Lexical Variation, Target Preference and Language Attitudes with Respect to American and British Englishes: A Mixed Methods Study of Undergraduate Students of English in Finland

Ossi Tuominen

MA Thesis

English, Degree Programme for Language Learning and Teaching

School of Languages and Translation Studies

Faculty of Humanities

University of Turku

May 2020

UNIVERSITY OF TURKU

School of Languages and Translation Studies / Faculty of Humanities

TUOMINEN, OSSI: Lexical Variation, Target Preference and Language Attitudes with Respect to American and British Englishes: A Mixed Methods Study of Undergraduate Students of English in Finland

_____

The legitimacy of the English used by non-native speakers is an is an ongoing debate. This debate is intertwined with the question of which variety of English should be used as the model in formal education. This study intends to provide information for both discussions by examining the relationship of the performance and the attitudes of Finnish undergraduate students of English (N = 40) with respect to American English (AmE) and British English (BrE). To this end, three variables are examined: 1) *varietally distinct lexical choices*, 2) *target preferences*, 3) *language attitudes* (on the dimensions of *status and competence* and *social attractiveness*). These variables enlighten the position of the two major varieties of English in Finland, and their interrelations provide information on whether the students use these varieties as instruments of identity construction.

Lexical choices were studied through elicited translation, whereas target preference and language attitudes were measured on 5-point Likert scales. Lexical choices were generally American, but BrE was slightly favoured as a target variety. Nevertheless, the two variables correlated significantly. BrE was evaluated clearly higher on the scale of status and competence and almost equally with AmE on the scale of social attractiveness. Perceived social attractiveness proved to be a significant predictor of target preference, but status and competence did not. Correlation of social attractiveness and target preference, along with the conducted interviews, suggest that the students' target preferences reflect their identity. The correlation of the target choice and actual production indicates that the students successfully use AmE and BrE to manifest that identity. The results suggest that formal English tuition should support students in attaining their self-selected target variety. They also support the legitimacy of non-native English by demonstrating that the lexical variation of non-native speakers can be socially motivated.

Keywords: EFL, ELF, Lexical Variation, Language attitudes, Sociolinguistics, SLA, Identity, American English, British English

Table of contents

Tables and Figures

Tables

Figures

Abbreviations
AmE: American English
BrE: British English

1 Introduction

First impressions are important. This oft-repeated adage holds true for the subjective evaluation of almost anything, humans and theses included. However, it subtly undermines the importance of second, third and nth impressions. People readjust their first impressions in accordance with another proverb: never judge a book by its cover. In spoken interaction, the "cover" tends to correspond to pronunciation due to its frequency-induced salience (Watt 2007, 6). Almost all, if not all, words include sounds that are subject to sociolinguistic variation, i.e. they give subtle hints of the speaker's geographic origin, social status, education, etc. (Llamas, Mullany and Stockwell 2007, xv). Conversely, most words used in this sentence, for example, hardly carry such connotations.

Yet, beyond the "cover" of phonology, words can also work as sociolinguistic indicators. For example, "excuse me?", "pardon?" and "what?" all have essentially the same pragmatic meaning in a certain context, but people make inferences of their collocutor's social status based on them (Fox 2004). Perhaps more obviously, there is a plethora of words that are typical of a certain geographical variety. If someone talks about taking "a lift to their flat" instead of "an elevator to their apartment", one assumes the person British rather than American.

In addition to categorising others, words are used to construct one's own *social face*, or "a kind of mask, designed on the one hand to make a definite impression upon others, and on the other to conceal the true nature of the individual", as defined by Carl Jung (1953, 190) In similar vein, Ludwig Wittgenstein, a contemporary of Jung, coined the term *language game* to refer to the ways in which social reality is constructed by varying the use of language in different settings ([1953] 1968). In other words, people construct an identity through the linguistic choices that they make. I am interested in whether this also applies to the use of foreign languages and, more specifically, foreign words.

Since Jung and Wittgenstein, there has been a myriad of sociolinguistic studies examining the ways in which social reality and language intertwine. Unfortunately, vocabulary has not received much attention. Labov (1983, 29) lamented this neglect as he considered the *word* "the most central element in the social system of communication". Almost forty years later, there are still very little studies on lexical variation and the topic is routinely omitted from introductory books of sociolinguistics.

Another rather neglected area of variationist studies is L2 variation, or variation in languages other than one's native language. Durham (2014, 14) ascribes this to sociolinguists not having been interested in non-native speakers and researchers of second language

acquisition not having been interested in *sociolinguistic competence*. In recent years, the two fields have converged to a degree, and the sociolinguistic competence of non-natives, i.e. their approximation of the native sociolinguistic patterns (Durham 2014, 21), has received some attention. However, I find the notion that L2 users' sociolinguistic competence is determined by their approximation of native patterns problematic. There is a great deal of evidence of L2 speakers purposefully deviating from natives in order to convey their social identity (Hansen Edwards 2008, 272). Unfortunately, this evidence stems almost exclusively from immigrants, i.e. L2 users living in L2 surroundings.

Given that, at present, English is mostly spoken outside a native setting (Seidlhofer 2011, 2), it is remarkable that *English as Foreign Language* (EFL), i.e. English learned outside a native setting (Quirk 1972 3-4), has not received attention as an instrument of identity construction. The reason is clear, however. It has not been considered a possibility. According to Durham (2014, 18), the general belief among sociolinguists is that L2 users' variation is either free or L1 influenced. Even native speakers' linguistic identity was considered a mere result of their social affiliations until the 1980s (Dyer 2007, 103).

The most obvious socially motivated decision with regard to EFL is the choice of the variety used as a model of one's own English idiolect. If this choice is based on the attitudes one associates with the various varieties of English, it is an act that constructs one's self-identity: "I think variety X is Y. I want to be Y, therefore, I want to sound like X." Sociolinguists, however, are more interested in the actual manifestations of identity, that I alluded to earlier with reference to Jung and the social face. This requires something tangible for others to see or, better yet, hear. In other words, Y must be conveyed to others by actually performing X.

I said earlier that almost all evidence of constructing collective identity by linguistic choices in an L2 stems from studies on immigrants. The reason for the modifier "almost" is that I found one study relating attitudes, target choice, and actual performance in an EFL setting. Rindal (2010) examined whether Norwegian upper secondary school students orient their pronunciation to their preferred target variety in a statistically significant way. The result was that, despite generally American phonology, they did (Rindal 2010, 247). They also said they make evaluations of their peers based on the American and British features in their pronunciation (Rindal 2010, 252) and produced significantly more BrE phonemes in a formal situation (Rindal 2010, 248), because they considered it the more formal variety. In sum, they clearly use English to construct their social identity.

The present study is quite similar to that of Rindal (2010), but, unfortunately, the scope of it did not permit the setting of formal and informal contexts. Furthermore, the participants are undergraduate students of English and, instead of phonology, I examine varietally distinct lexemes by means of an elicited translation task. Throughout the paper, I use *lexeme*, *word* and *lexical choice* interchangeably, but it is technically lexemes that I am talking about, i.e. the group of words related through inflection (e.g. *study* and *studies* are two *words* of a single *lexeme*). This is an important distinction as I do not intend to make any claims about phonology, which is an inherent part of knowing a *word*. Conversely, *lexeme* is more abstract and does not entail a phonological aspect (Haspelmath 2002).

As noted above, lexemes/words are understudied in sociolinguistics, but my main motive for studying them in this paper, instead of the more common phonological focus, is that the result is much more uncertain. After all, given that English proficiency in Finland and Norway is roughly equal (Education First 2019) and the participants of the present study are on a higher educational level, one would expect them to be at least equally adept at incorporating desired varietal features into their English idiolect. However, I suspect that varietally distinct vocabulary is more difficult to emulate due to the abovementioned salience of phonology. Consequently, a positive correlation between lexical choices and target preference would constitute an even stronger argument for EFL users' sovereign agency as English users.

I also examine whether attitudinal factors correlate with target preferences. Both variables are obtained through evaluations of American English (AmE) and British English (BrE) on a Likert scale. I hypothesised that Finnish undergraduate English students (N = 40) would follow the European tendency of associating AmE with *social attractiveness* and BrE with *status and competence* (e.g. Rindal 2010; Ladegaard and Sachdev 2006). However, it was much more uncertain, which dimension would have a greater impact on target preferences. To establish some tentative causal relationship to both correlations (lexical choices / target preference; language attitudes / target preference), I discuss the results from the point of view of Labov's (2001, 16) principles of language change and complement the quantitative data with interviews of the participants of the present and previous studies.

If the participants of this study are capable of using varieties of English as building blocks of their social identity, it would have repercussions on the didactic decisions regarding the models of English used in formal education in Finland. Such a result could also be used as an argument in favour of the legitimacy of *English as a Lingua Franca* (ELF), or English between speakers of different native languages (Seidlhofer 2011, 7), which is an ongoing debate.

Before going into details on the present study, I dedicate two chapters to previous research in order to provide the reader with the necessary tools for placing the present study in its proper context. In Chapter 2, I first give an overview of sociolinguistic variation and then proceed to the more specific areas of interest of the present study, viz. the impact of language attitudes on variation in general (subsection 2.1) and among EFL users in particular (subsection 2.2). I conclude the section with an overview of lexical variation and what is known about it in EFL contexts, particularly in Finland (subsection 2.3). In section 3, I justify limiting the study to AmE and BrE and provide a historical context for the different attitudes they evoke (3.1) as well as their different lexicons (3.2).

I start the empirical part of the paper by describing and justifying the methods of data collection and sampling in section 4. The obtained data are presented and analysed in section 5. Finally, in section 6, I conclude with an overview of the study and its possible practical implications, as well as suggest some further avenues of research that would shed more light on the issues surrounding the use of English in the pluricentric world of today.

2 Sociolinguistic Variation

As stated in the introduction, one fundamental characteristic of language is the potential to say things in different ways. The traditional paradigm has only acknowledged phonological, lexical and morphosyntactic variation (Meecham and Rees-Miller 2001), but in recent years variationist research has been extended also to pragmatics (Félix-Brasdefer and Koike 2012) and semantics (Hasan and Webster 2009). Much of this variation stems from socially motivated choices, conscious or not, which convey information on the speaker's age, gender, class, etc. (Llamas, Mullany and Stockwell 2007, xv). This connection of variation and social reality is studied by the field of sociolinguistics. Indeed, some academics even set the bounds of the field there and describe sociolinguistics as "the study of language variation" (Wardhaugh 2010, 11).

Such a view is criticised by others for ignoring the sociological side of the field, i.e. not paying enough attention to what the value judgments associated with the possible variables are based on. For example, Fairclough (2015, 6) argues that "Sociolinguistics is strong on 'what?' questions (what are the facts of variation) but weak on 'why?' questions (why are the facts as they are?)". Beyond establishing "the facts of variation", the present study does ask "why is one variety favoured over the other?" with regard to both lexical choices and target preference. For example's sake, let us say that a speaker mainly uses BrE vocabulary because s/he wants to sound British, and s/he wants to sound British because she considers it more sophisticated. This answers the two "why?" questions this study presents. Nevertheless, both are on an intermediary level in relation to the type of "why?" questions referred to by Fairclough. He is calling for analyses on a deeper level, or, alluding to the previous example: "Why is BrE considered more sophisticated?" While not central to this study, such questions are touched upon in interviews of the participants and, also, by addressing some of the historical background and present-day roles of AmE and BrE in subsection 3.1.

Another important distinction that needs to be made to frame this study is the one between *intraspeaker* and *interspeaker* variation. Interspeaker variation is variation on a group level, e.g. Finnish undergraduate students of English. Intraspeaker variation is concerned with variation within the language of a single person. This division epitomises the distinction between a *sociolect* and a *register* (Halliday 1978). A sociolect is a social dialect, a cluster of linguistic features shared by a group that is socially united in some way (ibid.). For example, if BrE was strongly favoured among Finnish academia and a participant of this study would consistently use BrE because of this, s/he could be argued to belong to a linguistic group whose sociolect is BrE, or a non-native approximation thereof. However, if s/he only resorted to BrE

features in an academic setting, this would constitute intraspeaker variation in which BrE is employed as an academic register.

So far, I hope to have established that language attitudes cause people to use certain variables over others. Labov (2001, 16) posits this and two other causes for variation leaning toward particular variables: (1) the principle of least effort, (2) the principle of density, and (3) the principle of imitation. Essentially these principles mean that people (1) avoid difficult features, (2) sound like the people they interact with, (3) sound like the people they admire (ibid.). One must, however, note that Labov was talking about phonological variation, but the same principles apply to morphology and syntax (Wardaugh 2010, 196). Unfortunately, once again, lexical variation is not accounted for. Nevertheless, I am confident that this omission is reflective of the general neglect of lexical variation, and not of the same principles not applying to it.

After all, there are well-documented phenomena of lexical variation in L2 use that match Labov's principles. For example, crosslinguistic influence, such as bias for words that are either cognates (Otwinowska 2015, 66), or phonologically compatible (Otwinowska 2015, 65) with a language the speaker already knows, is a clear example of Labov's principle of least effort at work. Regarding the principle of density, EFL users rely heavily on the media for language input, at least in Finland (Sjöholm 2000, 129). While Labov's (2001, 228) argument, that true interaction is required to change linguistic behaviour, and media input is mostly inconsequential, is founded on a great deal of evidence, it is made from the point-of-view of an L1 researcher. In an EFL context, where much of the language acquisition takes place incidentally through media (Forsman 2000, 167), it is rather clear that it affects lexical variation (Forsman 2000, 172–3).

Examining whether the third cause for variation, imitation, pertains to EFL users, specifically university level students of English in Finland, is the crux of the present study. The purpose of this subsection has been, primarily, to establish the framework for the niche that the principle of imitation occupies in the field of sociolinguistics. Such attitude-induced variation is considered in more detail in the following two subsections, first, in general terms, and then, as it relates to EFL users. In the final subsection of this section, I discuss lexical variation, first, in general terms and, finally, with regard to EFL contexts and Finland.

2.1 Language Attitudes and Variation

As established, imitating the people one admires is one of the major driving forces underlying variation (Labov 2001, 16; Tagliamonte 2012, 36). Kristiansen, Garrett and Coupland (2005,

10) even argue that *subjective factors* are more influential in language variation and change than *objective factors*. In this dichotomy, objective factors constitute all factors that are imposed on an individual by circumstances, ranging from human physiological and mental architecture to one's social environment (ibid.). These essentially correspond to Labov's principles of least effort and density. Subjective factors, on the other hand, relate to Labov's principle of imitation. They are those of *attitudes, social stereotypes* and a myriad of other concomitant terms stemming from different traditions (Kristiansen, Garrett and Coupland 2005, 15).

While Kristiansen, Garrett and Coupland make a convincing case for the primacy of subjective factors, their argument is made from an L1 point-of-view, and it is questionable whether it holds for L2. Indeed, one might expect L1 influence (principle of least effort) and L2 exposure (principle of density) to mostly determine variation among L2 users. After all, making a sociolinguistically motivated choice between various alternatives, be it phonemes, morphemes or lexical items, requires a very high level of linguistic competence. As mentioned earlier, sociolinguists have generally thought L2 variation to be random or L1 influenced (Durham 2014, 18). This view is discussed in the following subsection, and, of course, it's confirmation or refutation is the focus of the empirical part of the study.

I have chosen to refer to the socio-psychological processes underlying imitation by the term *attitude*, as defined by Gardner (1985, 9): "an evaluative reaction to some referent or attitude object, inferred on the basis of the individual's beliefs or opinions about the referent". In sociolinguistics, this "attitude object" tends to be a language, or a variety of one, and the study of such *language attitudes* is common in sociolinguistics. Typically, this is done by having a sample population grade the examined varieties on attitude scales that are grouped into two or three semantic categories. For example, Carranza and Zahn (1975) posited the categories of *status* and *solidarity*, which entailed attitude scales such as *wealthy*/*intelligent* for status and *kind*/*trustworthy* for solidarity.

While several other categories have been suggested, such as *dynamism* (Zahn and Hopper 1985) and *linguistic quality* (Ladegaard 1998), the aforementioned "[…] two particularly salient evaluational categories account for most of the variance" (Edwards 1999, 102). Indeed, these two categories are the only constants across all the reviewed models, although there are nominal differences. For example, Hopper (1977) first called them *competence* and *likability*, and later, in collaboration with Zahn (1985), *superiority* and *attractiveness.* For this study, I borrow the terms *status and competence* and *social attractiveness* from Rindal's study (2010), as it is the closest to my own that I have come across.

