis, the difference being due to chance, cannot be rejected on the basis of the Mann-Whitney test.

Next, a Mann-Whitney test was used to compare the pronunciation scores of two groups that were made based on how much English music the participants had reported listening to: those who listened to English music more than average and those who listened less. The test yielded the results seen in Tables 11 and 12.

Table 11 Pronunciation score mean ranks and listening to English music

	Listening group	n	Mean Rank	Sum of Ranks
Pronunciation score	Higher engagement	15	23.17	347.50
	Lower engagement	29	22.16	642.50
	Total	44		

As seen in Table 11 above, the mean ranks of the pronunciation scores by the group that had reported listening to more English music and the group that had reported listening to less English music were 23.17 for the former and 22.16 for the latter. The difference is small, yet those who reported listening to more English music had scored slightly higher.

Table 12 Comparison of pronunciation scores and listening to English music

	Pronunciation score
Mann-Whitney U	207.50
Wilcoxon W	642.50
Z	-.248
Asymp. Sig. (2-tailed)	.804

Grouping variable: Higher/lower engagement in listening to English music.

What can be seen in Table 12 is that there did not seem to be a significant difference between the two listening groups based on a Mann-Whitney test (U = 207.50, p = 0.804), and the null hypothesis is again retained meaning it cannot be suggested that the pronunciation skills of the two groups would be different.

A Mann-Whitney test was also conducted to compare the pronunciation scores of two groups that were made based on how much the participants had reported singing in English: those who had reported singing more than average and those who had reported singing less than that. The mean ranks of the pronunciation scores by the groups can be seen in Table 13 below.

Table 13 Pronunciation score mean ranks and singing in English

	Singing group	n	Mean Rank	Sum of Ranks
Pronunciation score	Higher engagement	17	25.18	428.00
	Lower engagement	27	20.81	562.00
	Total	44		

The pronunciation score mean rank of the group that sang more in English was 25.18, while the mean rank of the group that sang less was 20.81. This difference was noticeably greater than that of the two listening groups and indicates that those participants who had reported singing more in English might be better at pronouncing English than those who had reported singing less. Table 14 shows the significance of the same Mann-Whitney test.

Table 14 Comparison of pronunciation scores and singing in English

	Pronunciation score
Mann-Whitney U	184.00
Wilcoxon W	562.00
Z	-1.097
Asymp. Sig. (2-tailed)	.273

Grouping variable: Higher/lower engagement in singing in English.

From the Mann-Whitney test it can be concluded that, as with the listening groups, the two singing groups were not significantly distinct in their pronunciation scores (U = 184.00, p = 0.273), and the null hypothesis is thus retained. This means that based on the result it cannot be assumed that the group who sang more in English would be evidently better at pronouncing English.

6.2.4 Additional questions related to musical activities

Eventually, the pronunciation scores of the participants were also compared to a few other music-related variables that were not included in the musical index. These variables were singing along to English songs, singing in English without music, paying attention to song lyrics, and paying attention to the pronunciation of the lyrics. The results of these correlation tests are presented in Table 15 below.

Table 15 Pronunciation scores and additional questions related to musical activities

		Pronunciation score (overall)	Accuracy score	Fluency score	Intelligibility score
Singing along to English songs	Pearson Correlation	.315*	.246	.294	.355*
	Sig. (2-tailed)	.037	.108	.052	.018
	N	44	44	44	44
Singing in English without music	Pearson Correlation	.519**	.402**	.526**	.541**
	Sig. (2-tailed)	<.001	.007	<.001	<.001
	N	44	44	44	44
Paying attention to lyrics	Pearson Correlation	.298*	.268	.292	.284
	Sig. (2-tailed)	.049	.078	.055	.062
	N	44	44	44	44
Paying attention to pronunciation of lyrics	Pearson Correlation	.245	.251	.196	.254
	Sig. (2-tailed)	.109	.101	.201	.096
	N	44	44	44	44

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

A statistically significant positive correlation was found between singing along to English songs and the overall pronunciation score (r = 0.315, p = 0.037). Singing along was also linked to higher intelligibility scores (r = 0.355, p = 0.018), but it did not correlate significantly with accuracy (r = 0.246, p = 0.108) or fluency (r = 0.294, p = 0.052). A particularly significant positive correlation seemed to exist between the pronunciation scores and singing English songs without music. This activity produced a statistically significant correlation at the 0.01 level to the overall pronunciation score (r = 0.519, p<0.001), as well as to all three individual pronunciation areas. The correlation was strongest to intelligibility (r = 0.541, p<0.001), second strongest to fluency (r = 0.526, p<0.001), and third to accuracy (r = 0.402, p = 0.007). Paying attention to English song lyrics correlated positively with the overall pronunciation score (r = 0.298, p = 0.049), yet it did not seem to be significantly linked to any individual pronunciation area. Perhaps surprisingly, paying attention to the pronunciation of the song lyrics did not yield any statistically significant correlations.

6.3 Second research question

The second research question set out to discover how the musical activities of the language learner correlate with the three criteria of good pronunciation: accuracy, fluency, and intelligibility. This question was approached quantitatively by conducting a Pearson Correlation test in order to examine the relationship between the participants' musical activities and their accuracy, fluency, and intelligibility scores irrespective of one another. The purpose of this procedure was to find out whether there would be significant dissimilarities between the different pronunciation areas in relation to listening to English music and singing in English. Examined first were the musical activities together, the musical index, and the three criteria of good pronunciation. The results of these correlation tests are gathered in Table 16 below.

Table 16 Accuracy, fluency, and intelligibility in relation to musical index

		Musical index
Accuracy score	Pearson Correlation	-.351*
	Sig. (2-tailed)	.019
	n	44
Fluency score	Pearson Correlation	-.366*
	Sig. (2-tailed)	.015
	n	44
Intelligibility score	Pearson Correlation	-.359*
	Sig. (2-tailed)	.017
	n	44

*. Correlation is significant at the 0.05 level (2-tailed).

As seen in Table 16, a statistically significant correlation (r = -0.351, p = 0.019) was found at the 0.05 level between pronunciation accuracy and the musical index of the English learners. As the accuracy scores grew, the engagement in musical activities increased, indicating there was a link between better pronunciation accuracy and more time spent on musical activities in English. A comparison between pronunciation fluency and the musical index also yielded a statistically significant correlation (r = -0.366, p = 0.015); as the fluency scores grew, the engagement in musical activities increased, which suggests that better fluency and higher engagement in musical activities were linked. A statistically significant correlation was also revealed at the 0.05 level when pronunciation intelligibility scores were compared to the musical index (r = -0.359, p = 0.017), indicating that increased intelligibility was linked higher engagement in musical activities. The correlation to the musical index was most significant for

pronunciation fluency and least significant for accuracy, however, the differences were only marginal. This suggests that listening to English music and singing in English together might contribute to the accuracy, fluency, and intelligibility of the learners' English pronunciation.

The accuracy, fluency, and intelligibility of the participants' pronunciation was also inspected in relation to the two musical activities separately. Table 17 below presents the results of the Pearson correlation tests between the three separate pronunciation scores and listening to English music, and the three separate pronunciation scores and singing in English.

Table 17 Accuracy, fluency, and intelligibility in relation to single musical activities

		Listening to English music	Singing in English
Accuracy score	Pearson Correlation	-.272	-.306*
	Sig. (2-tailed)	.074	.044
	n	44	44
Fluency score	Pearson Correlation	-.285	-.317*
	Sig. (2-tailed)	.061	.036
	n	44	44
Intelligibility score	Pearson Correlation	-.271	-.318*
	Sig. (2-tailed)	.075	.036
	n	44	44

*. Correlation is significant at the 0.05 level (2-tailed).

As seen above in Table 17, once the correlation tests were run between listening to English music and the three individual pronunciation scores, no significant correlations were revealed. While the scores grew, no corresponding tendency could be seen in the engagement levels in listening to English music. The p-value closest to significant was obtained by fluency ($r = -0.285$, $p = 0.061$), whereas accuracy ($r = -0.272$, $p = 0.074$) and intelligibility ($r = -0.271$, $p = 0.075$) both yielded a similar result. However, the tests revealed a statistically significant correlation every time the relationship between singing in English and an individual pronunciation score was tested. While the pronunciation scores grew, so did the engagement levels in singing in English. The correlation was strongest to intelligibility ($r = -0.318$, $p = 0.036$) and fluency ($r = -0.317$, $p = 0.036$), and slightly less powerful for accuracy ($r = -0.306$, $p = 0.044$), but the differences were quite small.

6.4 Third research question

The third research question was interested in the indirect factors that could make English music a beneficial means for learning English pronunciation. The hypothesized factors included motivational and affective factors, such as the learners' motivation to learn English and their feelings and attitudes towards the English language and its native speakers. These were believed to be positively altered by English music. The question was also concerned with the learners' emotions and mood, the subjective learning experience, as well as the extent of English music in the participants' everyday lives. This question was approached quantitatively by contrasting the pronunciation scores to the motivational and affective factors and by examining whether having an English-speaking (musical or other) idol would seem to have an impact on the participant's pronunciation, motivation, and attitudes, and thereby on the pronunciation learning indirectly. The question also looked into how the participants had responded to various statements related to the topics mentioned above.

First presented will be the results related to English-speaking idols and role models, and their possible indirect influence on pronunciation learning. 43.2% (n = 19) of the participants reported having an English-speaking idol or role model. The participants were allowed to specify what their idol was known for, and out of the 18 people who did, eight specified that their idol was a musician. Other answers included, for instance, actors, athletes, and youtubers. The mean pronunciation scores of the groups with and without English-speaking idols or role models (both musicians and others) are presented in Figure 7 below.

Figure 7 Pronunciation scores with and without English-speaking idols
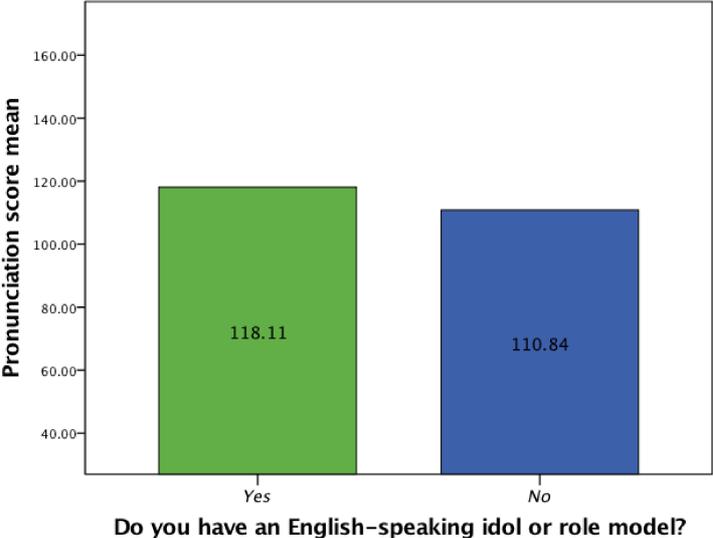
Figure 7 shows that there is a small difference in the mean pronunciation score of the groups with and without English-speaking idols or role models: those who had one had scored 118.11 on average, whereas those who did not have one had scored 110.84. This would indicate that having an English-speaking idol might be linked to better English pronunciation. To find out whether this difference is significant, an Independent-Samples Mann-Whitney U test was used to compare the mean pronunciation scores. The results of the test are presented below in tables 18 and 19.

Table 18 Pronunciation score mean ranks and having an English-speaking idol

	English-speaking idols	n	Mean Rank	Sum of Ranks
Pronunciation score	Yes	19	25.50	484.00
	No	25	20.22	505.00
	Total	44		

As seen in Table 18, the pronunciation score mean ranks of the groups with and without English idols were 25.50 for the former and 20.22 for the latter, indicating that those participants who had reported having an English-speaking idol or role model had scored better in the pronunciation test. The significance of this difference is shown in Table 19.

Table 19 Comparison of pronunciation score and having an English-speaking idol

	Pronunciation score
Mann-Whitney U	180.50
Wilcoxon W	505.50
Z	-1.351
Asymp. Sig. (2-tailed)	.177

Grouping variable: English-speaking idols yes/no.

As it turns out, there was no statistically significant difference between the groups even though the group with an English-speaking idol or role model had scored seemingly better (U = 180.50, p = 0.177). This means that the null hypothesis is retained: no suggestion can be made based on this result that the pronunciation skills of the participants would differ depending on whether they have an English-speaking idol or role model.

Next, the relationship between the pronunciation scores and the participants' motivation to learn English was examined by conducting a Pearson Correlation test between the

51

participants' pronunciation score and the points they had received from the question "I want to learn more English". The test yielded the result presented in Table 20 below.

Table 20 Correlation between pronunciation score and English learning motivation

		I want to learn more English.
Pronunciation score	Pearson Correlation	.444**
	Sig. (2-tailed)	.003
	n	44

**. Correlation is significant at the 0.01 level (2-tailed).

As seen in Table 20 above, the correlation between the pronunciation score and wanting to learn more English was statistically significant at the 0.01 level ($r = 0.444$, $p = 0.003$), meaning higher motivation was clearly linked to higher pronunciation scores.

The English-speaking idols and role models were hypothesized to be one of the factors that could increase the learners' overall motivation to learn English, which could in this manner also improve their pronunciation, thus the English learning motivation of the groups with and without English-speaking idols was investigated. This was done by conducting an Independent-Samples Mann-Whitney U test. Table 21 below shows the mean rank of points received from the question "I want to learn more English" based on whether the participants had reported having an English-speaking idol or role model.

Table 21 English learning motivation and having an English-speaking idol I

	English-speaking idol or role model (n = 19)	No English-speaking idol or role model (n = 25)
Mean rank	23.66	21.62
Sum of ranks	449.50	540.50

Table 21 above reveals the mean ranks of 23.66 and 21.62 for the groups meaning that the rank was slightly higher for the group who had reported having an English-speaking idol or role model than for the group that did not have one. Table 22 below shows the significance of the same test.

Table 22 English learning motivation and having an English-speaking idol II

	I want to learn more English.
Mann-Whitney U	215.50
Wilcoxon W	540.50
Z	-.611
Asymp. Sig. (2-tailed)	.541

Grouping variable: English-speaking idols yes/no.

As seen above in Table 22, the difference between the two groups was not statistically significant (U = 215.50, p = 0.541), which means that the null hypothesis is retained. Based on this sample, having an English-speaking idol and role model did not seem to make the participants more motivated to learn English and thereby a better pronouncer.

In order to examine the participants' attitudes towards the native speakers of English, another Independent-Samples Mann-Whitney U Test was conducted. It was used to compare how the group with no English-speaking idols and the group with English-speaking idols had replied to the question about the niceness of English-speakers. Table 23 below shows the mean rank of points received from the question "Native speakers of English are usually nice" by the groups with and without English-speaking idols or role models.

Table 23 Attitudes towards English-speakers and having an English-speaking idol I

	English-speaking idol or role model (n = 19)	No English-speaking idol or role model (n = 25)
Mean rank	27.68	18.56
Sum of ranks	526.00	464.00

As demonstrated by Table 23, the mean ranks for the groups with and without English-speaking idols or role models were 27.68 for the former and 18.56 for the latter. The mean rank was considerably higher for the group with an English-speaking idol than for the group without one. This suggests that having an English-speaking idol could potentially improve the learners' attitudes towards the native-speakers of English in general. The significance of the Mann-Whitney test is presented in Table 24.

Table 24 Attitudes towards English-speakers and having an English-speaking idol II

	Native speakers of English are usually nice.
Mann-Whitney U	139.00
Wilcoxon W	464.00
Z	-2.459
Asymp. Sig. (2-tailed)	.014

Grouping variable: English-speaking idols yes/no.

A statistical significance of p = 0.014 was discovered in the test. Thus, the null hypothesis that the distribution of answers is the same across both groups can be rejected and replaced with the assumption that the groups indeed differed in their ideas about English speakers; native speakers of English were considered nicer by those who had an English-speaking idol or role model than those who did not have one. This implicates that English-speaking idols might influence pronunciation learning in the sense that they improve the learners' attitudes towards English-speakers, which could affect language learning in a positive way. Other types of affective factors shall be discussed in the following four paragraphs.

