

**FINNISH LISTENERS' PERCEPTIONS OF OTHER
NON-NATIVE ENGLISH SPEAKERS' ACCENTS**

Lily Obeda

Master's Thesis

Applied Language Studies for the Changing Society

Department of Languages and Communications

University of Jyväskylä

November 2020

UNIVERSITY OF JYVÄSKYLÄ

Faculty Humanities and Social Sciences	Department Languages & Communications
Author Lily Obeda	
Title Finnish Listeners' Perceptions of Other Non-Native English Speakers' Accents	
Subject Applied Language Studies for the Changing Society	Level Master
Month and year December 2020	Number of pages 64 (including appendices)
Abstract <p>As the role of English language broadens in the world, it is important to note how different users of English perceive their language peers in second language settings, and to understand what factors go into those ratings. While much research has focused on native speakers rating non-native speakers, an increasing amount of research is being directed toward non-native listeners' assessments of non-native speech.</p> <p>This thesis examines how Finnish listeners assess the English language use of their peers from a variety of different language backgrounds, and tries to identify possible factors that contribute to how Finns perceive a "good" accent in English. A survey looking at comprehensibility and accentedness ratings in relation to valence is the basis of this work, which comprises both qualitative and quantitative questions.</p> <p>The study finds that there are some moderate effects for both accentedness and comprehensibility in relation to valence, with no strong patterns for native language background. Additionally, the study finds that Finnish listeners give better ratings in valence to Finnish speakers than to other European speakers.</p>	
Keywords perception, accent, accentedness, ESL, comprehensibility, valence	
Depository University of Jyväskylä	
Additional information	

TABLE OF CONTENTS

TABLE OF CONTENTS	2
1. INTRODUCTION	4
2. BACKGROUND	6
2.1 Learning English as a Second Language	6
2.2 English in Finland	10
3. CENTRAL CONCEPTS	12
3.1 Accent	12
3.2 Foreign accent	13
3.3 Intelligibility, comprehensibility, and accentedness	15
3.4 Valence	16
3.5 Mutual intelligibility benefit and own accent preference	17
4. PREVIOUS RESEARCH	17
4.1 Accent perception studies	17
4.1.1 NS ratings of NNSs	18
4.1.2 NNS ratings of NNSs	21
5. PRESENT STUDY	23
5.1 Research questions	23
5.2 Survey	24
5.2.1 Speaker Files	25
5.2.2 Survey participants	27
6. RESULTS	27

6.1 Qualitative questions	27
6.1.1 What does a good accent sound like?	29
6.1.2 What does a bad accent sound like?	30
6.1.3 What makes an accent easy to understand?	32
6.2 Quantitative data	34
6.2.1 Interrater reliability	34
6.3 Accentedness, comprehensibility, and valence	35
6.4 Own accent preference	40
7. DISCUSSION	45
7.1 Qualitative data	45
7.1.1 “Good” accents	45
7.1.2 “Bad” accents	47
7.2 Comprehensibility	49
7.3 Quantitative data	51
8. CONCLUSION	52
9. APPENDICES	55
9.1 Survey questions	55
9.2 Table 4 - Speaker file data	56
9.3 Table 5 - Responses to qualitative questions	57
10. REFERENCES	60

1. INTRODUCTION

You say [iːðəʃ], I say [aɪðəʃ], you say [niːðəʃ], I say [naɪðəʃ]; Let's call the whole thing off!
(Gershwin & Gershwin, 1937)

Accent matters. The above song lyric from a classic American song highlights the divisive nature of language varieties: Small differences in pronunciation can affect listeners' perceptions of speech, from single phoneme differences as above up to much more complex suprasegmental features of language. Pronunciation and accent are very salient in oral language, While both interlocutors in the example above are native English speakers, speaker-listener pairs with at least one non-native interlocutor face similar perceptual judgements.

The global influence of English is growing (Eberhard, Simons & Fennig, 2019), and as it does, so too does the number of English speakers worldwide. Broadly, the field of sociolinguistics has long been interested in the ways that people use language to interact with each other, and one manifestation of that is through accented speech research. There have been a number of ways that sociolinguists have addressed issues of perception, including folk linguistics, qualitative socioculturally-rooted analysis, and more quantitative laboratory methods. In many subclades, perceptual judgements of this accented speech have historically come from native speakers. Much research suggests that these native speakers tend to rate accented, non-native speech with less positive ratings on a variety of social features when compared to native speakers (see Gluszek & Dovidio, 2010 for a review). However, in more recent years, there has been increasing interest in how non-native listeners rate non-native speech. English increasingly continues to be spoken in lingua franca situations, and non-native listener-speaker pairs are increasingly common. It is important for a thorough understanding of these communication practices, their norms, and the variation that comes with different linguistic and cultural backgrounds to include non-native listeners rating non-native speakers in accent studies.

As it continues to expand, this field has explored a variety of language backgrounds as speakers and listeners, and has noted that there is variation among different first-language (L1) groups in their ratings of speech (Foote & Trofimovich, 2016). Additionally, some research suggests there may be an effect for shared first language listener-speaker pairs (Bent & Bradlow, 2003; Foote & Trofimovich, 2016). Due to the variable nature of this field, there is still much research to be done in order to not only match the depth of research in native listener research, but also to expand and deepen knowledge in individual language contexts.

As the field of non-native speakers rating other non-native speakers broadens, there are ample areas where research is still sparse, one of them being the Finnish listener context. I have chosen this area to look at not only for the purpose of adding to the body of research including this population, but also for practical reasons: I am in Finland, and have previous experience with Finnish users of English. This research focuses on a quantitative-heavy mixed methods approach to look at language status and the concept of a foreign accent, in particular studying the way that accents that non-native speakers of English are perceived by other non-native speakers of English. Specifically, this research looks at how native Finnish-speaking listeners rate other Europeans' speech in English. This focus is in line with that previous research tradition on accent perception studies, with a focus on *accentedness*, i.e. the approximation of, and closeness to, a native-like accent (Munro et al., 2006); and *comprehensibility*, i.e. the ease of being understood by a listener (Munro et al., 2006). These factors are compared against *valence*, or goodness of a stimulus, in this case, foreign accent (APA, n.d.). The first two ratings are related but distinct ways of measuring listener perceptions of accented speech (Trofimovich & Isaacs, 2012), while the latter functions to quantify the degree and strength of these subjective ratings. That is to say, this research asks the questions of whether Finnish listeners are partial to accentedness or comprehensibility more as means of rating non-native speakers, and to what degree those variables are related to ratings of goodness. Additionally, this research addresses whether an own accent effect exists for Finnish listeners listening to Finnish speakers, and whether that will result in higher ratings for Finnish speakers.

This research this thesis provides is intended to expand the research in accent perceptions with non-native listener ratings of non-native speakers, and specifically addresses the Finnish

listener perspective. I will first discuss English pronunciation education in general and English education in Finland, as well as provide a review of some of the major literature in the field of accent perception studies, with both studies done with native and non-native listeners of non-native speech. The study based off of this background knowledge uses a mixed-methods approach via a survey, asking both qualitative and quantitative questions about comprehensibility, accentedness, and valence. For this thesis, both simple analysis of qualitative responses and statistical methods to assess correlations for quantitative data are used to try and gain an initial picture about Finnish listeners' perceptions on a small scale.

2. BACKGROUND

2.1. Learning English as a Second Language

As of 2019, there are estimated to be roughly 379 million native speakers (NSs) of English in the world (Eberhard, Simons & Fennig, 2019). However, as the sociopolitical power of English grows, often being referred to as a *lingua franca*, or a common language, so too do the number of English learners, or non-native speakers (NNSs). Although estimates vary widely and are constantly changing, there are estimated to be roughly 1.1 billion people who now functionally use English as a second language (Eberhard, Simons & Fennig, 2019). This means that NNSs of English outnumber NSs approximately three to one.

The language learning process is highly variable on a global scale, however, there are a number of common challenges that second language learners of English face in regard to accent features that are generalizable. While the variety of teaching methods and goals are many, in terms of accent, many academic courses that include oral skills have an element of pronunciation teaching. Pronunciation is an important oral skill area, which Derwing and Munro define as “ways in which speakers use their articulatory apparatus to create speech” (2015). This definition is intentionally broad, and means that pronunciation encompasses both specific and small aspects of speech, called segmental features, meaning phonemes like single vowels or consonants— e.g. differentiating between /i/ and /ɪ/ — as well as suprasegmental features, or prosody. Prosody is

an overarching term that includes broader language features like sentence stress, intonation, rhythm and tone (Derwing & Munro, 2015). With this definition in mind, pronunciation, speaking broadly, is nearly equivalent to accent. Historically, teaching of pronunciation has often been tied to a standard English variety, thus encouraging students to objectify native-like pronunciation as a major goal in oral communication. In more recent memory, this paradigm is changing, with a broader acceptance for communicative language teaching practices (Levis, 2005), and standards that reflect realistic pronunciation outcomes (CEFR, citation). However, pronunciation obstacles for learners are theoretical, procedural, and practical in nature, meaning students may face a wide variety of challenges at all stages of learning English.

Firstly, English pronunciation is highly variable. As there is no one standard spoken variety of English, multiple standard varieties exist, each equally as valid as an English variety as the others. For example, Standard American English (SAE) and Standard British English (SBE; traditionally referred to as Received Pronunciation) are both widely accepted high-prestige, generally standard varieties of English which differ in pronunciation. While there are a number of ways that these varieties differ, one major example is that SAE is rhotic and SBE is non-rhotic, meaning that the phoneme /r/ is both pre- and postvocalic in SAE, but is only prevocalic in SBE (e.g. the word “car” would sound like /kɑr/ in SAE and /kɑ:/ in SBE). Even including only other predominantly Anglophone countries’ varieties of English, like standard or general Canadian, Irish, Australian, or New Zealand English, students face an overwhelming breadth of choice. As of the publication of this study, there is no commonly accepted Standard International English variant. There are also dialectal and sociolectal variations to consider, which are less standardized and may vary in indexicality and prestige, but are equally available and valid ways of speaking English.

Compounding on the logistical difficulties of learning pronunciation, teaching of pronunciation is often “orphaned” (Gilbert, 2010) in classrooms of English: forgotten about on a broader scale, with a minority of teachers noting that they have comprehensive pronunciation curriculum in their classes (Gilbert, 2010). This issue is a global one, with research documenting these challenges on a broad scale. For instance, Sharatol Ahmad Shah, Othman, & Senom (2017) found that the Malaysian English as a Second Language (ESL) teachers they interviewed had

little explicit teacher education in teaching pronunciation skills, and had the perception that pronunciation teaching should not be specifically taught, leaving pronunciation integrated in other oral skills. Uchida & Sugimoto (2019) found that while Japanese ESL teachers had confidence in their pronunciation skills in general, many wished for more training. Bringing the research contexts closer to my own topics, the EPTiES project looked at seven European contexts (Henderson et al., 2012; Tergujeff, 2013), which used an online survey to ask teachers about their own pronunciation, training, knowledge of learner goals, and English variety preferences. Broadly speaking, teachers from all seven countries felt that while they had positive feelings about their own English use, their training in how to teach pronunciation was rather limited. Henderson et al. (2012) found that while teachers from a variety of countries, including Finland, France, Germany, Macedonia, Poland, Spain, and Switzerland gave high ratings of the importance of English in general, at an average of 4.66 on a 5-point scale, their perceptions and experience with pronunciation was highly variable. The overall ratings for the importance of pronunciation had an average of 3.77 on a 5-point scale, with teachers from different countries having different overall perspectives on this topic. Some respondents, like the Spanish teachers, had strong feelings toward the lack of specific pronunciation teaching done, while others, like the macedonian teachers, saw pronunciation teaching as a more integrated skill toward an overall goal of communication. Overall, though, teachers felt that they were not as prepared as possible for teaching pronunciation. For example, 19 French teachers noted they had little or no teacher training in pronunciation teaching, another 19 noted some phonetics courses during their university period, and 9 had training from outside sources once they were already qualified teachers. Henderson et al. (2012) notes that while phonetics or pronunciation courses for the teacher are useful, they are not explicit teacher training for how to teach pronunciation to others. Many teachers from other countries also noted some degree of theoretical coursework during their degree, like a phonetics or phonology course. As Finland was one of the countries included in this survey, it is pertinent to dive more deeply into these results. Tergujeff (2013) found that Finnish teachers also did not give high ratings for their teacher training in pronunciation, in that they had good training for their own pronunciation, but training for teaching to others was limited. Among Europeans overall, they had the highest average rating for the teacher training

they received in pronunciation, but even this was only above the central rating, at 3.16. These studies show that even if teaching pronunciation in English may be the goal, the pedagogical basis for teaching that pronunciation is still not effectively put into place on a broad scale.

Finally, the most insurmountable difficulty for learners may be that the acquisition of native-like proficiency, including a native-like accent, is a statistical improbability, with most proficient speakers having low odds, even in ideal conditions, to gain a native-like accent (Birdsong, 2007; Flege, Munro, & MacKay, 1995). This means that although the goal in many cases may be a native-like accent or a perception of “non-accentedness,” that goal is simply unattainable for most learners. Because of this reality, some students have shifted their goals to focus on more attainable ways of improving their language, focusing on other aspects of perceived proficiency (Smith & Nelson, 2006). Levis (2005) notes that speakers are often caught between these two contradictory paradigms: the nativeness principle, which asserts the ultimate goal of native-like proficiency for a language learner, and the more realistic intelligibility principle, which asserts that language learners should simply strive to be understood, and acknowledges that having features of a foreign accent does not inherently prohibit this process. While the focus is often on learners, this paradigm begins with teachers, who then may pass on this unattainable goal to their students. Both Levis (2005) and Munro & Derwing (2005) note that pronunciation teaching is still often based on a teacher’s own intuitive notions of what promotes intelligibility the most, and not on research of what features really contribute to intelligibility. Additionally, although the nativeness principle has not been the primary teaching focus for many years, it still affects pronunciation education, as many existing textbooks continue to promote the nativeness principle (Levis, 2005). For an example in practice, in the study about Japanese teachers mentioned above, the teachers also felt pressure to model a variety of English that they felt that they themselves were not able to achieve (Uchida & Sugimoto, 2015). This shift of focus toward the intelligibility principle and away from the nativeness principle is gaining popularity in English as a lingua franca circles, but is in direct opposition to traditional perceptions of language education. However, language does not exist in a theoretical and practical vacuum, and sociocultural considerations also need to be made. This means that although perceptions and ideals of English use may be changing, the majority of Canadian L2

English students still seek out native-like pronunciation in their language goals (McCrocklin & Link, 2016), with one study finding that 95 out of 100 English learners questioned in Canada would prefer to sound native if they could (Derwing, 2003). This may be in part due to the context, wherein Canada is a predominantly Anglophone country and English proficiency is highly valued. For instance, in contrast to this perception, Tergujeff (2013) found that the Finnish students that were surveyed had responses in line with the intelligibility principle, seeking fluency and intelligibility.

2.2. English in Finland

While I can not discuss the specific English education the speakers in my study have had as that data was not collected, I can more generally discuss the English teaching that happens in the Finnish context to more deeply understand the listener data in this study.