The related concept of *prestige* adds an element of social value to the evaluational dimensions of attitudes (Rickford and Eckert 2001, 2). Traditionally prestige has been rather synonymous with status and competence, but these days most sociolinguists differentiate the terms. For example, one might evaluate BrE to have a great deal of status and competence, but also consider it snobbish and pretentious, thus rendering it stigmatised rather than prestigious.

This phenomenon is characterised by the distinction between *overt* and *covert prestige*. Overt prestige is essentially status and competence granted to a variety on a societal level. In other words, it is the variety that holds power and is usually considered the standard. On the other hand, as Labov (1966, 108) noted, the less powerful strata of society should be perfectly able to use the overtly prestigious standard variety. He deduced that they did not want to do so and coined the term covert prestige for "an equal and opposing prestige for informal, working-class speech" (ibid.). So, to further underline the difference between the two, clearly interlinked, divisions: overt prestige versus covert prestige is based on the established societal power structures and the consequent linguistic standard, while status and competence versus Social attractiveness is based on individual evaluations of a given variety.

2.2 Language Attitudes and Variation among EFL Users

In the 1970s, people started to reject the view of English as a uniform entity, and several models emerged to account for the pluricentric role of English as a world language (Buschfeld, Kautzsch and Schneider 2018, 17). The seminal one, and one that is still widely used (ibid.), was that of Quirk et al. (1972, 3–4), which posited the concepts of English as a Native Language (ENL), English as a Second Language (ESL), and English as a Foreign Language (EFL). ENL refers to countries, such as the United Kingdom and the United States, where English is the first language of the majority of people. In ESL countries, on the other hand, English is not the first language of most people, but does have important functions in the society, be it as a lingua franca or as a language of education, politics, etc. The term ESL is also used to refer to L2 English used in an English-speaking country (Nordquist 2019a). In EFL countries, English does not have much of an intranational role but is mostly restricted to international communication (Quirk et al. 1972, 4–5).

This model, however, is far from perfect. Mair (2013, 259), for example, maintains that English has become pluricentric to a degree that any tripartite model is too simplistic. Buschfeld, Kautzsch and Schneider (2018, 18) also lament the simplicity, pointing out that it does not account for both ESL and EFL occurring in the same country, or countries that are otherwise in an intermediary or transitional position. Sure enough, the Nordic countries whose

native languages share a relation with English, i.e. all but Finland, are already argued to be transitioning from EFL countries to ESL countries (Pitkänen-Huhta and Hujo 2012, 282). I find that a similar argument could also be made for Finland, as English has become so pivotal that its command is a prerequisite for full engagement in the society (ibid.). Consequent to the critism, more dynamic models have been proposed, such as Mair's (2013) World System of Englishes. In this study, I use the more established categorisation by Quirk et al (1972) and discuss Finland and other Nordic countries as EFL countries. However, it is good to note the criticism and remember that the use of English in the Nordic countries is so prevalent that the data gathered from them is not easily applied to all EFL countries.

Traditionally, sociolinguistic variation has only been studied within ENL countries. Eventually the so-called *World Englishes* started to garner recognition as legitimate varieties of English in the 1980s primarily due to the work of Braj Bihari Kachru (1982). Even more recently, as EFL speakers have outnumbered ENL and ESL speakers (Seidlhofer 2011, 2), EFL varieties, or, more precisely, ELF as their collective manifestation, have started to evoke attention and claims of legitimacy. This pre-eminence of ELF has even inspired arguments for its use as a model in formal education (Jenkins 2002; Jenkins, Cogo and Dewey 2011).

The natural consequence of the elevated status of ELF is the expansion of variationist research into EFL contexts. Although, to be precise, variation among EFL users does have a long history of scientific research, but the focus has been on what is often called type I variation as opposed to type II variation (Howard, Mougeon and Dewaele 2013, 341). The former refers to variation that is related to learning, such as L1-influence and sequences of acquisition. Such variation is not socially, but linguistically motivated.

Type II variation, on the other hand, conforms to native speakers' variation patterns and is often called sociolinguistic competence (Howard, Mougeon and Dewaele 2013, 341; Durham 2014, 16). For greater terminological transparency, instead of Type I and Type II variation, one could also talk about *Learning Related Variation* (LRV) and *Target-Based Variation* (TBV) (Durham 2014, 16). According to Durham (2014, 18), TBV has been neglected in research due to focus on less advanced learners. The focus has been on errors and EFL variation has been generally considered either free or L1 influenced (ibid.).

In some sense, the focus of the study at hand is on TBV, as I compare the words used by the participants to their preferred native varieties. On the other hand, I am not convinced that EFL users only goal is to emulate native patterns. Firstly, some EFL users "[…] wish to avoid native accents and use a *neutral* variety of English" (Rindal and Piercy 2013, 211). Secondly, as foreshadowed in the introduction, Rindal (2010) found that some Norwegian adolescents not

only incorporate AmE and BrE phonology in their general idiolect in a socially motivated way but also use BrE as a formal register (Rindal 2010, 246–8).

Socially motivated deviations from native patterns among ESL users have already been studied quite extensively, although, again, the focus has been on phonology. The overarching results are summarised well Hansen Edwards:

> Two major finding emerge from the research on social factors and variation in L2 phonology. One finding is that learners are active agents in choosing not only what and how they use their L2, but also in choosing the L2 target, and therefore what they acquire of the L2. (Hansen Edwards 2008, 272)

Given that EFL users are likely less affected by any single variety, it seems logical that they would also have more flexibility in employing native varieties as suits their needs. As the only study on the subject in an EFL context (Rindal 2010) supports this view, it seems entirely possible that the prevalent intraspeaker variation between AmE and BrE in European EFL countries, or the *Mid-Atlantic* variety (Mollin 2006, 46), is often socially motivated.

The several studies conducted across Europe demonstrate quite consistent language attitudes towards AmE and BrE. For example, studies conducted on university level students of English in Spain (Carrie 2017), Denmark (Ladegaard and Sachdev 2006) and Poland (Waniek-Klimczak and Klimczak 2005) all found that a clear majority of the participants preferred BrE over AmE as a target variety, justifying it with aspects of status and competence, such as *professionalism* (Carrie 2017, 439) and *intelligence* (Ladegaard and Sachdev 2006, 100). Waniek-Klimczak and Klimczak did not study the attitudes underlying target preference, but another study on Polish university level students of English indicates that they too associate BrE with aspects of status and competence, such as *prestige* and *scholarliness* (Janicka, Kul and Weckwerth 2005, 255).

AmE, in turn, was associated with social attractiveness in all three countries (Carrie 2017; Ladegaard and Sachdev 2006; Janicka, Kul and Weckwerth 2005), but this did not sway as many of the students to choose it as a target variety. Upper secondary school students in Finland (Forsman 2000) and Norway (Rindal 2010) manifested similar attitudes, but the perceived informality and social attractiveness of AmE caused more of them to favour it as a target, although, among Rindal's participants BrE was still slightly more popular. In the words of Rindal:

[…] US and GB aimers share attitudes towards American and British varieties of English, but disagree about whether the evaluations are positive or negative, and consequently make opposing language choices. (Rindal 2010, 251)

Notably, none of these studies supports Bayard et al's (2001, 44) prediction that AmE hegemony will cause it to take over as the variety of overt prestige in EFL countries. However, to be fair, most of these studies are already rather old and Bayard et al. did not predict a timeframe for the change. Young people in Finland generally preferring AmE over BrE (Leppänen et al. 2011, 73) may forecast such a change in attitudes. Alternatively, it may reflect a decreasing importance of status and competence. The present study should provide some insight into this.

2.3 Lexical Variation

Lexical variation can be divided into three different types: *semasiological*, *onomasiological* and *contextual variation* (Geeraerts 1994, 3–4). *Semasiological variation* has to do with variation of referents of words, or the objects and concepts that they denote. For example, in BrE, *pants* can refer to *trousers* or *underwear*. Onomasiology takes the opposite approach and examines the words that can be used to describe a referent. For example, a certain type of garment can be referred to as *skinny jeans, jeans, or pants*, depending on the desired specificity.

Geeraerts maintains that onomasiology relates particularly to such "conceptually distinct lexical categories" (ibid.) and posits the category of *contextual variation* to denote variation between synonymous lexical items, whose variation depends on "contextual factors such as the formality of the speech situation, or the geographical and sociological characteristics of the participants in the communicative interaction" (ibid.). Obviously, such variation falls within the field of sociolinguistics (Geeraerts 1994, 82) and the term *contextual variation* seems to differ from what I have thus-far called *sociolinguistic variation* only due to the difference between the traditions of sociolinguistics and lexicology, much like the abovementioned interdisciplinary terminological variation in denoting *attitudes*, *social stereotypes*, etc. (Kristiansen, Garrett and Coupland 2005, 15).

Again, there is little in the way of research into sociolinguistically motivated lexical variation in general, and into lexical variation in an EFL context in particular. I have already referred to Forsman's study that found that media plays a significant role in language acquisition (2000, 167) and directs lexical choices toward AmE (2000, 172–3). However, I find that there is a methodological problem in her data elicitation, likely consequent to varietal

variation not being central to her study. The problem is that the participants were asked to choose between AmE and BrE words presented to them (ibid.), which may result in choices the participants would not make naturally, i.e. if not given alternatives.

Gonçalves et al. (2018, 11), in turn, studied the vocabulary in tweets sent from 30 different countries around the world and observed a great deal of Americanisation, even among ESL countries, most of which are former British colonies. However, the phenomenon was more marked in the 21 EFL countries (ibid.). Exact numerical data is not provided, but based on the presented graph (ibid.), the vocabulary of the Finnish tweets was among the most Americanised, and the only EFL countries with a more AmE vocabulary were Sweden (almost even with Finland), Portugal, Brazil and Mexico (ibid.). The two latter ones were the only EFL countries included from the American continent and geopolitical proximity to the United States is a likely factor in the Americanness of their vocabulary.

Unfortunately, I find that there are some methodological shortcomings also in the study by Gonçalves et al. Firstly, Twitter is a problematic source, as there is little information on the sample population. For example, it is not known how many of the tweets in Finland are sent by Finnish L1 speakers. Nevertheless, the strong inclination toward AmE is compelling, as British people outnumber Americans in Finland both as residents (Statistics Finland 2017, n.p.) and as tourists (Statistics Finland 2018a, n.p.). In other words, native users of AmE and BrE should skew the data toward BrE rather than AmE.

Furthermore, the pairs of words were chosen based on their occurrences in the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA), but they were not weighted in any way, even though the frequencies of the words relative to their counterparts vary, and, thus, some words represent their respective variety more explicitly than others. Also, the chosen distinctive words were not controlled for the context in which they occurred in the corpora, or if they were, it is not explicitly stated. For example, one distinctive pair is the board game *checkers* (AmE) / *draughts* (BrE). Yet, in the corpora used in the study, these words can occur in such contexts as "fact checkers" and "draughts of air". Naturally, they lose their varietal distinctiveness in such contexts.

Nevertheless, despite my methodological objections, the results do seem robust enough to establish the global Americanisation of English vocabulary and Finland being in the vanguard of this phenomenon. The present study should be able to validate the results to some degree, even though the sample is different, and, consequently, the results are not entirely comparable.

3 American English and British English

Focusing on only two of the several native varieties of English may seem simplistic, but, given the demonstrably marginal role of the other varieties in Finland, I believe it is justified. In their extensive (N = 1284) *National Survey of English Language in Finland*, Leppänen et al. (2011, 71) found that, among the general population in Finland, BrE was the most appealing variety according to 39.6% of the respondents and AmE according to 35.9%. Only 3.8% preferred the third most popular ENL variety, Irish English. In addition, the AmE dominated media has a fair amount of BrE material, but other varieties are scarce, at least on TV (Finnpanel 2020a), which remains clearly the most popular medium of audio-visual entertainment in all age groups (Finnpanel 2020b). Furthermore, teaching materials in Finland are over 90% AmE or BrE (Ylönen 2007, 35) and it is "very rare" (Tergujeff 2012, 40) that teachers use other native models.

In the following subsection, I outline the history behind this bipolar situation and some of the possible reasons for the different attitudes they evoke. I start with a global point-of-view, which I then focalise to the Finnish context. In subsection 3.2, I provide some historical context and details on the lexical differences between AmE and BrE.

3.1 Power and Language Attitudes

As posited in subsection 2.1, linguistic prestige is intertwined with social prestige (Pearce 2007, 146) and power (Bonfiglio 2002, 23). Indeed, the present hegemony of English has been established through political, economic, social and military power of the British Empire and, later, by those of the United States. Conversely, established linguistic hegemony has been argued to sustain the political and cultural *status quo* (Canagarajah and Ben Said 2011, 387). Other varieties of English do not stem from such positions of power, historical or current, that they would have the same kind of influence, perhaps excluding their immediate geopolitical surroundings.

For much of the history of the English-speaking world, power has lied within the upper echelons of society in London and it comes as no surprise that their variety of English has been considered the standard. Any deviations from it, be they geographically or socially motivated, have been deemed inferior (Milroy 2007, 138). Of course, such attitude raised some early objections in the United States, perhaps most notably by Noah Webster (1789, 20), the *primus motor* of standardised AmE: "Pride and prejudice incline men to treat the practice of their neighbors with some contempt." Although Webster himself fell for contempt of BrE (Pelanda

2011, 451), the quote above is in the spirit of the present paradigm that there are several equally valid standards (Penhallurick 2003, 191).

Nevertheless, the status and competence associated with BrE, at least in Europe, (e.g. Carrie 2017, Ladegaard and Sachdev 2006, Janicka, Kul and Weckwerth 2005), and particularly its use as the model in formal education (Tergujeff 2012, Rindal 2010, Ladegaard and Sachdev 2006), suggests that it is still the variety of overt prestige. These attitudes constitute an interesting phenomenon considering that the British Empire has disintegrated, and power in its various forms has crossed the Atlantic. Surely, attitudes take a long time to change, and the overt prestige of BrE may well be vestigial and on its way out, as predicted by Bayard et al. (2001, 44), but so far there is little evidence to that effect.

It seems that the American hegemony exerts a different kind of power to that of the British Empire, and, consequently, a different kind of prestige. Before the Second World War, the United States had been on a rise for some time, but it truly cemented its position as a global superpower in the wake of it (Chari 2007). This roughly coincided with televisions and record players becoming commonplace in regular households in the western world and soon enough the American way of life found its way into people's living rooms. This made a lot of people familiar with American culture and American English, which I would argue, has granted it covert prestige. This is supported by AmE being generally rated high on the dimension of social attractiveness.

In Finland, the rise of American popular culture coincided with an ideological and cultural shift toward the west "with the English language as a symbol of modern westernisation" (Leppänen et al. 2011, 17). Sure enough, as the American music and films pervaded Finland, the popularity of English increased and, by the 1960s, it surpassed German as the most popular foreign language in school (Birkstedt 2015, 478). The influence of the mostly AmE popular culture further increased in the 1960s as TV programmes started to be subtitled (Leppänen et al. 2011, 18). Yet, formal teaching emphasises BrE to this day (Ylönen 2008, Tergujeff 2012), although the official policy is to promote understanding of English in its various forms and no mention is made of any single variety in the curricula of comprehensive education (National Curriculum 2014) or upper secondary education (National Curriculum 2019).

3.2 Lexical Differences

As necessitated by the novel flora, fauna and native cultures, new words started to emerge in the English language as soon as English-speaking colonists started to settle the Americas. Indeed, even before the famous Mayflower arrived on the North American continent, Alexander

Gil recorded in his *Logonomia Anglica* that *Maize* and *canoe* had entered English from indigenous Caribbean languages (Gil in Penhallurick 2003, 53). The languages of other European settlers also added to the vocabulary of English in the Americas (Penhallurick 2003, 54). Whether also some African words, such as *banjo* and *bogus* (Bailey 2004, 11), were adopted from the slaves is a debated issue, but, in any case, the impact would have been very small as their languages were highly stigmatised and actively suppressed (Healey 2011, 160).

The new words denoting new phenomena hardly made AmE very distinct from BrE as the words were necessarily employed on both sides of the Atlantic as needed. However, shortly, the geographical distance caused the English spoken on the American continent to diverge further from its origins and distinctly American words started to substitute existing British words. Naturally, as all living language varieties do, the English spoken on the British Isles was also changing and some of these changes were not adopted in America. For example, in Britain, *fall* fell out of use in favour of *autumn,* but remained in use in America (Mencken 1921, 67).