First of all, 95.4% of the participants (n = 42) found listening to music relaxing, 72.7% (n = 32) even agreeing strongly to this statement. 79.5% (n = 35) felt like music helps them deal with their personal feelings. Perhaps accordingly, song lyrics mattered to 84.1% (n = 37) of the participants, and one listed relatable lyrics as one of the reasons why they have a musician as their role model. 86.4% (n = 38) also reported learning English song lyrics easily. 95.5% (n = 42) of the participants agreed that music gets them in a good mood, out of which 68.2% (n = 30) agreed strongly. 56.9% (n = 25) reported enjoying singing, and a statistically significant positive correlation was found between the pronunciation score and enjoying singing (r = 0.334, p = 0.027). Whether the participants considered themselves a good singer did not seem to be linked to the pronunciation scores (r = 0.042, p = 0.785).

The participants' subjective learning experience was also examined. 31.8% (n = 14) thought English pronunciation is difficult to learn because English contains difficult sounds, and 25% (n = 11) felt like English pronunciation is difficult to learn because words are pronounced differently than they are written. 36.4% (n =16) felt like not enough pronunciation is being practised during English classes. At the same time, 70.5% (n = 31) of the participants thought that it is important to sound as native-like as possible, and 29.5% (n = 13) had accent goals. 88.6% (n = 39) of the participants agreed that music has helped them learn English, and as many as 90.9% (n = 40) felt like music has helped them learn English pronunciation. Listening to English music was considered a good way to learn English pronunciation by 95.5%

(n = 42), and singing in English was thought to be a good way to practise pronunciation by 93.2% of the participants (n = 41). 15.9% (n = 7) reported sometimes experiencing class room anxiety during English classes, whereas 52.3% (n = 23) disagreed strongly to this statement. School exercises that include music were considered pleasant by 70.4% (n = 31) of the participants.

Finally, to concisely demonstrate the extent of music in the participants' everyday lives, typical situations in which the they reported listening to music are gathered below into Table 25.

Table 25 Where and when participants typically listen to music

Where do you listen to music?	n	Percent (%)	Examples
Home	33	75.0	"In the morning", "When I'm going to bed", "While I'm doing homework", "While I'm playing video games", "By my vinyl player"
Travel	31	70.5	"On my way to school", "On the bus", "In the car"
Hobbies	9	20.5	"At the gym", "While I'm drawing"
School	6	13.6	"During breaks", "During boring classes"
With friends	4	9.09	"Outside with friends", "At my friends' house"
Everywhere	4	9.09	"Whenever it's possible", "Almost wherever you can wear headphones"

As seen in Table 25, the most popular place for listening to music was home: 75% (n = 33) reported listening to music at home. The answers included various situations, such as "In the morning", "When I'm going to bed", "While I'm doing homework", "While I'm playing video games", and "By my vinyl player". The second most common listening situation was while travelling. This was mentioned by 70.5% (n = 31) of the participants. The third most popular answer was hobbies: 20.5% (n = 9) reported listening to music while, for example, working out or drawing. Music was also being listened to at school – 13.6% (n = 6) of the participants reported listening to music either "During breaks" or "During boring classes". 9.09% (n = 4) reported listening to music with friends at their house or when outside together. The same percentage said they tend to listen to music basically everywhere, "Whenever it's possible" or "Almost wherever you can wear headphones". Note that the percentages may overlap because the same people might have mentioned multiple places and situations.

This section has provided a multitude of answers regarding the three research questions established in Section 5 by presenting the quantitative results of the study question by question. In the following section, Section 7, the results introduced previously will be discussed further and tied more closely to the research questions as well as to the theoretical framework that was covered in Sections 2–4.

7 Discussion

In this section, the results of the present study that were presented previously will be discussed and interpreted from the perspective of the three research questions presented comprehensively in Section 5 together with the theoretical framework laid out in Sections 2–4. The section will address each research question individually in Sections 7.1, 7.2, and 7.3, after which Section 7.4 shall discuss the limitations of the study.

7.1 Pronunciation score and musical activities in English

The first research question was concerned with how listening to English music and singing in English in one's spare time could explain the variation in the native-speaker evaluations of the English learners' pronunciation. Based on previous knowledge, it was hypothesized that the activities together could provide the language learner with the natural and variable L2 input and output that is necessary for pronunciation acquisition to take place. The pronunciation scores of the participants were higher than expected, and the scores they obtained were surprisingly consistent. Either the participants were very proficient in the English language, or the native-speakers were soft in their evaluations, as found by Tominaga (2011, 52). The evaluators were also accustomed to hearing non-native English at least on a monthly basis. There was little variation between how the participants scored in the three different pronunciation areas. This might, to some extent, be due to the evaluators not being able to tell the areas apart, or simply to the different areas corresponding quite well with each other.

First of all, listening to English music was hypothesized to provide the English learner with comprehensible input, which according to Krashen's Input Hypothesis (1977) is the single most important source of L2 learning. The Input Hypothesis was based on the principal that L1 and L2 learning are similar in many ways, and what is known about L1 learning provided further support for the hypothesis that listening to English language input in the form on English music could enhance pronunciation learning (see de Carvalho et al. 2018, 18; Moon, Lagercrantz, and Kuhl 2013; Partanen 2013). English music was hypothesized to be beneficial for pronunciation learning also based on several studies on adult L2 learners that have managed to demonstrate that the speech-related motor areas of the brain are activated during speech perception (Peltola et al. 2017; Watkins, Strafella, and Paus 2003; Fadiga et al. 2002). Also Lieberman's Motor Theory (1985) rests on a similar assumption. The activation of the motor areas suggests that mere listening to the English language without productive exercises could result in English pronunciation acquisition by altering the brain's nervous system. Listening to English music was also believed to be advantageous for the development of the English phoneme categories,

the hypothesis of which was based on the Perceptual Assimilation Model (Best, 1995). In addition, being familiar with the phonemic form of the word first was believed to minimize the effect of the misleading written cue for Peltola et al. (2015) found that orthographical cues altered the learners' pronunciation towards the cue.

Contrary to the hypothesis, no statistically significant correlation was revealed between listening to English music and the pronunciation score. What is noteworthy about the result, nevertheless, is that the p-value was considerably close to being significant (p = 0.054). This means that a tendency might exist between listening to English music and the English pronunciation skills, yet it would require further investigation of a larger sample to find the potentially statistically significant result. There was a small difference in the mean pronunciation scores of the group that listened to more English music and the group that listened to less English music, but this difference was not statistically significant either (U = 207.50, p = 0.804). The music listening habits of the participants were quite uniform, which resulted in the groups being different in size, and this may have distorted the result.

Importantly, the quality of listening seemed to have an impact on the participants' pronunciation, which is in line with the Noticing Hypothesis (Schmidt 1990, 149). According to Schmidt, noticing is a necessary condition for language learning, and this claim was partially supported by the present study as well: paying attention to English song lyrics while listening to music was linked to higher pronunciation scores (p = 0.049) but paying attention to the pronunciation of the lyrics in particular was not. The result was surprising yet makes sense once given more thought; if one does not pay attention to the lyrics, their pronunciation cannot be registered either, but if attention is paid to the lyrics, it is impossible not to hear (at least subconsciously) how the words are pronounced. Learning pronunciation through English music during one's spare time would qualify as implicit learning (Ellis 1994, 1) in the sense that the learners are normally not trying to learn English by listening to music or singing in English – music is consumed mainly because it is enjoyable. Still, attention is often paid to the music, simply not with the intention of learning pronunciation.

When it comes to singing in English, one needs to remember that pronunciation is a highly motoric skill (Pennington 1996, 20–22), and mastering it requires hours of physical training, which can be uninteresting or unpleasant for the L2 learner. Singing in English was hypothesized to act as a pleasant, unintentional means of training the vocal tract muscles and, therefore, to be beneficial for learning English pronunciation. The idea that a link would exist between singing and pronunciation learning success is not new in itself – several studies have indicated this to be true (Alisaari and Heikkola 2017; Chen 2016; Christiner and Reiterer 2013).

Singing in English was hypothesized to provide the learner with versatile opportunities to produce English language output since song lyrics naturally contain tongue twisters and minimal pairs and could function as "listen and repeat" exercises. All of these techniques were listed by Celce-Murcia, Brinton, and Goodwin (1996, 8–9) as techniques that have been found effective and have traditionally been used in English pronunciation teaching. Song lyrics also tend to contain stories and rhymes that were thought to make songs interesting, memorable, and easily repeatable L2 material. Moreover, Ludke, Ferreira, and Overy (2014) suggested that melody could act as a facilitator for language learning. Chen (2016) found that memorizing song lyrics helped the learners pronounce words with more ease. In the present study, the majority of participants reported learning English song lyrics easily, which suggests that music could, indeed, provide L2 material that is easy to remember.

In accordance with the hypothesis, singing in English yielded a statistically significant correlation with the pronunciation scores (p = 0.028). The better the participants' English pronunciation scores were, the more they had reported singing in English. Particularly strongly linked to the pronunciation score was singing in English without music (p<0.001). One possible explanation for this is that one has to concentrate more as well as to pay more attention to their own pronunciation when there is no music playing in the background. One can also hear themselves better, move at their own pace, and is able to self-correct on the go. Singing along to English songs also showed a statistically significant connection to the pronunciation score (p = 0.037). Here it can be pondered whether singing in English actually influences the learner's pronunciation positively or if the learner sings because they already know how to pronounce the lyrics – this could after all work in both ways. On the other hand, there was no statistically significant difference between the pronunciation scores of the group that sang more in English and the group that sang less, although the group that had reported singing more had a higher mean rank. Again, the groups were of different size due to the distribution of answers to the questions concerning the frequency of singing in English, and this may have affected the results.

Examining the two musical activities separately is tricky in the sense that those who sing in English are also very likely to listen to English music, whereas those who do not sing only engage in one of the activities: listening. This means that there is no group that only sings and does not listen to music, indicating that the people who sing are likely to enjoy the benefits of both listening to English music and singing in English. The musical index, listening to English music and singing in English combined, was significantly linked to the pronunciation score (p = 0.011). In other words, higher levels of engagement in listening to English music and singing in English were linked to higher pronunciation scores among the participants. This is

in accordance with the original hypothesis that the activities together would result in enhanced English pronunciation. Again, there was a difference in the mean pronunciation scores of the higher and lower engagement group, but it was not statistically significant (U = 176.00, p = 0.131). Based on this result it cannot be said that higher than average engagement in English musical activities would be linked to better English pronunciation skills. It is possible, though, that this result was due to there being too little variation within the groups and the sample size not being large enough.

To briefly conclude the findings related to the first research question, it is fair to suggest that listening to English music together with singing in English could have a positive impact on the English learners' pronunciation, and these activities are therefore encouraged. What can be concluded from the listening and singing results separately is that singing in English, with or without the original music playing in the background, was clearly linked to better pronunciation skills. As of now, the link between listening to English music and English pronunciation skills remains more unclear, though is likely to exist. Further research would be required, however.

7.2 Accuracy, fluency, and intelligibility in relation to musical activities

The purpose of the second research question was to examine the correlation between the three criteria of good pronunciation (Pennington 1996, 220) and the musical activities of the English learners. Accuracy, fluency, and intelligibility were all considered important components of pronunciation but according to Pennington (ibid.), intelligibility was the most important one. The intelligibility principle (Levis 2005, 370), according to which it is enough for the L2 speaker to be understood, is in line with Pennington. As international communication is a matter of great importance today, it is easy to see why intelligibility is considered the most important criterion, and this was also highlighted by the Communicative Approach (Celce-Murcia, Brinton, and Goodwin 1996, 10). Nonetheless, the significance of pronunciation accuracy and fluency should not be overlooked because they too contribute to intelligibility and should perhaps receive a little more attention than they currently do. People need to be more than just understood in English; they need to be considered an equal interlocutor. Derwing and Munro (2015, 2) made a point that listeners may sometimes evaluate non-native speakers negatively simply because of their heavily accented speech. Some of the native-speaker evaluators of the present study also found strong accents and "pronouncing words as they are spelt" distracting. Some fluency features, such as excessive pausing and pausing between words, were also considered distracting. Surprisingly, over 70% of the participants thought that it is important to sound as native-like as possible, yet this is not necessarily the case either.

It was hypothesized that listening to English music and singing in English could increase the learners' pronunciation accuracy by repeatedly providing examples of English sounds and by offering opportunities to practise these sounds. English music was also hypothesized to improve the learners' fluency because in music sounds and words blend into each other instead of always being pronounced clearly and highly accurately, similarly to natural language (Lintunen 2014, 166). Also, according to Derwing and Munro (1995), practising vast entities of speech has been found to have the biggest impact on the speaker's comprehensibility. Music also presents words in a context, both meaning- and sound-wise, which was hypothesized to familiarize the learner with the pronunciation phenomena that occur in natural language (see e.g. Gómez González and Sánchez Roura 2016, 234–58; Celce-Murcia, Brinton, and Goodwin 2010, 163–75). By possibly increasing the accuracy and fluency of pronunciation, English music was also hypothesized to increase the overall intelligibility. These above-mentioned hypotheses were based on the theories and studies mentioned along the discussion of the first research question in this section (de Carvalho et al. 2018, 18; Alisaari and Heikkola 2017; Peltola et al. 2017; Chen 2016; Christiner and Reiterer 2013; Moon, Lagercrantz, and Kuhl 2013; Partanen 2013; Strafella and Paus 2003; Watkins, Fadiga et al. 2002; Best 1995; Lieberman, 1985; Krashen 1977).

The results were similar for all the three pronunciation areas. While the accuracy scores of the participants grew, their engagement in English musical activities also increased ($p = 0.019$), indicating that improved accuracy seemed to be linked to more time spent on listening to English music and singing in English. Higher fluency was also linked to higher engagement in musical activities ($p = 0.015$), meaning that those participants who had received more fluency points had also reported spending more time on the English musical activities than those who had received fewer fluency points. A statistically significant correlation was also found between intelligibility and the engagement in musical activities ($p = 0.017$). The correlation to the participants' musical index was the most significant for fluency and the least significant for accuracy yet the differences were marginal.

Singing in English also produced a statistically significant positive correlation to every pronunciation area on its own. Singing correlated least with the participants' pronunciation accuracy ($p = 0.044$), whereas intelligibility and fluency both yielded the same p-value in relation to singing in English ($p = 0.036$). Singing along to English songs was linked to higher intelligibility scores ($p = 0.018$) but did not correlate significantly with accuracy ($p = 0.108$) or fluency ($p = 0.052$). The relationship between accuracy and singing along seemed to be the weakest. A particularly strong correlation, on the other hand, was found between singing in

English without music and the individual pronunciation areas: the correlation was significant at the 0.01 level to all three areas, namely p<0.001 for fluency and intelligibility, and p = 0.007 for accuracy. In contrast, listening to English music did not seem to be linked to any of the individual pronunciation areas. The p-value closest to significant was obtained by fluency (p = 0.061), while the p-values for accuracy and intelligibility were slightly less significant (p = 0.074 and p = 0.075). Again, the p-values were not very far from being significant, meaning it is possible that a larger sample would actually reveal significant correlations.

In conclusion, pronunciation accuracy seemed to be the least related to the musical activities of the English learners. This could potentially be due to song lyrics resembling natural language where sounds are not pronounced clearly and accurately, and segmental features of the language can be difficult to pick up from sung English. Fluency correlated with the musical activities the most, yet the difference between fluency and intelligibility was marginal. English music might influence pronunciation fluency the most for the same reason it was believed to influence accuracy the least: because sung language provides a pronunciation model where words blend into each other like in natural language and flow forward. This could be helpful in acquiring prosodic features of the English language, such as rhythm and stress. Singing in English was linked to higher scores in every individual pronunciation area, whereas listening to English music did not seem to be linked to any of the individual areas. Again, this missing link calls for further investigation of a larger group of English learners.