Finland is officially a bilingual nation, with Finnish and Swedish as its official languages. According to the newest curriculum documents from the Finnish National Agency for Education, called *Opetushallitus* in Finnish, students learn three languages during the time in their compulsory education (2014). In practice, this usually means Finnish and Swedish, along with a non-national language, most commonly English. While it is not mandatory to study English as your first non-native language, in 2018, 90.1% of Finnish third graders studied English as their initial second language (SUKOL, n.d.). As of 2020, English language education begins for schoolchildren in the first grade, which includes students around 7-8 years old (Opetus- ja kulttuuriministeriö, 2018). Students continue to study English through the end of compulsory education in ninth grade, when students are 15-16 years old (Opetushallitus, 2014). Past compulsory education, the VARIENG project at the University of Helsinki and the University of Jyväskylä found that 94.5% of Finns self-reported learning English in upper secondary school as well (Leppänen, et al., 2011).

In terms of pronunciation teaching, Tergujeff (2013) finds that while there is some explicit teaching of segmental features in Finnish textbooks and classrooms, but implicit skills via oral tasks and activities are common. Additionally, Kopperoinen (2011) found that in two

major textbook series, most of the focus on pronunciation is on Anglophone countries, meaning a higher bias toward the nativeness principle, with less engagement with Englishes used on a more global scale, which would be more associated with the intelligibility principle. Tergujeff also (2013) finds that in surveys from teachers, the majority use RP or SAE as their pronunciation models, although a number use other models for receptive tasks.

English is also highly visible in Finland outside of the classroom, with English-language media being prevalent both in traditional media (television, music) as well as more informally (YouTube, chat rooms). According to VARIENG, 79.6% of Finns reported seeing or hearing English in their environment or surroundings (Leppänen et al., 2011). The same study found that 52.5% of respondents used English in their free time, with listening, reading, writing, and speaking English all being represented in free-time English use (Leppänen et al., 2011).

With English being both taught in school as well as salient to some degree in the environment, it comes as no surprise that Finland often ranks highly in standardized assessments of English proficiency. The 2019 English First English Proficiency Index (English First, 2019) ranked non-English speaking nations on overall language proficiency, and gave Finland a global rank of 7th most English-proficient nation. Previous rankings from the same source consistently have Finland in the top ten most proficient globally.

However, self-perceptions of language proficiency, especially regarding pronunciation, do not necessarily match up with official reports. VARIENG found that 18.2% of Finns surveyed found Finnish-accented English to be the least appealing variety of English, second in frequency only to Indian English (Leppänen et al., 2011). The same study found that in a hypothetical situation of listening to a “famous Finn” speak English, the more the theoretical person was suggested to approximate a poor accent, the more negative feelings were associated with the accent. Conversely, the closer in approximation to a native accent, the more positive the perceptions were (Leppänen et al., 2011). Finally, the study noted that 41% of Finns reported using English only when it was necessary (Leppänen et al., 2011). Finnish speakers also contend with a number of negative self-stereotypes relating to their pronunciation, including *tankero-englanti*, or tankero English, which comes from a former Prime Minister’s misappropriation of the word *dangerous* and is now associated with highly Finnish-accented

English; and *rallienglanti*, or rally English, which originates from highly Finnish-accented athletes, especially Formula 1 drivers, and has similar connotations to tankero English.

3. CENTRAL CONCEPTS

3.1. Accent

I have spoken generally about accents above, but it is critical to define what is really meant when discussing “accent.” Generally speaking, an accent is the manner in which a person pronounces, enunciates, or stresses their words in speech (Giles, 1970). While Giles hints at the interrelated nature of segmental and suprasegmental features, Moyer (2013) directly describes accent as a “veritable orchestra of intonation, pitch, rhythm, stress, pause, tempo, syllable duration, and elision, not to mention phonetic precision[.]” Although people often correlate “accentedness” with non-standard sound features, all speakers have an accent, including speakers who speak a standard variety of a language as their first language. As mentioned above, accent is simply the fusion of pronunciation features, both segmental and suprasegmental.

While there is a breadth of study on accent as a general concept, it is crucial to recognize that the field is divided into two major areas of research: one group focuses on language variation between native speakers (NSs), usually referred to simply as “accent” (e.g. a Scouse accent, a New York accent). This field is also often concerned with the “-lects” of a language, as in dialect, idiolect, and sociolect. Although every person has an accent in their native language, some may not be as willing to admit their accentedness, or be conscious of it at all (Lippi-Green, 1997).

The other group focuses on the ways that non-native speakers (NNSs) speak when using a second language, usually called a “foreign accent.” The Linguistic Society of America describes a “foreign” accent as being one that “occurs when a person speaks one language using some of the rules or sounds of another one” (Berner, n.d.). To clarify, this does not mean that everyone who uses this type of accent is a “foreigner” per se, but rather that multilingualism plays a key part in this type of accent. Citizenship or residency is not an indicator of language

ability or accentedness. This thesis concerns itself with this side, and focuses on the perception of foreign accent. For the purposes of this thesis, both “accent” and “foreign accent” will henceforth be used interchangeably to describe this concept of non-native accentedness qualities.

3.2. Foreign accent

In keeping with the discussion above, the focus on who has an accent will focus predominantly on those who have a foreign accent. Delineating exactly who falls into this category of “having a foreign accent” is a difficult task due to the fluid and often intersectional nature of language and identity, and is a question that has been modeled many times. The sociocultural implications of having a foreign accent are also important to consider.

One of the most popular models for the description of English spoken globally is attributed to Kachru (1992). He describes the English speaking world as being divided into three concentric circles: the “Inner Circle” includes Anglophone countries, like the United States, the United Kingdom, and Australia, where the majority of people speak English as a native language; the “Outer Circle” includes countries that were former Anglophone colonies, like India and Nigeria, where English is commonly spoken, and may be used as an administrative language, but is not necessarily the first or preferred language of the majority of the population; and finally, the “Expanding Circle”, which includes all other countries that have a population using English in some capacity, which then includes all countries not in the Inner or Outer Circle. According to Kachru’s model, Inner Circle speakers provide norms for English, including accent, and other English speakers from the Outer and Expanding circles are dependent on those norms. That is to say, speakers from those latter circles are considered to have some degree of a foreign accent, while those in the Inner Circle may have an accent, but it is not perceived as a foreign accent. In the case of this research, there are only speakers from the Inner Circle (monolingual English speakers from the United States) and the Expanding Circle (Finnish, Estonian, Russian, and Italian speakers) used as speaking examples. Following Kachru’s model, for the purposes of this research, I define the English speakers from the United States as native speakers (NSs) and the Finnish, Estonian, Russian, and Italian speakers as non-native speakers

(NNSs). In my own study, I did not use speakers from the Outer Circle. The reasoning behind this decision was manifold, but can ultimately be explained in a few ways: Firstly, it is consistent with many other studies that only use Expanding Circle and/or Inner Circle speakers; and secondly, the complex and nuanced analysis needed to include the Outer Circle is past the scope of this thesis and my own abilities.

With a defined group of speakers, it is now possible to shift to the sociocultural effects and implications of being a person with a foreign accent. Accent is one of many linguistic features that can provide salient social and indexical cues (Kozłowski, 2015). Language is socially encoded with a number of features that other interlocutors can interpret either consciously or subconsciously, to help a listener build a profile of the speaker. In terms of foreign accent, the most obviously salient social cue is out-grouping. That is to say, by using non-nativelike language features, the speakers are saliently grouped outside of the realm of “native” (Kozłowski, 2015). This outgrouping allows the possibility of cultural or linguistic stereotypes to come into play, where listeners may judge a variety of social variables based on their perceptions of language and accent, but also on perceptions of nationality, culture, and other related non-linguistic variables. This research area, called language attitudes, or sometimes more specifically accent attitudes, is well studied (Kozłowski, 2015; Rubin, 1990; Gluszek & Hanson, 2016). Additionally, there is another research tradition that focuses less on the social, para- and extralinguistic aspects of perception of accent, and focuses more on the linguistic aspects that give way to these social perceptions (Munro & Derwing, 1999; Munro et al. 2006; Kang et al., 2016; Trofimovich & Isaacs, 2012).

While some research shows positive correlations and stereotyping, like Lindemann (2005), which showed that American perceptions of Western European accents had higher solidarity ratings (e.g. friendliness) than non-Western European accents, the overall literature suggests a trend toward negative effects for general populations. Some negative social effects that have been correlated with foreign accent may be general, like having lower overall prestige, especially as compared to NSs (Moyer, 2013). Personal traits may also be rated more negatively, like intelligence and friendliness (Giles & Watson, 2013). These negative perceptions can have significant real-life consequences. For example, Rubin (1990) found that students associated

lower work capability with foreign accentedness. Studies like these will be more deeply investigated in the following chapter.

3.3. Intelligibility, comprehensibility, and accentedness

To understand the studies to be discussed in the next chapter, three critical concepts must be defined. These three variables are most often measured in this field of study, and make up the basis of a significant number of the studies presented.

The first concept is accentedness. Munro, Derwing, & Morton (2006) define accentedness as “the degree to which the pronunciation of an utterance sounds different from an expected production pattern.” In research, this feature is often rated on a scale of strength, for example from no accent to light accent, to heavy accent; that is to say, an accent perceived as “strong” likely deviates greatly from the expectation. In the case of foreign accent, many studies rely on the axiom that deviation from a native-like pronunciation will result in higher accentedness ratings, as the nativeness principle continues to be at odds with the intelligibility principle (Levis, 2005).

Comprehensibility is the subjective factor of message receipt that describes “the listener’s estimation of difficulty in understanding an utterance” (Munro et al., 2006). This factor is rated on a scale of easiness and difficulty. Very comprehensible speech is speech that can be understood easily by the listener, and very incomprehensible speech is speech that is difficult or impossible for the listener to understand, both as perceived by the listener.

Finally, intelligibility is the third and final rating factor. Intelligibility has been defined as “the extent to which a speaker’s utterance is actually understood” (Munro et al., 2006), making it a more objective factor of message receipt. This factor is distinct because while intertwined with accentedness and comprehensibility, it does not necessarily rely on them. This factor must be measured in a more in-depth way to test how well the listener actually understood the speech, for example by having the listener transcribe the speech they heard, as in Munro et al. (2006). It is often confused with comprehensibility, because the concepts look at different interpretations of the same activity: receiving speech. However, comprehensibility is a rating of how the listener

perceives the speech and is therefore subjective, whereas intelligibility is the more objective rating, looking solely at how much of the intended message ends up getting through, regardless of the ease or difficulty of understanding the speech. This may be measured in the percentage of speech correctly transcribed.

These three factors are most commonly used in the research tradition of perceptual accent ratings, as they provide different levels of information about how listeners perceive the speech of others. Additionally, these three factors are intertwined, with accentedness being in part correlated to the other two factors. For example, Derwing & Munro (2009) found that speakers with heavy accents may be assessed as still being highly intelligible, but highly unintelligible speakers will always be rated as having a heavy accent.

3.4. Valence

Valence is a term most commonly associated with psychology, and refers to the positive or negative value associated with a stimulus (APA, n.d.). In the case of accents, valence is most often studied in relation to language or accent attitudes, which looks at how listeners perceive different talker characteristics and the positive or negative associations based on that speech. Dragojevic et al. (2017) identifies both status (e.g. intelligence) and solidarity (e.g. friendliness) categories for description in this field.

While an infinite number of characteristics can have an associated valence, I am simply interested in the perceived goodness or badness of accent as a whole, without individual traits. This type of work has been established in the field of accent studies, usually using more sophisticated or covert methods of asking about valence directly, using matched-guise techniques to assess underlying perceptions of valence and traits (Dragojevic & Goatley-Soan, 2020; Dragojevic et al., 2017). To that end, for my own research, I will not be using a matched-guise technique for my study, as I am interested in the absolute basic questions of what makes an accent “good” to a listener, and in turn, what makes an accent “bad,” and will therefore use the most simplistic definition and direct operationalization of valence as possible.

3.5. Mutual intelligibility benefit and own accent preference

Finally, the last critical concepts are the mutual intelligibility benefit, also called interlanguage mutual intelligibility benefit (Bent & Bradlow, 2003), and the own accent preference effect, or accent familiarity effect (Kozlowski, 2015). Similarly to the above concepts, I am not using intelligibility or the intelligibility benefit in my own research, but it is pertinent both to the literature and to understand the overall effect of shared first language for listeners across different ratings.

Mutual intelligibility refers to how native language background may influence perceptions of intelligibility of speakers of both the same and different native language backgrounds (Bent & Bradlow, 2003). This means that listeners listening to a speaker with the same first language may have positively biased ratings for those speakers. However, evidence for this effect is contradictory, with some research finding an effect, like Bent & Bradlow (2003) and Major et al. (2002), which both found that listeners with the same first language background as the speakers had higher overall intelligibility ratings; and some studies finding no effect, like Munro et al. (2006) and Major (2007), which found no effect for listener benefit of familiar accents.

The own accent effect is a social in-grouping effect, and occurs from a young age (Kozlowski, 2015). This effect shows that listeners tend to have a preference for those who share their accent. This manifests itself in more positive ratings for NSs rating other NSs as compared to NNSs (Gluszek & Dovidio, 2010). For NNS raters, this means there may be a preference shown for raters from the same first language background. This feature is critical for my satellite question to assess whether the Finnish listeners' ratings were biased positively toward Finnish speakers speaking English.

4. PREVIOUS RESEARCH

4.1. Accent perception studies

Many different approaches have been tried to assess perception of non-native speaker (NNS) accents. Many existing studies, some of which will be discussed in the next subsection, heavily focus on native speaker (NS) perceptions of those NNS speakers, using NSs as a linguistic barometer. While this approach is logical to the end that Anglophone— and usually British or American English— pronunciation has historically been taught to students, it does not account for the changing demographics of global English speakers. Even if the “norm-providing” (Kachru, 1992) varieties continue to have a critical impact on how people perceive accent, NS-NNS speaker-listener relationships are not the only communicative pairs. When NNS-NNS pairs (or groups, for that matter) communicate, it is likely that while they have been exposed to many language norms, they are not likely to have the exact same experience with those norms. Because of this, both NS-NNS research, which can tell us about what the norms are and how they develop, as well as NNS-NNS research, which can tell us about the results of exposure to that normativity, and to differing perspectives on English and accents.

4.1.1. NS ratings of NNSs

To the end that this field of research is much more highly established, this review is only a small portion of the available research in this field. Having NS listeners assess the speech of NNSs has been established for many decades, as English as a Second Language (ESL) is a critical research area in the North American context. This has also meant that a large preponderance of NS-listener studies have used North American English language users. Research in this area shows that both sociocultural as well as linguistic information is taken into consideration by these speakers when assessing L2 English.

To begin, a seminal article by Rubin (1990) found that American undergraduate students had very negative attitudes toward East Asian teaching assistants, as they were often perceived as being significantly less comprehensible as compared to their white peers. Interestingly, when shown a photo of an Asian woman, even when a native English speaking woman was speaking, listeners perceived an accent (Rubin, 1990). This shows that some NSs may be using extralinguistic and social features like cultural knowledge, stereotypes, and past experiences to

assess accent in part, thus not relying completely on what is actually presented linguistically. To that end, accent strength may be skewed much more heavily in NS-NNS situations, meaning that NNSs, regardless of their actual accentedness, may be perceived as having non-native sounding accents regardless of their actual accentedness.