In subsection 2.3, I referred to Gonçalves et al. (2018, 11), who showed the global dominance of AmE not only in EFL and ESL countries, but also ENL countries that used to be a part of the British Empire. In addition, as the United States have established something of a cultural hegemony, its vocabulary has also started to influence Britain. Ever since audio-visual entertainment became a part of everyday life in the west, British people have adopted more and more Americanisms in their active vocabulary (Gould 2007, 25), much to the dismay of many language purists. Sure enough, it is this sort of "invasive" type of Americanism that is the popular notion of the meaning of the word (Algeo 1990, 129; Nordquist 2019b, n.p.).

Converse ly, Briticisms are also adopted into AmE, albeit, to a lesser degree (Algeo 1990). According to Kory Stamper, the associate editor of the Merriam-Webster dictionaries, the influence of the media is evident also in this phenomenon. As an example, she quotes the Harry Potter book series as a source of Briticisms that have gained ground in the United States (Stamper in Hebblethwaite 2012, n.p.). In the same article, Geoffrey Nunberg maintains that Briticisms are mainly used among the "educated elite" (Nunberg in Hebblethwaite 2012, n.p.). I did not find empirical evidence to substantiate this, but, curiously, the same view was expressed some hundred years ago by Mencken (1921, 167), who declared that Briticisms were typical "among folk of social pretensions". Conversely, in Britain, it seems to be the educated people that are most opposed to Americanisms. It seems that Americans and Britons have the same connotations about their own varieties as the rest of the world.

Varietally distinct vocabulary is particularly common in areas of life that developed after the British settling the American continent (e.g. *elevator* vs. *lift*, *highway* vs. *motorway*) and in

the more intimate, domestic spheres of life (e.g. *living room* vs. *sitting room*, *dessert* vs *pudding*) (Finegan 2004, 20). Overall, words that are relatively distinct to each variety are seemingly abundant in number, but they are less so in frequency. In the words of Finegan:

> A British reader of Time or Newsweek would note distinctly American expressions only a few times on any page, matching the few distinctly British expressions an American reader of The Economist would note. (Finegan 2004, 23)

While this makes adapting one's vocabulary to a target variety less meaningful in terms of imitating a target variety, I would argue that it also makes doing so more difficult and, consequently, more indicative of one's sociolinguistic ability. In the following I describe the empirical means employed in this study to examine 1) do the target choices of Finnish undergraduate students of English correlate with their language attitudes, and 2) how successful the students are at orienting their lexical variation toward their preferred target. If the answer to the former question is "yes", the attitudes that do correlate can provide insight into what kind of attributes the students want to attach to their ideal self (Dörnyei ja Chan 2013, 438). If such a connection is found, the latter research question will address whether this self-identity is manifested in the actual production. If it is, that would indicate that EFL can convey social signals with their linguistic variation, just like ESL users have been found to do.

4 Data Elicitation Methods

This section first introduces the sample population (participants) and discusses its relation to the target population, i.e. Finnish undergraduate students of English and, more tentatively, advanced EFL users in general. Then, the methods used to gather the data are discussed in the order they were employed. Subsection 4.2 describes the elicited translation task designed to prompt lexical choices between AmE and BrE alternatives. Subsection 4.3 presents the questionnaire used to measure target preferences and language attitudes with regard to AmE and BrE. Subsection 4.4 outlines the framework for the semi-structured interviews, whose purpose was to provide further nuance and reliability to the quantitative methods.

4.1 Participants

As pointed out in subsection 2.1, making sociolinguistically motivated lexical choices requires a great deal of linguistic competence. The most conveniently available group of people with such competence are the university level students of English. As future professionals of English, they can also give some insight into the varietal direction English is heading in Finland. This does somewhat undermine my interest on EFL users in general, as there is a myriad of underlying variables related to being a university level student of English. However, even if the participants are not a representative sample of all EFL users, they are an example of them. In other words, it is possible to say "based on the data **some** EFL users are able to…". Furthermore, to somewhat reduce the potential underlying variables, I focused on undergraduate students as they have not been influenced by the university for a long time.

I did, however, want the results to be applicable to Finnish undergraduate students of English in general for the abovementioned insight into the future of English in Finland. For this purpose, I had to consider how to obtain a representative sample. A completely random sample would have been ideal, as it eliminates possible underlying variables (Dörnyei 2007, 97). However, as the test of lexical choices (subsection 4.2) had to be administered in person, gathering data from several Finnish universities with an English department was not practically attainable. As such, it is possible that regional differences in attitudes and production relating to English affect the data, but I find it unlikely. Firstly, to my knowledge, there is no evidence of such regional differences in Finland. Secondly, there is a great deal of mobility within Finland, and it is common to move to another region to study (Statistics Finland 2010).

Consequently, the test and the questionnaire (subsection 4.3) were administered in the beginning of a lecture in one Finnish university. To obtain a maximally random sample within

that university, I chose a lecture of a BA-level course, which is a part of a set of five courses of which one must choose at least four. Consequently, if the course selections were random, 80% of all students could be expected to take this course. However, according to the lecturer, it has more enrolments than the other four courses. Also, some students are likely to participate in all five. In other words, a vast majority of the English students participate on this course, which reduces the chance of the sample being skewed by the students' course selections.

The test of lexical choices and the questionnaire on target preferences and language attitudes yielded altogether 49 responses. However, one respondent only produced 5/10 of the words that the test was designed to elicit, which I deemed to be insufficient for providing a reliable picture of her lexical choices. All the other respondents produced at least 7/10. Furthermore, eight respondents were excluded based on having spent more than one month in an English-speaking country. I wanted to focus on the EFL context, and their results were clearly affected by the extended language contact (Appendix 8). Nevertheless, their input was valuable for test validation, as their lexical choices deviated from the mean in a predictable way, i.e. residency in the USA seemed to promote AmE lexical choices and residency in the United Kingdom and its former colonies seemed to promote BrE lexical choices.

The mean age among the qualified participants was 21.8, ranging from 19 to 34. The gender distribution of the participants (29 females, 10 males, 1 other) reflected the ratio among language students in Finnish universities (79% female) (Statistics Finland 2018b, n.p.). Native language was not asked, but the test of lexical choices necessitated a near-nativelike command of Finnish.

4.2 Varietally Distinct Lexical Choices

I was allowed 15 minutes at the beginning of a lecture for data gathering. Obtaining a sufficient amount of distinctly AmE or BrE lexical production within this timeframe necessitated an extremely structured elicitation method, which compromised the naturalness of the data to a degree. I consider this an acceptable shortcoming as ethical considerations and the observer's paradox cause natural production to rarely be available for scientific research in any case. I found *elicited translation* (Doughty and Long 2003, 794) to be a suitable method for prompting an adequate number of varietally distinct lexemes in a short period of time.

The participants were asked to translate a sequence of ten short texts from Finnish into English. The texts included ten words with distinct equivalents in AmE and BrE (Appendix 1). To maximize the naturalness of the production, masking was an important consideration. Had the participants deciphered the purpose of the study, they would have likely paid special

attention to their lexical choices, which would have had an adverse effect on the reliability of the test. In order to distract the participants from speculating on the purpose of the study, the sequence was designed to form a narrative whole. Also, some of the texts did not include a varietally distinct word at all to make the purpose less obvious. Furthermore, each text was projected on a screen for one minute, which was found to be a tight, but sufficient, amount of time in the piloting phase. According to the piloting performed on two students, the prompted production was also suitable for analysis as both students produced 8/10 of the varietally distinct words. Masking also proved to be successful as the participants did not guess the purpose of the test when asked. Also, all five participants interviewed at a later stage said that the masking had been successful.

There were two considerations when selecting the words to be elicited. Firstly, they had to be of relatively high frequency in order to be easily accessible to the participants that were working under strict time constraints. Furthermore, high frequency words are, obviously, more common in natural production and, thus, they could be argued to be more representative of one's idiolect. Secondly, the words had to be truly varietally distinctive, i.e. relatively exclusive to their respective varieties as measured by their frequencies in the Corpus of Contemporary American English (COCA) (Davies 2008) and the British National Corpus (BNC) (Davies 2004). To establish how distinct each word is to its variety, every word was calculated for its frequency in the corpora relative to the corresponding word of the other variety.

To this end, I first had to ensure the synonymity of the words by reviewing the contexts in which they occurred. As unintended meanings appeared, I modified the search. For instance, the word *flat* was often used as an adjective in the corpora. Consequently, I searched for the sequence "my flat" (BrE) instead, which only yielded results that were synonymous with "my apartment" (AmE). I then proceeded to calculate the frequency of each sequence in its respective corpus in relation to its counterpart. For example, "my flat" occurs 164 times in BNC and the sequence "my apartment" occurs 34 times. Thus, the frequency of "my flat" in BNC is 164 / (164 + 34) = 0.828 = 82.8%. The corresponding figure for "my apartment" in COCA resulted in 94.1%. I then extrapolated the frequencies of these two sequences to apply to *flat* and *apartment* in general. The lowest frequency qualified for the test was *high street* (BrE) at 73.9%.

For a reliable representation of the distinctiveness, it was necessary to account for the frequencies of both words in a distinctive pair. For example, it could be argued that *high street* is very British, because it hardly ever occurs in the American corpus. Nevertheless, conversely, one could also argue that it is not very British, because the supposedly American *main street* is

also often used by the British. Thus, both arguments were accounted for, and the value of distinctiveness was obtained by adding the two percentages together and subtracting 1 in order to get a scale from 0 to 1. For example, the value of distinctiveness for *high street / main street* is 0.739 + 0.944 − 1.000 = 0.683.

To calculate the percentage of Americanness of the varietally distinct lexical choices of each respondent, I summed the values of distinctiveness of all the produced AmE words and divided that by the maximum score, i.e. the sum of all the values of distinctiveness. As mentioned in subsection 4.1, sometimes the participants did not produce either word of a distinctive pair. In such cases, the value of the omitted word pair was subtracted also from the potential maximum score. For example, participant #1 produced five AmE words, whose distinctive values add up to 3.713. There were three distinctive word pairs that she omitted altogether. The combined distinctive values of these (2.446) was subtracted from the potential maximum score (7.741). Thus, her percentage of AmE lexical choices is 3.713 / (7.741 − 2.446) = .701 = 70.1%. Given the dichotomous nature of the study, the corresponding percentage for BrE lexical choices is 29.9% (100.0% − 70.1%).

4.3 Target Preferences and Language Attitudes

This study also set out to explore the target preferences and the language attitudes as relates to AmE and BrE. These are, obviously, completely subjective variables, and the best way to determine them is to ask the participants. As stated in the previous section, masking the purpose of the test of lexical choices was of paramount interest and, therefore, these data were collected after administering the test.

Most previous studies I reviewed simply asked the students, which variety they aimed for (Waniek-Klimczak and Klimczak 2005; Ladegaard and Sachdev 2006; Rindal 2010, 2014), thus creating a binary division. To allow for the correlation of target preferences with other variables, I wanted to add an element of magnitude to them. To this end, I had the participants assess the statements "I would like to sound American" and "I would like to sound British" on a *Likert scale*, which is the most used scaling technique in questionnaires (Dörnyei and Taguchi 2010, 27). There are several variants, but I opted for the original version with five response categories to each statement (*strongly disagree, disagree, neither agree not disagree, agree, strongly agree*) (ibid.). Some researchers promote omitting the neutral alternative, as the less motivated respondents may choose it by default (Dörnyei and Taguchi 2010, 28). While this is possible, it is also possible to truly have a neutral stance, which is why I find eliminating the

neutral option a little unfair. I worried that some people might not respond at all, if their preferred option was not available

Language attitudes tend to be studied either with a matched guise technique (MGT) (e.g. Rindal 2010) or a verbal guise technique VGT (e.g. Ladegaard and Sachdev 2006; Carrie 2017). In both, the participants listen to and evaluate speakers of different varieties, but in MGT they unwittingly evaluate the same speaker(s) producing different varieties, arguably, removing the effect of voice and other idiolectal factors. These methods are lauded for their rater-reliability, as the varieties are evaluated indirectly, which is argued to bypass the preconceived attitudes toward the varieties (Kircher 2015, 206).

Alas, both methods would have been too time-consuming for the present study. Particularly, as I opine that one should use several speakers to collectively represent the varieties, as any single speaker cannot represent such macro-level varieties as AmE and BrE. Furthermore, I doubt that neither variety is stigmatised in Finland to a degree that would necessitate the indirect methods of MGT and VGT. Thus, contrary to MGT and VGT, the evaluated referents were the AmE and BrE varieties themselves, not speakers producing them. Akin to MGT and VGT, I provided the participants with certain adjectives, which were then graded based on how well they matched the raters' attitudes toward the referents. As with target preferences, I opted for a 5-point Likert scale ranging from *strongly disagree* to *strongly agree.*

In subsection 2.1, I posited status and competence and social attractiveness as the fundamental categories of language attitudes. In any questionnaire, it would be recommendable to use *multi-item scales* (Dörnyei and Taguchi 2010, 25), i.e. prompt evaluations of several attitudinal dimensions for each semantic category. For example, Rindal (2010, 246) used six components such as *reliability, likeability* and *generosity* to account for social attractiveness. Unfortunately, I only had 15 minutes for data gathering at the beginning of a lecture, and the piloting phase revealed that filling in the background information and completing the test described in the previous subsection almost depleted that time. As the relationship of lexical choices and target preferences was my primary interest, I chose to compromise on the attitude scales instead.

One possibility would have been to use the names of the categories as evaluative dimensions, but I suspect them to be too complex and technical to elicit a quick and intuitive response from the participants. Referring back to subsection 2.3, a semasiological study mapping the participants' perceptions of potential attitudinal scales would have been ideal, but that would have been a massive undertaking and, also, compromised the masking. Instead, I resorted to dictionary definitions, which lead me to choose the term *sophisticated* for status and

competence, as it incorporates social status and intellectual ability, which are both dimensions of this category (Rindal 2010, 249). According to Oxford English Dictionary *sophisticated* is characterised by a "[…] great deal of worldly experience and knowledge of fashion and culture" and being "aware of and able to interpret complex issues" (*MOT ODE*, s.v. "sophisticated," a.). For Social attractiveness, I decided to use *cool*, which was spontaneously brought up by one of Rindal's respondents.

> I would use American with adolescents and British with grown-ups. (...) When we hang out with friends (...) we don't want to use the British English we try to learn at school, we would rather do what we think is cool. (Rindal 2010, 254)

As an adjective, *cool* is defined as "fashionably attractive or impressive" (*MOT ODE*, s.v. "cool," a.), but also as "showing no friendliness towards a person or enthusiasm for an idea or project" (ibid.). While I was quite confident that the latter connotation would not occur to the participants when assessing varieties of English, I decided to add a scale of *pleasantness* to the dimension of social attractiveness. Even if the participants did not connote coolness with aloofness, pleasantness would complement it with a different, a "friendly and considerate" (*MOT ODE*, s.v. "pleasant," a.) type of social attractiveness.

4.4 Semi-structured Interviews

I wanted to complement the abovementioned quantitative methods by interviewing some of the participants. In addition to establishing some causal relation to the possibly emerging correlations, I hoped the interviews would confirm and validate some of the obtained data, particularly the attitude scales, whose validity was somewhat compromised due to temporal limitations. At the end of the questionnaire form (Appendix 2), I provided an opportunity for the participants to volunteer their e-mail for further research. Nine of the qualified participants volunteered for further research, and I was eventually able to schedule an interview with five of them. Luckily, these five participants covered the continuum from AmE to BrE quite well respective to all variables (Appendix 6). This increases the probability of all points-of-view being represented. The main shortcoming was the lack of an interviewee without a clear target preference for either variety.

The interviews were conducted approximately 8 months after the administration of the test and the questionnaire. The temporal distance and a completely different setting (in a café over a drink) to the first session should serve the triangulation of the data and add to the reliability of the responses, under the condition that the evaluations are consistent, of course.

The interviews were conducted in Finnish in order to not prime the answers with my own English idiolect.