7.3 Music-related indirect factors and English pronunciation

The third research question was concerned with the indirect factors that could potentially make English music a beneficial means for learning English pronunciation. These included the learner's motivation and the tightly intertwined affective factors (attitudes, emotions, and mood), the subjective experience of the learner regarding learning English, as well as the prevalence of English music consumption. According to Masgoret and Gardner (2003), motivation is a strong predictor when it comes to successful L2 learning. This was true in the sample of the present study as well: higher pronunciation scores were strongly linked to the desire to learn more English (p = 0.003). According to Chambers (1999), personally meaningful and relevant learning content promotes motivation, and music could function as this type of learning content. In the sample of the present study, song lyrics mattered to the majority and music was thought to help one deal with one's feelings by most of the participants, which suggests that people listen to music that has personal meaning to them. Relatable lyrics was also one of the reasons for considering an artist a role model, and it was also found by IFPI and

Teosto (2018) that young people are interested in artists they can relate to. Thus, it was hypothesized that having an English-speaking idol or role model would increase the L2 learners' integrative motivation (Gardner 1985) to learn English by promoting the learner's desire to identify with English speakers (Pietilä 2014, 50), which would result in higher pronunciation scores. This assumption also had to do with Dörnyei's (2009, 217–218) L2 Motivational Self System, according to which people have a desire to reduce the discrepancy between their actual and ideal selves. Nonetheless, this hypothesis was not supported by the data of the present study; those participants who had an English-speaking idol did not seem to be more motivated to learn English than those who did not have one (U = 215.50, p = 0.541). Although there was a difference in the mean pronunciation scores between the two groups, those who had an English-speaking idol scored better on average, it was not statistically significant (U = 180.50, p = 0.177). However, again, the groups were different in size, which may have affected the test result.

English-speaking idols were also thought to increase the learners' desire to learn English by improving the learners' attitudes towards native-speakers of English. Having positive attitudes towards the TL culture and its speakers was said to feed into the processes related to becoming proficient in a particular language (Singleton 2014, 92) and to positively influence the degree of L2 pronunciation proficiency achieved by the learner (Celce-Murcia et al. 2010, 35). According to Schumann's Acculturation Model (1986, 379–392), L2 learning success can be predicted by the extent and quality of contact between the learner and the TL and its culture. In the sample, native-speakers of English were considered nicer by those who had an English-speaking idol by than those who did not have one (U = 139.00, p = 0.014). However, less than half of the participants (43.2%) had reported having an English-speaking idol or role model, and roughly half of the idols mentioned were musicians. A much larger sample would be needed to examine the true relationship between English-speaking idols, English learning motivation, and English pronunciation.

Emotions were the third factor hypothesized to affect pronunciation learning indirectly. According to Krashen's Affective Filter Hypothesis (1982), the language learner's emotional state affects learning. The affective filter has to be low for the mind to be in a state receptive for new information. It has been found previously that music could make students more receptive to teaching and open to learning (Aguirre, Bustinza, and Garvich 2016; Sigurðardóttir 2011). School exercises that contain music were generally considered pleasant by the participants of the present study as well: only 29.6% disagreed. A vast majority of the participants agreed that music gets them in a good mood. The reason why school exercises were

less liked than music in general might have to do with the students not being able to choose the music themselves, or the fact that the activity is obligatory and taking place in a school setting. Ortis (2008, 216) explained that when people hear music, the serotonin levels of the brain rise producing a stress-relieving effect. A vast majority of the participants found listening to music relaxing, which suggests a low affective filter and a receptive mind state. Although Chen (2016) yielded promising results about the enjoyability of singing in pronunciation teaching, one cannot help but wonder whether it is fun for everyone. Slightly over half of the participants of the present study reported enjoying singing, and a positive correlation was found between this and the pronunciation score ($r = 0.334$, $p = 0.027$). Whether the participants considered themselves as a good singer was not linked to the pronunciation scores ($r = 0.042$, $p = 0.785$), which suggests that in order to benefit from English music in pronunciation learning, it may not matter if one can sing or not.

According to Ushioda (2015, 47), the subjective learning experience and learning environment function as external variables that influence the learner's motivation, development, and performance. More than a third of the participants felt like not enough pronunciation is being practised during English classes at school. This was in line with Tergujeff (2013). Some of the participants thought English pronunciation is difficult to learn because it contains difficult sounds, and some felt like the difficulty lies in the words being pronounced differently than they are written. However, most of the participants felt like music had helped them learn English, and almost everyone thought it had helped them learn English pronunciation specifically. What is more, listening to English music and singing in English were both considered good ways to practise and learn English pronunciation by the participants. Positive learning experiences and learning success can improve the learner's attitudes towards the TL (Singleton 2014, 92), which could further boost learning.

It was discovered by the Music Consumer Insight Report (IFPI 2018) that music is consumed everywhere and at every point of the day. The data suggested a similar listening trend: the most typical situations and places where the participants reported listening to music included travelling, school, hobbies, home, with friends, and – "everywhere". At home, music was listened to from waking up until going to bed and everywhere between. English music was listened to by the majority of participants on four to seven days a week, and most participants listened to it for one to four hours a day. Singing was less popular: it was done on two to three days a week and not more than two hours a day. However, this was to be expected for one would have to be very serious about singing in order to sing for multiple hours a day. This

constant exposure to the English language through music is a considerable source of English language input for many people also outside the sample of this study.

7.4 Limitations of the present study

As in any study, there are issues that could compromise its reliability and validity. Regarding the study at hand, demonstrating that the outcome is truly a function of the variables that were measured is challenging. English pronunciation is a skill that can be affected by various factors, so it is difficult to ensure that the result was not caused by an underlying factor. Most importantly, when a Pearson Correlation test is used as a measuring device, it can only ever be stated that two variables are linked and suggested that one may have affected the another, because it does not reveal any cause and effect relationships.

In the present study *English music* was defined as all music sung in the English language and no difference was made between native- and non-native singers. Also, all types of English music were considered and, therefore, it has to be mentioned that not every musical genre is equally suitable for pronunciation learning. The genre, origin, and topic of the songs are all important variables when it comes to evaluating how suitable music is for pronunciation learning purposes. For example, in country music lyrics play a central part and are often presented more clearly than in, say, extreme metal. The genres that were most popular among the participants were pop and hip hop, which are considered suitable because the lyrics are generally distinguishable. The origin of the music is noteworthy considering that English is spoken all around the world in hundreds of regional as well as social varieties. An L2 learner might end up acquiring elements from two completely different accents depending on whether they are exposed to American hip hop or British hip hop. The complete list of the artists and bands mentioned by the participants can be found in Appendix 5 for anyone interested in the type of music that led to the results of the present study.

The sample size of the study was rather small, meaning that the generalizability of the findings to a larger population is not ideal. To improve the external validity, a larger sample size would be required. There was less variation in the participants' pronunciation skills than expected. The musical habits of the participants were also more uniform than expected, which meant that grouping the participants based on their habits produced groups that were of different size. This made comparing the music listening habits of the participants challenging. Equal group sizes would be better for reliable comparisons between them to be made. Different kinds of grouping methods could have been tested on the sample but not in the limited scope of this study. To reveal possible underlying factors the participants were asked about their other

activities, too. While no significant correlation was found between the other activities (watching English tv-shows, movies, or videos, listening to English podcasts, and reading in English) and the pronunciation score, playing video games in English did produce a statistically significant correlation ($r = 0.344$, $p = 0.022$). What is also noteworthy is that musicality is strongly linked to greater L2 learning success (see Milovanov 2009; Nardo and Reiterer 2009; Slevc and Miyake 2006) so caution must be taken here: it is a potential underlying factor that could explain some of the results of the present study. Some participants might spend time with English music because they are musically talented while at the same time pronounce English well because they are also linguistically talented.

What decreases the reliability of the questionnaire is that it was not based on a previous one but constructed by the researcher based on her previous knowledge of the topic. This was done because no suitable previous questionnaire template could be found. Using a Likert scale to measure people's feelings towards certain matters is also always risky because information gathered via self-reporting is not always reliable and the answers might be embellished (Pietilä 2014, 49). No middle value was provided in order to improve the validity of the study, but this also means that if someone was completely neutral about one of the statements they still had to tend towards either end of the scale. Also, choosing convenient time-frames regarding the frequency of a certain activity so that the options would not overlap was challenging. The questions also provoked some semantic issues: semantically, "one day a week" and "couple of times a month" could mean the same thing depending on the interpretation made by the participants and despite the researcher's idea of "couple times a month" meaning less than four times. The main difference between these two options most likely lies in the regularity, perceived or actual; singing on one day a week sounds more habitual than singing on a couple of days a month, but this should preferably be made clearer in future research. Also, the options "Never", "Rarely", "Sometimes", "Often", and "Always" are open for subjective interpretation. Overall, the type of self-assessment where participants are asked to reflect on their language use and everyday habits can raise the question about its validity. It might be challenging for a person to estimate how much time they spend on an activity that is usually not measured chronologically. Without the possibility of conducting a longitudinal study, however, self-estimations are a convenient way of gathering information about the participants' everyday habits. 9th graders are also mature enough to be able to make these estimations.

Collecting natural speech samples is difficult, because the test situation always tends to change the performance in some ways. Without a doubt, this is a problem with clinically elicited data. In the recording situation of the PDT, the researcher's close presence may have affected

the participants' performance, namely the naturalness of the language that was produced. This is known as the *Hawthorne effect* (Dörnyei 2007, 53). The presence of a microphone and knowing that one is being recorded is also known to induce self-consciousness in the language learner, which could make the speech less natural (Ellis and Barkhuizen 2005, 27). Some of the participants were very nervous, which in some cases led to decreased fluency. Planning time, on the other hand, may have increased the accuracy of the performance (Ellis and Barkhuizen 2005, 22). The participants also knew that their pronunciation was being studied, which might have caused them to concentrate on it more than they normally would. When collecting learner language samples, there is also the *variability problem* (Ellis and Barkhuizen 2005, 7) that has to do with linguistic performance being inherently heterogenous. That is, the speaker's production can vary from time to time, and from one task to another. Someone who performed poorly in the PDT could have performed better at a different time due to multiple reasons, for instance current emotions or energy levels. Also, as mentioned earlier, there was music class next door, which might have distracted some of the students.

When it comes to evaluating non-native speech samples, a larger and more geographically varied group of native-speakers would ideally be needed. The concepts of accuracy, fluency, and intelligibility as defined in the context of this study were not familiar to the evaluators and understanding them may have been challenging. Also separating pronunciation from other English skills may have been difficult for the evaluators, thus the pronunciation of those speakers whose English was better overall was likely to be rated higher than the pronunciation of those who made, say, more grammatical errors. This time the researcher had no control over the evaluators since the evaluations took place in the evaluators' own countries. There is no way of knowing how well the evaluators followed the instructions given to them, and in an ideal situation the evaluations would have been more controlled and done in the presence of the researcher. Nevertheless, a decision had to be made between having a less controlled native-speaker jury and a having a more controlled L2 learner jury, and a native-speaker jury fit the purpose better despite the disadvantages.

This section has discussed the results of the present study in light of the research questions and the theoretical framework as well as addressed certain limitations that cannot be ignored when interpreting the results of the study. However, the methodological choices were made based on the researcher's best and most carefully debated knowledge. All in all, considering the reality and challenge of studying a topic such as this one, the study was successfully executed.

8 Conclusion

The present study set to discover the potential effects listening to English music and singing in English during one's spare time could have on the English learner's pronunciation. The issue was approached through three research questions, and a quantitative study was conducted on a group of 9th graders in Southwest Finland. The participants filled in a questionnaire, after which they provided a short speech sample in English. The pronunciation in the speech samples was then evaluated and rated by a native-speaker jury, and the pronunciation scores were compared to the participants' musical habits.

The results suggest that the English music people consume for enjoyment could be beneficial for the development of their English pronunciation. It was discovered that the two musical activities together, listening to English music and singing in English, were linked to better pronunciation skills. Singing in English seemed to be more closely related to better pronunciation than listening to English music, but the quality of listening was also relevant: paying attention to the lyrics was linked to higher pronunciation scores. Higher engagement in the musical activities was also linked to higher independent accuracy, fluency, and intelligibility scores. These above-mentioned links are likely to be due to the constant natural language input listening to English music provides the learner with, combined with the versatile and implicit motor training of the vocal tract muscles that takes place through singing in English. These are factors that can directly influence the learner's pronunciation.

In addition to the direct factors, it was found that some indirect factors could also have an influence on the pronunciation of the English learner. For example, a receptive mind state is likely to be achieved through listening to English music and singing in English; music was considered a relaxing "mood-booster" by the vast majority of the participants. Music could also help create a positive learning experience, and listening to English music and singing in English were thought to be good ways to learn and practise English pronunciation. What is more, the participants felt like English music had helped them learn English and English pronunciation especially. Also, native-speakers of English were considered nicer by those who had an English-speaking idol or role model, and positive attitudes towards the speakers of the target language are known to have a positive impact on L2 learning. The participants who had an English-speaking idol or role model seemed to pronounce English better than those who did not have one, yet the difference was not statistically significant and further research needs to be conducted.

The results of this study have both practical and educational implications. Knowing that the engagement in musical activities in English could influence one's English pronunciation

positively could encourage people to listen to more English music in their spare time, as well as to sing in English even if singing was not one of their strengths. English teachers could encourage their students to pay attention to the lyrics whenever they listen to English music and encourage them to sing their favourite songs without shame, home alone, for example. The teacher could also ask the students to pick songs that could be listened to during the English class – the fact that the songs are chosen by the students themselves should enhance the efficacy of the exercise. The students could then, for instance, try to write down the lyrics based on what they hear without looking them up online.

The present study has only scratched the surface of the relationship between the music people consume in their spare time and English pronunciation, and plenty remains to be studied. For future research, a larger sample size would be recommended so that more reliable comparisons could be made between groups that have different musical habits. Regarding a native-speaker evaluation, it should be kept as controlled as possible. It would also be highly recommendable to include the issue of musicality in the study, since it is such a potential underlying factor when pronunciation skills are being examined. In more advanced studies, attention could also be paid on the type of music the participants consume: are certain genres perhaps more beneficial than others? The incredibly complex question of musical idols as an indirect factor affecting pronunciation learning is also yet to be answered. Despite the evident difficulty of demonstrating cause and effect relationships, researchers of the future should keep exploring – even though the consumption of English music may not be the only factor influencing English learners' pronunciation, it is likely to be one of the factors.

References

Aguirre, Diego, Daisy Bustinza, and Mijail Garvich. 2016. "Influence of Songs in Primary School Students' Motivation for Learning English in Lima, Peru." *English Language Teaching* 9, no. 2: 178–91. Accessed 20 July 2018. http://dx.doi.org/10.5539/elt.v9n2p178.

Alisaari Jenni, and Leena Maria. Heikkola. 2017. "Songs and Poems in the Language Classroom: Teacher's Beliefs and Practices." *Teaching and Teacher Education* 63: 231–42. Accessed 2 April 2019. https://doi.org/10.1016/j.tate.2016.12.021.

Arnold, Jane, and H. Douglas Brown. 1999. "A Map of Terrain". In *Affect in Language Learning*, edited by Jane Arnold, 1–27. Cambridge: Cambridge University Press.

Barbeau, Elise B., Xiaoqian J. Chai, Jen-Kai Chen, Jennika Soles, Jonathan Berken, Shari Baum, Kate E. Watkins, and Denise Klein. 2017. "The Role of the Left Inferior Parietal Lobule in Second Language Learning: An Intensive Language Training Study." *Neuropsychologia* 98: 169–76. Accessed 8 June 2019. ScienceDirect.

Bartram, Dave. 2007. "Increasing Validity with Forced-Choice Criterion Measurement Formats." *International Journal of Selection and Assessment* 15, no. 3: 263–72. Accessed 9 May 2019. https://doi-org.ezproxy.utu.fi/10.1111/j.1468-2389.2007.00386.x.

Best, Catherine. 1995. "A direct realist perspective on cross-language speech perception." In *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues in Cross-language Speech Research,* edited by Winifred Strange, 167–200. Timonium, MD: York Press.

Boersma, Paul. 2012. "A constraint-based explanation of the McGurk effect." In *Phonological Explorations: Empirical, Theoretical and Diachronic Issues,* edited by Bert Botma and Roland Noske, 299–312. Berlin: Mouton De Gruyter.