To continue with social factors, Lindemann (2005) found that stereotypes play a large role in estimation of accent, and that those stereotypes were relatively consistent among the judging group members, who were US undergraduates. Lindemann (2005) used a map labeling task, where the participants were asked to think about international students at their university, and then given a world map and asked to describe the English spoken by those students. For part of the experiment the map was unlabeled, but was switched to a labeled map due to concerns about participants' global geographic knowledge. The results showed that places that the participants labelled "correct" were also usually labeled "friendly" and "pleasant" and that sociopolitical attitudes contributed to perceptions of goodness (Lindemann, 2005). As discussed above, many studies suggest overall negative attitudes toward NNSs, which often have components outside of strictly linguistic boundaries of accent. These types of studies reveal more intricate and intersectional power dynamics between NSs and NNSs, which may be crucial to understanding why NS-NNS and NNS-NNS ratings may be different in the end, as different groups have different relationships to other languages and cultures.

A review conducted by Gluszek & Dovidio (2010) found that the NS-NNS dichotomy has been studied in the literature in multiple country contexts, with NSs having a predominantly negative perception of NNSs. For instance, personal traits may be rated more negatively, like intelligence and friendliness (Giles & Watson, 2013). Following Lindemann's (2005) process, Dragojevic et al. (2017) found that when speaking English, Mandarin Chinese L1 speakers were rated lower on a variety of status categories by American listeners.

Outside of social and mixed social-linguistic factors, there has also been research regarding language features that affect ratings of accentedness and comprehensibility. One early study by Munro & Derwing (1995) found that NNSs' speech takes more time for NSs to process than other NS speech, but that degree of accentedness did not seem to be a factor in this effect. Another study, also by Munro & Derwing (1999), found a number of interesting correlations

between accent, comprehensibility, and intelligibility. Firstly, listeners gave harsher scores when rating accent than when rating comprehensibility. Additionally, correlations between the three factors were all of moderate strength, but comprehensibility scores were a better predictor comparing against intelligibility scores, and not accentedness. And, perhaps most importantly, they found that the strength of accent does not necessarily result in reduced ratings for comprehensibility or intelligibility.

Another such article focusing on linguistic factors is Kang, Rubin & Pickering (2010), which looks at the role of suprasegmental features in assessing NNSs' comprehensibility. They found that suprasegmental features make up about 50% of variance in relation to comprehensibility ratings, showing that in addition to social factors, linguistic factors also contribute greatly to perceptions of NNS speech. Saito, Trofimovich, & Isaacs (2016) additionally find that in NS ratings of Japanese English L2 speakers, all linguistic domains played a part in comprehensibility, and segmentals were most related to ratings of accentedness. Trofimovich & Isaacs (2012) found that with both naive and trained Canadian raters, both accentedness and comprehensibility measures contributed to judgements of L2 English speech. Saito, Trofimovich, & Isaacs (2016) also note that in rating these factors, "comprehensibility appears to be related to segmental, prosodic, temporal, lexical, and grammatical aspects of L2 speech, while accentedness is mainly associated with pronunciation factors, particularly with segmental accuracy."

Finally, a study by Saito & Shintani (2016) looked at how native English speakers from Canada and Singapore rated comprehensibility for NNSs. They used these speakers as the Singaporeans were in a highly multilingual and multivarietal English environment, while the Canadians were in a monolingual predominantly monovarietal English environment. They found that listeners' linguistic backgrounds and setting led to different distributions in ratings of comprehensibility: Singaporean listeners were more lenient than Canadian listeners, and that while Canadian raters' scores were only significantly related to grammatical accuracy, Singaporean listeners' scores were correlated with both grammatical accuracy and lexical appropriateness (Saito & Shintani, 2016). This shows that the background and expectations of the speaker are important in ratings of comprehensibility.

4.1.2. NNS ratings of NNSs

To continue the discussion of relevant research is to move into the territory that I will be addressing in my own study: non-native listener ratings of other non-native speakers of English. Far fewer studies have addressed NNS perceptions of other NNSs as compared to NS-NNS studies, likely due to the more recent inclusion and interest in this work on a global scale due to the increased interest in English as a lingua franca. With that said, the studies that have been conducted in this area do show some interesting effects for NNS-NNS contexts which differ from the NS-NNS contexts. Many studies using NNSs have found that NNSs are equally good raters as NSs (Major, 2007; Gallardo de Puerto et al., 2015), making the possibilities in this field incredibly broad.

A few of the studies have looked at possible effects of native language or language familiarity in intelligibility. As described in the previous chapter, evidence for this effect is mixed. For instance, Munro et al. (2006) found that some NNS listeners may find speakers with the same native language to be more intelligible. However, when considering listeners with different native language backgrounds overall, the NNS listeners produced similar intelligibility ratings to each other, meaning that the a speaker's intelligibility could reliably be assessed by listeners from a variety of language backgrounds with little mutual intelligibility benefit (Munro et al. 2006). Similarly, Major (2007) found that when listening to Brazilian Portuguese speech, both NSs of Portuguese and NNSs of Portuguese with a range of familiarity with Portuguese rated groups of both NNSs and NSs of Portuguese with similar ratings, meaning there was no perceived benefit for the NSs when listening to the speech. On the other side of this, however, there has been some work that does show the possibility of the existence of this benefit. For instance, Bent & Bradlow (2003) found that when listeners from a variety of backgrounds were given intelligibility tasks where the speakers were both NSs and NNSs of English, with some speakers with the same background and others without, NNSs with the same L1 language background were given equal intelligibility scores to the NSs. This represents a matched intelligibility benefit. Additionally, Bent & Bradlow (2003) found a mismatched intelligibility

benefit, wherein NNSs rated high-proficiency NNS talkers with a different L1 background as not only equal in intelligibility, but sometimes greater. Finally, Major et al. (2002) found that Chinese and Japanese speakers showed no intelligibility benefit, but Spanish speakers did show a small advantage.

On the other hand, Foote & Trofimovich (2016) looked at the own accent effect, which they call mutual comprehensibility. This is complementary to the mutual intelligibility benefit, but is based on the subjective rating of comprehensibility. They tested how Hindi, Mandarin, and French L1 speakers rated the comprehensibility of Hindi, Mandarin, and French L2 English speakers. They found that different groups of speaker-listener pairs were associated with different speech measures (e.g. when rating Hindi speakers, Mandarin listeners' ratings were associated with segmental errors, word stress errors, and intonation) (Foote & Trofimovich, 2016). They also found that Mandarin listeners gave higher ratings of Mandarin speakers, but other groups did not give significant preference to those speakers with the same language background. This suggests that there is variance in speaker-listener pairings, and that each language background may provide unique information about the listener group, and about the existence of a mutual comprehensibility benefit for different language populations. This is supported by Kang, Vo, & Moran (2016), who find that language background, along with language learning experience, affects assessments of accented speech.

Episcopo (2009) found that within the category of accentedness, NNS listeners were more likely to use accent strength as a standard as compared to native-likeness, with those who were perceived as being less regionally accented in a general sense being seen as more comprehensible, not those with a more native-like accent. Interestingly, Episcopo (2009) also noted that when it came to character ratings, NNSs were more likely to associate positive characteristics with those speakers who had more native-like accents, and not those who were more comprehensible. With that said, although the basic demographics of the two cohorts were similar save for native language background, Lindemann's (2005) NS cohort readily provided and fed into stereotypes and assumptions, while Episcopo's (2009) NNS cohort were much more reticent to attribute character traits to speakers on the whole. It is difficult to say what factors may contribute to this finding, but may show that NNSs have different perceptions due to

cultural or linguistic background, which supports the notion of testing different populations to see whether there are any generalized effects, or if the effects are highly specific to each language group.

5. PRESENT STUDY

5.1. Research questions

The goal of this study was to expand on the growing body of work on NNS-NNS accent perception, focusing specifically on the Finnish context. To that end, I developed a few research questions to guide this work. My main, general research question addresses a necessary first question, as it begins to define the context and provide some insight into the listeners' perceptions:

- a) What features do Finnish listeners perceive as important when quantifying a “good” foreign accent?

To assess this, comprehensibility and accentedness were chosen as factors to compare against valence, which are factors from previous research. Finding the answer to this question would make any future research easier, as well as provide some superficial knowledge about the perceptions of the Finnish listeners. Although discussed above in the broader research, intelligibility will not be tested as the speaker files are all the same short text and would be easily transcribed after a few iterations regardless of what the listener perceived. Using the features above, the broad research question is then branched specifically into the a parallel question for this study:

- b) Do Finnish listeners predominantly link comprehensibility, accent strength, both, or neither to determine a “good” accent?

This question addresses each measured factor, and considers all possible combinations of comprehensibility and accentedness ratings. An additional satellite question also arose during the process of creating the research and would be simple enough to test within the method of study proposed. This question reads:

- c) Is there an own accent preference effect for Finnish listeners rating Finnish speakers in any category?

The first two questions will be operationalized through direct questions about the three features, and then analyzed for possible statistical correlations using Pearson's r correlations among other statistical tests like Fleiss' kappa and mean ratings to test the validity and reliability of the data. The final question will be operationalized in the same data analysis with two-way t -tests to see if the overall means of groups scores for Finnish speakers differ from the other NNS groups.

5.2. Survey

To gather the data, an internet survey was sent out. A copy of the survey questions can be found in the Appendices section. The survey is a mixed qualitative and quantitative survey, with an emphasis on quantitative data, and using qualitative data to more thoroughly inform the analysis.

The survey asked three qualitative opinion questions at the beginning to gauge some very basic information about language attitudes, namely how they qualitatively perceive good and bad accents, and what they believe contributes to comprehensibility. After this portion, the quantitative portion was presented. The quantitative portion included sound files paired with three questions referring to each sound file. Each sound file was presented in its own "speaker block," so only one sound file, along with the respective questions, was available at a time. Each speaker block had instructions for the listener to listen to the sound file and answer the three questions. These questions rated the speech on three dimensions: comprehensibility, accentedness, and valence; one question for each. The questions were the same for each speaker

block. In total, there were ten speaker blocks, which were randomly presented to the listener to try and negate a training effect.

The rating scales for the questions in each speaker block were from 1-7, where 1 = very easy to understand and 7 = very difficult to understand for comprehensibility; 1 = no accent and 7 = very strong accent for accentedness; and 1 = very good and 7 = very bad for valence. While linear, integer-based scales are commonly used in these kinds of studies, a 1-7 scale has not been historically as common in this area, though Kang et al. (2010) used a 1-7 scale for comprehensibility ratings. More common are 1-9 scales (Munro et al., 2006), and 1-5 scales (Episcopo, 2009). I chose to use a 1-7 scale for practical reasons. In some small-scale pilot testing ($n = 3$), the listeners found the 9-point scale difficult to understand and appropriately weigh their responses, and noted that the 1-9 scale did not fit very well on one of my users' screens due to the software I have used. I did not want to limit the listeners' responses to only a 5-point scale, so I compromised in the middle and chose a 7-point scale.

5.2.1. Speaker Files

The speaker files were sourced from the George Mason University Speech Accent Archive (SAA) (Weinberger, 2015). The SAA is open source via Creative Commons license and in part intended for research, and the sound files are all available online.

Each speaker file in the survey is a man saying a short passage in English. The passage is the same for each speaker. Ten total speaker files were chosen, two from each speaker group. Two were chosen to try to mitigate any idiolectal effects of individual speaker or sound file variations. Eight of ten speakers are NNSs of English from various first language backgrounds, with the final two being NSs of American English, used as a sort of control. I was able to control for some speaker variables, including factors like limiting age between 18-30, the same as the target listener group; considering length of residence in an English speaking country, which I tried to limit to less than three months; and other language ability, particularly trying to eliminate dual native language background. The only feature that was not fully controllable due to the available samples was age, with two of the speakers being older than the desired limits. The

non-ideal speaker ages were 43 and 46 years old. The speech samples come with limited information, one of the restrictions being the lack of clarity on whether multiple language knowledge was a result of having multiple native languages or whether the languages were learned later in life. Due to this limitation, there is some variation in overall language abilities, with some monolingual and some multilingual speakers. Also, a number of the speakers come from multilingual contexts. For example, Finnish is a bilingual country and some degree of Swedish proficiency is likely for most speakers. When there were more than two speakers that fit the criteria, two were chosen at random from the possible choices. For the American English speakers, enough sample choice made it possible to limit to male speakers, 18-30, monolingual, who speak a predominantly standard variety of American English, with no features that were obviously salient to any regional dialect nor any sociolect. The American speakers' standardness was only assessed by me, both by parameters of region and overall sound. All speakers' data files were also considered for audio quality, eliminating any speaker whose speech file had a noticeably low quality recording or audio interference (white noise, obstructive shuffling sounds, etc.) The recording quality of all speech samples was not the same. The data on each speaker can be found in the Appendices in Table 4.

The language groups of Finnish, Estonian, Russian, and Italian also have some reason to why they were chosen as the other groups. For the Finnish speakers, I was interested to see if there would be any kind of own accent effect, as discussed in the above chapter. Estonian and Russian speakers were chosen due to their immigration patterns to Finland. As of 2019 (Statistics Finland, 2020), Estonians are the largest non-citizen nationality group in Finland. Russians come in second place. Therefore, it is likely that Finnish listeners may be familiar with Estonian and Russian accents, or have some background knowledge of these populations. It is also pertinent to note that Estonian is in the Finnic branch of the Uralic language family along with Finnish, and as such there are phonological similarities between the languages. If the own accent effect exists in part because of phonological features, we would expect to see somewhat similar ratings for Estonians as well. On the other hand, Russian is a Slavic language, and does not have a particularly similar phonology to Finnish. Finally, Italians were chosen as a likely more unfamiliar accent, with Italians having low immigration patterns to Finland, coming in at 23rd

most common in 2019 (Statistics Finland, 2020), and being part of the Romance language family.

5.2.2. Survey participants

To comply with GDPR restrictions on personal data, all data were anonymously collected. To participate in the survey, participants had to acknowledge and confirm that they were a native Finnish speaker and thus fit the language background criteria; between 18-30 years old, limiting them to within approximately one generation; and that they had not studied, nor were they currently studying English as a major subject in higher education. The final qualification was to eliminate, as much as possible, non-naive listeners who may have had greater than average knowledge about English accents and English pronunciation training. No other participant criteria were used. There were a total of 23 participants who completed the survey in full, and only their data were used.

Survey participants were recruited within practical limitations. Recruitment posts were published via social media posts and university email lists, as well as WhatsApp groups and word-of-mouth. No direct recruiting was used, in that no potential participants were contacted one-to-one and asked to participate, so that anonymity of potential participants would be preserved. Recruiting took place in September and October of 2020.

6. RESULTS

In total, 34 respondents either partially or fully filled out the survey past the consent page, with 23 participants getting to 100% completion. For the data analysis of this thesis, only these complete responses (n=23) were considered in order to keep training effects similar, as well as rule out any possible erroneous data patterns.