Methods of interviewing vary with respect to structuredness. Highly structured interviews are essentially questionnaires filled in with the guidance of an interviewer. This is an appropriate method for a quantitative content analysis as it elicits responses that are comparable across various interviewees (Dörnyei 2007, 136). At the other end of the spectrum, one finds unstructured interviews, that only have a general theme and let the respondent guide the direction of the interview. This is a good approach for an exploratory research into an unknown topic or for exploring a phenomenon on a very profound level (Dörnyei 2007, 136–7).

In the present case, some structure was necessary as there were topics, I wanted all the respondents to address. On the other hand, I wanted to gain added value from a more qualitative approach and give the respondents an opportunity to lead the interview in new directions. After all, there is always a possibility that the researcher fails to consider all relevant perspectives. Consequently, I opted for a semi-structured interview, which is a compromise between the two extremes and the most common type in applied linguistic research (Dörnyei 2007, 137). Accordingly, I had a framework for the interviews (Appendix 7), but I was prepared to follow the respondents' lead should they orient the interview in an unexpected direction.

While the focus of the study is quantitative and the first five questions/tasks of the interview were highly structured, I apply a qualitative content analysis, or *latent level analysis* (Dörnyei 2007, 246) to the open-ended questions. In other words, I systematically look for patterns relating to the central themes of the study but, in the end, the emergence of these patterns is subject to my interpretation of what the interviewees meant to say.

As mentioned, the primary goal of the interviews was to increase the validity and reliability of the data through triangulation. To this end, the dimensions of status and competence and social attractiveness of AmE and BrE were addressed more thoroughly to see whether this matched the participants evaluations of *sophistication, coolness* and *pleasantness,* that represented these dimensions in the original questionnaire. In the beginning of the interview, I had the participants evaluate AmE and BrE in terms of two more attitudinal scales: *competence* and *likability*, which were Hopper's (1977) labels for status and competence and social attractiveness, respectively. I also asked the interviewees to describe AmE and BrE in their own words.

I also wanted to confirm two assumptions I had made regarding the attitudinal evaluations. Firstly, I assumed that phonology would be the most salient aspect of any variety and that the evaluations would, consequently, be largely based on that. Secondly, I assumed

that the participants would relate AmE and BrE phonology to *General American* (GA) and *Received Pronunciation* (RP) and not any of the other possible sub-varieties. I was careful to phrase the questions in a neutral way that would not affect the responses (Appendix 7, questions 6, 7, 12 and 15).

With the analysis of the results in mind, I also wanted to address the exposure to the two varieties. As established in subsection 3.1, none of the qualified participants had spent any significant time in an English-speaking country. Also, regular contact with native speakers was rare among all participants and none of the interviewed ones reported such contact. Therefore, I focused on the varieties encountered through media and formal education (Appendix 7, questions 21–26).

I piloted the interview with the same two students that I used for the questionnaire and their responses to both attitudinal scales were consistent with those given in the questionnaire, i.e. both evaluated BrE much more favourably on the status and competence dimension (sophistication and competence), but only slightly more favourably with regard to social attractiveness (pleasantness/coolness and likability). This validated the chosen attitudinal scales to some degree. Also, their responses to the open-ended questions indicated that they had understood the questions as intended.

5 Results and Analyses

In this section, I present the data whose elicitation was described in the previous section. The interviews are an exception to this as their results are discussed when relevant, mostly in section 6. In the following three subsections, I introduce the variables (varietally distinct lexical choices in 5.1, target preferences in 5.2, and attitudinal scales in 5.3) mainly by means of descriptive statistics and graphs. In the last two subsections, where the datasets are related to one another, I employ more sophisticated statistical methods (rank order correlation of lexical choices and target preferences in 4.4 and ordinal regression of target preferences and attitudinal scales in 4.5). Throughout, I seek to justify the applied statistical methods with reference to the type of data examined and methodological literature.

5.1. Lexical Variation between AmE and BrE

The test of lexical choices included 10 varietally distinctive pairs of words. The AmE variant was more popular 6/10 times, the BrE variant 3/10 times, and 1 distinctive pair was even as *holiday* and *vacation* were both used 14 times. The 40 qualified participants produced 309 of the 400 possible varietally distinct words. 66.0% of them were AmE and 34.0% were BrE. However, as described in subsection 3.1, each pair of words was weighted for its varietal distinctiveness to account for the use of AmE words by the British and vice versa. Table 1 presents the central tendencies and range of the weighted lexical choices in the tested vocabulary.

Table 1 Varietally distinct lexical choices – Central tendencies and range

N = 40	Mean	Median	Minimum	Maximum	Standard deviation
American English	68.7%	69.3%	23.9%	100.0%	18.5%
British English	31.3%	30.7%	0.0%	76.1%	18.5%

The results are quite consistent with the measures presented in the previous paragraph. The slightly higher percentage of Americanness means that the produced AmE words were generally a little more distinctive than the BrE words. The dichotomous nature of the study is evident in that the sum of the AmE and BrE means and medians is 100.0%. Consequently, I only provide a visualisation of the AmE distribution in Figure 1, as the corresponding BrE graph would just be its mirror image.
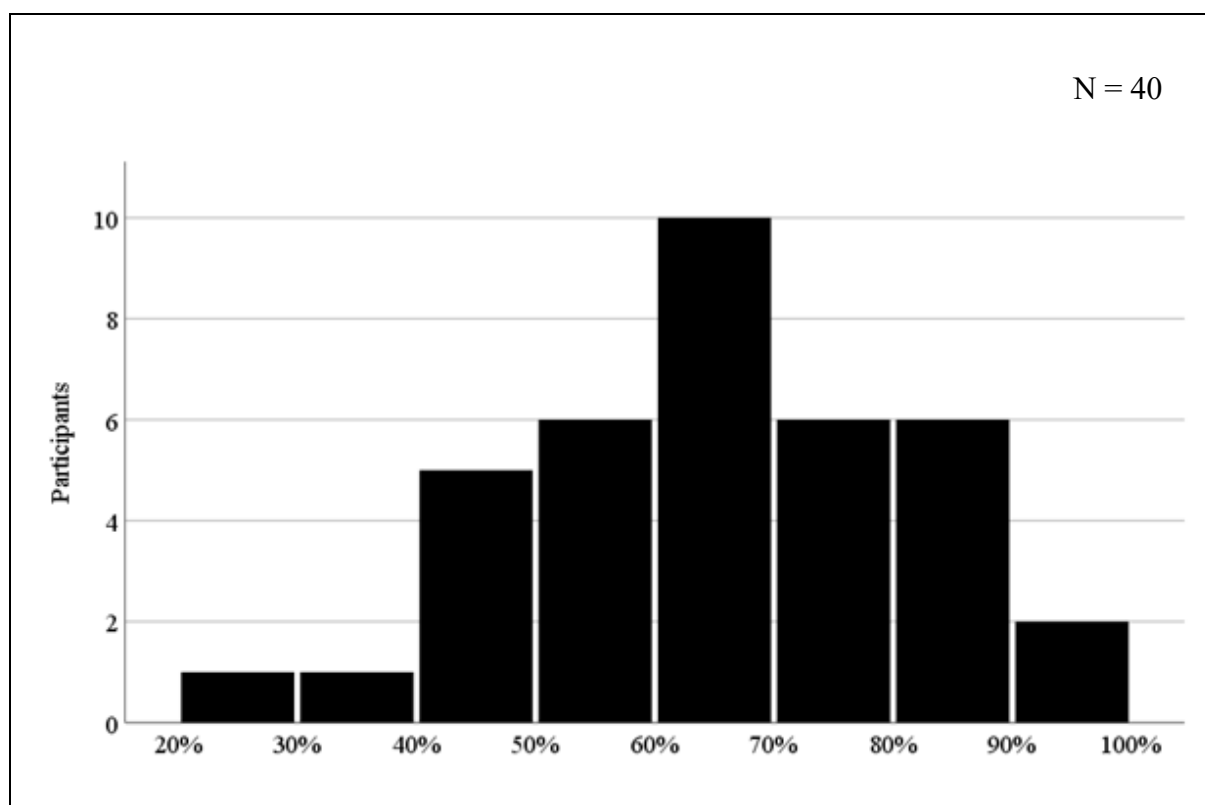
Figure 1 American English lexical choices – Histogram

Each column shows the number of participants whose percentage of AmE lexical choices falls between the corresponding figures on the X-axis. The distribution is clearly concentrated around the mean and seems to fall on the bell curve of normal distribution (Dörnyei 2007, 27). In addition to visual assessment, I ran a test of normality to establish it mathematically. Of the two recommended models for sample sizes below 50 (Elliot and Woodward 2007, 4) Shapiro-Wilk test is often recommended due to its greater power in comparison to the alternative Kolmogorov-Smirnov test (Thode 2002; Steinskog, Tjøstheim, and Kvamstø 2007). The null hypothesis of Shapiro-Wilk is that the data are normally distributed, meaning that if the p-value is higher than .05, normal distribution can be assumed (Elliot and Woodward 2007, 6). As calculated on the IBM SPSS 25 software (henceworth SPSS), the p-value for the data at hand resulted in .528 and the results can, thus, be assumed normally distributed. This being the expected distribution of a large sample in social sciences (Dörnyei 2007, 27) renders further reliability to the result.

In summary, the lexical choices were roughly 2/3 American by all measures. According to the most sophisticated method, the weighted percentage of lexical choices, the mean percentage of AmE lexical choices among the 40 qualified participants was 68.7%. The corresponding figure for BrE is 31.3% (= 100.0 – 68.7%).

5.2 Target Preferences vis-à-vis AmE and BrE

As stated in subsection 4.3, target preferences were measured on a 5-point Likert scale. For the purposes of analysis, the responses to the statements "I would like to sound American" and "I would like to sound British" were quantified by assigning strong disagreement with value 1 and strong agreement with value 5. As the values cannot be assumed equidistant, i.e. one cannot be sure that 4 (*agree*) differs from 3 (*neither agree nor disagree*) as much as it does from 5 (*strongly agree*), calculating the mean may be misleading. For such ordinal data the median and the mode are more appropriate measures of central tendency.

Table 2 AmE and BrE as target varieties – Central tendencies

N = 40	Median	Mode
"I would like to sound American"	3	2; 4
"I would like to sound British"	4	4

Evaluations on a 5-point Likert scale from *strongly disagree* (1) to *strongly agree* (5)

The central tendencies suggest that attitudes toward BrE as a target variety are a little more favourable. Also of note is that the distribution of the responses to "I would like to sound American" is bimodal, with 2 (disagree) and 4 (agree) being the most popular responses. In figure 2, one can see a visualisation of the distributions.
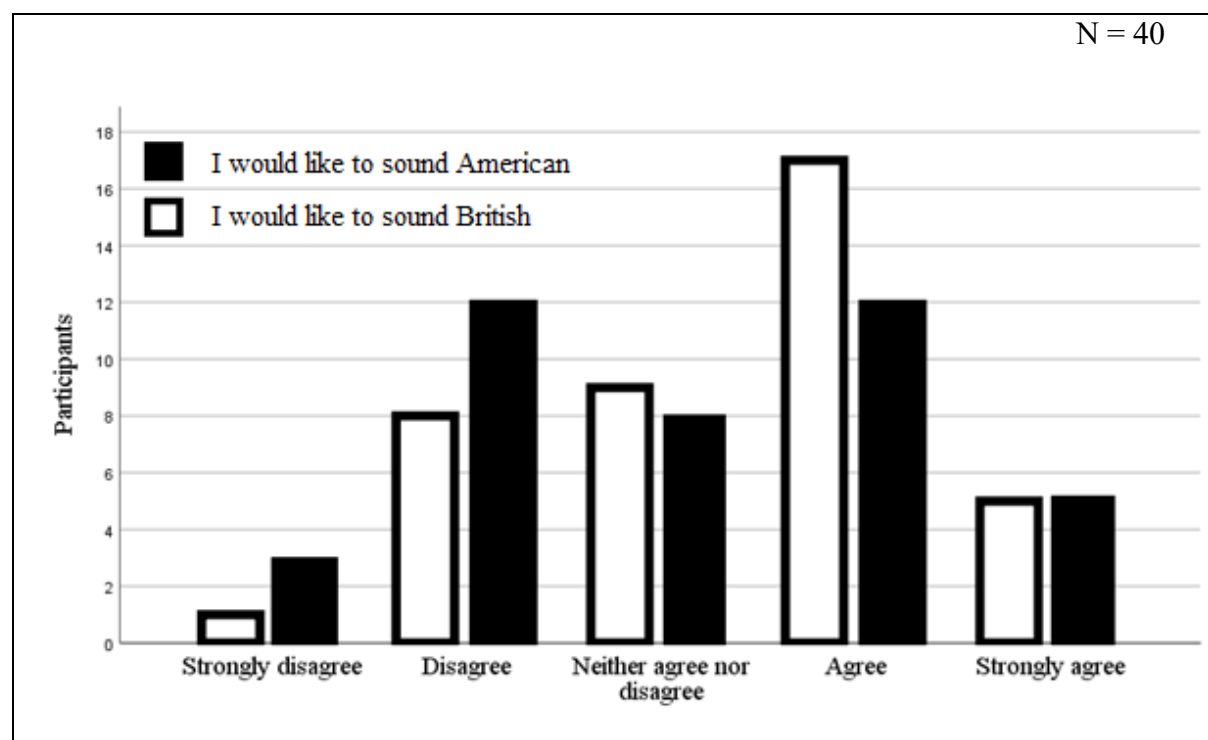
Figure 2 Evaluation of American and British Englishes as target varieties – Bar chart

In addition to being bimodal, AmE is quite evenly distributed around the neutral option. BrE, on the other hand, is clearly skewed toward the positive end of the scale. Studies conducted elsewhere in Europe have also found BrE to be the preferred target variety among university students in Spain (Carrie 2017), Poland (Janicka, Kul and Weckwerth 2005) and Denmark (Ladegaard and Sachdev 2006). In Finnish context, Leppänen et al. (2011??) found BrE to be favoured among university educated people, which arguably applies also to the participants of the present study, even though they were still studying at the time of research.

However, the difference in preferences in the present study is notably small and testing its statistical significance is in order. As the two sets of data are ordinal, they cannot be assumed normally distributed. In such a situation, nonparametric tests are usually applied (Dörnyei 2007, 227) In the present case, as the two sets of data come from the same participants, the typical method would be the Wilcoxon test (Dörnyei 2007, 230). However, some statisticians prefer a *Van der Waerden Test*, which entails normalising the data and then applying parametric methods, as they are more powerful than the non-parametric ones (Darlington 1996; Gilman 2015). Just like in the Wilcoxon test, the actual figures in the data are replaced by rank numbers, but they are also transformed further into z-scores in order to get a normal distribution. After these measures, parametric methods can be applied (Gilman 2015). Consequently, I used SPSS to perform a paired t-test, the parametric equivalent of Wilcoxon (Dörnyei 2007, 230), on the transformed data and found the difference between the evaluations of AmE and BrE as target varieties to be statistically insignificant: $t(39) = -.99$, $p = .329$. In other words, the result cannot be extrapolated beyond the sample population.

While the abovementioned results illustrate the attitudes toward AmE and BrE as target varieties, they do not provide a full picture. It is also necessary to find out how the attitudes toward AmE and BrE relate to one another, in other words, whether the participants that would like to sound American would not want to sound British and vice versa. To this end, I contrasted the evaluations of AmE and BrE into a single variable, henceforth *contrasted target preference*, by subtracting the values of AmE target preference from the corresponding BrE values. Again, the values are derived from a 5-point Likert scale, so the contrasted values range from –4 (= 1 – 5), which indicates the strongest possible preference for AmE over BrE, to 4 (= 5 – 1), which indicates the strongest possible preference for BrE over AmE. Obtaining a single variable by contrasting the target preferences also makes it more straightforward to correlate target preference to lexical choices and language attitudes in subsections 4.4 and 4.5, respectively.

Figure 3 Contrasted target preference – Bar chart

Notes: (−4) represents the strongest possible evaluation of American English over British English.
4 represents the strongest possible evaluation of British English over American English.

The distribution of contrasted target preferences in Figure 3 illustrates that there is generally a preference one way or the other. There were only 5 participants whose contrasted target preference score was 0 indicating that they evaluated AmE and BrE equally as target varieties. It seems that the participants with a BrE preference generally feel stronger about their target pereference as values from 2 to 4 are more common than the respective negative values. On the other hand, it is more common to slightly favour AmE (−1) than to slightly favour BrE (1).