Bongaerts, Theo, Chantal van Summeren, Brigitte Planken, and Erik Schils. 1997. "Age and Ultimate Attainment in the Pronunciation of a Foreign Language." *Studies in Second Language Acquisition* 19, no. 4: 447–65. Accessed 11 June 2019. ResearchGate.

Brown, H. Douglas. 2004. *Language Assessment: Principles and Classroom Practices.* San Francisco, CA: Pearson Longman.

Carroll, Susanne E. 1999. "Putting 'Input' in Its Proper Place." *Second Language Research* 15: 337–88. Accessed 8 June 2019. https://doi.org/10.1191/026765899674928444.

Celce-Murcia, Marianne, Donna Brinton, and Janet Goodwin. 1996. *Teaching Pronunciation: A Reference for Teachers of English to Speakers of Other Languages.* Cambridge: Cambridge University Press.

Celce-Murcia, Marianne, Donna Brinton, and Janet Goodwin. 2010. *Teaching Pronunciation: A Course Book and Reference Guide.* 2nd ed. Cambridge: Cambridge University Press.

Chalmers, Lynne, Myrna R. Olson, and Joyce K. Zurkowski. 1999. "Music as a Classroom Tool." *Intervention in School and Clinic* 35, no. 1: 43–52. Accessed 19 July 2018. https://doi.org/10.1177/105345129903500108.

Chambers, Gary. 1999. *Motivating language learners.* Clevedon, UK: Multilingual Matters.

Chen, Ai-Hwa. 2016. "Perspectives upon Integrating Music into Freshman English Pronunciation Training Classroom." *Asian Journal of Education and Training* 2, no. 1: 16–22. Accessed 11 November 2018. https://files.eric.ed.gov/fulltext/EJ1148345.pdf.

Christiner, Markus, and Susanne M. Reiterer. 2013. "Song and Speech: Examining the Link Between Singing Talent and Speech Imitation Ability." *Frontiers in Psychology* 4: 874. Accessed 15 June 2019. https://doi.org/10.3389/fpsyg.2013.00874.

Cockburn, Victor. 1991. "The Uses of Folk Music and Songwriting in the Classroom." *Harvard Educational Review* 61, no. 1: 71–80. Accessed 20 May 2019. https://doi.org/10.17763/haer.61.1.k74t252544t00710.

de Carvalho, Alex, Isabelle Dautriche, Séverine Millotte, and Anne Christophe. 2018. "Early perception of phrasal prosody and its role in syntactic and lexical acquisition." In *The Development of Prosody in First Language Acquisition,* edited by Pilar Prieto and Núria Esteve-Gibert,* 18–35. Amsterdam: John Benjamins Publishing Company.

de Jong, Kenneth J. 2018. "Sensitivity to Foreign Accent." *Acoustics Today* 14, no. 2: 9–16. Accessed 20 July 2019. https://acousticstoday.org/wp-content/uploads/2018/06/Sensitivity-to-Foreign-Accent-Kenneth-J.-de-Jong.pdf.

Delmonte, Robert. 2017. "*Music Consumption: The Overall Landscape." AudienceNet.* Accessed 13 May 2019. https://musicbiz.org/wp-content/uploads/2017/09/MusicConsumptionTheOverallLandscape_AudienceNet.pdf.

Derwing, Tracey M. 2018. "The efficacy of pronunciation instruction." In *The Routledge Handbook of Contemporary English Pronunciation,* edited by Okim Kang, Ron I. Thompson, and John M. Murphy, 320–34. Abingdon, Oxon: Routledge.

Derwing, Tracy M., and Murray J. Munro. 2005. "Second Language Accent and
    Pronunciation Teaching: A Research-Based Approach." *TESOL Quarterly* 39, no. 3:
    379–97. Accessed 2 June 2019. https://doi.org/10.2307/3588486.

Derwing, Tracy M., and Murray J. Munro. 2013. "The Development of L2 Oral Language
    Skills in Two L1 Groups: A 7-Year Study". *Language Learning* 63, no. 2: 163-85.
    Accessed 25 April 2018. https://doi.org/10.1111/lang.12000.

Derwing, Tracy M., and Murray J. Munro. 2015. *Pronunciation Fundamentals: Evidence-
    based Perspectives for L2 Teaching and Research.* Amsterdam: John Benjamins
    Publishing Company.

Dörnyei, Zoltán. 2007. *Research Methods in Applied Linguistics: Quantitative, Qualitative,
    and Mixed Methodologies.* Oxford: Oxford University Press.

Dörnyei, Zoltán. 2009. *The Psychology of Second Language Acquisition.* Oxford: Oxford
    University Press.

Dörnyei, Zoltán, and Peter Skehan. 2003. "Individual differences in second language
    learning." In *The Handbook of Second Language Acquisition*, edited by Catherine
    Doughty and Michael H. Long, 598–630. Malden, MA: Blackwell.

Dörnyei, Zoltán, and Tatsuya Taguchi. 2010. *Questionnaires in Second Language Research:
    Construction, Administration, and Processing.* 2nd ed. New York: Routledge.

Eagly, Alice H., and Shelly Chaiken. 1993. *The Psychology of Attitudes*. Orlando, FL:
    Harcourt Brace Jovanovich College Publishers.

Educational Testing Service (ETS). 2019. "About the TOEIC Speaking and Writing Tests."
    Accessed 13 November 2018. https://www.ets.org/toeic/test-takers/speaking-
    writing/about.

Ellis, Nick C., ed. 1994. *Implicit and Explicit Learning of Languages.* London: Academic
    Press.

Ellis, Rod, and Gary Barkhuizen. 2005. *Analysing Learner Language*. Oxford: Oxford
    University Press.

EnglishClub. 2019. "TOEIC Practice – Speaking." Accessed 13 November 2018.
    https://www.englishclub.com/esl-exams/ets-toeic-practice-speaking.htm.

Fadiga, Luciano, Laila Craighero, Giovanni Buccino, and Giacomo Rizzolatti. 2002. "Speech
    Listening Specifically Modulates the Excitability of Tongue Muscles: a TMS Study."
    European Journal of Neuroscience 15: 399–402. Accessed 21 May 2019.
    https://doi.org/10.1046/j.0953-816x.2001.01874.x.

Gardner, Robert C. 1985. *Social Psychology and Second Language Learning: The Role of Attitudes and Motivation.* London: Edward Arnold.

Gholipour, Bahar. 2018. "What is a normal heart rate." Accessed 30 October 2018. https://www.livescience.com/42081-normal-heart-rate.html.

Gilbert, Judy B. 2010. "Pronunciation as Orphan: What Can Be Done?" Accessed 15 June 2019. https://www.tesol.org/docs/default-source/new-resource-library/pronunciation-as-orphan-what-we-can-do-about-it-.pdf?sfvrsn=0.

Gómez González, María de los Ángeles, and Teresa Sánchez Roura. 2016. *English Pronunciation for Speakers of Spanish: From Theory to Practice.* Berlin: De Gruyter Mouton.

Heikkola, Leena. 2018. "Ääntäminen kehittyy laulun sanoja lausumalla." [Pronunciation develops by reciting song lyrics]. MA Thesis, University of Turku.

IFPI. 2018. "Music consumer insight report." *AudienceNet.* Accessed 13 May 2019. https://www.ifpi.org/downloads/Music-Consumer-Insight-Report-2018.pdf.

Iverson, Paul, Valerie Hazan, and Kerry Bannister. 2005. "Phonetic Training with Acoustic Cue Manipulations: A Comparison of Methods for Teaching English /r/-/l/ to Japanese Adults." *Journal of the Acoustical Society of America* 118, no. 5: 3267–78. Accessed 28 January 2019. https://doi.org/10.1121/1.2062307.

Kang, Okim, and April Ginther. 2018. *Assessment in Second Language Pronunciation.* London: Routledge.

Keltner, Dacher, and Paul Ekman. 2000. "Emotion: An Overview". In *Encyclopedia of Psychology,* edited by Alan E. Kazdin, 162–67. Oxford: Oxford University Press.

Krashen, Stephen D. 1982. *Principles and Practice in Second Language Acquisition.* Oxford: Pergamon Press.

Lee, Angela Y., and Brian Sternthal. 1999. "The Effects of Positive Mood on Memory." *Journal of Consumer Research* 26, no. 2: 115–27. Accessed 8 June 2019. https://doi.org/10.1086/209554.

Lennon, Paul. 1990. "Investigating fluency in EFL: A Quantitative Approach." *Language Learning* 40, no. 3: 387–417. Accessed 21 April 2019. https://onlinelibrary-wiley-com.ezproxy.utu.fi/doi/epdf/10.1111/j.1467-1770.1990.tb00669.x.

Levis, John. M. 2005. "Changing Contexts and Shifting Paradigms in Pronunciation Teaching." *TESOL Quarterly* 39, no. 3: 369–77. Accessed 20 February 2019. ResearchGate.

Levis, John. M. 2006. "Pronunciation and the Assessment of Spoken Language." In
Spoken English, TESOL and Applied Linguistics: Challenges for Theory and Practice,
edited by Rebecca Hughes, 245–70. London: Palgrave Macmillan.

Liberman, Alvin M. 1985. "The Motor Theory of Speech Perception Revised." *Cognition* 21,
no. 1: 1–36. Accessed 3 June 2019. https://doi.org/10.1016/0010-0277(85)90021- 6.

Lieberman, Philip, and Sheila E. Blumstein. 1988. *Speech Physiology, Speech Perception,
and Acoustic Phonetics*. Cambridge Studies in Speech Science and Communication.
Cambridge: Cambridge University Press.

Lintunen, Pekka. 2014. "Ääntämisen oppiminen ja opettaminen." [Pronunciation learning and
teaching]. In *Kuinka kieltä opitaan,* edited by Päivi Pietilä and Pekka Lintunen, 165–
87. Helsinki: Gaudeamus.

Litovsky, Ruth. 2015. "Development of the Auditory System." *Handbook of Clinical
Neurology* 129: 55–72. Accessed 13 September 2018. https://doi.org/10.1016/B978-0-
444-62630-1.00003-2.

Long, Mike. 2013. "Maturational constraints on child and adult SLA." In *Sensitive periods,
Language Aptitude, and Ultimate L2 Attainment,* edited by Gisela Granena and Mike
Long, 3–37. Amsterdam: John Benjamin Publishing Company.

Lord, Gillian. 2009. "Podcasting Communities and Second Language Pronunciation."
Foreign Language Annals 41, no. 2: 364–79. Accessed 8 June 2019.
https://doi.org/10.1111/j.1944-9720.2008.tb03297.x.

Ludke, Karen M., Fernanda Ferreira, and Katie Overy. 2014. "Singing Can Facilitate Foreign
Language Learning." *Memory and Cognition* 42, no. 1: 41–52. Accessed 11 May
2019. ResearchGate.

Macdonald, Doris, George Yule, and Maggie Powers. 1994. "Attempts to Improve English L2
Pronunciation: The Variable Effects of Different Types of Instruction." *Language
Learning* 44, no. 1: 75–100. Accessed 3 September 2018.
https://doi.org/10.1111/j.1467-1770.1994.tb01449.x.

MacIntyre, Peter D., S.P. Mackinnon, and Richard Clément. 2009. "The baby, the bathwater,
and the future of language learning motivation research." In *Motivation, Language
Identity and the L2 Self,* edited by Zoltán Dörnyei and Ema Ushioda, 43–65. Bristol:
Multilingual Matters.

Magnan, Sally S., and Barbara A. Lafford. 2012. "Learning through immersion during study
abroad." In *The Routledge Handbook of Second Language Acquisition,* edited by
Susan M. Gass and Alison Mackey, 525–40. Abingdon, Oxon: Routledge.

Masgoret, Anne-Marie, and Robert. C. Gardner. 2003. Attitudes, motivation, and second language learning: A meta-analysis of studies conducted by Gardner and Associates. Language Learning 53, no. 1: 123–63. Accessed 3 May 2018. https://doi.org/10.1111/1467-9922.00212.

Mathers, Marshall B., Jeffrey I. Bass, Kevin D. Bell, Anne J. Dudley, Malcolm R. A. McLaren, and Trevor C. Horn. 2002. "Without Me". On *The Eminem Show* [CD], track 10. Santa Monica, CA: Aftermath Records.

McIntosh, James. 2018. "What is serotonin and what does it do?" *Medical News Today*. Accessed 30 October 2018. https://www.medicalnewstoday.com/kc/serotonin-facts-232248.

Meneses, Alfredo, and Gustavo Liy-Salmeron. 2012. "Serotonin and Emotion, Learning and Memory." *Reviews in the Neurosciences* 23, no. 5–6: 543–553. Accessed 28 May 2019. https://doi.org/10.1515/revneuro-2012-0060.

Milovanov, Riia. 2009. "Musical Aptitude and Foreign Language Learning Skills: Neural and Behavioural Evidence About Their Connections." In *Proceedings of the 7th Triennial Conference of European Society for the Cognitive Sciences of Music (ESCOM)*, edited by Jukka Louhivuori et al., 338–42. Jyväskylä: University of Jyväskylä. Accessed 8 June 2019. https://jyx.jyu.fi/bitstream/handle/123456789/20935/1/urn_nbn_fi_jyu-2009411285.pdf.

Miranda, Dave. 2013. "The Role of Music in Adolescent Development: Much More Than the Same Old Song." *International Journal of Adolescent and Youth* 18, no. 1: 5–22. Accessed 9 August 2019. https://doi.org/10.1080/02673843.2011.650182.

Mishan, Freda. 2005. *Designing Authencity into Language Learning Materials*. Bristol: Intellect Books.

Moon, Christine, Hugo Lagercrantz, and Patricia K. Kuhl. 2013. "Language Experienced in Utero Affects Vowel Perception After Birth: A Two-Country Study." *Acta Paediatrica* 102, no. 2: 156–60. Accessed 18 September 2018. https://doi.org/10.1111/apa.12098.

Moreno, Sylvain, Ellen Bialystok, Raluca Barac, E. Glenn Schellenberg, Nicholas J. Cepeda, and Tom Chau. 2011. "Short-term Music Training Enhances Verbal Intelligence and Executive Function." *Psychological Science* 22, no. 11: 1425–33. https://doi.org/10.1177/0956797611416999.

Moyer, Alene. 1999. "Ultimate Attainment in L2 Phonology: The Critical Factors of Age, Motivation, and Instruction." *Studies in Second Language Acquisition* 21, no 1: 81–108. Accessed 11 June 2019. https://doi.org/10.1017/S0272263199001035.

Munro, Murray J., and Tracy. M. Derwing. 1995. "Foreign Accent, Comprehensibility and Intelligibility in the Speech of Second Language Learners." *Language Learning* 45: 73–97. Accessed 30 October 2018. https://doi.org/10.1111/j.1467-1770.1995.tb00963.x.

Møller, Aage R. 2009. *Malleable Brain: Benefits and Harm from Plasticity of the Brain*. Hauppauge, NY: Nova Science Publishers.

Nardo, Davide, and Susanne M. Reiterer. 2009. "Musicality and phonetic language aptitude." In *Language talent and brain activity: Trends in applied linguistics,* edited by Grzegorz Dogil and Susanne M. Reiterer, 213–55. Berlin: Mouton De Gruyter.

North, Adrian C., and David J. Hargreaves. 1999. "Music and Adolescent Identity." *Musical Education Research* 1, no. 1: 75–92. Accessed 9 August 2019. https://doi.org/10.1080/1461380990010107.

Ortis, Jana. 2008. "The Effects of Music, Rhymes and Singing on the Classroom Environment." *Teaching the Child in Front of You in a Changing World,* Conference Proceedings, 214–21. Accessed 8 June 2019. http://archives.evergreen.edu/masterstheses/Accession89-10MIT/2008MITMastersProjects.pdf.

Partanen, Eino, Teija Kujala, Risto Näätänen, Auli Liitola, Anke Sambeth, and Minna Huotilainen. 2013. "Learning-induced Neural Plasticity of Speech Processing Before Birth". *PNAS* 110, no. 37: 15145-50. Accessed 18 September 2018. https://doi.org/10.1073/pnas.1302159110.