6.1. Qualitative questions

The three questions at the beginning of the survey were grouped into themes for the answer analysis. The questions were all prefaced with the statement “Imagine a non-native speaker of English.” After this statement, the questions, in order, were “What does a good accent sound like?”; “What does a bad accent sound like?”; and “What makes an accent easy to understand?”. Each question is discussed separately below. Answers for this section were generally short and direct, but some responses went into a bit more detail about their opinions. Most participants gave more than one language feature, so the number of data points is greater than the number of participants. Most data will be aggregated, but some answers of interest will be highlighted to try to explain more about the ideas of this cohort in the discussion section. The complete set of answers can be found in Table 5 in the Appendices section. Overall, there was a general pattern to the answers for the first two questions, wherein whatever they said was good, they gave a direct opposite for what was bad and did not expand. For example, if the participant answered that being easy to understand was good, their response for what made an accent bad was being difficult to understand. This was expected due to the open-ended format with little prompting. With that said, as the participants could input as much text as they liked, a number of respondents included multiple variables.

While there is some variation in the categories, the overall categories I organized the answers into are accentedness features, with specific (e.g. naming a language or type of language) and general accentedness categories (e.g. pronunciation, segmentals); comprehensibility, which deals with all other factors of language (e.g. tone, other suprasegmentals); and an “other” category, meant to represent answers that were not readily categorizable into either previous category, because of broadness or inspecificity. These categories are based on Saito, Trofimovich, & Isaacs (2016), who note that “comprehensibility appears to be related to segmental, prosodic, temporal, lexical, and grammatical aspects of L2 speech, while accentedness is mainly associated with pronunciation factors, particularly with segmental accuracy.” Thus, for my categorization, it can be said that the general accentedness category contains segmental features, while the comprehensibility category contains suprasegmental features.

6.1.1. What does a good accent sound like?

The responses to the first question can be grouped broadly into the categories of accentedness features, with subcategories of specific accent features and general accent features, and comprehensibility features. For the first question, which was about what makes an accent good, the participants had a variety of responses. There were a total of 62 identified responses.

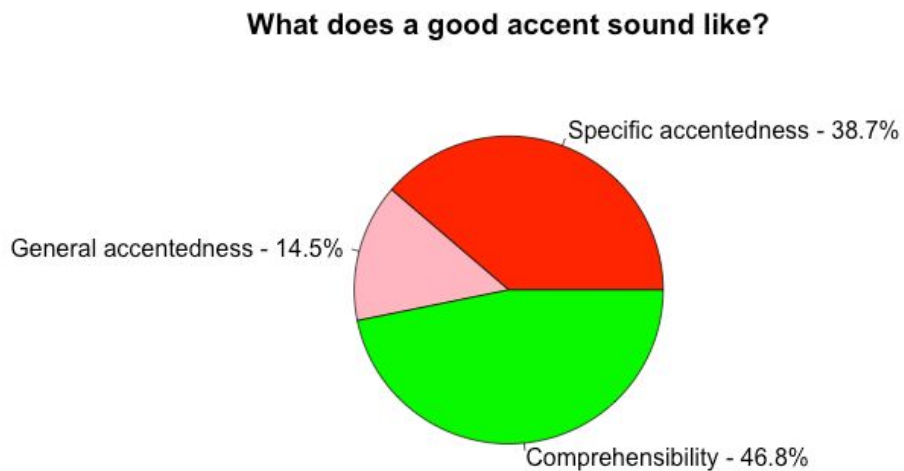
Features in the accentedness category were broad, and can be subdivided into two groups: overall accent, and specific language features. In total, accentedness with both subcategories included had 33 responses, making up 53.2% of all responses for this question. In the first subcategory for specific accent features, there were 24 total responses, making up 38.7% of the total responses. The topics were approximation of native accent in general (n=2); specific accent approximation, namely British (n=4), American (n=4), South African (n=1) and Dutch (n=1); general accent neutrality (n=3), sounding “natural” (n=2), and having an accent that was not noticeable or salient (n=3); and having any kind of accent (n=1), or an accent that incorporates the native language of the speaker (n=3). There are responses that directly contradict each other, e.g. having an accent that sounds native-like versus embracing the first language features in English, but the majority of responses deal with native-like or specific approximation of Anglophone accents, or a sense of a “neutral” or non-specific accent. As for the other subgroup, general language features, there were 9 responses, making 14.5% of the responses. This category included more overarching linguistic topics, like accurate pronunciation (n=6), with both vowels (n=1) and consonants (n=2) being specifically addressed. Pronunciation was most often not directly associated with a specific accent in the same answer, but a general concept of “goodness.”

In the comprehensibility category, there were 29 responses, making up 46.8% of the total responses. Features included ease of comprehensibility (n=9), with some responses noting overall comprehensibility (n=2), and one response that understanding every word was important (n=1); speed (n=4), with both non-fast (n=3) and fast (n=1) responses being noted; general fluency (n=4); and tone (n=1) intonation (n=1), and overall clarity of speech (n=3). This category

was described less than accentedness features overall, and has more variation in perception of goodness, e.g. whether fast speech or non-fast speech is better for comprehensibility.

Figure 1 below shows these responses by category. As noted above, accentedness features overall, including both general and specific, make up a slight majority of responses. By itself, comprehensibility makes the largest classification group, with just under half of the responses.

Figure 1. Pie chart representing percent of responses in each category. Specific accentedness, general accentedness, and comprehensibility are noted in individual colors.



6.1.2. What does a bad accent sound like?

For the second question about what makes an accent bad, as discussed above, the responses are generally the direct counter features to the positive traits. For this question, answers can be divided into the same categories of accentedness with subdivisions for specific and general features and comprehensibility, with an additional overflow category of answers that fit into neither category or were unclear as to what they meant. There were a total of 43 responses for this question, which is notably less than the 62 answers for goodness.

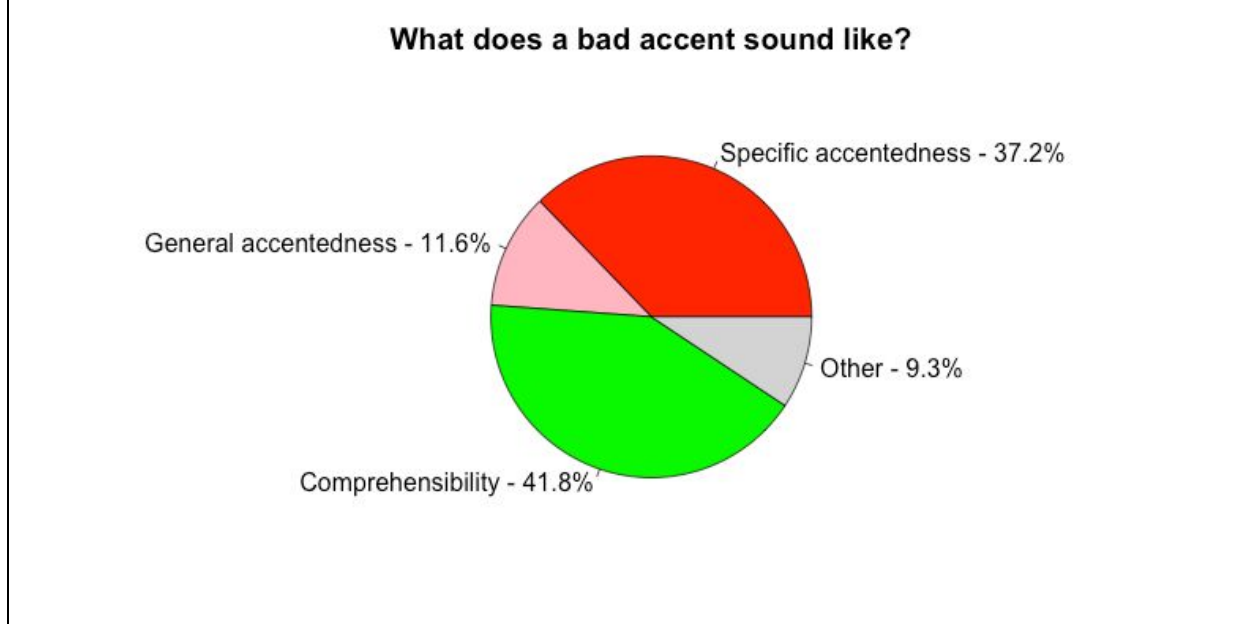
For accentedness overall, there were 21 responses, or 48.8% of the total. For the specific accentedness category, there were 16 responses, making up 37.2% of all responses. Responses included having a strong (n=3) or unnatural (n=1) accent; having an Italian (n=1), Spanish (n=1), French (n=1), “fake” British (n=1), Rally English (n=2), or an identifiable accent (n=6). In the general category there were only 5 responses, or 11.6%, with poor pronunciation (n=4), specifically of consonants (n=1), being submitted. There were more varied responses for specific locations or language backgrounds than in the previous question.

In the comprehensibility category, there were 18 responses total, making 41.8% of the total answers. Being unclear (n=1); being hard to understand (n=12), specifically being hard to understand overall (n=1); intonation (n=1); and high listener effort to understand the speaker (n=3) were noted. This category had the largest agreement of the cohort, with overall comprehensibility being specifically mentioned by about half of the respondents.

In the final category, all answers that seemed to not fit into a previous category were grouped. There were 4 answers, or 9.3% of the total responses. They predominantly include descriptions that would need more in-depth discussion with participants to understand what features these descriptions are based on. The responses included sounding like a “try hard” (n=1), and sounding “hard” (n=1), “sharp” (n=1), or “lazy” (n=1).

Figure 2 below shows the responses for what the participants believe contribute to what a “bad” accent sounds like. In total, both accentedness categories make up nearly half of the responses, similarly to the question about “good” accents. Comprehensibility makes up less of the responses than in the previous question. However, the additional category of “other” contributes a significant percent of responses, which are uncategorizable into either accentedness and comprehensibility.

Figure 2. Pie chart representing percent of responses in each category. Specific accentedness, general accentedness, and comprehensibility are noted in individual colors.



6.1.3. What makes an accent easy to understand?

For the third question, which was about features that relate to greater overall comprehensibility, many of the previous answers were echoed. There were also a similar number of answers, with 47 responses being reported. Similarly to the above ratings, these answers were divided into accentedness, with specific and general subcategories; comprehensibility; and an “other” category, to catch answers that were either ambiguous or did not directly fit into previous categories.

Overall, there were 31 responses for accentedness, making up 66% of the total responses. For the specific accent features category, there were only 8 responses, or 17% of the total. Participants responded that neutral or natural speech (n=3), standard pronunciation (n=1), American (n=2), and British (n=2) contributed to comprehensibility. For the general subcategory, there were 23 total responses, or 49% of all responses. Clear pronunciation (n=17),

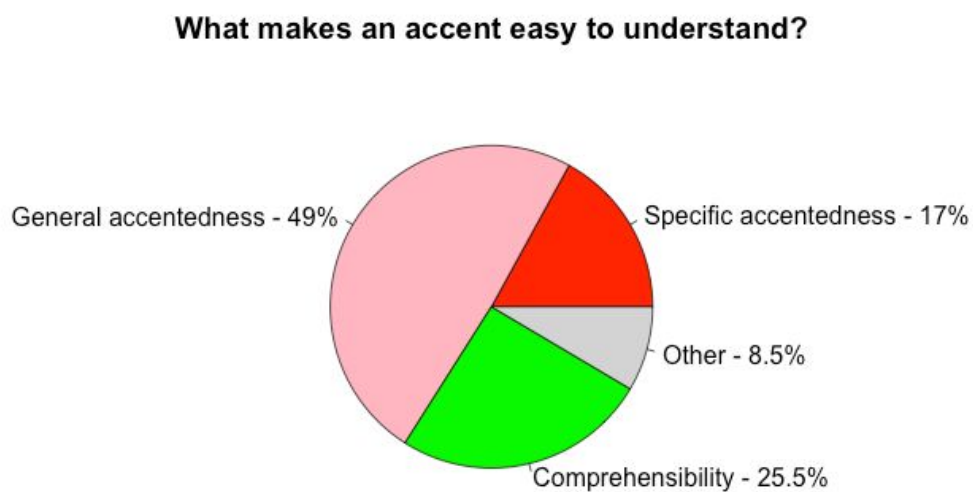
specifically of both vowels (n=3) and consonants (n=2); and familiarity with the accent being spoken (n=1) were noted.

For comprehensibility, there were 12 responses, making 25.5% of the total. Understandability (n=1), speed (n=7), intonation (n=2), and speech clarity (n=2) were noted by participants.

For the “other” category, there were 4 responses, or 8.5% of the total. The responses were about language level (n=2) and specifically vocabulary(n=1), as well as having a “smooth” sound (n=1).

Figure 3 shows responses to the third question, which was about what makes an accent easy to understand. It is in particular contrast to Figures 1 and 2, with general accentedness representing about half of the responses, predominantly due to the large volume of answers that referenced pronunciation specifically. Specific accentedness features were noted about half as much as in the previous questions. Comprehensibility only makes up about a quarter of responses, which is also less than previous questions. Finally, the “other” category again makes up a small portion of the responses that were uncategorizable.

Figure 3. Pie chart representing percent of responses in each category. Specific accentedness, general accentedness, and comprehensibility are noted in individual colors.



6.2. Quantitative data

Quantitative analysis was done in both Qualtrics Stats iQ program and in RStudio. Basic statistical analysis, like calculating Pearson's r and completing t-tests, was done in Qualtrics Stats iQ, while graphs and interrater reliability ratings were coded in RStudio. For the purposes of statistical analysis and matching other research (Munro & Derwing, 1999; Munro et al., 2006; Trofimovich & Isaacs, 2012), the data has been coded as interval data and not ordinal data.

6.2.1. Interrater reliability

First, it is important to note that for this data, interrater reliability was poor, which means that overall, listeners did not necessarily agree with each other. This effect was found both overall, as well as across the three questions types asked. This result came from Fleiss' kappa tests, which is an intraclass rating method that tests the level of agreement between raters on a number of values. Its range is between 0 to 1, with ratings closer to 1 indicating better agreement in a similar way to Pearson's r . Previous studies that use Fleiss' kappa (Munro & Derwing, 1999; Munro et al., 2006) find general agreement above 0.8, which indicates good reliability. Table 1 below shows both the Fleiss' kappa ratings and the p-values associated with them for this study. The table shows the overall interrater reliability across all quantitative questions, as well as a breakdown of each category. Each category includes all speakers (American English, Finnish, Estonian, Russian, and Italian). The overall agreement was 0.109 overall, which indicates broadly poor rater agreement. Each individual category is weaker than the overall agreement, with ratings of 0.066 for comprehensibility, 0.026 for accentedness, and 0.074 for valence. All p-values are well within significant bounds ($\alpha = 0.05$). These results indicate that raters had very wide ranges of responses when rating the same speaker, with little agreement. The greatest reliability was with overall ratings without respect to category, but the effect remains incredibly weak.

Table 1. Fleiss kappa values assessing interrater reliability for all rating categories overall, comprehensibility, accentedness, and valence ($\alpha = 0.05$).

	Overall	Comprehensibility	Accentedness	Valence
Fleiss' kappa	0.109	0.066	0.026	0.074
p-value	0.000	<0.001	0.015	<0.001

6.3. Accentedness, comprehensibility, and valence

First, to look at the three categories overall, I created box plots to see the overall distribution of the data, partially in response to the low interrater reliability measures. I expected to see overall better (i.e. lower accentedness, higher comprehensibility, and higher goodness) ratings for the NSs, but rating trends of all other speakers were not presumed. Some variability was to be expected due to natural differences in speech features (Munro & Derwing, 1999). The data notably has some outliers, as the cohort was small.