To summarise this subsection, BrE is slightly favoured among the 40 participants, but the difference is small and statistically insignificant (*p* = .329), i.e. it cannot be extrapolated to a wider population. Contrasting the target preferences indicates that most participants do have a preference one way or the other and the collectively quite even result is due to two contradicting groups (19 BrE, 16 AmE, 5 neutral). Notably, according to all five interviewed participants, the evaluations of AmE and BrE as target varieties strongly emphasise phonology and evaluating AmE and BrE essentially means evaluating GA and RP, respectively.

5.3 Language Attitudes toward AmE and BrE

This study also set out to explore how the participants view AmE and BrE on the well-established attitudinal dimensions of status and competence and social attractiveness. As mentioned in the relevant subsection on methodology (4.3), I had to restrict the attitude scales measuring these two dimensions to one for status and competence (sophistication) and to two for social attractiveness (coolness and pleasantness).

As with target preferences, I contrasted the evaluations of AmE and BrE by subtracting the AmE values from the BrE values, resulting in a scale from (–4) (strong inclination to AmE) to 4 (strong inclination to BrE). The central tendencies of the contrasted values are presented in Table 3.

Table 3 Contrasted attitudes toward AmE and BrE – Central tendencies

N = 40	Median	Mode
Sophistication	2	2
Coolness	0	0
Pleasantness	0	0

Notes: All variables are based on a 9-point scale.
 (–4) represents the strongest possible evaluation of American English over British English.
 4 represents the strongest possible evaluation of British English over American English.

BrE is clearly deemed more sophisticated than AmE, which, according to Van der Waerden Test, is a statistically significant result $t(39) = -11.37$, $p < .001$. The median ratings of the other attitudinal scales were equal, resulting in contrasted values 0. The little differences there were on the scales of coolness and pleasantness proved statistically insignificant (p-values: .287 and .345, respectively). However, the differences observed among the sample population are presented in the following figures, which display the distributions of the contrasted attitudinal evaluations of sophistication, coolness and pleasantness of AmE and BrE.
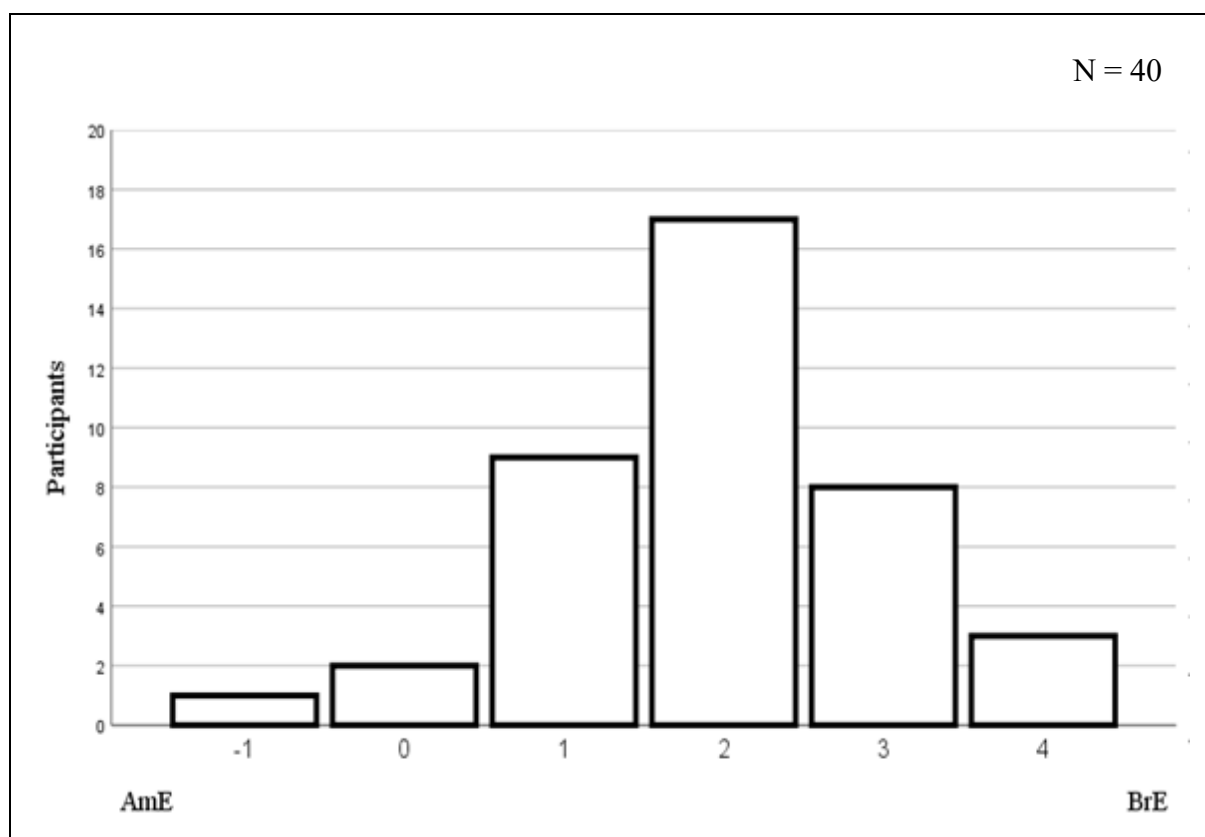
Figure 4 Contrasted evaluations of sophistication – Bar chart

Notes: (−4) represents the strongest possible evaluation of American English over British English.
     4 represents the strongest possible evaluation of British English over American English.
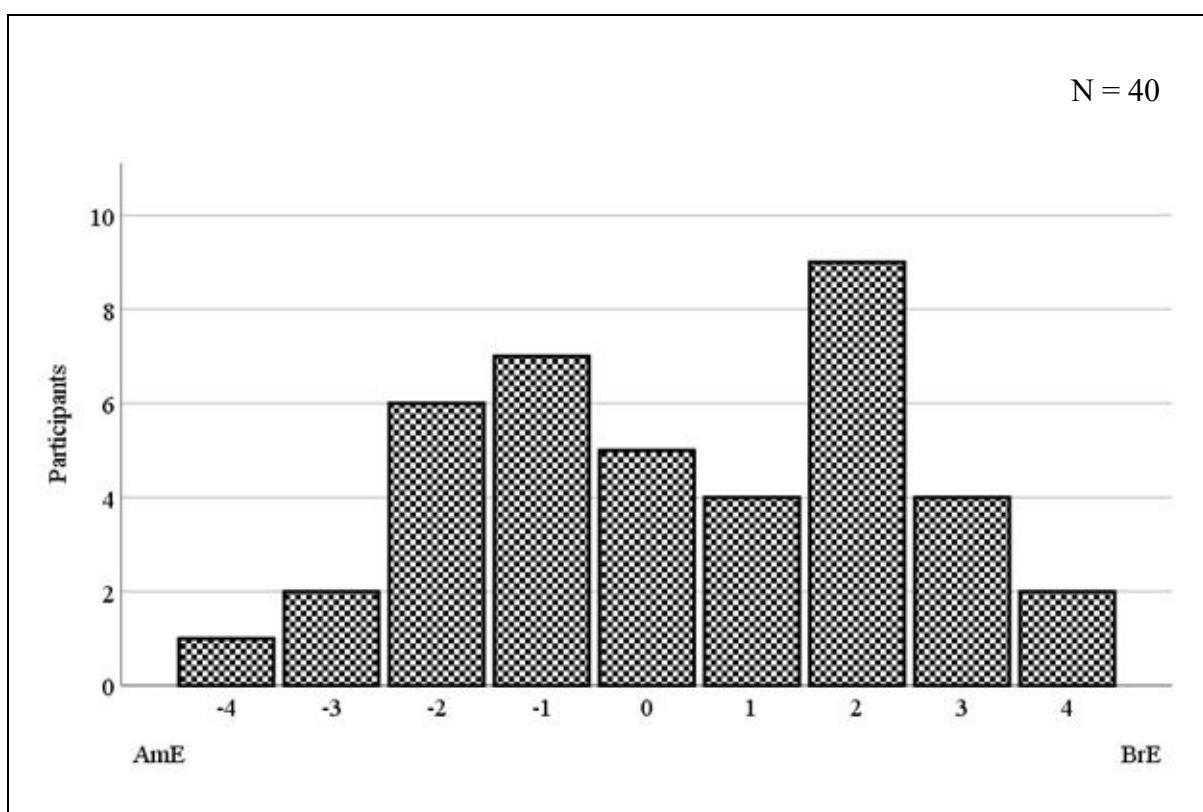

Figure 5 Contrasted evaluations of coolness – Bar chart

Notes: (−4) represents the strongest possible evaluation of American English over British English.
     4 represents the strongest possible evaluation of British English over American English.

Figure 6 Contrasted evaluations of pleasantness – Bar chart

Notes: (−4) represents the strongest possible evaluation of American English over British English.

     4 represents the strongest possible evaluation of British English over American English.

The evaluations of coolness are skewed slightly toward AmE and those of pleasantness slightly toward BrE. Nevertheless, in both cases, a neutral stance was clearly the most common one. This is in contrast to target preferences, in which the collectively rather neutral results were due to two opposite "camps" levelling the scores.

In summary, BrE was rated significantly higher on the scale of sophistication, representing status and competence. As the participants were almost unanimous in this respect, it also suggests that BrE remains the variety of overt prestige, which, again, means status granted on a societal level. The evaluations of AmE and BrE on the scales of coolness and pleasantness, representing social attractiveness, manifested only small differences that proved statistically insignificant.

5.4 Correlation of Lexical Choices and Target Preference

As demonstrated in subsections 5.1 and 5.2, the participants' strongly favour AmE lexicon despite BrE being more popular as a target variety. In this subsection, I explore whether there is, nevertheless, a correlation. It may be useful to start with a visual presentation of the two data sets' relationship.
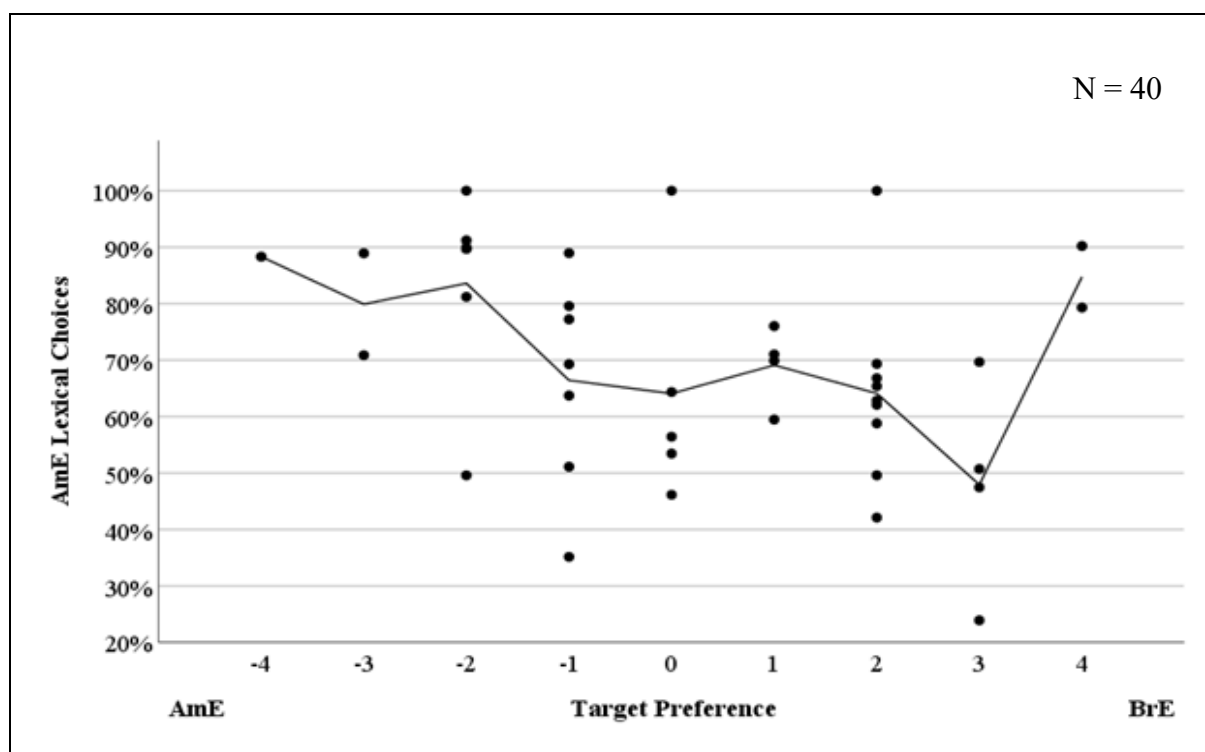
Figure 7 Target preference and American English lexical choices – Scatterplot
Notes: (–4) represents the strongest possible evaluation of American English over British English.
    4 represents the strongest possible evaluation of British English over American English.

The dots represent individual participants and the line indicates the mean score of AmE lexical choices for each group of participants with the same target preference. The line descends in a rather linear fashion, suggesting that AmE lexical choices decrease as the Target Preference for BrE grows, until there is a clear jump at the end (target preference value 4). It is indeed odd that the participants with the maximal preference for BrE over AmE scored clearly above the total mean in terms of AmE lexical choices (79.3% and 90.2% versus 68.7%). Such outliers could, of course, be anomalies resultant of the small sample size. The other target preference groups also have a lot of deviations from the fit line. Whether this is a problem for the significance of the correlation, and whether there, in fact, is a correlation, can be determined through statistical means.

As the target preference is an ordinal variable, parametric tests are not pertinent. The standard non-parametric method for establishing a correlation between two variables is the *Spearman's rank order correlation,* or *Spearman's rho (s_r)*, which correlates the ranks of the data instead of their actual values (Dörnyei 2007, 230).

Table 4 Correlation of target preference and lexical choices

Spearman's rho (s_r)	Significance (p)
.347	.028

The result in Table 4 indicates that 12.0% ($s_r{}^2 = .347^2 = .120$) of the variance is explained by the relationship of the variables, as opposed to other factors or randomness (Dörnyei 2007, 224). The p-value (.028), in turn, indicates that there is a 2.8% chance of the result being due to mere chance. As the p-value is clearly below the standard 5% threshold for statistical significance in social sciences (Dörnyei 2007, 210), the result can be applied to Finnish undergraduate students of English with confidence. In other words, students make lexical choices in accordance with their self-selected targets, at least to some degree.

5.5 Ordinal Regression Analysis of Target Preference and Attitudes

I also set out to explore how well target preference is predicted by the status and competence and the social attractiveness assigned to AmE and BrE. As I have assigned a single scale (sophistication) to account for status and competence and two (coolness and pleasantness) to account for social attractiveness, there are altogether three independent variables and one dependent variable (target preference).