Peltola, Kimmo U., Henna Tamminen, Paavo Alku, and Maija S. Peltola. 2015. "Non-native Production Training with an Acoustic Model and Orthopraphic or Transcription Cues." *18th International Congress of Phonetic Sciences.* Accessed 18 May 2018. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0236.pdf.

Peltola, Kimmo U., Paavo Alku, and Maija S. Peltola. 2017. "Non-native Speech Sound Production Changes Even with Passive Listening Training." *Linguistica Lettica* 25: 158–172. Accessed 18 May 2018. https://research.utu.fi/converis/getfile?id=29817132&portal=true.

Pennington, Martha. 1996. *Phonology in English Language Teaching: An International Approach.* Harlow: Pearson Education Limited.

Pietilä, Päivi. 2014. "Yksilölliset erot kielen oppimisessa." [Individual differences in language learning]. In *Kuinka kieltä opitaan*, edited by Päivi Pietilä and Pekka Lintunen, 45–67. Helsinki: Gaudeamus.

Raviv, Amiram, Daniel Bar-Tal, Alona Raviv, and Asaf Ben-Horin. 1996. "Adolescent Idolization of Pop Singers: Causes, Expressions, and Reliance." *Journal of Youth and Adolescence* 25, no. 5: 631–50. Accessed 11 August 2019. https://link.springer.com/article/10.1007/BF01537358.

Reber, Arthur S. 2003. "Implicit learning". In *Encyclopedia of Cognitive Science,* edited by Lynn Nadel, 486–91. London: Nature Publishing.

Rizzolatti, Giacomo, and Laila Craighero. 2004. "The Mirror-Neuron System." *Annual Review of Neuroscience* 27, no. 1: 169–92. Accessed 21 April 2018. https://doi.org/10.1146/annurev.neuro.27.070203.144230.

Rogerson-Revell, Pamela. 2018. "English vowels and consonants." In *The Routledge Handbook of Contemporary English Pronunciation*, edited by Okim Kang, Ron Thomson, and John Murphy. Abingdon, Oxon: Routledge. Accessed 26 August 2019. EBSCOhost.

Schmidt, Richard W. 1990. "The Role of Consciousness in Second Language Learning." *Applied Linguistics* 11, no. 2: 129–58. Accessed May 23, 2019. http://nflrc.hawaii.edu/PDFs/SCHMIDT%20The%20role%20of%20consciousness%20in%20second%20language%20learning.pdf.

Schumann, John H. 1986. "Research on the Acculturation Model for Second Language Acquisition." *Journal of Multilingual and Multicultural Development* 7, no. 5: 379–92. ResearchGate.

Scott, Sophie K. 2017. "Perception and Production of Speech: Connected, But How?". In *Speech Perception and Spoke Word Recognition*, edited by Gareth Gaskell and Jelena Mirkovic, 23–36. Abingdon, Oxon: Routledge. Accessed 27 August 2019. ProQuest Ebrary.

Segalowitz, Norman, and Jan H. Hulstijn. 2005. "Automaticity in bilingualism and second language learning." In *Handbook of Bilingualism: Psycholinguistic Approaches*, edited by Judith F. Kroll, and A. M. B. De Groot, 317–88. Oxford: Oxford University Press.

Sheeran, Edward C., and Jake N. Gosling. 2011. "Wake Me Up". On + [CD], track 5. London: Warner Music UK Ltd.

Sigurðardóttir, Drífa. 2011. "Language learning through music." Bachelor of Education, University of Iceland. Accessed 7 October 2018. https://skemman.is/bitstream/1946/12591/1/B.Ed.%20Thesis.%20Language%20learning%20through%20music.%20Dr%C3%ADfa%20Sigurðardóttir.pdf.

Singleton, David. 2014. "How Do Attitude and Motivation Help in Learning a Second Language?" In *Key Topics of Second Language Acquisition,* edited by Vivian Cook and David Singleton, 90–102. Bristol: Multilingual Matters.

Skehan, Peter. 1996. "Second language acquisition research and task-based instruction." In *The Challenge and Change in Language Teaching*, edited by Jane Willis and Dave Willis, 17–30. Oxford: Heinemann.

Slevc, L. Robert, and Akira Miyake. 2006. "Individual Differences in Second-Language Proficiency: Does Musical Ability Matter?" *Psychological Science* 17, no. 8: 675–81. https://doi.org/10.1111/j.1467-9280.2006.01765.x.

Sloboda, J. A. 1993. "Musical ability." In *The Origins and Development of High Ability,* edited by Gregory R. Bock and Kate Ackrill, 106–118. Chichester: Wiley.

Sundman, Marketta. 2014. "Kuinka kielioppia opitaan ja opetetaan." [How grammar is learned and taught]. In *Kuinka kieltä opitaan*, edited by Päivi Pietilä and Pekka Lintunen, 114–37. Helsinki: Gaudeamus.

Suomi, Kari, Juhani Toivanen, and Riikka Ylitalo. 2008. *Finnish Sound Structure: Phonetics, Phonology, Phonotactics and Prosody.* Oulu: Oulu University Press.

Swain, Merrill. 1995. "Three functions of output in second language learning". In *Principle and Practice in Applied Linguistics*, edited by Guy Cook and Barbara Seidlhofer, 125–44. Oxford: Oxford University Press.

Szyszka, Magdalena. 2011. "Foreign Language Anxiety and Self-perceived English Pronunciation Competence." *Studies in Second Language Learning and Teaching* 1, no. 2: 283–300. Accessed 1 June 2019. ResearchGate.

Tatham, Mark, and Katherine Morton. 2011. *A Guide to Speech Production and Perception.* Edinburgh: Edinburgh University Press.

Tervonen, Kari. 2017. "Music consumption by age group: What do people listen to, what do they like – and why? How Finns listen to music and use TV content online*." IFPI and Teosto.* Accessed 13 May 2019.

https://www.teosto.fi/sites/default/files/files/Research_Report_How_Finns_listen_to_ music_and_use_TV_content_online_2017.pdf.

Tervonen, Kari. 2018. "Kuinka suomalaiset kuuntelevat musiikkia 2018: Miten uudet tähdet syntyvät?" *IFPI and Teosto*. Accessed 13 May 2019. https://www.teosto.fi/sites/default/files/files/Teosto_IFPI_tutkimus_Kuinka_suomalai set_kuuntelevat_musiikkia_2018.pdf.

Tergujeff, Elina. 2012. "English Pronunciation Teaching: Four Case Studies from Finland." *Journal of Language Teaching and Research* 3, no. 4: 599–607. Accessed 9 November 2018. ResearchGate.

Tergujeff, Elina. 2013. "English Pronunciation Teaching in Finland." Doctoral Dissertation, University of Jyväskylä. Accessed 17 June 2019. https://jyx.jyu.fi/handle/123456789/41900.

TheTalkList. 2016. "TOEIC Picture Question Examples." Accessed 13 November 2018. http://www.thetalklist.com/article/Examples_TOEIC_Part_2.pdf.

Ushioda, Ema. 2015. "Context and Complex Dynamic Systems Theory." In *Motivational Dynamics in Language Learning*, edited by Zoltán Dörnyei, Peter D. MacIntyre, and Alastair Henry, 47–54. Bristol: Multilingual Matters.

Ushioda, Ema, and Zoltán Dörnyei. 2012. "Motivation." In *The Routledge Handbook of Second Language Acquisition,* edited by Susan M. Gass and Alison Mackey, 396–409. Abingdon, Oxon: Routledge.

Watkins, Kate E., Antonio P. Strafella, and Tomán Paus. 2003. "Seeing and Hearing Speech Excites the Motor System Involved in Speech Production". *Neuropsychologia* 41, no. 8: 989–94. Accessed 24 October 2018. ScienceDirect.

Wolfe, Patricia. 2001. *Brain Matters: Translating Research into Classroom Practice.* Alexandria, VA: Association for Supervision and Curriculum Development.

Englanninkielinen musiikki ja ääntäminen

Tämä kysely on osa Turun Yliopiston englannin kielen oppiaineen pro gradu -tutkielmaani varten tehtävää tutkimusta, joka koskee englanninkielisen musiikin mahdollisia vaikutuksia englantia toisena tai vieraana kielenä opiskelevien englannin ääntämiseen. Tutkimus on luottamuksellinen ja anonyymi, mikä tarkoittaa, ettei henkilökohtaisia tietojasi tulla luovuttamaan kenellekään ulkopuoliselle ja vastauksiasi käytetään ainoastaan tutkimustarkoitukseen. Tutkimustuloksia ei myöskään ole mahdollista yhdistää sinuun. Luethan kysymykset huolella sekä vastaat niihin rehellisesti. Happy answering!

TAUSTAKYSYMYKSET

1. Etunimi:

2. Luokka:

○ 9a

○ 9c

3. Englannin kielen oppimäärä:

○ A1

○ A2

4. Ikä:

5. Sukupuoli:

◯ Mies

◯ Nainen

◯ Muu

6. Maa, jossa olen syntynyt:

7. Kansallisuus:

8. Äidinkieli/-kielet:

9. Muut kielet, joita osaan:

_____

_____

_____

_____

_____

10. Kieli/kielet, joita käytän eniten:

11. Viimeisin englannin arvosanani:

12. Ikä, jolloin aloitin englannin opiskelun:

13. Olen viettänyt pidempiä aikoja (= yli kolme viikkoa) jossakin englanninkielisessä maassa:

○ Kyllä

○ Ei

14. Vastasit edelliseen kysymykseen kyllä, joten kerro a) missä, b) miksi ja c) kuinka kauan:

_____

_____

_____

_____

_____

OSA I/IV: Ota kantaa seuraaviin väittämiin asteikolla 1–6. Kysymyksillä halutaan kartoittaa mielipiteitäsi ja omia kokemuksiasi aiheista, joten oikeita ja vääriä vastauksia ei ole olemassa.

1 = täysin eri mieltä
2 = eri mieltä
3 = jokseenkin eri mieltä
4 = jokseenkin samaa mieltä
5 = samaa mieltä
6 = täysin samaa mieltä

15. Englanninkielinen kulttuuri kiinnostaa minua.

	1	2	3	4	5	6	
Täysin eri mieltä	○	○	○	○	○	○	Täysin samaa mieltä

16. Haluan oppia lisää englantia.

	1	2	3	4	5	6	
Täysin eri mieltä	○	○	○	○	○	○	Täysin samaa mieltä

17. Englannin natiivipuhujat (= henkilö, jonka äidinkieli on englanti) ovat yleensä mukavia.

	1	2	3	4	5	6	
Täysin eri mieltä	○	○	○	○	○	○	Täysin samaa mieltä

18. Pidän englannin puhumisesta.

	1	2	3	4	5	6	
Täysin eri mieltä	○	○	○	○	○	○	Täysin samaa mieltä

19. Tunnistan englanninkielisestä puheesta erilaisia murteita ja/tai aksentteja (esim. brittiaksentti, Amerikan aksentti tai Australian aksentti).

	1	2	3	4	5	6	
Täysin eri mieltä	○	○	○	○	○	○	Täysin samaa mieltä

20. On tärkeää kuulostaa mahdollisimman paljon englannin natiivipuhujalta (= henkilö, jonka äidinkieli on englanti).

	1	2	3	4	5	6	
Täysin eri mieltä	○	○	○	○	○	○	Täysin samaa mieltä

21. Olen hyvä ääntämään englantia.

	1	2	3	4	5	6	
Täysin eri mieltä	○	○	○	○	○	○	Täysin samaa mieltä

22. Haluan oppia paremmaksi englannin ääntäjäksi.

	1	2	3	4	5	6	
Täysin eri mieltä	○	○	○	○	○	○	Täysin samaa mieltä

23. Englannin ääntämistä on vaikea oppia, koska sanat lausutaan eri tavalla kuin kirjoitetaan.

	1	2	3	4	5	6	
Täysin eri mieltä	○	○	○	○	○	○	Täysin samaa mieltä

24. Englannin ääntämistä on vaikea oppia, koska englannissa on vaikeita äänteitä.

	1	2	3	4	5	6	
Täysin eri mieltä	○	○	○	○	○	○	Täysin samaa mieltä

25. Englannintunneilla harjoitellaan tarpeeksi ääntämistä.

	1	2	3	4	5	6	
Täysin eri mieltä	○	○	○	○	○	○	Täysin samaa mieltä

26. Välillä minua jännittää englannintunnilla.

	1	2	3	4	5	6	
Täysin eri mieltä	○	○	○	○	○	○	Täysin samaa mieltä

27. Koulutehtävät, joissa on mukana musiikkia, ovat mukavia.

	1	2	3	4	5	6	
Täysin eri mieltä	○	○	○	○	○	○	Täysin samaa mieltä

28. Musiikin kuunteleminen on rentouttavaa.

	1	2	3	4	5	6	
Täysin eri mieltä	◯	◯	◯	◯	◯	◯	Täysin samaa mieltä

29. Musiikki saa minut hyvälle tuulelle.

	1	2	3	4	5	6	
Täysin eri mieltä	◯	◯	◯	◯	◯	◯	Täysin samaa mieltä

30. Laulujen sanoilla on minulle merkitystä.

	1	2	3	4	5	6	
Täysin eri mieltä	◯	◯	◯	◯	◯	◯	Täysin samaa mieltä

31. Opin helposti englanninkielisten laulujen sanoja.

	1	2	3	4	5	6	
Täysin eri mieltä	◯	◯	◯	◯	◯	◯	Täysin samaa mieltä

32. Musiikki auttaa minua käsittelemään tunteitani.

	1	2	3	4	5	6	
Täysin eri mieltä	◯	◯	◯	◯	◯	◯	Täysin samaa mieltä

33. Musiikki on auttanut minua oppimaan englantia.

	1	2	3	4	5	6	
Täysin eri mieltä	◯	◯	◯	◯	◯	◯	Täysin samaa mieltä

34. Musiikki on auttanut minua oppimaan englannin ääntämistä.

	1	2	3	4	5	6	
Täysin eri mieltä	○	○	○	○	○	○	Täysin samaa mieltä

35. Englanninkielisen musiikin kuunteleminen on hyvä tapa oppia englannin ääntämistä.

	1	2	3	4	5	6	
Täysin eri mieltä	○	○	○	○	○	○	Täysin samaa mieltä

36. Englanniksi laulaminen on hyvä tapa harjoitella englannin ääntämistä.

	1	2	3	4	5	6	
Täysin eri mieltä	○	○	○	○	○	○	Täysin samaa mieltä

37. Tykkään laulamisesta.

	1	2	3	4	5	6	
Täysin eri mieltä	○	○	○	○	○	○	Täysin samaa mieltä

38. Osaan laulaa hyvin.

	1	2	3	4	5	6	
Täysin eri mieltä	○	○	○	○	○	○	Täysin samaa mieltä

39. Olen hyvä englannissa.

	1	2	3	4	5	6	
Täysin eri mieltä	○	○	○	○	○	○	Täysin samaa mieltä

OSA II/IV: Valitse sinulle parhaiten sopiva vaihtoehto.

40. Kun kuuntelen englanninkielistä musiikkia, kiinnitän huomiota laulujen sanoihin.

○ En koskaan

○ Harvoin

○ Joskus

○ Usein

○ Aina

41. Kun kuuntelen englanninkielistä musiikkia, kiinnitän huomiota siihen, millä tavalla laulaja lausuu sanoja.

○ En koskaan

○ Harvoin

○ Joskus

○ Usein

○ Aina

42. Kun kuuntelen englanninkielistä musiikkia, erotan erilaisia aksentteja (esim. brittiaksentti, Amerikan aksentti tai Australian aksentti), vaikka en osaisikaan nimetä niitä.

○ En koskaan

○ Harvoin

○ Joskus

○ Usein

○ Aina

43. Kun kuuntelen englanninkielistä musiikkia, laulan mukana.

○ En koskaan

○ Harvoin

○ Joskus

○ Usein

○ Aina

44. Jos englanninkielinen kappale jää soimaan päähäni, lauleskelen sitä myös ilman musiikkia.

○ En koskaan

○ Harvoin

○ Joskus

○ Usein

○ Aina

OSA III/IV: Valitse vaihtoehto, joka parhaiten kuvaa seuraaviin toimintoihin KESKIMÄÄRIN käyttämääsi aikaa vapaa-ajalla.