Overall distribution of accentedness ratings can be seen in Figure 4. Looking at the data generally, accentedness medians did not exceed 4, which was the center of the provided scale. Additionally, the rating of 7 was not used for any speaker. Overall ratings for the American English NS speakers was, on average, concentrated closer to the lower end of the scale than for the NNSs; that is to say they were rated as having generally little or no accent. One speaker (AME1) had a median rating at the bottom of the scale, a 1; and the other (AME2) had a median rating of a 2. Ratings for AME1 were quite in a quite compressed, positively skewed distribution, while the distribution for AME2 was broader, as indicated by the length of the box. Ratings of both speakers were contained in the 1-5 range. For the Finnish speakers, there were median ratings of 4 (FIN1) and 3 (FIN2). Ratings for these speakers were broader, with a range of 1-6. The distribution of the data was relatively similar between FIN1 and FIN2, although FIN1 was skewed toward the high end of the scale. The Estonian speakers had the highest median accentedness ratings across all speaker groups, at 3 (EST1) and 4 (EST2). The range of

responses was from 1-6 for EST1 and 1-5 for EST2. The distributions were rather broad for both speakers, with a particularly noticeable negative skew for EST2. The Russian speakers both had a median score of 3 (RUS1, RUS2), with ranges of 1-6 for RUS1 and 1-5 for RUS2. Distribution of the data was quite centrally condensed for RUS1, while RUS2, had very normal distribution. Finally, the Italian speakers had relatively low median ratings at 2 (ITA1) and 3 (ITA2) respectively. There was a range of 1-4 for ITA1, and 1-5 for ITA2, which is more condensed than even AME1, although the skew was negative, trending toward the center.

Figure 4. Box plot of accentedness ratings for each speaker. Same color indicates same L1 background.

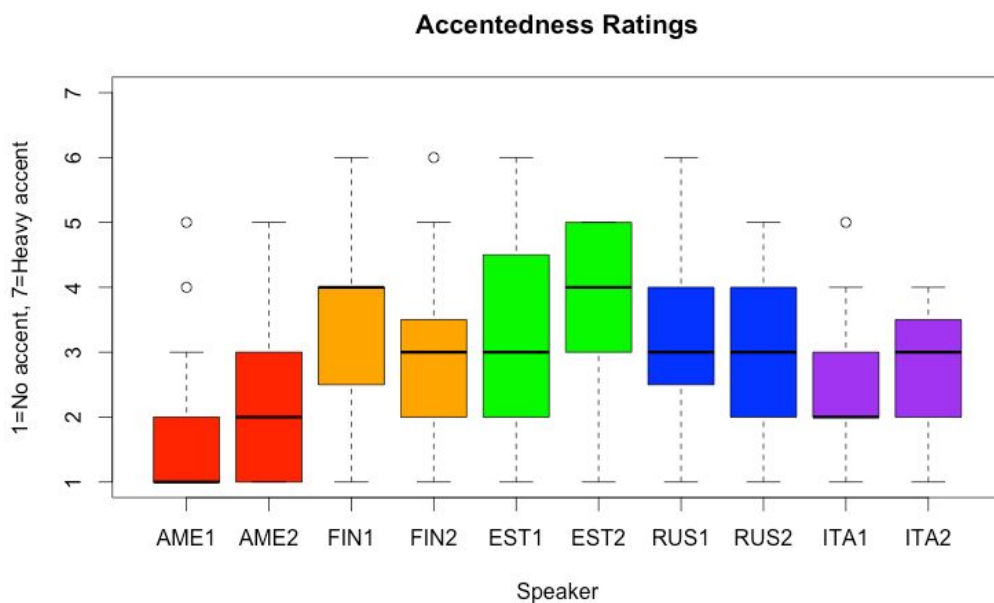


Figure 5 shows comprehensibility ratings for the speakers. The same names for the speakers were used as in Figure 4. Overall, no median ratings were above a 3, and, just as above, the value 7 was not used by raters at all. For the American English speakers, the median rating for AME1 was a 1, with a range of 1-2, and the median rating for AME2 was also a 1, with a range of 1-4. The distribution of AME1 is highly condensed, while AME2 remains highly

positively skewed, but much more board than AME1. For the Finnish speakers, FIN1 had a median rating of 3, and a range of 1-5. FIN2 had a median of 1 and a range of 1-3. FIN1 had a relatively broad and normal distribution, while FIN2 has the same distribution as AME2. The Estonian speakers EST1 and EST2 both had median ratings of 2, but EST1 had a range of 1-5 while EST2 had a range of 1-4. EST1 had a positively skewed, but rather broad rating range, while EST2 had a similar distribution as FIN2 and AME2, with the median being higher than the others. The Russian speakers had medians of 1 for RUS1 and 3 for RUS2. RUS1 had the same distribution as AME2 and FIN2, while RUS2 had a very normal distribution similar to FIN1. Finally, the Italian speakers had median ratings of 2 for ITA1 and 3 for ITA2. ITA1 had a positively skewed distribution and a range of 1-4, while ITA 2 had a relatively normal distribution, but had the broadest range, from 1-6.

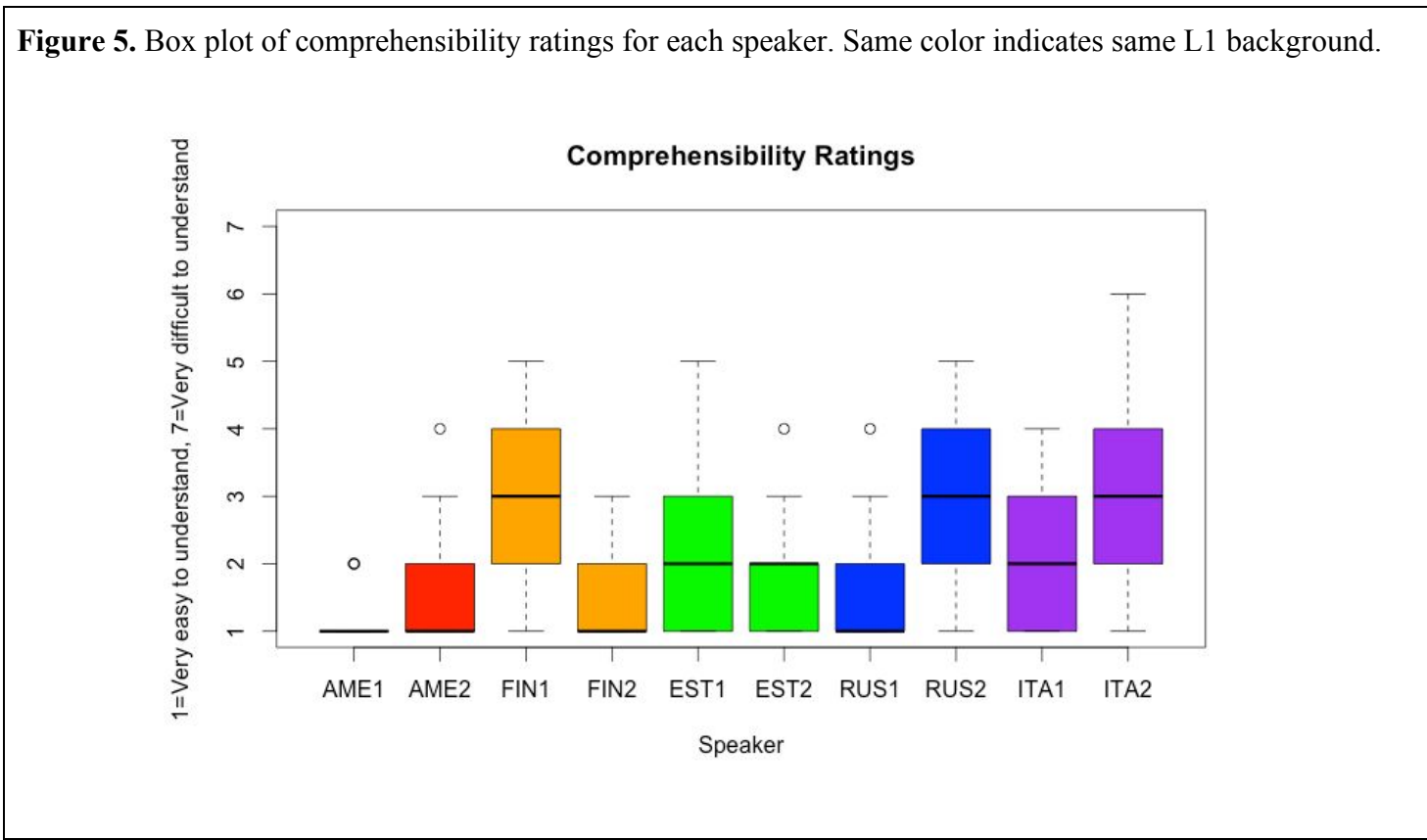
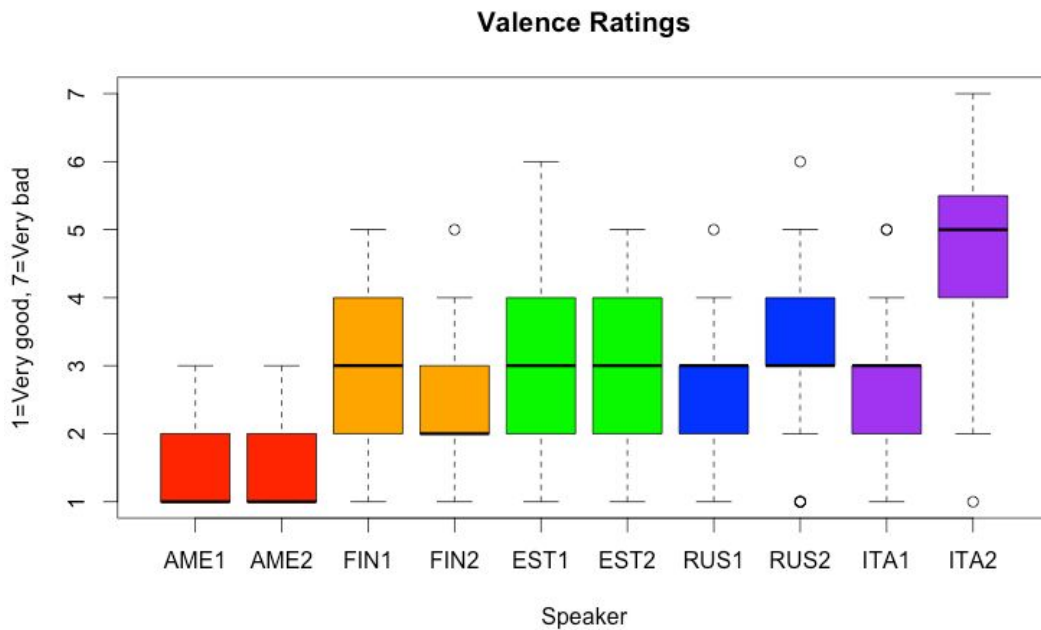


Figure 6 shows the responses for valence for the speakers. Valence was the only question among the three where listeners used the rating of 7. Both American English speakers had the same

median, 1, as well as the same range and distribution, which was positively skewed and in the 1-3 range. Among NNSs, Finnish speakers had the lowest combined medians, with median ratings of 3 for FIN1 and 2 for FIN2. FIN1 had normal distribution, with a range of 1-5, and FIN2 had a positively skewed distribution with the same range. For Estonian speakers, EST1 and EST2 had median ratings of 3, with similar distributions. The ranges were 1-6 for EST1 and 1-5 for EST2. Russian speakers had somewhat unusual results, as both RUS1 and RUS2 had 3 as their median rating, but their data was skewed in different directions. Both speakers' data was highly compressed, with a positive skew and range of 1-5 for RUS1, and a negative skew and range of 1-6 for RUS2. Finally, the Italian speakers had very different medians, with ITA1 at 3, and ITA2 at 5. Ranges were 1-5 for ITA1 and 1-7 for ITA2, with negative skews for both.

Figure 6. Box plot of valence ratings for each speaker. Same color indicates same L1 background.



To assess correlation between accentedness to valence and comprehensibility to valence, basic statistical tests were conducted, namely Pearson's r calculations. This measure was chosen as there was only one listener group and only two comparable variables, so ANOVA was

unsuitable for this purpose. Similar analysis methods can be found in, for example, Munro & Derwing (1999). Values were calculated for each speaker for correlation between comprehensibility and valence (CV) and accentedness and valence (AV). In Table 2 below, all correlations that are statistically significant ($p < 0.05$) are lightly highlighted.

Table 2 below shows the Pearson's r and p -values for both AV and CV. For relationships with statistical significance, the effect was moderate for nearly all combinations, with relationships between 0.437 and 0.741. For the American English speakers, CV had a moderately statistically significant relationship for both speakers, and a weaker but still statistically significant correlation of AV for American 2. For the Finnish speakers, correlation was moderate overall, but with statistically significant relationships for CV for Finnish 1 and AV for Finnish 2. The Estonian speakers have a similar pattern to the Finnish speakers, with Estonian 1 having a moderately statistically significant effect of AV while Estonian 2 has a significant effect of CV. Both Russian speakers had statistically significant relationships in both CV and AV, with a large correlation in CV for Russian 2. For the Italians, there are no statistically significant relationships for either CV or AV for Italian 1, while both CV and AV were moderately statistically significant for Italian 2.

Table 2. Pearson's r values for comprehensibility and accentedness rated against valence for all speakers ($\alpha = 0.05$).

	Comprehensibility		Accentedness	
	Pearson's r	p -value	Pearson's r	p -value
American 1	0.681	< 0.001	0.437	0.037
American 2	0.714	< 0.001	0.195	0.373
Finnish 1	0.519	0.011	-0.001	0.996
Finnish 2	0.382	0.072	0.553	0.006
Estonian 1	0.385	0.070	0.614	0.002
Estonian 2	0.614	0.002	-0.125	0.569

Russian 1	0.512	0.013	0.667	< 0.001
Russian 2	0.828	< 0.001	0.672	< 0.001
Italian 1	0.196	0.371	0.325	0.130
Italian 2	0.741	0.0000520	0.560	0.00544

6.4. Own accent preference

With the data visualization from the box plots generally visually inconclusive, I decided to look more deeply into the average ratings of the speakers for the language features to see if there were any statistically significant differences. To assess whether there was a specific effect for the Finnish speakers' ratings as compared to the other speakers, the data was divided into Finns and Non-Finns groups. The American data was excluded from the Non-Finns group and from the analysis in general, as they were the only NS group and were used predominantly as a light control. First, I looked at the means for each speaker to see if there was a visibly noticeable difference in the means of ratings between Finnish speakers and the non-Finnish speakers. Americans are included in the table and graphs to have a complete visual data set. The table below shows ratings for each speaker as well as both speakers combined, for each language feature: comprehensibility, accentedness, and valence. Values have been rounded to the hundredths place. The NSs had expectedly low ratings, while the NNSs had generally higher means in all categories.

Table 3. Mean response values for comprehensibility, accentedness, and valence for each speaker as well as both speakers of the same L1 combined. As per the rating scale, 1 = very easy to understand/no accent/very good; 7 = very difficult to understand/heavy accent/very bad.