Again, I want to start with a visual as it may make it easier to perceive the consequent numerical data. The following scatterplots relate each individual attitudinal dimension with target preference.
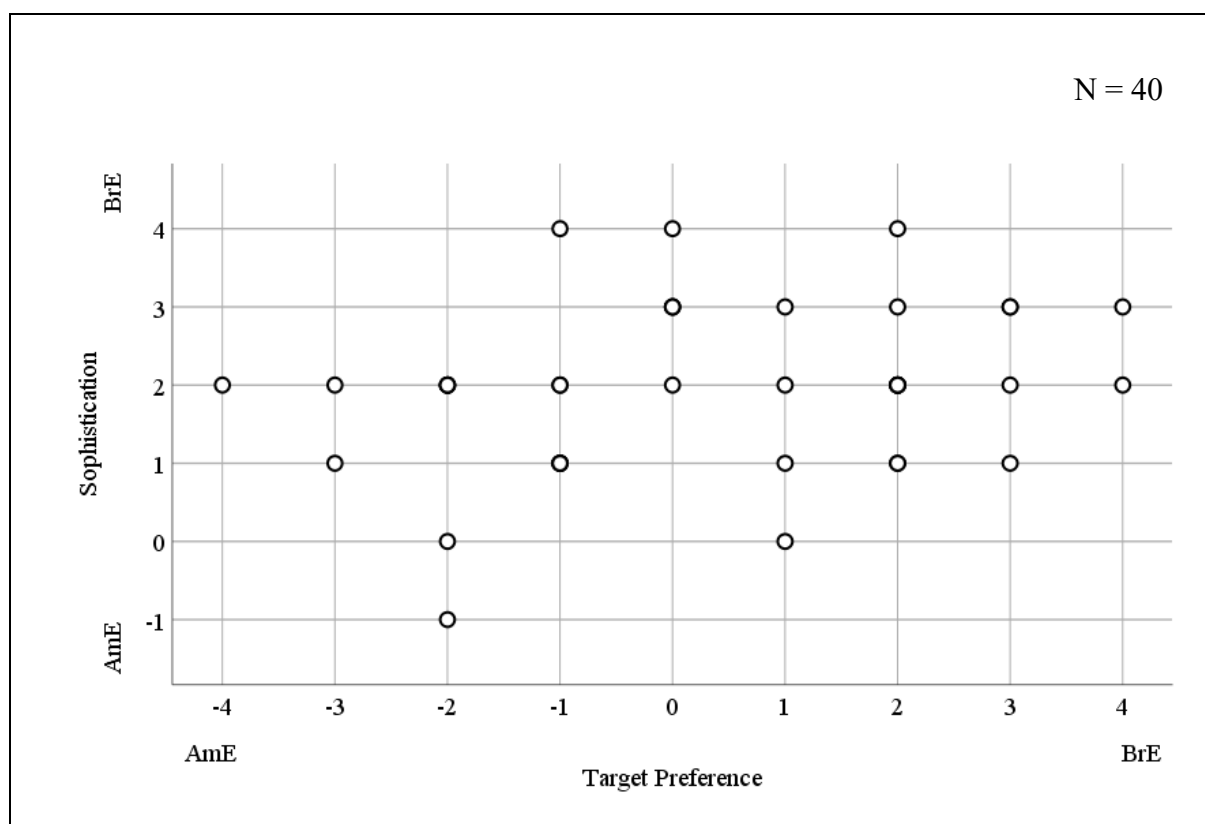
Figure 8 Contrasted Evaluations of Sophistication and Target Preference – Scatterplot
Notes: (−4) represents the strongest possible evaluation of American English over British English.
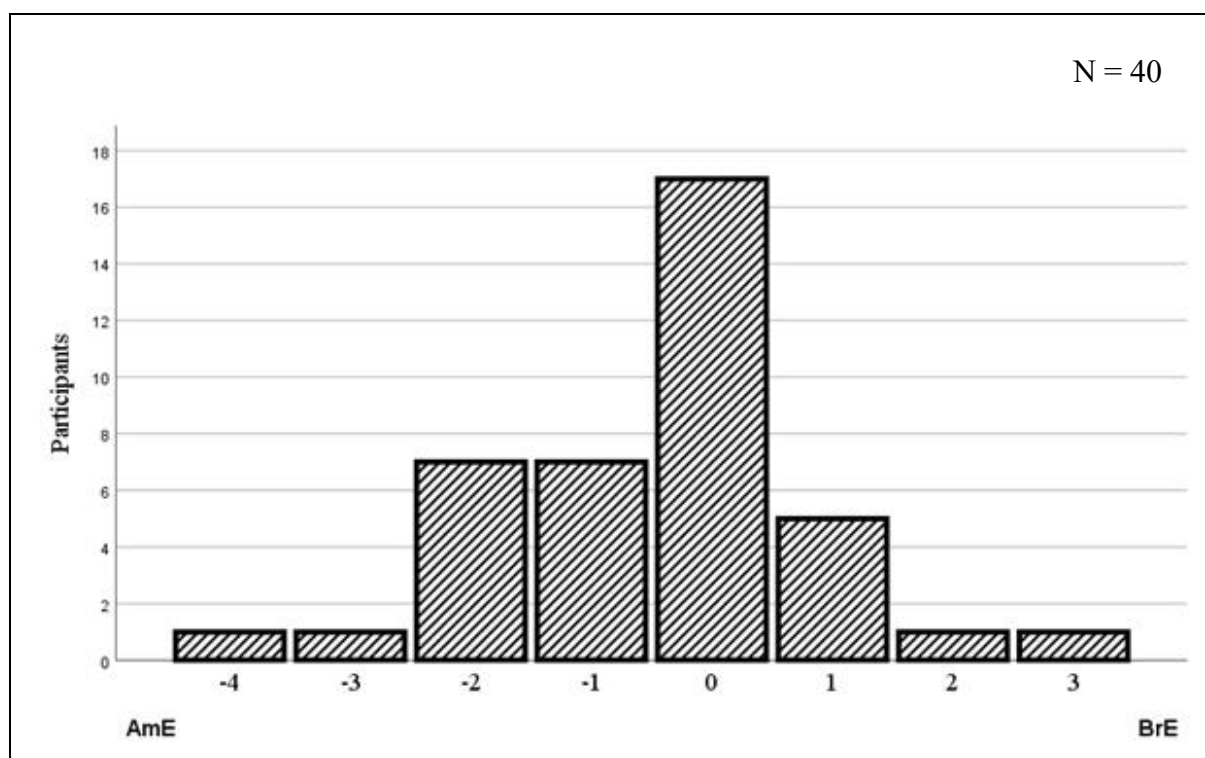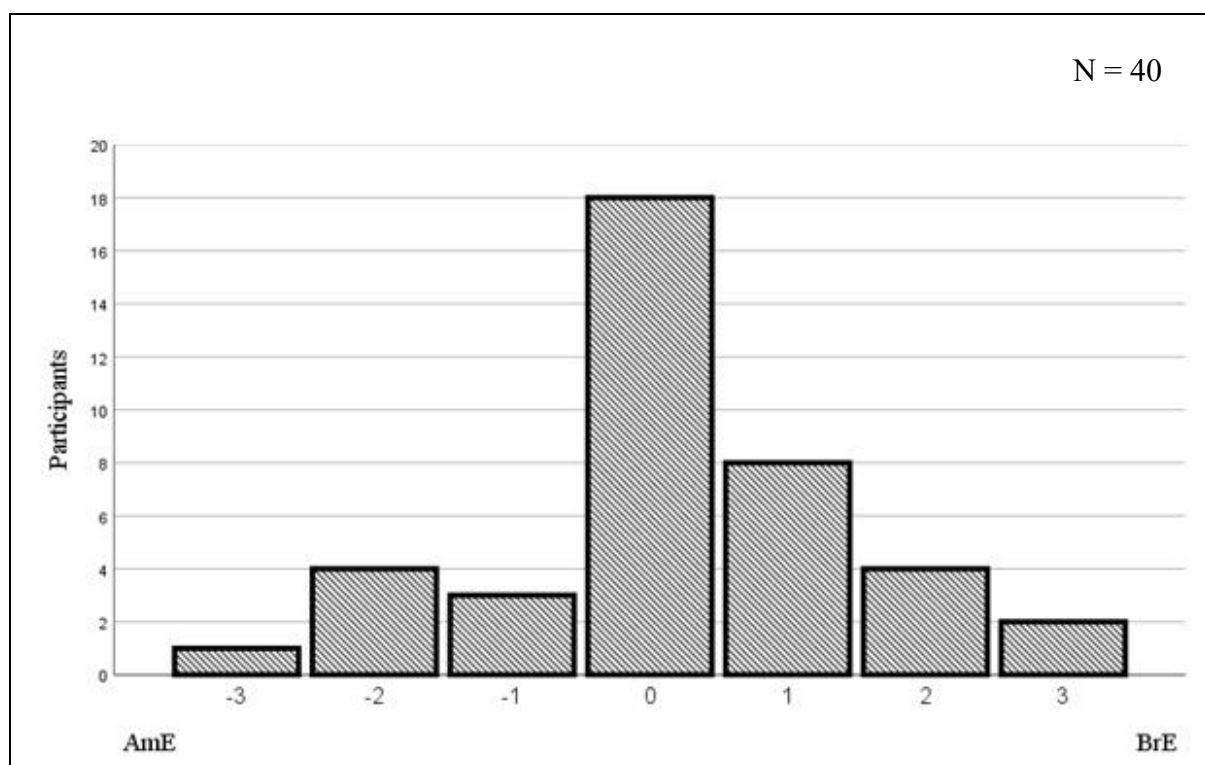   4 represents the strongest possible evaluation of British English over American English.

As established in subsection 4.3 there is a general consensus that BrE is more sophisticated. However, this opinion seems to be slightly stronger among those participants that favour BrE as a target preference. Only one participant with an AmE target preference estimated BrE 3 or 4 points higher than AmE, whereas among those with an BrE there were six. Please note that responses that were identical on both scales are stacked and some of the dots represent more than one respondent.
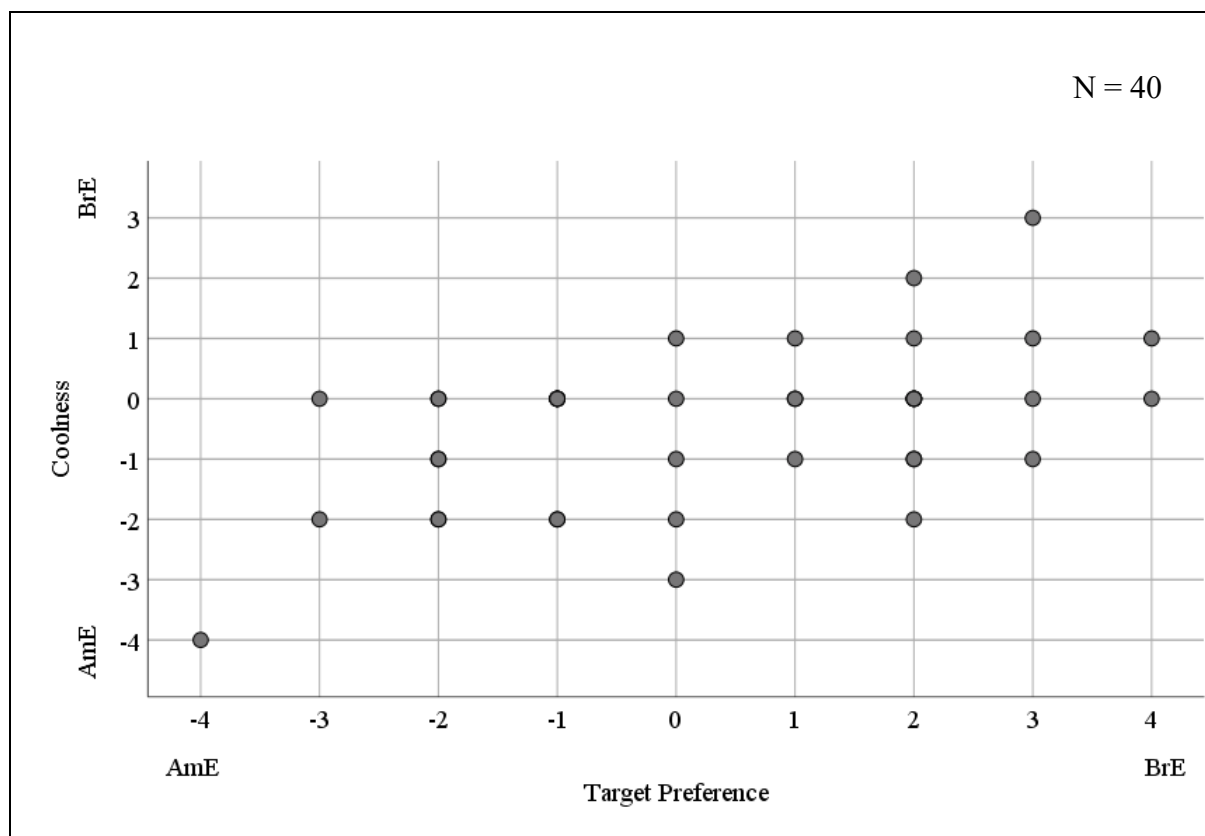
Figure 9 Contrasted Evaluations of Coolness and Target Preference – Scatterplot
Notes: (−4) represents the strongest possible evaluation of American English over British English.
  4 represents the strongest possible evaluation of British English over American English.

Here one can see a much stronger upward tendency and it seems rather clear that there is a positive correlation.
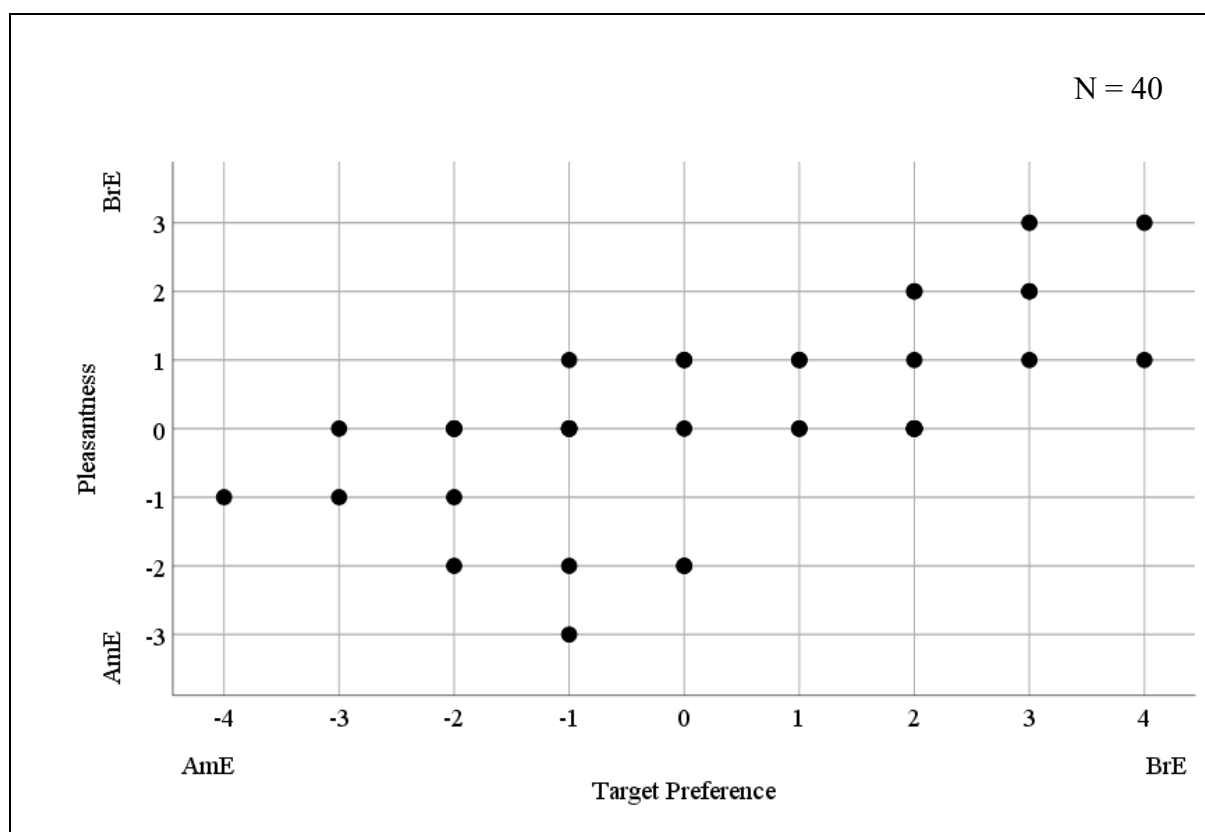
Figure 10 Contrasted Evaluations of Pleasantness and Target Preference – Scatterplot

Notes: (−4) represents the strongest possible evaluation of American English over British English.
       4 represents the strongest possible evaluation of British English over American English.

As with coolness, the evaluations of pleasantness seem to clearly correlate with target preferences. However, this must be confirmed mathematically. Regression analysis is a good method for examining multiple correlations (Dörnyei 2007, 226) and determining how well the independent variables (*sophistication, coolness* and *pleasantness*) predict the values of the dependent variable (*target preference*). More precisely, the ordinal nature of the data at hand necessitates an ordinal regression analysis (Osborne 2015, 408).

The conditions that the data must satisfy for this method to be valid are 1) the dependent variable must be in order, and 2) "the effects of any (and all) independent variables are the same regardless of what two groups are being compared" (ibid.). The first assumption is met as data obtained using a Likert scale are ordinal by nature. The second assumption can be tested on SPSS with the *test of parallel lines* (Osborne 2015, 409). The null hypothesis of this test is that the model is indeed valid across all response categories, so, if the significance is greater than .05, the null hypothesis stands, and the second condition is fulfilled (ibid.). For the present model, SPSS yielded a significance of .97 (> .05) and, thus, the assumptions of an ordinal regression analysis are met.

When running an ordinal regression on SPSS it by default provides a calculation of how well the model with the predictor variables fits the data as opposed to an *intercept only* model, which does not account for any predictors. (ibid). In the present case there was a significant improvement in fit ($p < .001$). In other words, the model works and should yield meaningful results.