45. Kuuntelen englanninkielistä musiikkia:

○ 6–7 päivänä viikossa

○ 4–5 päivänä viikossa

○ 2–3 päivänä viikossa

○ Yhtenä päivänä viikossa

○ Muutamana päivänä kuukaudessa

○ Yhtenä päivänä kuukaudessa

○ Harvemmin kuin joka kuukausi

46. Arvioi, kuinka paljon aikaa käytät päivässä englanninkielisen musiikin kuunteluun niinä päivinä kun kuuntelet englanninkielistä musiikkia?

◯ Yli 8 tuntia

◯ 7–8 tuntia

◯ 5–6 tuntia

◯ 3–4 tuntia

◯ 1–2 tuntia

◯ Alle tunnin

47. Laulan englanniksi:

◯ 6–7 päivänä viikossa

◯ 4–5 päivänä viikossa

◯ 2–3 päivänä viikossa

◯ Yhtenä päivänä viikossa

◯ Muutamana päivänä kuukaudessa

◯ Yhtenä päivänä kuukaudessa

◯ Harvemmin kuin joka kuukausi

48. Arvioi, kuinka paljon aikaa käytät päivässä englanniksi laulamiseen niinä päivinä kun laulat englanniksi?

◯ Yli 8 tuntia

◯ 7–8 tuntia

◯ 5–6 tuntia

◯ 3–4 tuntia

◯ 1–2 tuntia

◯ Alle tunnin

49. Katson englanninkielisiä ohjelmia, elokuvia tai videoita tms.:

○ 6–7 päivänä viikossa

○ 4–5 päivänä viikossa

○ 2–3 päivänä viikossa

○ Yhtenä päivänä viikossa

○ Muutamana päivänä kuukaudessa

○ Yhtenä päivänä kuukaudessa

○ Harvemmin kuin joka kuukausi

50. Arvioi, kuinka paljon aikaa käytät päivässä englanninkielisten ohjelmien, elokuvien tai videoiden tms. katseluun niinä päivinä kun katsot niitä?

○ Yli 8 tuntia

○ 7–8 tuntia

○ 5–6 tuntia

○ 3–4 tuntia

○ 1–2 tuntia

○ Alle tunnin

51. Kuuntelen englanninkielisiä podcasteja tai radiota tms.:

○ 6–7 päivänä viikossa

○ 4–5 päivänä viikossa

○ 2–3 päivänä viikossa

○ Yhtenä päivänä viikossa

○ Muutamana päivänä kuukaudessa

○ Yhtenä päivänä kuukaudessa

○ Harvemmin kuin joka kuukausi

52. Arvioi, kuinka paljon aikaa käytät päivässä englanninkielisten podcastien tai radion tms. kuunteluun niinä päivinä kun kuuntelet niitä?

○ Yli 8 tuntia

○ 7–8 tuntia

○ 5–6 tuntia

○ 3–4 tuntia

○ 1–2 tuntia

○ Alle tunnin

53. Pelaan videopelejä englanniksi:

○ 6–7 päivänä viikossa

○ 4–5 päivänä viikossa

○ 2–3 päivänä viikossa

○ Yhtenä päivänä viikossa

○ Muutamana päivänä kuukaudessa

○ Yhtenä päivänä kuukaudessa

○ Harvemmin kuin joka kuukausi

54. Arvioi, kuinka paljon aikaa käytät päivässä videopelien pelaamiseen englanniksi niinä päivinä kun pelaat niitä?

○ Yli 8 tuntia

○ 7–8 tuntia

○ 5–6 tuntia

○ 3–4 tuntia

○ 1–2 tuntia

○ Alle tunnin

55. Luen englanniksi (esim. kirjoja, blogeja, lehtiä, uutisia jne.):

○ 6–7 päivänä viikossa

○ 4–5 päivänä viikossa

○ 2–3 päivänä viikossa

○ Yhtenä päivänä viikossa

○ Muutamana päivänä kuukaudessa

○ Yhtenä päivänä kuukaudessa

○ Harvemmin kuin joka kuukausi

56. Arvioi, kuinka paljon aikaa käytät päivässä englanniksi lukemiseen niinä päivinä kun luet englanniksi?

○ Yli 8 tuntia

○ 7–8 tuntia

○ 5–6 tuntia

○ 3–4 tuntia

○ 1–2 tuntia

○ Alle tunnin

57. Teen vapaa-ajallani jotakin muuta, jossa käytän englannin kieltä:

○ 6–7 päivänä viikossa

○ 4–5 päivänä viikossa

○ 2–3 päivänä viikossa

○ Yhtenä päivänä viikossa

○ Muutamana päivänä kuukaudessa

○ Yhtenä päivänä kuukaudessa

○ Harvemmin kuin joka kuukausi

58. Jos vastasit tekeväsi englanniksi jotakin muuta, kerro mitä:

59. Arvioi, kuinka paljon aikaa käytät päivässä edellä mainitsemaasi tekemiseen niinä päivinä kun teet kyseistä asiaa?

◯ Yli 8 tuntia

◯ 7–8 tuntia

◯ 5–6 tuntia

◯ 3–4 tuntia

◯ 1–2 tuntia

◯ Alle tunnin

OSA IV/IV: Vastaa seuraaviin kysymyksiin.

60. Muistatko, että englannintunnilla olisi joskus käytetty musiikkia?

◯ Kyllä

◯ Ei

61. Minkälaista englanninkielistä musiikkia kuuntelet eniten? Mainitse vähintään kolme artistia/bändiä:

_____

_____

_____

_____

_____

62. Missä kuuntelet musiikkia (esim. koulumatkoilla, kotona, harrastuksessa jne.)?

_____

_____

_____

_____

_____

63. Onko sinulla englanninkielistä idolia tai esikuvaa?

○ Kyllä

○ Ei

64. Vastasit edelliseen kysymykseen kyllä, joten kerro, mistä idolisi/esikuvasi on tunnettu (hän on näyttelijä, urheilija, muusikko, kirjailija, vloggaaja tms.). Voit halutessasi mainita myös tämän nimen:

_____

_____

_____

_____

_____

65. Kun puhut englantia, pyritkö puhumaan jollakin tietyllä aksentilla (esim. brittiaksentti, Amerikan aksentti tai Australian aksentti)?

○ Kyllä

○ Ei

66. Vastasit edelliseen kysymyseen kyllä, joten kerro millä aksentilla pyrit puhumaan:

67. Tunnetko henkilökohtaisesti jonkun englantia äidinkielenään puhuvan, jonka kanssa kommunikoit suullisesti englanniksi?

○ Kyllä

○ Ei

68. Vastasit edelliseen kysymykseen kyllä, joten kerro a) kuinka usein olette tekemisissä b) missä yhteydessä ja c) mistä kyseinen henkilö on kotoisin. Jos tunnet usemman englanninkielisen henkilön kerro siitä, jonka kanssa olet eniten tekemisissä:

_____
_____
_____
_____
_____

69. Pelaatko videopelejä, joissa keskustelet englanninkielisessä voice chatissa?

○ Kyllä

○ Ei

70. Vastasit edelliseen kysymykseen kyllä, joten kerro a) mitä pelejä pelaat ja b) kuinka usein:

_____
_____
_____
_____
_____

71. Haluaisitko sanoa vielä jotakin tieyttyyn kysymykseen tai kyselyn aiheeseen liittyen?

_____
_____
_____
_____
_____

Appendix 2 Questionnaire English translation

English music and pronunciation

This questionnaire is a part of a study conducted for a Master's Thesis at the University of Turku. The study is concerned with the possible effects English music might have on the English pronunciation of English second or foreign language learners. The study is confidential and anonymous, meaning your personal information will not be shared with any outside party, and that your answers will only be used for the purposes of this study. It is also not possible for anyone to trace the results back to you. Please read the questions carefully and answer truthfully. Happy answering!

BACKGROUND QUESTIONS

1. First name: _____

2. Class: 9a / 9c

3. Syllabus: A1 / A2

4. Age: _____

5. Gender: Male / Female/ Other

6. Country I was born in: _____

7. Nationality: _____

8. Native language(s):

9. Other languages I know: _____

10. Language(s) I use the most: _____

11. Last English grade: _____

12. Age I started learning English: _____

13. I've spent longer periods of time (= more than three weeks) in an English-speaking country:

Yes / No

14. If yes, please state a) where, b) why, and c) how long: _____

PART I/IV: Please take a stand on the following statements on a scale from 1–6. These questions are used for mapping your own opinions and experiences so there are no right and wrong answers.

1 = Strongly disagree
2 = Disagree
3 = Slightly disagree
4 = Slightly agree
5 = Agree
6 = Strongly agree

15. English culture interests me.

1   2   3   4   5   6

16. I want to learn more English.

1   2   3   4   5   6

17. Native speakers of English (= person whose mother tongue is English) are usually nice.

1   2   3   4   5   6

18. I like speaking English.

1   2   3   4   5   6

19. I recognize different dialects/accents (eg. British, American or Australian) from English speech.

1   2   3   4   5   6

20. It is important to sound as native-like as possible.

1   2   3   4   5   6

21. I'm good at pronouncing English.

1   2   3   4   5   6

22. I want to become better at pronouncing English.

1   2   3   4   5   6

23. English pronunciation is difficult to learn because words are pronounced differently than they are written.

1   2   3   4   5   6

24. English pronunciation is difficult to learn because English has difficult sounds.

1   2   3   4   5   6

25. Enough pronunciation is being practiced during English classes.

1   2   3   4   5   6

26. Sometimes I feel nervous during an English class.

1   2   3   4   5   6

27. School exercises that include music are nice.

1   2   3   4   5   6

28. Listening to music is relaxing.

1   2   3   4   5   6

29. Music gets me in a good mood.

1   2   3   4   5   6

30. Song lyrics matter to me.

1   2   3   4   5   6

31. I learn English song lyrics easily.

1   2   3   4   5   6

32. Music helps me deal with my feelings.

1   2   3   4   5   6

33. Music has helped me learn English.

1   2   3   4   5   6

34. Music has helped me learn English pronunciation.

1   2   3   4   5   6

35. Listening to English music is a good way to learn English pronunciation.

1   2   3   4   5   6

36. Singing in English is a good way to practice English pronunciation.

1   2   3   4   5   6

37. I like singing.

1   2   3   4   5   6

38. I can sing well.

1   2   3   4   5   6

39. I'm good at English.

1   2   3   4   5   6

PART II/IV: Choose the option that best suits you.

40. When I listen to English music I pay attention to lyrics.

Never / Rarely / Sometimes / Often / Always

41. When I listen to English music I pay attention to how the singer pronounces words.

Never / Rarely / Sometimes / Often / Always

42. When I listen to English music I recognize different accents (eg. British, American, or Australian accent) even if I couldn't name them.

Never / Rarely / Sometimes / Often / Always

43. When I listen to English music I sing along.

Never / Rarely / Sometimes / Often / Always

44. If an English song gets stuck in my head," I sing it even without music.

Never / Rarely / Sometimes / Often / Always

PART III/IV: Choose the option that best describes the time you spend on the following activities during your spare time ON AVERAGE.

45. I listen to English music:

6–7 days a week

4–5 days a week

2–3 days a week

One day a week

Couple of days a month

One day a month

Less than every month

46. Estimate how much time you spend on listening to English music on the days you listen to English music:

Over 8 hours

7–8 hours

5–6 hours

3–4 hours

1–2 hours

Less than an hour

47. I sing in English:

6–7 days a week

4–5 days a week

2–3 days a week

One day a week

Couple of days a month

One day a month

Less than every month

48. Estimate how much time you spend on singing in English on the days you sing in English:

Over 8 hours

7–8 hours

5–6 hours

3–4 hours

1–2 hours

Less than an hour

49. I watch English tv-shows, movies, or videos etc.:

6–7 days a week

4–5 days a week

2–3 days a week

One day a week

Couple of days a month

One day a month

Less than every month

50. Estimate how much time you spend on watching English tv-shows, movies, or videos etc. on the days you watch them:

Over 8 hours

7–8 hours

5–6 hours

3–4 hours

1–2 hours

Less than an hour

51. I listen to English podcasts or radio etc.:

6–7 days a week

4–5 days a week

2–3 days a week

One day a week

Couple of days a month

One day a month

Less than every month

52. Estimate how much time you spend on listening to English podcasts or radio etc. on the days you listen to them:

Over 8 hours

7–8 hours

5–6 hours

3–4 hours

1–2 hours

Less than an hour

53. I play video games in English:

6–7 days a week

4–5 days a week

2–3 days a week

One day a week

Couple of days a month

One day a month

Less than every month

54. Estimate how much time you spend on playing video games in English on the days you play those video games:

Over 8 hours

7–8 hours

5–6 hours

3–4 hours

1–2 hours

Less than an hour

55. I read in English (eg. Books, blogs, magazines, news etc.):

6–7 days a week

4–5 days a week

2–3 days a week

One day a week

Couple of days a month

One day a month

Less than every month

56. Estimate how much time you spend on reading in English on the days you read in English:

Over 8 hours

7–8 hours

5–6 hours

3–4 hours

1–2 hours

Less than an hour

57. I do something else in my spare time where I use English:

6–7 days a week

4–5 days a week

2–3 days a week

One day a week

Couple of days a month

One day a month

Less than every month

58. If you do something else in English, please state what: _____

59. Estimate how much time you spend on the above-mentioned activity on the days you engage in this activity:

Over 8 hours

7–8 hours

5–6 hours

3–4 hours

1–2 hours

Less than an hour

60. Do you remember English music being used in an English class?

Yes / No

61. What kind of English music do you listen to the most? Mention at least three artists/bands: _____

62. Where do you listen to music (eg. on your way to school, at home, at a hobby etc.)? _____

63. Do you have an English-speaking idol or role model?

Yes / No

64. If you answered yes, please state what the person is known for (they are an actor, a musician, an author, a vlogger etc.). You can also mention their name if you like: _____

65. When you speak English, do you aim to speak with a specific accent (eg. British, American, or Australian accent)?

Yes / No

66. I you answered yes, please state which accent you are aiming at:

_____

67. Do you personally know someone whose native language is English and who you communicate with orally in English?

Yes / No

68. If you answered yes, please state a) how often you are in touch, b) in what context, and c) where this person is from. In case you know more than one person, concentrate on the one you are in touch with the most.

_____

69. Do you play video games where you communicate in English via a voice chat?

Yes / No

70. If you answered yes, please state a) which games you play, and b) how often:

_____

71. Is there anything else you would like to say regarding a certain question or topic?

_____

Appendix 3 Picture description task

Spoken part

*First, you have **one minute** to read the sentences and to take a look at the picture.*

Once the recorder is on...

1. Say your name

2. Read out loud

Please read the following sentences out loud. Remember to speak loudly enough!

a) My favourite thing about Christmas is getting to eat a lot of food.

b) I like chocolate and gingerbread cookies the most.

c) I've never been a huge fan of rice pudding – nor has my sister.

d) This year I'm wishing for a new pair of boots because my old ones are broken.

e) My father would like to get me a jacket he saw on television the other day.

3. Describe the picture

*Please describe in **full sentences** what is happening / what you can see in the picture. You may speak freely – there are no right and wrong answers. You have **30 seconds**, try to keep talking the whole time. It doesn't matter what you say as long as you say something!*

Examples of how to start:

In this picture...

There is...