	Comprehensibility	Accentedness	Valence
American 1	1.13	1.74	1.35

American 2	1.43	2.17	1.52
American total	1.28	1.96	1.44
Finnish 1	2.87	3.61	3.22
Finnish 2	1.52	3	2.35
Finnish total	2.19	3.31	2.79
Estonian 1	2.22	3.09	3.13
Estonian 2	1.83	3.78	3
Estonian total	2.03	3.44	3.07
Russian 1	1.74	3.13	2.61
Russian 2	2.91	3	3.17
Russian total	2.32	3.07	2.89
Italian 1	2.22	2.52	2.78
Italian 2	2.78	2.96	4.57
Italian total	2.5	2.74	3.68

Figure 7 reinterprets the data from Figure 4, this time showing the mean values instead of median values. These graphs are visual representations of Table 3, found above, and exact values can be found there. The data is presented similarly to the box plots above, with each color representing one language background, and in this case, grey representing both speakers from the same language background. The average comprehensibility rating for each American English speaker on their own as well as together is lower than any mean of the NNSs, at mean ratings of 1.13 and 1.43 for individual speakers, and 1.28 overall. Finnish speakers FIN1 and FIN2 have significantly different ratings, over a point apart at 2.87 and 1.52 respectively, with a combined mean of 2.19. Ratings for the Estonian speakers are closer between the two, at 2.22 and 1.83 for EST1 and EST2 respectively, and a combined mean rating of 2.03. Russian speakers RUS1 and RUS2 had mean ratings of 1.74 and 2.91 and a combined mean of 2.32, with similar rating gaps

to the Finns between speakers. The Italian speakers, ITA1 and ITA2, had the highest combined mean comprehensibility ratings, with 2.22 and 2.78 for each speaker, and a combined mean of 2.5.

Figure 7. Mean responses for comprehensibility for each speaker. Each color is a language background, and grey represents both speakers from the previous language background averaged together.

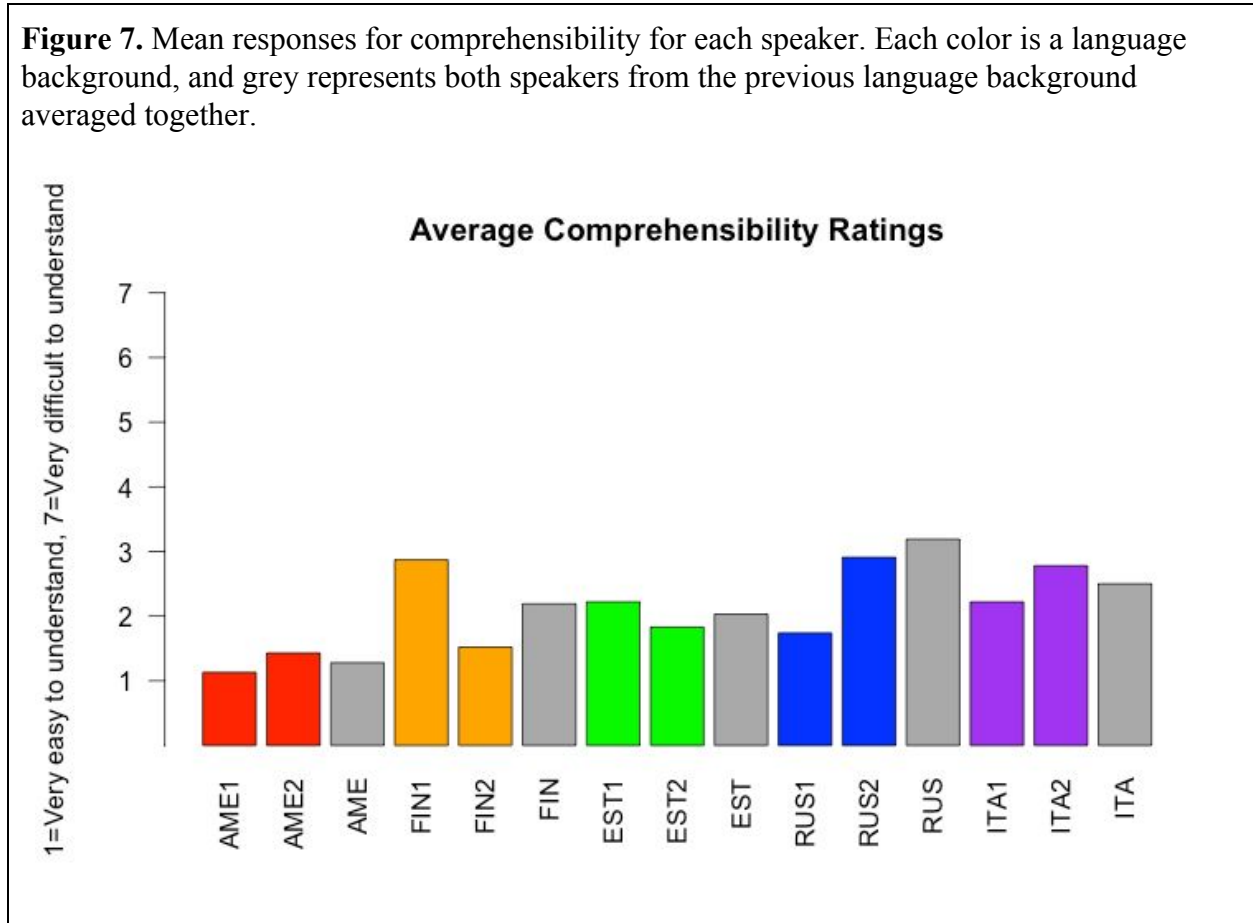


Figure 8 shows mean accentedness ratings for each speaker. Accentedness ratings were generally the highest among the three questions, with comprehensibility and valence lower. American English speakers AME1 and AME2 were given ratings of 1.74 and 2.17, with a combined average of 1.96. Finnish speakers FIN1 and FIN2 had comparably high ratings, at 3.61 and 3, and a combined mean score of 3.31. Estonians had the highest overall accentedness scores, at 3.09 for EST1 and 3.78 for EST2, and a combined score of 3.44. The Russian speakers had ratings of 3.13 and 3 for RUS1 and RUS2, and a combined score of 3.07. Finally, the Italian speakers had ratings of 2.52 and 2.96 for ITA1 and ITA2, and a combined mean accentedness score of 2.74.

Figure 8. Mean responses for accentedness for each speaker. Each color is a language background, and grey represents both speakers from the previous language background averaged together.

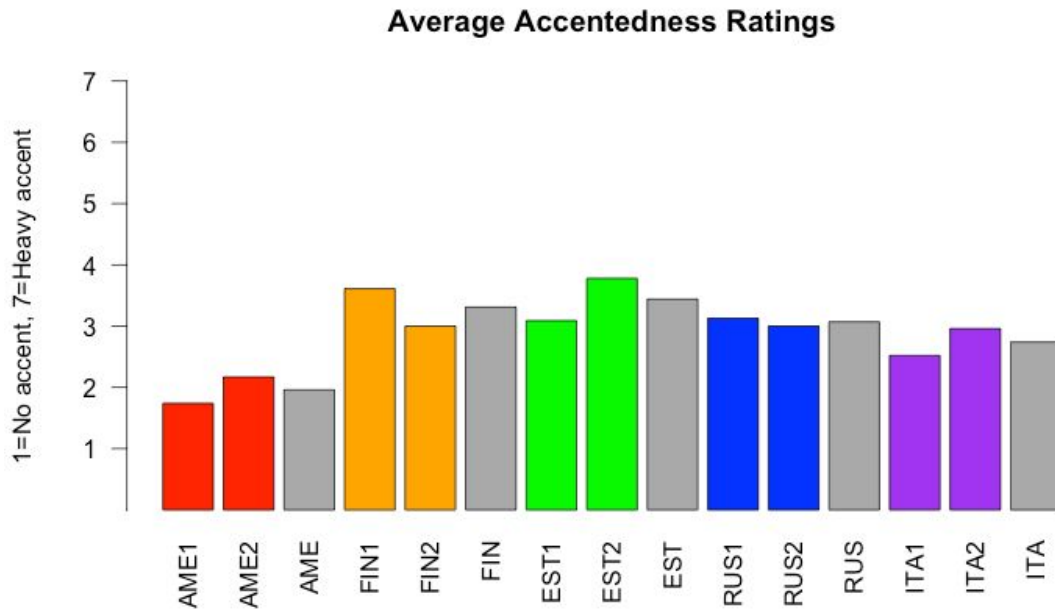


Figure 9 shows the mean valence responses for each speaker and language group. The American English speakers had a mean valence rating of 1.35 for AME1 and 1.52 for AME2, and a combined mean of 1.44. Finnish speakers had ratings of 3.22 for FIN1 and 2.35 for FIN2, and a combined mean of 2.79. Estonian speakers had ratings of 3.13 and 3 for EST1 and EST2 respectively, and an overall mean of 3.07. Russian speakers had ratings of 2.61 and 3.17 for RUS1 and RUS2, and a combined rating of 2.89. Finally, Italian speakers had the highest valence ratings, at 2.78 for ITA1 and 4.57 for ITA2. ITA2 is the only speaker to have a mean rating above the central “4” rating. Combined, their mean is 3.68.

Figure 9. Mean responses for valence for each speaker. Each color is a language background, and grey represents both speakers from the previous language background averaged together.

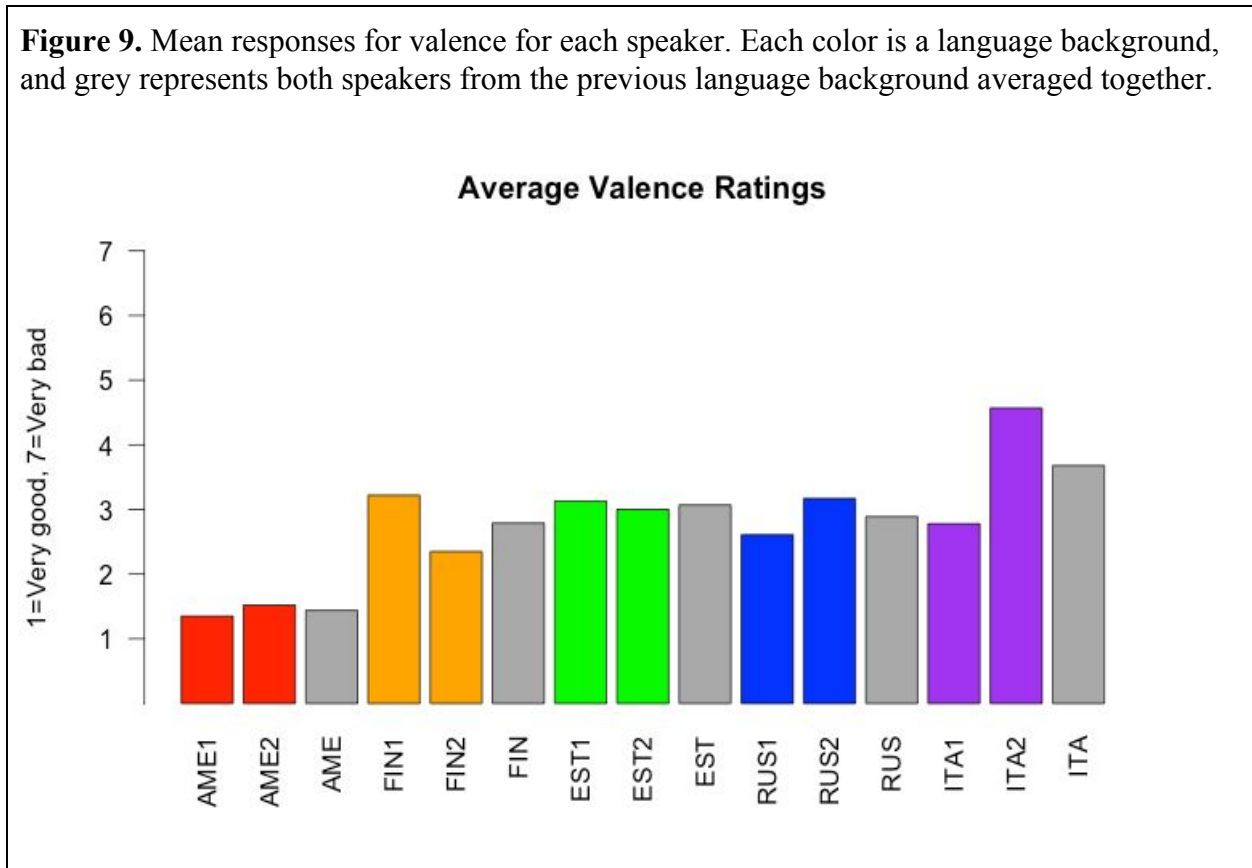


Table 4 shows the results of Welch’s t-tests, a two-tailed method. The t-tests were used to look at differences in the ratings of each group, so see if the means of the Finnish group were lower than the non-Finnish NNS group in a statistically significant way. As noted in Table 4 below, neither comprehensibility, at $t(75.06) = -0.56163$, $p = 0.576$; nor accentedness at $t(66.965) = 1.0328$, $p = 0.3054$ were significant. Valence was the only feature that was significantly different for Finns and Non-Finn NNSs, at $t(89.252) = -2.0845$, $p = 0.03997$, meaning that the participants rated Finnish speakers as having a more good accent compared to the Non-Finn NNSs.

Table 4. Welch’s t-test values for comprehensibility, accentedness, and valence ($\alpha = 0.05$).

	Comprehensibility	Accentedness	Valence
t-value	-0.56163	1.0328	-2.0845
df	75.06	66.965	89.252
p-value	0.576	0.3054	0.03997

7. DISCUSSION

7.1. Qualitative data

7.1.1. “Good” accents

Overall, the qualitative data was uniquely telling, as the participants had a wide variety of responses. To analyze the qualitative data, I parsed out the meaningful information from the responses and classified the responses simply into categories. As there were only 23 participants, I would like to clarify that none of my suggestions or analyses are definitive, but rather that they are general possibilities. Additionally, it is crucial to remember that these responses are culturally and linguistically situated and while they are specific to the Finnish context, the experiences of the participants may be highly varied and informed by their experiences. More in-depth discussion with participants or a broader cohort would be more effective for understanding the patterns and experiences of the listeners.

Generally speaking, goodness seemed to be associated relatively equally with accent features and comprehensibility features. It comes as no surprise that British and American standards were specifically referenced, as they are often most prevalent in English language pronunciation education in Europe (Henderson et al., 2012), and are often the most represented in textbooks in Finland (Kopperoinen, 2010). Tergujeff (2013) also notes that for both receptive and productive activities, the Finnish teacher surveyed used either standard British or American varieties most of the time. What did surprise me, however, was the specific mention of South

African and Dutch speech to be perceived as very good, both from the same participant. While more information would have to be gathered from the participant to understand where this perception came from, perhaps there is some positive personal or media experience with these language backgrounds that contributes to this response. In other specific groups, not having an accent, or not having an identifiable accent also came up. These answers may reflect the changing perceptions from the nativeness principle to the intelligibility principle in practice (Levis, 2005), where the participants acknowledge neutrality as a valid standard. These descriptions are generalized, and it would be interesting to know what neutrality means to these participants, as we know that everyone has an accent to some degree. Finally in this category, a small proportion of people noted that incorporating one's own first language phonological features into an accent is a good thing. This pushes the intelligibility principle even farther than neutrality, and proposes an accent identity and reclamation of foreign accented features as within acceptable limits. This is in direct contrast to McCrocklin & Link (2016) and Derwing (2003), where the vast majority of participants desired sounding native. This may be due to the fact that the above noted studies took place in Canada, an Anglophone country, and the NNSs in this study have identities that rely less on native-like approximation of English in European English as a lingua franca contexts. This is one reason it is important to consider the context, as my own results are more consistent with the results from Tergujeff's (2013) Finnish student cohort, who reported their goals as being related to comprehensible and fluent speech, not native-like speech.