Table 5 Ordinal regression analysis of target preference predicted by language attitudes

Variable	Sophistication	Coolness	Pleasantnes
Regression coefficient	.471	.678	.594
Significance (p)	.125	.017	.043

All variables are based on a 9-point scale between American English and British English.

The regression coefficients represent the expected change in the dependent variable (target preference) for each point changed in the respective independent variable. For example, a participant that evaluates the sophistication of BrE to be one point higher than AmE on the Likert scale is expected to have .471 point higher target preference for BrE. However, this is not a statistically significant result ($p = .125 > .05$). Coolness and pleasantness, on the other hand are both statistically significant. For each point more or less on the scale of coolness, target preference is expected to change by .678 points in the same direction. For pleasantness the respective figure is .594.

Intuitively it would seem that target variety is chosen mainly based on its perceived social attractiveness. However, the correlation does not help establishing the direction of causation. It may be that target preference is based on something else, and then reflected also on the perceived social attractiveness of the preferred variety. It is even possible that there is an underlying variable as the root cause of both target preference and social attractiveness. I try to shed some light on the causality via interviews and previous research in the following section.

6 Discussion and Conclusion

Three variables were examined regarding Finnish undergraduate students of English and their relationship with the American and British varieties of English: *varietally distinct lexical choices*, *target preferences*, and *language attitudes*. The results of each individual variable are mostly consistent with previous research, which serves to reaffirm the validity of the older studies and corroborate the results of the newer ones.

This study also set out to examine some of the correlations between these variables, the main contribution being the correlation of varietally distinct lexical choices and target preferences, which, to my knowledge has not been studied before. I also examined the relationship of target preference and language attitudes. Before elaborating on the interrelations of the variables, I discuss each one independently. Finally, I conclude by discussing the practical implications of the results and possible avenues for further research.

Lexical choices

The 40 participants produced 309 varietally distinct words, which were 66.0% AmE and 34.0% BrE. When weighted for distinctiveness, i.e. accounting for the occurrences of AmE words in the British corpus and vice versa, the respective figures are 68.7% and 31.3%. This is in line with Forsman's (2000, 172–3) finding that Finnish upper secondary students use primarily AmE vocabulary. Gonçalves et al. (2018, 11), in turn, found Finland to be a part of a global trend of Americanisation with respect to vocabulary. Out of the 30 countries in the study, only the United Kingdom, Ireland and India produced more BrE than AmE vocabulary, and in the case of India, the result was a virtual tie (ibid.).

According to Gonçalves et al. (ibid.), Finland is in the vanguard of this global trend. Most countries ranking ahead of Finland in the use of AmE vocabulary are either on the American continent or, in the case of the Philippines, have a historical connection to the United States. Only Portugal clearly yielded more AmE vocabulary than Finland without a clear underlying connection to the United States. Looking at the European EFL countries, there seems to be a correlation between subtitling audio-visual entertainment and using AmE vocabulary. Given the dominance of AmE in the media, this seems to suggest media influencing EFL users' vocabulary.

Target Preferences

The strong American influence on the English used in Finland seems to manifest itself also in the attitudes toward AmE as a model. Even though BrE was the preferred target for 47.5% of

the participants, as opposed to 40.0% that preferred AmE, the difference was not statistically significant as per Van der Waerden test ($p = .329$). Previous studies conducted on European university level students of English have found a much clearer preference for BrE. Among the Danish respondents of Ladegaard and Sachdev (2006, 101), 55% aimed for BrE pronunciation and 25% for AmE pronunciation. In Poland, the respective figures were roughly 65% and 20% (data in graph form without exact figures) (Waniek-Klimczak and Klimczak 2005, 242). In Finland, Leppänen et al. (2009, Table 15a.4., n.p.) found that among people with a university degree BrE was favoured by 50.0% and AmE by 32.3%.

Leppänen et al. did have a slightly different approach, as they examined the general attractiveness of the varieties, not necessarily their appeal as a personal model. This, however, is likely a minor distinction. Notably, the results of Leppänen et al. are a little closer to the present study than those obtained in Denmark and Poland. This may be another indication of Finland being particularly Americanised among European EFL countries. On the other hand, in this small sample of studies, the more recent the study the more popular AmE is as a target, and the underlying reason for the different results could also be temporal. This interpretation is further supported by young age correlating with preference for AmE over BrE in Finland (Leppänen et al. 2011, Table 15a.2., n.p.). It would also be in line with the prediction of Bayard et al. (2001, 44) that AmE will bypass BrE as the preferred target variety.

Finally, clear as the result of the present study may seem in terms of AmE and BrE as target varieties, one should bear in mind that its dichotomous nature may be misleading. It seems likely that it caused the participants to evaluate the two varieties primarily in relation to one another, without consideration for other native varieties or ELF. Indeed, some of the interviewed participants expressed a clear target preference in the questionnaire (Appendix 6), but in the interview everyone stressed intelligibility as the main target. The interviews also revealed that evaluating AmE and BrE as target varieties essentially meant evaluating GA and RP pronunciation models.

Language Attitudes

I also set out to investigate the attitudes attributed to the two varieties. The participants clearly favoured BrE on the scale of sophistication, but estimated AmE and BrE roughly equally with respect to coolness and pleasantness. This suggests that, if sophistication is accepted as a good indicator of status and competence, and coolness and pleasantness of social attractiveness, the results are consistent with the results of previous studies conducted on undergraduate students of English in Spain (Carrie 2017), Denmark (Ladegaard and Sachdev 2006) and Poland

(Janicka, Kul and Weckwerth 2005), as well as upper secondary school students in Norway (Rindal 2010) and Finland (Forsman 2000).

As a measure of post-validation, I had the five interviewed participants evaluate two more terms (*competent* and *likeable*). Evaluations of competence corresponded very well to sophistication and evaluations of likeability corresponded to pleasantness. Only one evaluation deviated by more than one point, as participant #45 evaluated the competence of AmE two points higher than its sophistication. Such consistent results lend some validity to the attitudinal scales used in the questionnaire.

In subsection 3.1, I attributed some of the status of BrE in Europe to English originating in Britain and the vestigial prestige of the British Empire. The interviews also revealed an association with the history of the stratified class system in Britain, which seems to promote the status and competence of BrE. When asked about the first things that come to mind about BrE, participant #15 named "Downton Abbey", a TV-series about early 20th century aristocracy in England. Both BrE aimers and one AmE aimer (#25) also described BrE as sounding either "posh" or "upper class" (the latter is my translation of Finnish *yläluokkainen*). The same interviewees also considered BrE more "refined" (my translation of Finnish *hieno*).

AmE was called the "standard" (my translation of the roughly synonymous Finnish words *standardi* and *vakio*) variety in Finland by two interviewees (#8 and #25) and "laid-back" (my translation of Finnish *rento*) by one (#25). Generally, AmE seemed to evoke much less emotive responses than BrE. It was viewed as the variety that one acquires, if no special effort is made to adopt another variety. This is, of course, supported by the generally AmE vocabulary among the participants.

This is an interesting finding given the emphasis on BrE in formal education (Ylönen 2007; Tergujeff 2012). Also, the interviewees said that their teachers had primarily spoken and taught BrE, and that most lecturers in the university speak BrE while not one speaks AmE. The media was unanimously viewed as the reason for the hegemony of AmE in Finland. Forsman (2000) and Sjöholm (2000) also noted the power of media over formal education. I suspect that this effect has only increased in the twenty years that has elapsed since their studies.

Correlation of Lexical Choices and Target Preference

As far as I am aware, sociolinguistically motivated lexical variation has not been studied among EFL users. The closest study I was able to find was that of Rindal (2010) on the relationship of target preference and pronunciation among Norwegian 17–18-year-old upper secondary school students. The difference in the level of formal education is somewhat mitigated, firstly, by the

participants in Rindal's study being in "a prestigious school in Oslo, where academic results […] were among the highest in the area" (Rindal 2010, 250), and, secondly, by the participants of the present study being generally first and second year students, judging by the course they attended and their low mean age (21.8). Also, the general circumstances are rather similar, as the media in both countries is subtitled rather than dubbed and BrE is the most common model in formal education (Rindal, 2010, 241; Tergujeff 2012, 40; Ylönen 2007, 35).

Among the participants in Rindal's study, target Preferences were 42.1% AmE and 57.9% BrE. The corresponding figures in the present study, excluding the five neutral responses, were 45.7% and 54.3%. Also, the observed varietally distinct features were distributed almost equally. Rindal found that 67.2% of the observed phonemes were AmE, whereas the present study resulted in a collectively 68.7% AmE vocabulary. I posited earlier that the frequency-induced salience of varietally distinct phonological features should make them easier to incorporate to one's idiolect. Consequently, the results of the two studies being almost identical may reflect the difference in the level of formal education. The sociolinguistic salience of phonology is also supported by all interviewed participants agreeing that sounding American or British almost exclusively depends on pronunciation. Only one participant (#8) said that he pays some attention to vocabulary when trying to emulate his preferred variety.

Yet, seemingly in spite of not trying, the participants with a BrE preference produced significantly less AmE features. As the present study measured target preference on a Likert scale, I was also able to determine the correlation between target preferences and lexical choices, which proved to be relatively weak, but statistically significant, $s_r = .347$, $p = .028$. In other words, 12.0% ($.347^2$) of the variation of each variable can be attributed to the other. It is worth noting that categorising the words examined in this study as either AmE or BrE is an idealisation. On average, the words examined in this study are used in the corpus of the other variety (BrE in COCA and AmE in BNC) 1/9 times. In other words, even perfectly adopting the target variety patterns would not result in a 100% correlation and the .347 correlation is more meaningful than it seems at first glance. Also, all participants do not wish to emulate all aspects of their target variety but choose the features that suit their needs. For example, both participants that aimed for BrE considered some AmE words (#15: *elevator*; #45: *apartment* and *vacation*) more "refined" or "formal" (my translations of Finnish *hieno* and *muodollinen*) than their BrE counterparts and preferred using them.

One possible underlying factor for the correlation could be contact with native speakers. However, this was clearly not a significant factor in the present study, as the participants that had spent over a month in English-speaking countries were eliminated and the rest reported

only scarce contact with native speakers. Furthermore, the latter performed opposite to what one might expect. At least weekly contact with AmE speakers was reported by 6 participants and BrE speakers by 1. The participants with regular AmE contact produced AmE words below the mean (59.0%) and the one with BrE contact close to it (69.3%).

In subsections 2 and 2.2, I argued in favour of media's influence to variation in an EFL setting, and one possible explanation for the correlation would be that the participants consume more media of their target variety, thus, acquiring and activating vocabulary of that variety. The interviews lend some support to this, as the two interviewees whose lexical choices were less AmE than the mean (Appendix 6) asserted that they consumed a lot of BrE media. While these two participants also said that they are probably more exposed to AmE media, the other three participants unequivocally stated that they consume more AmE media and made no mention of BrE media. All of them made lexical choices clearly above the mean (Appendix 6).

It could also be that, whatever the media influence, the participants make use of the lexicon of their target variety unconsciously, even if they do not consider it important enough for conscious monitoring. The prerequisite of this is, of course, awareness of the varietally distinct vocabulary. The interviewed participants clearly met this prerequisite as they consistently knew, which of the tested words were AmE and which ones were BrE. Only *main street* vs *high street* and *counterclockwise* vs *anticlockwise* caused doubt to two of the interviewees (Appendix 7, question 2).

I have so far discussed exposure and attitudes as the causes of variation. These correspond to Labov's (2001, 16) socially motivated principles of density and imitation, respectively, but I must also account for the third principle, the linguistically motivated principle of least effort. Finegan and Biber (2001, 261) refer to the same phenomenon as *internal constraints* and dissect it further into the *clarity* and *ease mandates*. As relates to the present study, the clarity mandate could cause the AmE *main street* to be favoured over BrE *high street* simply because of the linguistic opacity of the latter term. After all, the street is not necessarily more elevated, or *high*, than any other. As for the ease mandate, some participants could have used *line* because they were unsure of how to spell *queue*. In the context of the present study, L1 influence could also be a phenomenon related to the *ease mandate*. However, I doubt that it is much of a factor given that none of the studied words are neither borrowed into Finnish nor cognates with their Finnish equivalent. In fact, I do not think that any of the factors discussed in this paragraph have a significant impact on the results as there is no reason to believe they would consistently affect one variety more than the other.

Correlation of Target Preference and Language Attitudes

Ordinal regression analysis revealed that sophistication, which represented the dimension of status and competence, was not a significant predictor variable of target preference ($p = .125$). This was to be expected, as the participants were quite unanimous in their evaluations that BrE is more sophisticated than AmE, yet, the two were valued rather equally as target varieties. On the other hand, coolness and pleasantness, which represented social attractiveness, were deemed significant predictors of target preference. Both the target preferences and the attitude scales were measured on contrasted Likert scales with 9 level. For each point of coolness given to one variety over the other on this scale, target preference is expected to be .678 higher as well ($p = .017$). Similarly, for each point of pleasantness target preference is expected to be .594 points higher ($p = .043$).

The intuitive causal relation would be that the perceived social attractiveness causes the participants to favour one variety over the other, but that is not necessarily the case. Previous European studies (Janicka, Kul and Weckwerth 2005; Ladegaard and Sachdev 2006; Carrie 2017) have found that university students of English typically base their target preference on language attitudes but sometimes also on practical considerations, such as ease of acquisition and perceived future benefits of speaking a certain variety. It is possible that some Finnish students also adopt a target preference for one of these reasons and then reflect it also on the perceived social attractiveness of their preferred variety.

Janicka, Kul and Weckwerth (2005, 254) found that the pragmatic justifications were typical to those students that preferred AmE as the target, whereas BrE evoked more emotional responses. As noted above, there was a similar tendency among the students interviewed for this study. All three interviewees with AmE target preference justified it with familiarity, which might be attributed to the "ease" category. Participant #25 also associated AmE with social attractiveness, calling it "more laid back" (my translation of Finnish "rennompaa"). The two interviewees with a BrE preference justified it exclusively with affective reasons. Both said that BrE is "more refined" (my translation of Finnish *hienompaa*). Participant #15 also considered BrE to be more beautiful (my translation of Finnish *kauniimpaa*). This difference in emotional responses is likely reflected in that the participants with BrE target preference generally rated it quite high in relation to AmE, whereas those with AmE preference rated AmE only slightly higher than BrE. Even when prompted, the participants did not see any future benefits in emulating one variety over the other.

In conclusion, based on the other European studies quoted above, as well as the five interviews conducted for this study, it is likely that language attitudes are the main reason for

the target choices of undergraduate students in Finland. However, status and competence attributed to each variety does not seem to be very important to Finnish undergraduates. Both measured dimensions of social attractiveness, on the other hand, clearly correlated with target preference. Nevertheless, familiarity and ease of acquisition cannot be excluded as reasons for target selection. There is some tentative evidence that such pragmatic considerations are particularly typical to the students with an AmE target preference.

Having stressed the tentative nature of the causal relationships suggested above, I should note that I opine the most salient finding of the study to be independent of the precise mechanisms behind the correlations. In my mind, the most salient discovery is the students' ability to use varieties of English to construct their self-image and social identity. From this point-of-view it is a moot point whether they align their preferred target variety with their personal values by modifying the target choice, the attitudes assigned to the target, or both. In any case, it is a matter of constructing one's ideal self (Dörnyei and Chan 2013, 438). The correlation between target preference and the use of varietally distinct lexicon indicates that they successfully send socially motivated signals with their lexical choices and, thus, reduce the gap between their actual and ideal selves (ibid.).

Practical implications

There are three findings that all have ramifications on the model of English employed in formal education in Finland.

1) Strong global and domestic Americanisation of English vocabulary
2) Increasingly favourable attitudes toward AmE as a target variety
3) Ability of students to use varieties of English for identity construction

At present there is no official guideline as to which variety teachers should teach, but rather that the students should learn about different varieties (National Curriculum 2014; National Curriculum 2019). Nevertheless, the most recent studies I found on the subject suggest that the tuition is still very much focused on BrE, both in terms of materials (Ylönen 2007), and models employed by the teachers (Tergujeff 2012). As the present study indicates that at least some advanced students use varieties of English to construct their identity, it would be ideal if they were given personalised guidance toward their self-selected targets.