I can see...
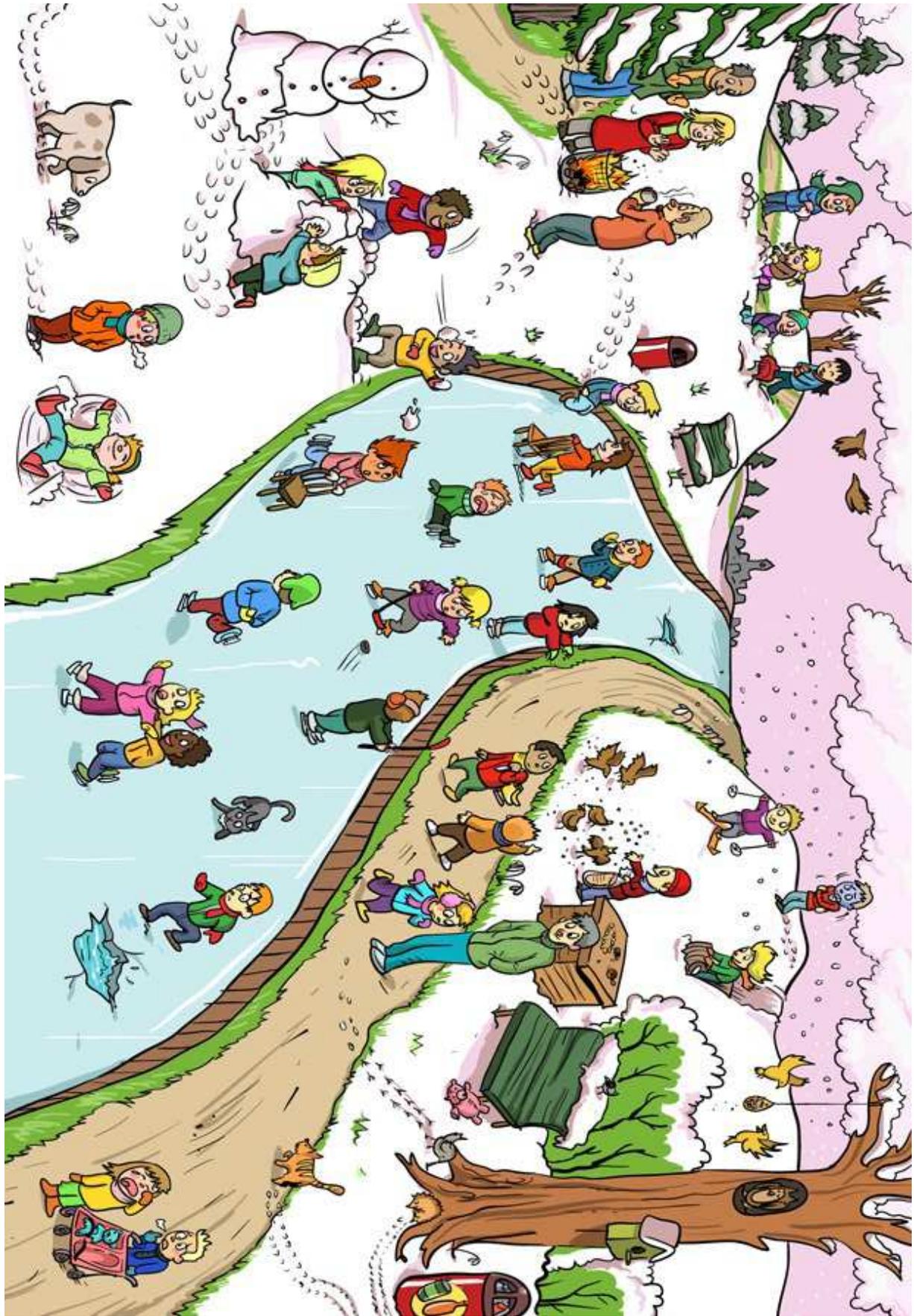
Picture used in the PDT, Aimée de Jongh (2008).

Appendix 4 Native-speaker evaluation

Spoken English Evaluation

Please read the instructions carefully.

Your task is to listen to the following speech samples by 45 non-native speakers of English and rate their PRONUNCIATION. Try to ignore all the other aspects of speech, e.g. ungrammatical utterances or odd word choices. Each sample is approximately 30 seconds long, and you are allowed to listen to each of them as many times as you wish. However, you don't have to listen to them more than once.

You will be evaluating three pronunciation criteria on the scales provided. It is extremely important you familiarize yourself with these criteria since you will be needing them throughout the questionnaire!

1. ACCURACY = How well the language is spoken in relation to its pronunciation rules, i.e. are English sounds and words pronounced correctly.

2. FLUENCY = How easily and naturally the speech flows without too many hesitations, pauses, and self-repetitions.

3. INTELLIGIBILITY = How understandable the speaker is, i.e. do they get their message across.

Before getting started:
- Make sure you're sitting in a comfortable position in a relatively quiet place.
- Don't do anything else at the same time, not even chew gum.
- Use headphones to maintain the quality of the audio.
- There will be three breaks, and it is important you take them as advised.
- Remember to click "Submit" after finishing the questionnaire!

This study is anonymous, meaning your name will not be published anywhere. Also, remember that the data is confidential, and it is strictly prohibited to play it to any outside party. If you have any questions or doubts at this point, please contact me.

Otherwise, happy evaluating!

Name of evaluator: *

Short-answer text

Nationality of evaluator: *

Short-answer text

Age of evaluator: *

Short-answer text

Do you study or have you studied linguistics at a university level? *

○ Yes

○ No

Speaker 1

Description (optional)

Untitled video

1a. How accurate is the speaker's English pronunciation? *

	1	2	3	4	5	6	
Not at all accurate	○	○	○	○	○	○	Extremely accurate

1b. How fluent is the speaker's English? *

	1	2	3	4	5	6	
Not at all fluent	○	○	○	○	○	○	Extremely fluent

1c. How intelligible is the speaker's pronunciation? *

	1	2	3	4	5	6	
Not at all intelligible	○	○	○	○	○	○	Extremely intelligible

(There were 44 pages following this template, titled Speaker 1, Speaker 2, Speaker 3 etc.)

Questions for evaluators

1. As a native speaker of English, what do you find the most distracting about non-native speakers' pronunciation (also outside of this questionnaire)?

Long-answer text

2. How often do you hear non-native English on average? *

○ Daily or almost daily

○ Weekly

○ Monthly

○ Less than monthly

○ Never

3. How did you find this task? Was it easy or difficult to evaluate the speakers?

Long-answer text

4. Is there something else regarding this questionnaire you would like to comment on?

Long-answer text

Appendix 5 Artists and bands mentioned by participants

Artist/band	Times mentioned
ASAP Rocky	1
Abba	1
Adele	1
Arctic Monkeys	1
Ariana Grande	3
Avicii	1
Beyoncé	2
Billie Eilish	2
Bloodbath	1
BONES	2
Boney M.	1
Cardi B	1
Charlie Baker	1
Chris Brown	1
CJ Fly	1
Dean Lewis	1
Dr. Dre	1
Drake	4
Dua Lipa	3
Ed Sheeran	4
Eminem	6
Five Finger Death Punch	1
Frankie Cosmos	1
G-Unit	1
Gabrielle Smith	1
George Ezra	1
Gorillaz	1
Grandson	1
Green Day	1
Ice Cube	1
Imagine Dragons	2
Jaden Smith	3
Joey Bada$$	2
Juice Wrld	2
Justin Bieber	1
K.A.A.N.	1
Kanye West	2
Lady Gaga	1
Led Zeppelin	1
Lil Mosey	1
Lil Peep	3
Lil Uzi Vert	1

Lil Xan	1
Logic	2
Machine Gun Kelly	1
Maroon 5	2
Melanie Martinez	1
Migos	1
My Chemical Romance	1
NF	2
Nicki Minaj	1
Nirvana	1
One Republic	1
Papa Roach	1
Pink	1
Post Malone	1
Queen	1
Rag'n'Bone Man	1
Rihanna	1
Risa Against	1
Rita Ora	1
Say Hi	1
Shakira	1
Shawn Mendez	3
Snoop Dogg	1
The Game	1
The Score	1
The Vamps	1
The Weeknd	1
Tom Walker	1
Travis Scott	5
Trippie Redd	1
Tyga	1
Tyler, The Creator	1
Vashti Bunyan	1
Wiz Khalifa	1
XXXTentacion	4
Zara Larsson	2

Johdanto

Musiikki on olennainen osa nykyihmisen päivittäistä äänimaailmaa. IFPI:n (2018) viimeisimmän kansainvälisen musiikinkuluttajaraportin mukaan ihmiset kuuntelevat musiikkia keskimäärin kaksi ja puoli tuntia päivässä, ja musiikkia saatetaan kuunnella mihin tahansa aikaan päivästä. Huomattava osa ihmisten kuuntelemasta musiikista on englanninkielistä – englanninkielinen pop ja englanninkielinen hip hop olivat viime vuoden suosituimmat musiikkilajit myös 13–18-vuotiaiden suomalaisnuorten keskuudessa (IFPI ja Teosto 2018). Tällä saattaa olla yllättäviä käytännön vaikutuksia, joista musiikkia huvin vuoksi kuunteleva ihminen ei ole lainkaan tietoinen.

Ääntäminen on englantia toisena tai vieraana kielenä puhuvalle tärkeä taito. Ensisijaisesti on tultava ymmärretyksi, mutta kuulijan on myös voitava suhtautua puhujaan tasavertaisena keskustelukumppanina, mitä vahva vieraskielinen aksentti saattaa vaikeuttaa. Tutkimuksia musiikin vaikutuksesta kielen ja ääntämisen oppimiseen on viime vuosikymmenten aikana tehty useita, ja niiden tulokset ovat osoittaneet, että etenkin laulaminen ja ääntäminen ovat yhteydessä toisiinsa (Alisaari ja Heikkola 2017; Chen 2016; Christiner ja Reiterer 2013), mutta myös pelkän vieraan kielen äänteiden kuuntelemisen avulla on onnistuttu parantamaan ääntämistä (Peltola et al. 2017). Aiemmissa tutkimuksissa on kuitenkin oltu lähinnä kiinnostuneita tietoisen ja tarkoituksenmukaisen harjoittelun vaikutuksista vieraan kielen ääntämisen oppimiseen. Tämän tutkimuksen lähtökohtana oli sen sijaan selvittää, voisiko englanninkielisellä musiikilla, jota ihmiset kuuntelevat ja laulavat arjessaan, olla samankaltaisia vaikutuksia heidän ääntämiseensä kuin tarkoituksenmukaisella harjoittelulla. Kysymykset, joihin tutkimus pyrki vastamaan, olivat seuraavat:

1. Millä tavoin englanninkielisen musiikin kuunteleminen ja laulaminen vapaa-ajalla voisi selittää eroja osallistujien ääntämiskokeesta saaduissa pisteissä?
2. Missä määrin ääntämisen tarkkuus, sujuvuus ja ymmärrettävyys ovat yhteydessä englanninkielisen musiikin kuuntelemiseen ja englanniksi laulamiseen?
3. Mitkä epäsuorat tekijät saattaisivat tehdä englanninkielisestä musiikista hyödyllistä ääntämisen oppimisen kannalta?

Tutkimuksen hypoteesina oli, että englanninkielisen musiikin kuunteleminen ja englanniksi laulaminen vaikuttaisivat englannin oppijan ääntämiseen positiivisesti, koska kyseiset

aktiviteetit muistuttavat luonnollista kielen oppimista (ks. de Carvalho et al 2018, 18; Partanen et al 2013; Moon, Lagercrantz ja Kuhl 2013, Krashen 1977) sekä tarjoavat useita hyödyllisiä työkaluja englannin ääntämisen oppimista ajatellen (ks. Heikkola 2018; Alisaari ja Heikkola 2017; Chen 2016; Ludke, Ferreira ja Overy 2014; Aguirre, Bustinza ja Garvich 2012; Sigurðardóttir 2011; Iverson, Hazan ja Bannister, 2005; Wolfe 2001, 162; Celce-Murcia, Brinton ja Goodwin 1996, 8–9, Krashen 1977). Kysymyksiä lähestyttiin kvantitatiivisen tutkimuksen metodein.

Teoreettinen tausta

Englanninkielisellä musiikilla tarkoitettiin tutkimuksessa kaikkea musiikkia, jossa lauletaan englanniksi riippumatta siitä, puhuuko laulaja äidinkielenään englantia. Vapaa-ajalla viitattiin aikaan, jolloin henkilö ei ole töissä tai koulussa, joskin musiikin kuunteleminen näissäkin paikoissa voitiin tietyissä tilanteissa laskea vapaa-ajan kunteluksi. Ääntämisellä tarkoitetaan sitä, miten puhuja lausuu sanoja ja sanaketjuja. Ääntäminen on mahdollista jakaa kahteen tasoon riippuen siitä, kuinka läheisesti sitä halutaan tarkastella (Lintunen 2014, 165). Segmentaalinen koostuu kielen pienimmistä osista, jotka muuttavat sanojen merkitystä. Jos esimerkiksi sanan "tikka" ensimmäinen äänne, /t/, muutetaan äänteeksi /h/, sanan merkitys muuttuu puuta hakkaavasta linnusta välillä kiusalliseksikin vaivaksi. Suprasegmentaaliseen tasoon puolestaan kuuluvat laajemmat kokonaisuudet, kuten kielen rytmi, paino ja intonaatio (Derwing ja Munro 2015, 5). Englannin ja suomen kielen suprasegmentaaliset piirteet ovat erilaisia, mikä saattaa ajoittain tuottaa ongelmia suomalaiselle englannin oppijalle. Suomen kielessä esimerkiksi sanapaino on normaalisti sanan ensimmäisellä tavulla (Suomi, Toivanen ja Ylitalo 2008, 75), kun taas englannissa painotus vaihtelee sana- ja lausekohtaisesti siten, että jotkin sanat jäävät kokonaan ilman painoa (Pennington 1996, 135).

Krashenin (1977) mukaan niin kutsuttu ymmärrettävä kielisyöte on kielenoppimisen tärkein yksittäinen lähde. Ymmärrettävällä kielisyötteellä tarkoitetaan kieltä, joka on hieman oppijan taitotason yläpuolella, mutta mahdollista ymmärtää esimerkiksi kontekstin avulla. Musiikin voidaan usein katsoa tarjoavan kielenoppijalle tämänkaltaista kielisyötettä. Kielisyötehypoteesi perustuu olettamukseen, että äidinkielen ja toisen kielen oppimisprosessit ovat monilta osin samankaltaisia. Lapset oppivat äidinkielensä äänteet tiedostamatta ensin kuuntelemalla muiden puhetta ja myöhemmin toistamalla kuulemaansa (ks. de Carvalho et al. 2018, 18; Moon, Lagercrantz ja Kuhl 2013; Partanen 2013). Tutkimus on osoittanut, että myös aikuiset kykenevät oppimaan äänteitä pelkästään niitä kuuntelemalla (Peltola et al. 2017), minkä uskotaan perustuvan puheen tuottamiseen liittyvien motoristen aivoalueiden

aktivaatioon puheen kuuntelemisen seurauksena. Tämän aktivaation ovat onnistuneet osoittamaan muun muassa Watkins, Strafella ja Paus (2003) sekä Fadiga et al. (2002). Huomiota on kuitenkin kiinnitettävä kielisyötteeseen; on kuunneltava eikä pelkästään kuultava. Schmidtin (1993) mukaan tämä on oppimisen edellytys. Myös Libermanin *Motor Theory* (1985) perustuu ajatukseen, että kuulijat tulkitsevat puhetta sen äänteiden tuottamiseen käytettyjen lihasten liikkeiden perusteella itse sitä tiedostamattaan. Edellä mainitut tutkimustulokset viittaavat siihen, että jo pelkkä englannin kielen kuunteleminen saattaa olla hyödyllistä sen äänteiden oppimisen kannalta. Tätä tukee myös havaintoassimilaatiomalli (Best 1995), joka korostaa kuuntelun tärkeyttä ääntämisen oppimiselle. Mallin mukaan kielenoppija kuuntelee uutta kieltä aluksi äidinkielensä foneemijärjestelmän kautta: jos jotakin äännettä ei ole oppijan omassa äidinkielessä, tulkitaan se siksi äidinkielen äänteeksi, jota se eniten muistuttaa. Esimerkiksi englannin kielen äänne /ɪ/ (sanassa ”bit”) saatetaan aluksi tulkita virheellisesti suomen kielen /i/-äänteeksi. Vaikka kielenoppimiseen liittyy sensitiivisiä kausia eli kausia, jolloin yksilö on erityisen herkkä tietynlaisille ärsykkeille (Long 2013, 4), myös aikuisen aivot muovautuvat oppimiskokemusten seurauksena. Tämän ovat osoittaneet muun muassa Barbeau et al. (2017). Kyse on aivojen plastisuudesta, jolla tarkoitetaan niiden kykyä muovautua yksilön kokemusten perusteella läpi elämän (Møller 2009, 27).