General language features consisted of pronunciation features, with specific mention of segmentals at the forefront. This is often how pronunciation is traditionally perceived, with focus on segmental features. Although the degree to which segmentals contribute to comprehensibility is still being studied, Kang et al. (2010) did find that listeners from different backgrounds focused on different aspects of speech (e.g. Arabic listeners attended more to suprasegmental features while Vietnamese listeners attended to segmental features more). It would make sense to assume the same is true for Finnish speakers, and that the combination of their educational and personal experience with language may tailor their perceptions. Similarly, as Saito, Trofimovich, & Isaacs (2016) note that for NS listeners, segmentals were most related to accentedness ratings

for Japanese L2 English speakers, whereas the focus on segmentals in this qualitative data was proportionately very small.

Finally, comprehensibility responses made up the other half of total responses for goodness. Ease of understanding was the most common response in this category, which may be associated with listeners' desire for a lesser share of communicative burden when functioning as the listener (Lippi-Green, 1997). While Lippi-Green (1997) focused on the power balance between NSs and NNSs, it would not be out of the realm of possibility that listeners and speakers in general tend to try and accept equal communicative burden in their native language, and that some listeners feel frustrated when that balance is violated in a multilingual setting where they feel they are carrying more burden than they would normally. In the same category, speed was also noted, but with contradictory responses. Three of the four responses said that not-fast speech was good, while one said that fast speech was good. The association between goodness and not overly quick speech is likely related to processing time, allowing the listener to comprehend, accommodate, and process the speech of the speaker. Faster speech may be associated with overall fluency and prosody. Munro & Derwing (2001) found that the “ideal” speed for NNSs was 4.1 syllables per second, so perhaps the listeners had different personal experiences with speech that was more distant from this prototype. Finally, clarity of speech was mentioned. Clarity may help listeners pick out smaller details in speech, especially in terms of pronunciation and word comprehension to help ease communication, so it is expected for some respondents to see general speech clarity as a positive aspect.

7.1.2. “Bad” accents

In assessing what made an accent “bad,” the overall distribution in the categories was different than for what made an accent “good,” in that fewer accentedness answers were given. Even though many participants gave direct opposites to their goodness answers, the additional answers supplemented this category. There were a number of language background specific responses as in the first question, including Italian, Spanish, French, “fake” British, and Rally English. The first three suggestions may reflect linguistic or cultural stereotypes of these

countries (Rubin, 1990; Lindemann, 2005) as poor English speakers, or having strong first language phonology influences to their speech. There may also be media experience with these accents, which are known to affect people's perceptions. For instance, some research by Lippi-Green (1997) has shown that in Disney movies, those characters who have non-native accents were considered villains or bad characters when compared to their native-accented peers. She also uses French-accented English as an example of how cultural stereotypes can be perpetuated through language use, as those characters who have French accents in particular are characterized as having stereotypical French associations, like being associated with food or love (Lippi-Green, 1997). Outside of these specifics, there was also mention of a "fake" British accent. While British was used as an example for a "good" accent, it is interesting to consider what makes a British accent "fake," and how perceptions of a non-authentic reduplication of an accent, perhaps due to going outside of the commonly accepted boundaries of language features, may also be perceived negatively. Finally, Rally English was mentioned by one participant. This was an expected response, as Rally English was described above as one of the more well known negative cultural stereotypes of Finnish speakers of English. Finally, there were some more general conceptions of accents, like having a strong, unnatural, or identifiable accent. Unnatural accent may be related to something like the "fake" British accent, wherein authenticity is highly valued. Strong was another feature to be expected, as it may relate to the communicative burden property discussed above. Identifiableness was an interesting category, as it directly contradicted some participants' ideas that integrating a foreign accent was positive, and reinforces the nativeness principle (Levis, 2005). This may be linked to accent strength as well, as a speaker that is identifiably from a specific language background is likely using a variety of language features that are salient to the listener.

As for general accentedness responses, just as few responses were noted. Similarly to the "good" ratings, specific mention of segmentals did not play as large of a role as might be thought based off of Saito, Trofimovich, & Isaacs (2016). This may be something that is different between NSs and NNSs, at least when NNSs are considering what they believe they rely on for accentedness ratings. The lack of features in this section may also suggest that while overall

goodness is associated with a fine-tuning of minute accent features, the minimum threshold for badness is quite low, with only the most salient or distracting features being contributed.

Comprehensibility categorized answers made up a large minority of answers, with about half of the participants agreeing on being hard to understand as a contributing feature to accent badness. This, of course, is nearly the definition of comprehensibility. That is to say, it is clear that while there were more data in accentedness overall, the highest agreement among participants was that low comprehensibility contributed to a “bad” accent. This was paired with low clarity, a direct opposite of the answer for “good,” and high listener effort to understand the speaker. This returns to Lippi-Green’s (1997) conceptions of communicative burden, in that listeners do not want to feel like they are using more effort to understand than what they are used to.

The final category was small, and was made up of answers that did not directly line up with a description of a linguistic feature. These descriptions were things like “try hard,” “hard,” “sharp,” and “lazy.” These words of course must be associated with some linguistic feature, but are too indefinite to really analyze in depth. For example, it would be hard to know whether “sharp” could refer to significant intonation, pronunciation of plosives, overall tone, pitch, specific accent dislike, or any number of other features.

7.2. Comprehensibility

The final question was about comprehensibility, and directly asked what things contributed to good comprehensibility. Categories for answers were the same as for previous questions.

In this category, accentedness ratings made up 66% of the total, far more than either goodness or badness categories. This is in direct contrast with Kang, Rubin & Pickering (2010), who found that suprasegmental features— in this survey grouped into comprehensibility features— make up about 50% of variance in relation to comprehensibility ratings. However, this is in part due to the large proportion of answers in the general accentedness category.

Far fewer participants, however, listed specific accent features directly relating to comprehensibility. American and British English variants were mentioned specifically, which is to be expected, not only as these are the most common standards in language education, but also as these standard varieties dominate global media in English. It is likely that these are the most familiar varieties, and more experience with an accent is associated with more expertise in understanding that accent (Kozlowski, 2015). Other features in this category were neutral or natural speech and standard pronunciation. These features suggest an idea of authenticity as above. “Standard English pronunciation” as one participant wrote is an interesting concept, but can be assumed that this means adherence to a specific variety’s standardized rules, and is more associated with notions of the nativeness principle (Levis, 2005).

In general accent features, clear pronunciation was mentioned by nearly three quarters of the cohort, making it a significant feature in listeners’ perceptions of comprehensibility. This suggests that in Finnish participants’ self-ratings, pronunciation features make up a significant portion of perceived comprehensibility. I would hypothesize that answers relate less to a basic notion of pronunciation, like Giles (1970) where individual phonemes and words are of greatest concern, but that they more likely are associating pronunciation with Moyer’s (2013) examples of an “orchestra” of features, incorporating both segmentals and suprasegmentals. Interestingly, one participant also noted that familiarity made an impact in their perception of comprehensibility:

Also if you understand/know/recognice [sic] the native language of the speaker and recognise how they use their own language behind the accent, it is easier to understand compared to listening to a person who's [sic] own language you are not used to hearing.

This would tend to support the notion that familiarity does have an effect for at least some Finnish listeners, as supported by Kozlowski (2015).

For comprehensibility, far fewer features fell into these categories representing suprasegmentals, with speed, speech clarity, and intonation being noted, which follows a similar pattern from the goodness question.

In the nondescript category, having a high language level, having a broad vocabulary, and having a “smooth” sound were noted. The first two responses suggest a general leveling, that more advanced language users are more likely to be comprehensible. These are generally sensible statements, because as language users learn new things, they tend to improve overall in multiple areas of language due to exposure and experience. The word “smooth” is relatively difficult to decipher, and may reference fluency, clarity, or a combination of features that lessens the communicative burden.

7.3. Quantitative data

In the quantitative category, the first measures category, interrater reliability, was incredibly low. Other similar studies that use Fleiss kappa as an interrater reliability measure like Munro & Derwing (1999) and Munro et al. (2006) have ratings above $k=0.90$, so having such weak agreement is highly unexpected. In terms of accent, comprehensibility, and valence ratings, there were some statistically significant effects. Previous research (Varonis & Gass, 1982; Munro & Derwing, 1999) has shown that for NS listeners, comprehensibility and valence were highly correlated. The data collected for this study showed moderate correlations for seven out of ten speakers on those grounds. This data also showed six out of ten speakers were related with accentedness and valence. These results match, to some degree, with Trofimovich & Isaacs’ (2012) study of native speakers, wherein both accentedness and comprehensibility measures contributed to judgements of L2 English speech. This may reflect a pattern among listeners of all language backgrounds of having a complex understanding of NNSs’ speech. Additionally, this likely reflects results from Foote & Trofimovich (2016), in that Finnish listeners may have different balances of these categories depending on the language background of the speaker. More in-depth looks at these speaker-listener language pairs would reveal if there was indeed a pattern.

One other feature of note was that ratings of comprehensibility between the two NSs were rather different, as well as comprehensibility scores between NNSs with the same L1 background, showing no overall pattern. However, I have no reason to believe that this is

particularly unusual. Munro & Derwing (1999) explain this variation, in that “nonpathological native speech may vary in comprehensibility because of such factors as rate of speech, speech clarity, voice quality, and word choice.” While this explanation is certainly valid for the NSs, it is likely that similar features, perhaps along with accentedness features, are also as a result of natural variation in individuals’ speech as well as possible recording quality.

While individual speaker variation between the speakers was expected, the variation in ratings within speakers was not. As mentioned previously, there was overall low interrater reliability, which suggests a more complex phenomenon than I was able to measure. It may be that different listeners were relying on different measures when assessing accentedness, comprehensibility, and overall valence. This was also reflected in the qualitative data, wherein there was a large variety of responses to each question. To that end, while the quantitative data can of course provide some insight into the experiences of the listeners, this is a major limitation and the results must be contextualized with this poor reliability in mind.

8. CONCLUSION

The purpose of this thesis was to contribute to the growing body of research in sociolinguistics and accent attitudes by addressing the Finnish context and its relationships with other speakers of English. Three main research questions were addressed in this study: Firstly, what features do Finnish listeners perceive as important when quantifying a “good” foreign accent? Based on previous research in other populations, comprehensibility, accentedness, and valence were selected as possible factors. Both qualitative and quantitative data showed approximately even preference for both comprehensibility and accentedness, with mixed data from the quantitative portion. Secondly, do Finnish listeners predominantly link comprehensibility, accentedness, both, or neither to determine a “good” accent? Both qualitative and quantitative results show a mixed effect, with both features contributing to some extent to valence, with no overwhelming or statistically significant preference for one over the other. Results showed that individual speakers had different patterns of CV or AV significance. Finally, is there evidence of a first language preference effect? The quantitative data showed a preference

in the goodness of accent for Finns when compared to other NNSs, but no effect for comprehensibility or accentedness preference.

For the qualitative data overall, both goodness and badness of accent answers were approximately even with respect to accentedness and comprehensibility features. With that said, overall accent goodness seemed to skew more toward accentedness than overall accent badness. In part, this is due to a notable consensus about low comprehensibility being associated with badness. This may suggest that there are different thresholds for goodness and badness within accent and comprehensibility. For example, it may be that to go from “bad” to somewhere in the middle, comprehensibility is key, but at higher levels, to go from fine to “good,” accentedness is the more crucial feature. This data shows the complexity of the phenomenon, even among a relatively small cohort. Additionally, in the comprehensibility question, a large proportion of answers related to accentedness features. This supports the idea that these features are interrelated.

The quantitative data were mixed in terms of outcome. There were some statistically significant effects for both accentedness and comprehensibility in relation to valence, but no generalizable patterns between speakers of the same language. This may suggest that there are certain talker features that are contributing to listeners’ understanding that were not measured and not controlled in this study. Additionally, interrater reliability was poor, so the results are volatile at best. A larger and more carefully curated population could have helped this issue. Finally, there was some evidence that Finnish listeners showed a preference for Finnish speakers in valence. This supports previous research as discussed, and shows that even within a small cohort, there is some evidence for the own accent preference.

More research with more participants are needed in this area to help more thoroughly answer the questions above and questions like these, and to understand how Finnish listeners contribute to our overall understanding of second language accent attitudes, sociolinguistic and sociocultural understandings of language and perception, and the global influence of English and English education in specific contexts.

This study as a whole has a number of possible directions to move in, and improvements for this study as a whole. To directly address this study, there are a number of things I would do

differently: firstly, I have no data on my participants. Although there were age, native language, and education criteria, I imagine that the responses were skewed due to the methods of gathering participants. As I used mainly university channels of communication, it is likely that my cohort was younger than the theoretical average, more educated than the average Finn, and may have more experience with a variety of accents through contact with international students like me. Of course, there was also no control for the participants' listening environment, how much time they spent doing the survey and listening to the samples, or similar environmental factors. Secondly, the number of participants was also unideal. Due to time restrictions, the data needed to be analyzed as it was. An ideal number would of course be as high as possible, but a goal number of 50 responses likely would have been analyzed for more solid and telling patterns. Finally, the Speech Accent Archive, while an invaluable tool, does have variation in the quality of speech samples available. This meant that not all sound tokens may not have had equal quality and clarity due to the number of available speech files. Additionally, since I did not use any pre-tests to vet the speakers, the listeners may have perceived the speakers as being highly dissimilar or of varying language ability, which may have impacted the ratings.

As for future directions, there is room to do more targeted studies toward Finns' perceptions of a variety of accents, full matched-guise studies to assess innate bias toward certain language backgrounds, or even more ethnographic studies to assess speakers' lived experiences of accent in a variety of situations. For me personally, the qualitative data was very compelling. I imagine that a mixed methods study with an integrated pre- and post-survey interview would greatly inform a greater quantity of quantitative data.

All in all, this research, while inconclusive due to low reliability among a relatively small group of participants for the task presented, found some evidence that reflects current knowledge of accent perception and proves some insights into the Finnish context. Listener preferences for both accentedness and comprehensibility were found in both quantitative and qualitative data with no strong skew toward either. Some statistically significant relationships were found, both between accentedness and valence as well as comprehensibility and valence, with no particular language pattern. Additionally, some evidence of a language preference effect was found for valence of Finnish listeners rating Finnish speakers.

9. APPENDICES

9.1. Survey questions

The survey questions were seen in “blocks” of questions, denoted below in boxes.

CONSENT

By clicking the button below, I acknowledge that I am between 18-30 years old, speak Finnish as a native language, and do not or have not studied English as a major subject in higher education.

I understand that my data is anonymously collected, and that my answers will not be used for any reason other than research.