Koska ääntäminen on pitkälti motorinen taito, opittavan kielen kuunteleminen yksin ei riitä sen oppimiseen. Kielenoppijan tarvitsee sekä kuulla kyseistä kieltä (*input*) että tuottaa sitä itse (*output*) (Lintunen 2014, 175). Puhetta tuotetaan säätelemällä ilman kulkua äänihuulten läpi sekä suun ja kielten lihasten avulla, mikä vaatii puhujalta äärimmäisen tarkkaa hienomotoriikkaa (Pennington 1996, 20–22). Lisäksi kielen ääntämisessä ei ole kyse pelkästään yksittäisten äänteiden tai sanojen ääntämisestä, vaan pitkälti sanaketjujen tuottamisesta. Naapuriäänteet vaikuttavat yksittäisten äänteiden toteutumiseen, ja luonnollisessa englanninkielisessä puheessa esiintyykin useita ilmiöitä, joiden seurauksena äänteet saattavat muuttua, jotain saatetaan lisätä tai jopa poistaa kokonaan (ks. Celce-Murcia, Brinton ja Goodwin 2010, 163–75; Gómez González ja Sánchez Roura 2016, 234–58). Laulettu kieli muistuttaa puhuttua siinä määrin, että samat ilmiöt kuuluvat myös englanninkielisessä musiikissa. Ääntämistä voidaan harjoitella esimerkiksi kielivoimistelulorujen (*tongue twisters*), minimiparien sekä kuuntele ja toista -harjoitusten avulla (Celce-Murcia, Brinton ja Goodwin 1996, 8–9). Kielivoimistelulorut ovat haastavia äänne- tai sanakokonaisuuksia, joita toistetaan nopeasti. Minimiparit puolestaan ovat sanapareja, jotka eroavat toisistaan vain yhden äänteen verran ja niiden avulla voidaan harjoitella äänteiden välisiä kontrasteja, kuten esimerkiksi sanojen ”ship” (/ʃɪp/) ja ”sheep” /ʃiːp/) erottamista toisistaan. Kuuntele ja toista -harjoituksissa

kielen oppijat kuuntelevat sanoja tai lauseita ja pyrkivät toistamaan ne mahdollisimman tarkasti. Tarkemmin ajateltuna englanniksi laulaminen muistuttaa huomattavasti näitä ääntämisharjoituksia. Puhetta ja laulua tuotetaankin pohjimmiltaan samalla tavalla (Christiner ja Reiterer 2013).

Englannin oppijan tavoitteena saattaa olla hyväksi englannin ääntäjäksi tuleminen. Kaiken kaikkiaan "hyvän ääntämisen" määritteleminen on kuitenkin haasteellista. Puhujan on ensisijaisesti tultava ymmärretyksi, mikä on nykyään lähtökohtana myös englannin opetuksessa. Tämä noudattaa niin kutsuttua ymmärrettävyysperiaatetta (Levis 2005, 370). Pelkkä ymmärretyksi tuleminen ei kuitenkaan aina riitä – joskus puhujaan saatetaan suhtautua negatiivisesti pelkästään tämän vahvan vieraskielisen aksentin ja/tai vaikeasti ymmärrettävän puheen takia (Derwing ja Munro 2015, 2). Natiiviusperiaatteen mukaan ääntämisen tarkkuus ja virheettömyys on puolestaan ensisijaisen tärkeää (Levis 2005, 370). Realistinen tavoite saattaisi kuitenkin löytyä näiden kahden periaatteen välimaastosta, sillä harva saavuttaa natiivipuhujan kaltaista ääntämistä (Celce-Murcia, Brinton ja Goodwin 1996, 8). Pennington (1996, 220) kehitti kolme hyvän ääntämisen kriteeriä helpottamaan ääntämisen arviointia. Kriteerit ovat ääntämisen tarkkuus (*accuracy*), sujuvuus (*fluency*) ja ymmärrettävyys (*intelligibility*). Ääntämisen tarkkuudella tarkoitetaan sitä, kuinka hyvin kielenoppija ääntää kieltä suhteessa sen sääntöihin (Skehan 1996, 23) eli ääntääkö tämä kieltä niin sanotusti oikein. Ääntämisen sujuvuudella taas viitataan puheen yleiseen soljuvuuteen (Derwing 2018, 321) ja sujuvuutta käytetään joskus myös synonyymina hyvälle ääntämiselle (Pietilä ja Lintunen 2014, 22). Ymmärrettävyydellä puolestaan viitataan siihen, kuinka hyvin kuulija ymmärtää puhujan viestin (Derwing ja Munro 2015, 5). Kouluissa ääntämistä ei aina harjoitella tarpeeksi (Tergujeff 2013), ja esimerkiksi Suomessa kielioppi ja sanasto saavat usein osakseen ääntämistä enemmän huomiota. Lordin (2008, 364) mukaan useat kielenopettajat olettavat oppilaiden oppivan kielen ääntämisen itsenäisesti. Englannin ääntämisen kannalta olisi hyödyllistä tutustuttaa oppija sanojen foneettiseen muotoon ennen kirjoitettua muotoa, sillä kirjoitusasun on todettu johtavan kielenoppijaa harhaan (Peltola et al. 2015).

Ihmiset ovat tottuneet pitämään musiikkia pelkästään kulttuurisena tai taiteellisena ilmiönä, mutta musiikin prosessoinnin mekanismit ovat yhteydessä useisiin muihin aivojen perustoimintoihin, kuten tunteisiin, muistiin ja kieleen (Wolfe 2001, 160–161). Kun ihminen kuulee musiikkia, aivojen serotoniinin tuotanto lisääntyy, mikä muun muassa vähentää koettua stressiä (Ortis 2008, 216). Stressi haittaa uusien asioiden omaksumista; se on niin kutsuttu affektiivinen filtteri (Krashen 1982), joilla tarkoitetaan oppimista häiritsevää psykologista estettä. Musiikkia kuunnellessa mieli on avoinna uuden oppimiselle, ja musiikin onkin todettu

helpottavan kielten ja ääntämisen oppimista (Chen 2016; Ludke, Ferreira ja Overy 2014; Wolfe 2001, 162; Christiner ja Reiterer 2013; Aguirre, Bustinza ja Garvich 2016; Sigurðardóttir 2011). Musiikin kuunteluun vapaa-ajalla liittyy lisäksi kuunneltavan musiikin omavalintaisuus, jolla saattaa myös olla vaikutusta oppimiseen. Nuoret ovat kiinnostuneita artisteista, joihin voivat samaistua (IFPI ja Teosto 2018), ja muusikot toimivat roolimalleina ja idoleina nuorille vaikuttaen heihin sosiaalisen oppimisen kautta (Raviv et al. 1996, 632). On mahdollista, että englanninkielinen idoli motivoi englannin oppijaa harjoittelemaan englannin kieltä. Motivaatio on vahva ennustava muuttuja kielen oppimisen kannalta (Masgoret ja Gardner 2003) ja sen avulla voidaan pitkälti selittää, miksi toiset kielenoppijat ääntävät kieltä paremmin kuin toiset (Bongaerts et al. 1997; Moyer (1999). Myös Dörnyein (2009, 217–218) ihanteellisen minän (*ideal self*) käsite liittyy idoleihin ja roolimalleihin: ihanteelliseen minään voi liittyä piirteitä, joita idolilla on. Englanninkielinen idoli saattaa myös parantaa oppijan englanninpuhujia koskevia asenteita, mikä omalta osaltaan tukee paremmaksi kielen puhujaksi tulemista (Singleton 2014, 92; Celce-Murcia et al. 2010).

Aineisto ja menetelmät

Tutkimuksessa tarkasteltiin vapaa-ajalla kuunnellun ja lauletun englanninkielisen musiikin mahdollisia vaikutuksia englannin oppijoiden englannin ääntämiseen. Tutkimus toteutettiin turkulaisessa yläkoulussa syksyllä 2018 ja siihen osallistui yhteensä 44 vapaaehtoista yhdeksäsluokkalaista. Tutkimukseen kuului kaksi osaa: suomenkielinen kysely ja englanniksi toteutettu kuvailutehtävä. Kysely koostui sekä monivalintakysymyksistä että tarkentavista avoimista kysymyksistä, ja osallistujat saivat vastata siihen omilla puhelimillaan. Kysymykset koskivat osallistujien kielitaustaa, heidän ajatuksiaan ja tuntemuksiaan englannin kielestä, sen puhujista ja sen oppimisesta, sekä osallistujien musiikinkuuntelu- ja laulutottumuksia. Suomeksi kysely löytyy kokonaisuudessaan Liitteestä 1. Kuvailutehtävässä osallistujat lukivat ensin ääneen viisi englanninkielistä lausetta lämmitelläkseen, minkä jälkeen heidän tuli puhua kolmenkymmenen sekunnin ajan piirretystä kuvasta (ks. Liite 3). Tehtävä nauhoitettiin, ja puhenäytteet arvioi myöhemmin yhdeksänhenkinen englannin natiivipuhujista koostuva raati kolmen yksittäisen kriteerin perusteella. Arviointikriteerit olivat ääntämisen tarkkuus, sujuvuus ja ymmärrettävyys, ja ne perustuivat Penningtonin (1996, 220) "hyvän ääntämisen" kriteereihin. Arviointi tapahtui Likert-asteikolla yhdestä kuuteen, ja arvioista koostettiin kullekin osallistujalle sekä ääntämisen kokonaispisteet että yksittäisiin kriteereihin perustuvat pisteet laskemalla yhteen raadin tekemät Likert-arviot. Jokainen osallistuja sai myös oman niin kutsutun musiikki-indeksin, joka muodostettiin yhdistämällä neljän yksittäisen kysymyksen

vastaukset. Nämä kysymykset koskivat osallistujien englanninkielisen musiikin kuuntelua ja laulamista vapaa-ajalla (ks. Liite 1, kysymykset 45–48). Aineiston analysointiin käytettiin Pearsonin korrelaatiokerrointa, jonka avulla tarkasteltiin osallistujien ääntämispisteiden ja englanninkielisen musiikin parissa vietetyn ajan suhdetta, sekä Mann-Whitney U -testiä, jolla vertailtiin toisiinsa osallistujista koostettuja ryhmiä, joista toinen vietti vähemmän ja toinen enemmän aikaa englanninkielisen musiikin parissa.

Tutkimuksen tulokset

Tutkimuksen tulokset osoittivat, että englanninkielisen musiikin kuunteleminen ja laulaminen vapaa-ajalla saattaa olla hyödyllistä englannin ääntämisen oppisen kannalta. Yhdessä kyseiset aktiviteetit olivat merkitsevästi yhteydessä parempaan ääntämiseen (r = -0.379, p = 0.011). Enemmän aikaa englanninkielisen musiikin parissa viettävien ääntämispisteet olivat keskimäärin korkeammat kuin vähemmän aikaa englanninkielisen musiikin parissa viettävillä, mutta ero ei ollut tilastollisesti merkitsevä. Tulokset olivat samansuuntaisia kun kuuntelemista ja laulamista tarkasteltiin erikseen. On mahdollista, että kyseiset tulokset johtuivat esimerkiksi otoksen pienestä koosta. Korrelaatioiden kannalta etenkin englanniksi laulaminen vaikutti olevan yhteydessä parempaan englannin ääntämiseen, mikä tukee aikaisempien tutkimusten tuloksia. Laulamiseen käytetyn ajan ja ääntämiskokeesta saatujen pisteiden välinen korrelaatio oli tilastollisesti merkitsevä (r = -0.331, p = 0.028). Yhteys oli merkitsevä myös musiikin mukana laulamisen ja hyvän ääntämisen välillä (r = 0.315, p = 0.037), ja erityisen vahva yhteys löytyi ilman musiikkia laulamisen ja hyvän ääntämisen väliltä (r = 0.519, p<0.001). Myös englanninkielisen musiikin kuuntelemisen ja hyvän ääntämisen välillä vaikutti olevan jonkinlainen tendenssi, joskaan korrelaatio ei ollut aivan tilastollisesti merkitsevä (r = -0.292, p = 0.054). Kuuntelemisen laatu vaikutti tuloksiin siten, että huomion kiinnittäminen laulun sanoihin oli merkitsevästi yhteydessä parempaan ääntämiseen (r = 0.298, p = 0.049). Sen sijaan huomion kiinnittämisellä itse laulun sanojen ääntämiseen ei ollut yhteyttä ääntämisen hyvyyteen. Englanninkielisen musiikin parissa vietetty aika oli merkitsevästi yhteydessä myös yksittäisiin ääntämisen osa-alueisiin, tarkkuuteen, sujuvuuteen ja ymmärrettävyyteen. Yhteys oli vahvin ääntämisen sujuvuuteen (r = -0.366, p = 0.015) ja heikoin tarkkuuteen (r = -0.351, p = 0.019), mutta erot osa-alueiden välillä olivat marginaalisia.

Tulokset osoittivat myös, että englannin oppijat kokivat musiikin vaikuttavan positiivisesti heidän mielialaansa, minkä ansiosta mieli on oppimiselle otollisessa tilassa. Valtaosa tutkimuksen osallistujista koki englanninkielisen musiikin auttaneen heitä oppimaan englantia (88,6%) ja erityisesti englannin ääntämistä (90,9%). Koska musiikkiin liittyy vahvasti

myös idolien käsite, tutkimuksessa oltiin kiinnostuneita englannin oppijoiden englanninkielisistä esikuvista. Tulosten mukaan ne, joilla oli englanninkielinen esikuva, vaikuttivat ääntävän englantia paremmin kuin ne, joilla sellaista ei ollut, joskaan tulos ei ollut tilastollisesti merkittävä. Tulos saattaa selittyä osittain esimerkiksi sillä, että ne osallistujat, joilla oli englanninkielinen esikuva, pitivät englantia äidinkielenään puhuvia mukavampina kuin ne, joilla ei ollut englanninkielistä esikuvaa. Tämä mahdollisesti vaikuttaa asenteisiin englannin kieltä kohtaan ja sitä kautta myös englannin oppimismotivaatioon. Aihe vaatii kuitenkin lisää tarkastelua.

Lopuksi

Pähkinänkuoressa tämän tutkimuksen pohjalta voidaan sanoa, että englanninkielisen musiikin kulutus arjessa ja vapaa-ajalla saattaa vaikuttaa englannin ääntämisen oppimiseen sekä suorasti että epäsuorasti. Englanninkielisen musiikin kuunteleminen ja englanniksi laulaminen huvin vuoksi voivat olla hyödyllisiä aktiviteetteja englannin ääntämisen oppimisen kannalta suorasti, koska ne tarjoavat oppijalle luonnollista kielisyötettä sekä monipuolisia mahdollisuuksia ääntämisen tahattomaan harjoitteluun. Musiikilla itsellään on vaikutusta ihmisen mielialaan; se muun muassa rentouttaa ja saa paremmalle tuulelle. Tämä puolestaan mahdollistaa sen, että mieli on avoinna uusien asioiden oppimiselle, millä saattaa olla vaikutusta ääntämisen oppimiseen epäsuorasti. Englanninkieliset esikuvat ja sitä kautta suhtautuminen englantia äidinkielenään puhuviin saattaa myös vaikuttaa englannin oppimiseen. Näiden tulosten perusteella voidaan todeta, että vaikka syy- ja seuraussuhteiden esittäminen tämänkaltaisissa tutkimuksissa onkin haasteellista, englanninkielinen musiikki on todennäköisesti ainakin yksi englannin oppijoiden ääntämiseen vaikuttavista tekijöistä. Toivon mukaan tieto siitä, että huvin vuoksi tapahtuvalla englanninkielisen musiikin kuuntelulla ja laulamisella saattaa olla pysyviä kielitaidollisia vaikutuksia voisi kannustaa ihmisiä viettämään entistä enemmän aikaan kyseisten aktiviteettien parissa, etenkin jos haluaa tulla paremmaksi englannin ääntäjäksi. Myös opettajat voisivat kannustaa oppilaitaan kiinnittämään huomiota englanninkielisten laulujen sanoihin myös vapaa-ajalla sekä laulamaan lempikappaleitaan – vaikka yksin kotona.