For this survey, you will first answer a few short questions about your opinions. Then, you will listen to some sound files and rate some aspects of the speech. Please make sure that your sound is on and that you are in a quiet place.

The survey should take about 15 minutes.

Please click here you you accept the terms:

I understand, and I agree

QUALITATIVE QUESTIONS

1. Imagine a non-native speaker of English. In your opinion, what does a good accent sound like?
[short answer space]
2. Imagine a non-native speaker of English. In your opinion, what does a bad accent sound like?
[short answer space]
3. Imagine a non-native speaker of English. In your opinion, what makes an accent easy to understand?
[short answer space]

Ten SPEAKER blocks were presented in random order. Each had a replayable, pausable audio file. The 1-7 scale was presented as individual bubbles with numbers above each one, with qualifiers at each end of the scale.

SPEAKER

[audio file]

Please listen to the speaker and then answer the questions below.

1. How easy was it to understand the speaker?

[1=Very easy 2 3 4 5 6 7=Very difficult]

2. How strong was the speaker's accent?

[1=No accent 2 3 4 5 6 7=Very heavy accent]

3. How good was the speaker's accent?

4.

[1=Very good 2 3 4 5 6 7=Very bad]

9.2. Table 4 - Speaker file data

Native Language	Other Languages Spoken (not including English)	Age	Birthplace	Length of Residence in Anglophone Country (in years)
English	none	30	West Jordan, UT, USA	30
English	none	19	Libertyville, IL, USA	19
Finnish	Swedish	20	Helsinki, Finland	0
Finnish	Swedish, Estonian	46*	Tampere, Finland	0.3
Estonian	German, Russian	27	Viljandi, Estonia	0
Estonian	German, Russian	43*	Parnu, Estonia	0.1
Russian	Hebrew, Italian	23	Moscow, Russia	0
Russian	German	30	Moscow, Russia	0
Italian	German	23	Teramo, Italy	0
Italian	none	24	Rome, Italy	0.1

* = indicates not fitting ideal criteria

9.3. Table 5 - Responses to qualitative questions

Each separated section represents one user's response; questions are denoted above the answers. Answers have not been edited, so any errors are the writers' own.

- Imagine a non-native speaker of English. In your opinion, what does a good accent sound like?

Natural. Can reflect British or American accents if fluent.
Easy to understand, not fast, accurate pronunciation of consonants is important.
Easy to understand, not really noticeable
When you understand every word
British accent
Close to american accent
Neutral, with cleanly and clearly pronounced words
It sounds almost like a native speaker's, but gives a hint of your own native language.
Smooth and quick
Easy to understand in bigger picture
All accents are beautiful in their own way and just different variations to the language, but I'd still say it's important to pronounce words at least somewhat recognizably. It's difficult to converse with someone, if their speech is difficult to understand due to their accent. It's perfectly fine if there are strange words here and there, but if it is more like two people speaking two different languages it's a bit problematic at least when it comes to being able to have conversations and understand each other.
Neutral. One can't locate the accent.
Clear to understand
It has clear pronunciation and the rythm of speech is not too fast.
Neutral, banal. Think of someone from South Africa or the Netherlands.
Easy to understand, no noticeable foreign accent
Good accent sounds like English. I don't mind hearing an accent in speech as long as I can make out at least most of what they say.
Exact pronunciation, professional and fluent tone of speaking
It sounds like an american accent
Clear, easy to understand and comes easy for the speaker
It lacks the characteristics of their native language. The intonation is correct
Clear, with vowels and consonants pronounced as close to British/American English as possible
It sounds easy to understand and it is not fast and the pronunciation is clear.

- Imagine a non-native speaker of English. In your opinion, what does a bad accent sound like?

Hard to understand and "try-hard"
Very fast, letter pronunciation is incorrect.
So strong that it is hard to understand
It's something you don't understand very well
Italian accent
Unnatural (pretended) british english
It sounds kind of lazy in a way, like they are not even trying to pronounce the words correctly or they're trying to fit english words into their native language and way of speaking
It is really difficult to follow or understand.
Sharpish
Difficult to understand the whole topic of discussion
Difficult or impossible to follow what they are saying.
Very strong and some words are pronounced so strangely that one can not understand the speaker
Fast and not undersandable
If the speaker's pronunciation is affected by their own language too much, it is may be difficult to understand.
Oversaturated with their native language's pronunciation to the point of being hard to understand.
You have to focus to understand the words. It is easy to guess where the speaker is from.
If I have to make an effort to figure out what the speaker said, it's a bad accent.
Slow and little bit "hard"
It sounds like rallienglish
So thick that words are hard to hear/understand, when words become "mushy" or like porridge.
Wrong intonation. Wrong pronouncing. Either rally english or french accent
Unclear, too many "correct" sounds are substituted for the way they would pronounce their own language
It is hard to speak English with the person if you can hear his/hers mother language while she/he speaks.

- Imagine a non-native speaker of English. In your opinion, what makes an accent easy to understand?

Clear pronunciation
The speed of speech, word pronunciation.
Clear prononciation, being neutral
If you can speak the way everybody understands you

Clear articulation
Clear articulation
Clear and neutral pronunciation is probably the most important. If accent includes the way the speaker chooses to use words in a sentence and their understanding of grammar, then that makes a difference too.
It is slow enough, the tone doesn't go too up or down.
Smooth and not too fast
Good pronunciations
Clear(ish) pronunciation, not mumbling, moderate velocity of speech
No strong intonations
Pronunciation
It's easy to understand when the pronunciation is similar to standard English pronunciation.
A native language that doesn't intervene (much) with a speaker's English.
Words are separated from each other by stress or pauses. Vowels sound about right.
The words must be pronounced mostly right. For example if the speaker struggles to pronounce some consonants, they should cover them with consonants that they are comfortable with and also sound roughly the same. In my opinion it makes an accent harder to understand if the vowels aren't pronounced exactly or mostly right, consonants aren't that important.
Speaking with a good speed, exact but natural pronunciation
It's easy to understand if the speaker can articulate well and has a wide vocabulary.
Usually the better you try to pronounce all the letters, the easiest it is to understand (compared to skipping a lot of vocals or changing them). Also if you understand/know/recognize the native language of the speaker and recognize how they use their own language behind the accent, it is easier to understand compared to listening to a person who's own language you are not used to hearing.
As much american or english as possible. E.g. Spanish and French accents are probably the hardest to understand as they are unclear and the letters in the words are swallowed.
Clarity, with vowels and consonants pronounced as close to British/American English as possible (same answer as "good" accent because good and easy understand go hand in hand)
If you try to speak slow and you say words clearly (learn the differences between English and your mother languages). Don't highlight anything while you speak if you are not sure who the word should be pronounced.

10. REFERENCES

- Bent, T., & Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *Journal of the Acoustical Society of America*, 114 (3), 1601-1609.
- Birdsong, D. (2007). Nativelike pronunciation among late learners of French as a second language. In S. Bohn, & M. Munro (Eds.), *Language Experience in Second Language Learning: In honor of James E. Flege* (pp. 99-116). Amsterdam: John Benjamins.
- Berner, B. (n.d.). Why Do Some People Have an Accent? The Linguistic Society of America. Retrieved from <https://www.linguisticsociety.org/content/why-do-some-people-have-accent>
- Crystal, D. (1985) How many millions? The statistics of English today. *English Today*, 1, 7–9.
- Derwing, T. M. (2003). What do ESL students say about their accents? *The Canadian Modern Language Review*, 59, 547– 566.
- Derwing, T. M. & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly*, 39, 379–97.
- Derwing, T. M., & Munro, M. J. (2009). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching*, 42, 476–490.
- Derwing, T. M. & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. Amsterdam: John Benjamins.
- Dragojevic, M., Giles, H., Beck, A.C. & Tatum N.T. (2017) The fluency principle: Why foreign accent strength negatively biases language attitudes, *Communication Monographs*, 84:3, 385-405.
- Eberhard, D., Simons, G., and Fennig C. (eds.). 2019. *Ethnologue: Languages of the World*. Twenty-second edition. Dallas, Texas: SIL International.
- English First (2019). *EF English Proficiency Index: A Ranking of 100 Countries and Regions by English Skills*. PDF file. Retrieved from: <https://www.ef.com/~/media/centralefcom/epi/downloads/full-reports/v9/ef-epi-2019-english.pdf>

- Emotional valence. In *APA Dictionary of Psychology*. Retrieved from <https://dictionary.apa.org/emotional-valence>
- Episcopo, S. (2009). *Non-native speaker attitudes toward non-native English accents* (Unpublished master's thesis). University of Texas, Austin, Texas.
- Flege, J. E., Munro, M. J., & Mackay, I. R. A. (1995) Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America*, 97, 3125–3134.
- Foote, J., & Trofimovich, P. (2018). Is it because of my language background? A study of language background influence on comprehensibility judgments. *Canadian Modern Language Review*, 74, 253–278.
- Gallardo del Puerto, F., García Lecumberri, M-L. & Gómez Lacabex, E. (2015). The assessment of foreign accent and its communicative effects by naïve native judges vs. experienced non-native judges. *International Journal of Applied Linguistics*, 25 (2), 202–224.
- Gershwin, G. & Gershwin I. (1937). Let's Call the Whole Thing Off [Recorded by Fitzgerald, E. & Armstrong, L]. On *Ella and Louis Again*. New York, New York: Verve Records.
- Gilbert, J. B. (2010). Pronunciation as orphan: What can be done? *TESOL SPLIS*, 7(2), 1-5.
- Giles, H. (1970). Evaluative reactions to accents. *Educational Review*, 22, 211-227.
- Gluszek, A., & Dovidio, J. F. (2010). The way they speak: A social psychological perspective on the stigma of nonnative accents in communication. *Personality and Social Psychology Review*, 14, 214–237.
- Gluszek, A. & Hansen, K. (2016). What does speaking with a foreign accent mean? Content analysis of newspaper articles. 10.13140/RG.2.1.1912.3281.
- Henderson, A., Frost, D., Tergujeff, E., Kautzsch, A., Murphy, D., Kirkova-Naskova, A., Waniek-Klimczak, E., Levey, D., Cunningham, U. & Curnick, L. 2012. The English pronunciation teaching in Europe survey: Selected results. *Research in Language* 10(1), 5–27.
- Kachru, B. (1992). *The Other Tongue: English across cultures*. University of Illinois Press.

- Kang, O., Vo, S. C. T. & Moran, M. K. (2016). Perceptual judgments of accented speech by listeners from different first language backgrounds. *The Electronic Journal for English as a Second Language*, 20, 1–24.
- Kozłowski, A. (2015). The Influence of Accents on Social Perception. *Inkblot*, 4, 12-16.
- Leppänen, S., Pitkänen-Huhta, A., Nikula, T., Kytölä, S., Törmäkangas, T., Nissinen, K., Kääntä, L., Räisänen, T., Laitinen, M., Koskela, H., Lähdesmäki, S. & Jousmäki, H. (2011). National Survey on the English Language in Finland: Uses, meanings and attitudes. *Studies in Variation, Contacts and Change in English* (5). Helsinki: VARIENG. <http://www.helsinki.fi/varieng/series/volumes/05/evarieng-vol5.pdf>
- Lippi-Green, R. (1997) *English with an accent: Language, ideology, and discrimination in the United States*. London: Routledge.
- Lindemann, S. (2005). Who speaks broken English?: US undergraduates' perceptions of nonnative English. *International Journal of Applied Linguistics*, 15 (2), 187-212.
- Major, R., Fitzmaurice, S., Bunta, F., & Balasubramanian, C. (2002) The effects of nonnative accents on listening comprehension: Implications for ESL assessment. *TESOL Quarterly*, 36, 173–190.
- Major, R. (2007). Identifying a foreign accent in an unfamiliar language. *Studies in Second Language Acquisition*, 29, 539–56.
- McCrocklin, S. & Link, S. (2016). Accent, Identity, and a Fear of Loss?: ESL Students' Perspectives. *The Canadian Modern Language Review / La revue canadienne des langues vivantes* 72(1), 122-148. University of Toronto Press.
- Mitterer, H., Eger, N.A., & Reinisch, E. (2020). My English sounds better than yours: Second-language learners perceive their own accent as better than that of their peers. *PloS One*, 15(2), E0227643.
- Moyer, A. (2013). *Foreign accent: The phenomenon of non-native speech*. Cambridge: Cambridge University Press.
- Munro, M., & Derwing, T. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech*, 38, 289–306.

- Munro, M. J., & Derwing, T. M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 49, 285–310.
- Munro, M. J., & Derwing, T. M. (2001). Modeling perceptions of the accentedness and comprehensibility of L2 speech: The role of speaking rate. *Studies in Second Language Acquisition*, 23, 451–468.
- Munro, M.J., Derwing, T.M., & Morton, S.L. (2006). The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition*, 28(1), 111–131.
- Opetushallitus [Finnish National Agency for Education] (2014). *Perusopetuksen opetussuunnitelman perusteet [Core curriculum for basic studies]* 2014. Helsinki. Retrieved from: http://www.oph.fi/download/163777_perusopetuksen_opetussuunnitelman_perusteet_2014.pdf.
- Opetus- ja kulttuuriministeriö [Ministry of Education and Culture] (2018). *A1 -kielen opetuksen varhentaminen [Making A1 language teaching earlier]*. PDF file. Retrieved from: <https://minedu.fi/kielten-opetuksen-varhentaminen>
- Rubin, D. L. (1990). Effects of accent, ethnicity, and lecture topic on undergraduates' perceptions of non-native English-speaking teaching assistants. *International Journal of Intercultural Relations*, 14, 337-353.
- Sharatol Ahmad Shah, S., Othman, J. & Senom, F. (2017). The pronunciation component in ESL lessons: Teachers' beliefs and practices. *Indonesian Journal of Applied Linguistics*, 14.
- Smith, L., & Nelson, C. (2006). World Englishes and Issues of Intelligibility. In B. Kachru, Y. Kachru, & C. Nelson, *The Handbook of World Englishes* (pp. 428- 445). Malden: Blackwell Publishing.
- Statistics Finland (2020). Biggest nationality groups 2019 [Graph]. Retrieved from http://vertinet2.stat.fi/VERTI/graph/viewpage.aspx?ifile=quicktables/Maahanmuuttajat/kansa_2&lang=1&gskey=2
- Suomen kieltenopettajien liitto ry (SUKOL) [Federation of Foreign Language Teachers in Finland]. *Tilastotietoa kielivalinnoista [Statistics about language choices]*. n.d. Retrieved from: https://www.sukol.fi/liitto/uutiset/tilastotietoa/tilastotietoa_kielivalinnoista

- Trofimovich, P., & Isaacs, T. (2012). Disentangling accent from comprehensibility. *Bilingualism: Language and Cognition*, 15(4), 905-916.
- Uchida, Y., & Sugimoto, J. (2019). Non-native English teachers' confidence in their own pronunciation and attitudes towards teaching: A questionnaire survey in Japan. *International Journal of Applied Linguistics*, 30(1), 19-34.
- Varonis, E. M., & Gass, S. (1982). The comprehensibility of non-native speech. *Studies in Second Language Acquisition*, 4, 114–136.
- Weinberger, Steven. (2015). *Speech Accent Archive*. George Mason University. Retrieved from <http://accent.gmu.edu>