Judging by the, admittedly few, interviews conducted, this is particularly true with regard to phonology, as that seems to be by far the most salient aspect of linguistic varieties. This is supported by Zuegler (1988, 34): "[…] phonology is a domain within which one's identity is

expressed […]". Vocabulary, in turn, seems to be less important for identity construction, and it may, thus, be justified to take a more pragmatic approach and emphasise intelligibility in teaching. However, it is not clear what is the best way of doing this. In the following, I outline three alternatives:

1) Embrace the Americanisation. The present study adds to the evidence on the global Americanisation of English lexicon. It might, therefore, be advisable to comply with the general trend and teach primarily AmE vocabulary.
2) Keep the *status quo* favouring BrE. If most students learn AmE from the media in any case, learning BrE in school should add to the depth of their vocabulary.
3) Take an ELF approach. Accept the variation between native varieties and teach the internationally most frequent variables, e.g. based on the Vienna-Oxford International Corpus of English.

On a global scale, the participants of this study can also be considered representative of EFL users in general. Along with the Norwegian adolescents studied by Rindal (2010), they make a strong case for some EFL users using English varieties to construct their social identity and to achieve their social goals in a way that does not necessarily correspond to the sociolinguistic variation of any native variety. Furthermore, in both this study and that of Rindal and Piercy (2013) there were some participants that did not aim at any native variety. These to me are compelling arguments in favour of acknowledging EFL speakers as autonomous users of English. Sociolinguists used to view also L1-speakers' linguistic identity as something imposed on them by the social group they were a part of (Dyer 2007, 103). Finally, in the 1980s people's active agency in the construction of their identity was acknowledged (Dyer 2007, 104-5). Later still, this has been extended to ESL users (Hansen Edwards 2008, 272). The present study indicates that it is time to grant EFL users the same consideration.

Limitations of the present study

The major methodological shortcoming of the present study is the already mentioned scarceness of attitudinal scales, which future studies can easily remedy by reserving a little more time for administering the questionnaire. Another possible improvement would be accounting for word frequencies in the test of varietally distinct lexical choices. For example, the relatively infrequent AmE word *counterclockwise* was used much more than the corresponding BrE variant *anticlockwise* (22 versus 2 tokens), which means that it has a relatively significant impact on the percentages of AmE and BrE lexical choices. Arguably, *autumn*, the most used BrE word in the test (17 tokens versus 9 tokens of *fall*), should have more weight than

counterclockwise as it is likely to be used much more in natural production. Thus, possible future studies could improve upon my methodology by weighting the values of distinctiveness also in accordance with the frequencies of use.

The distinctive pairs *anticlockwise/counterclockwise* and *autumn/fall* were also the ones most frequently omitted altogether, 16 and 14 times respectively. I suspect that due to the relative infrequency of the words *anticlockwise* and *counterclockwise* all students either did not know them or were not able to retrieve them in time. A possible reason for the omission of *autumn/fall* was that the corresponding part in the Finnish text was *syysloma* (*autumn holiday* [my translation]), which is not a concept in English-speaking countries. Consequently, literal translation may have seemed weird to the participants, many of whom translated it only as *vacation* or *holiday*. More comprehensive piloting would have undoubtedly helped to avoid such problems.

Suggestions for future studies

Even though I maintain that the abovementioned ramifications can be based on just some EFL users' making sociolinguistically motivated choices between varieties of English, the scope of the phenomenon requires much more research. Indeed, Rindal's (2010) was the only other study I was able to find that related target preference to performance. Furthermore, to my knowledge, it is the only study that has examined intraspeaker varietal variation that is context-dependent, i.e. employing different varieties of English for different styles and registers. EFL variation should also be studied from the point-of-view of Speech Accomodation Theory, i.e. whether EFL users accommodate their English to that of their collocutor.

Also, the possible systematicity of the variation in any given context should be examined. The data of the present study is too restricted to reliably establish any patterns, but it does suggest that the variation is not equal among the contrastive pairs of words. For example, despite the lexical choices being collectively very American, the BrE *autumn* was considerably more popular than the AmE *fall* (17 versus 9 tokens). On the other hand, *high street* (BrE) was not used once, except for one participant that had lived in England for several years and was, consequently, excluded from the final analysis. Studying the frequency of each varietally distinct lexeme for each referent in ELF would help construct an intelligibility-driven ELF curriculum.

Furthermore, the characteristics of ELF should be studied in different contexts beyond differences relating to ENL varieties. After all, the features that the speakers from various EFL backgrounds employ for mutually intelligible communication do not always stem from any

native model. For example, the relatively irregular morphology of English is often regularised, e.g. *approvement** pro *approval* (Seidlhofer 2011, 103). I would expect great many of the EFL users producing such non-standard forms to still have several options for expressing themselves. In other words, they exhibit intraspeaker variation, which should be studied and not simply assumed random.

Sources

Algeo, John. 1990. "The Briticisms Are Coming! How British English Is Creeping into the American Language." *Journal of English Linguistics* 23, no. 1-2: 123–140.

Bailey, Richard W. 2004. *Language in the USA: Themes for the Twenty-first Century*, edited Edward Finegan and John R. Rickford. "American English its Origins and History." Cambridge: Cambridge University Press. 3–17.

Bayard, Donn, Cynthia Gallois, Ann Weatherall and Jeffery Pittam. 2001. "Pax Americana? Accent attitudinal evaluations in New Zealand, Australia and America." *Journal of Sociolinguistics* 5, no. 1: 22–49. Accessed 17 April 2020. Wiley Online Library.

Birkstedt, Johanna. 2015. "Suomalaisen englanninopetuksen lyhyt historia." [A Short History of English Teaching in Finland.] In *Englannin aika: Elävän kielen kartoitusta* [The time of English: Mapping a living language.], edited by Terttu Nevalainen, Matti Rissanen and Irma Taavitsainen. Accessed 26 April 2020. https://helda.helsinki.fi/handle/10138/153953

Bonfiglio, Thomas Paul. 2002. *Race and the Rise of Standard American*. Accessed 6 March 2020. De Gruyter.

Buschfeld, Sarah, Alexander Kautzsch and Edgar W. Schneider. 2018. *Modeling World Englishes: Assessing the Interplay of Emancipation and Globalization of ESL Varieties*, edited by Sandra C. Deshors and Alexander Kautzsch. Accessed 17 April 2020. EBSCOhost.

Canagarajah, Suresh, and Selim Ben Said. 2011. "World Englishes." In *The Routledge Handbook of Applied Linguistics*, edited by James Simpson. 393–407. Accessed 18 April 2020. EBSCOhost.

Carrie, Erin. 2017. "'British Is Professional, American Is Urban': Attitudes Towards English Reference Accents in Spain." *International Journal of Applied Linguistics* 27, no. 2: 427–447. Accessed 17 April 2020. Wiley Online Library.

Chandra Chari. 2007. *War, Peace and Hegemony in a Globalized World: the Changing Balance of Power in the 21st Century* New York, NY: Routledge.

Darlington, Richard B. 1996. "A normal-scores alternative to the Wilcoxon test." *Richard B. Darlington's homepage*, 09/96. Accessed 16 April 2020. http://node101.psych.cornell.edu/Darlington/normscor.htm.

Davies, Mark. 2004. "British National Corpus." *Brigham Young University's collection of corpora*. Accessed 25 September 2019. https://corpus.byu.edu/bnc/.

Davies, Mark. 2008. "The Corpus of Contemporary American English." *Brigham Young University's collection of corpora*. Accessed 25 September 2019. https://corpus.byu.edu/coca/.

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

Dörnyei, Zoltán, and Letty Chan. 2013. "Motivation and Vision: An Analysis of Future L2 Self Images, Sensory Styles, and Imagery Capacity Across Two Target Languages." *Language Learning* 63, no. 3: 437–462.

Dörnyei, Zoltán, and Tatsuya Taguchi. 2009. *Questionnaires in Second Language Research: Construction, Administration, and Processing*. 2nd ed. Abingdon, Oxon: Routledge.

Doughty, Catherine, and Michael H. Long. 2003. *The Handbook of Second Language Acquisition*. Malden, MA: Blackwell Publishing.

Durham, Mercedes. 2014. *Acquisition of Sociolinguistic Competence in a Lingua Franca Context*. Blue Ridge Summit, PA: Multilingual Matters. Accessed 20 January 2020. De Gruyter.

Education First. 2019. *EF English Proficiancy Index.* Accessed 31.3.2020.
    https://www.ef.com/wwen/epi/.

Edwards, John. 1999. "Refining Our Understanding of Language Attitudes." *Journal of
    Language and Social Psychology* 18.1: 101–110. Accessed 22 February 2020.
    https://doi.org/10.1177/0261927X99018001007.

Félix-Brasdefer, J. César, and Dale April Koike. 2012. *Pragmatic Variation in First and
    Second Language Contexts: Methodological Issues*. Amsterdam: John Benjamins.

Finegan, Edward and Douglas Biber. 2001. "Functionally Motivated Situational Variation."
    In *Style and Sociolinguistic Variation*, edited by John R. Rickford and Penelope Eckert,
    235–267. Cambridge: Cambridge University Press.

Finegan, Edward, and John R. Rickford. 2004. *Language in the USA: Themes for the Twenty-
    first Century*. Cambridge: Cambridge University Press.

Finnpanel. 2020a. *TV-ratings in Finland, week 16/2020.* Accessed 20 April 2020.
    https://www.finnpanel.fi/tulokset/tv/vko/top/2020/16/

———. 2020b. *Television katselu Suomessa vuonna 2019.* [Television Viewing in Finland in
    Year 2019]. Accessed 20 April 2020.
    https://www.finnpanel.fi/lataukset/tv_vuosi_2020.pdf

Forsman, Liselott. 2000. "Effects of media input on the incidental learning of English and on
    linguistic attitudes among Finland-Swedish high-school students." In *Perspectives on
    Language and Communication in Multilingual Education*, edited by Kaj Sjöholm and
    Anna-Lena Østern, 167–81. Vasa: Åbo Akademi University.

Fox, Kate. 2004. *Watching the English: The Hidden Rules of English Behaviour*. London:
    Hodder & Stoughton.

Gardner, Richard C. 1985. *Social Psychology and Second Language Learning: the Role of
    Attitudes and Motivation*. London: Arnold.

Garrett, Peter. 2007. "Language Attitudes." In *The Routledge Companion to Sociolinguistics*,
    edited by Carmen Llamas, Louise Mullany, and Peter Stockwell. Abingdon: Routledge.
    116–122.

Geeraerts, Dirk, Stefan Grondelaers, and Peter Bakema. *The Structure of Lexical Variation:
    Meaning, Naming, and Context.* Cognitive Linguistics Research 5. Accessed 19 April
    2020. De Gruyter.

Gilman, Andrew. 2015. "Don't Do the Wilcoxon." *Statistical Modeling, Causal Inference,
    and Social Science,* 13 July 2015. Accessed 27.3.2020.
    https://statmodeling.stat.columbia.edu/2015/07/13/dont-do-the-wilcoxon/

Gonçalves, Bruno, Lucía Loureiro-Porto, José Ramasco, and David Sánchez. 2018. "Mapping
    the Americanization of English in Space and Time." *arXiv.org* 13, no. 5. e0197741.
    Accessed 13 October. http://search.proquest.com/docview/2073955616/.

Gould, Jonathan. 2007. *Can't Buy Me Love: The Beatles, Britain and America.* London:
    Portrait.

Halliday, Michael Alexander Kirkwood. 1978. *Language as Social Semiotic: The Social
    Interpretation of Language and Meaning*. London: Arnold.

Hansen Edwards, Jette G. 2008. "Social factors and variation in production in L2 phonology."
    In *Phonology and Second Language Acquisition*, edited by Jette G. Hansen Edwards
    and Mary L. Zampini. John Benjamins Publishing Company, 2008. Accessed 22 April
    2020. ProQuest Ebook Central. –279

Hasan, Ruqaiya, and Jonathan J. Webster. 2009. *The Collected Works of Ruqaiya Hasan: 2,
    Semantic Variation Meaning in Society and in Sociolinguistics*. London: Equinox.

Haspelmath, Martin. 2002. *Understanding Morphology.* London: Arnold.

Healey, Joseph F. 2011. *Race, Ethnicity, Gender, and Class: The Sociology of Group Conflict
    and Change* 5th ed., 2010/2011 update. Thousand Oaks, CA: Pine Forge Press.

Hebblethwaite, Cordelia. 2012. "Britishisms and the Britishisation of American English." *BBC*, 27 September 2012. Accessed 21 April 2020. https://www.bbc.com/news/magazine-19670686.

Howard, Martin, Raymond Mougeon and Jean-Marc Dewaele. 2013. "Variationist Sociolinguistics." *The Oxford Handbook of Sociolinguistics*. Edited by Bayley, Robert., Richard Cameron, and Ceil Lucas. Oxford: Oxford University Press. 340–359.

Janicka, Katarzyna, Malgorzata Kul and Jaroslaw Weckwerth. 2005. "Polish Students Attitudes to Native English Accents as Models for EFL Pronunciation." In *English Pronunciation Models: A Changing Scene*, edited by Katarzyna Dziubalska-Kołaczyk and Joanna Przedlacka. Bern: Peter Lang. 251–292.

Jenkins, Jennifer. 2002. "A Sociolinguistically Based, Empirically Researched Pronunciation Syllabus for English as an International Language." *Applied Linguistics* 23, no. 1: 83–103. Accessed 19 April 2020. Oxford Journals.

Jenkins, Jennifer, Alessia Cogo, and Martin Dewey. 2011. "Review of developments in research into English as a lingua franca." *Language Teaching* 44, no. 3: 281–315. Accessed 19 April 2020. ProQuest.

Jung, Carl G. and Herbert Read. 1953. *The Collected Works: 7, Two Essays on Analytical Psychology*. London: Routledge & Kegan Paul.

Kachru, Braj B. 1982. *The Other Tongue: English Across Cultures*. Urbana, Ill: University of Illinois Press.

Kircher, Ruth. 2015. "The Matched-Guise Technique." *Research Methods in Intercultural Communication*. Hoboken, NJ, USA: John Wiley & Sons, Inc. 196–211. Accessed 21.3.2020. EBSCOhost.

Kristiansen, Tore, Peter Garrett, and Nikolas Coupland. 2005. *Subjective Processes in Language Variation and Change*. Copenhagen: C. A. Reitzels forlag.

Labov, William. 1966. *The Social Stratification of English in New York City*. Washington D.C.: Center for Applied Linguistics.

Labov, William. 1983. "The Boundaries of Words and Their Meanings." In *Variation in the Form and Use of Language: a Sociolinguistics Reader*, edited by Ralph W. Fasold. Washington, D.C.: Georgetown University Press. 29–62.

Labov, William. 2001. *Principles of Linguistic Change. Volume 2, Social Factors.* Oxford: Blackwell.

Ladegaard, Hans J. 1998. "National stereotypes and language attitudes: The perception of British, American and Australian language and culture in Denmark." *Language and Communication* 18, no. 4: 251–274. Accessed 25 April 2020. Science Direct.

Ladegaard, Hans J. and Itesh Sachdev. 2006. "'I Like the Americans... But I Certainly Don't Aim for an American Accent': Language Attitudes, Vitality and Foreign Language Learning in Denmark." *Journal of Multilingual and Multicultural Development* 27, no. 2: 91–108. Accessed 25 April 2020. https://www.researchgate.net/publication/249024233_'I_Like_the_Americans_But_I_C ertainly_Don't_Aim_for_an_American_Accent'_Language_Attitudes_Vitality_and_For eign_Language_Learning_in_Denmark

Leppänen, Sirpa, Anne Pitkänen-Huhta, Tarja Nikula, Samu Kytölä, Timo Törmäkangas, Kari Nissinen, Leila Kääntä, Tiina Räisänen, Mikko Laitinen, Päivi Pahta, Heidi Koskela, Salla Lähdesmäki and Henna Jousmäki. 2009. *National Survey on the English Language in Finland: Uses, Meanings and Attitudes.* Helsinki: Varieng. Accessed 25 April 2020. http://www.helsinki.fi/varieng/series/volumes/05/evarieng-vol5.pdf.

Llamas, Carmen, Louise Mullany, and Peter Stockwell. 2007. "Introduction." In *The Routledge Companion to Sociolinguistics*, edited by Carmen Llamas, Louise Mullany, and Peter Stockwell. Abingdon: Routledge. xv–xix.

Mair, Christian. 2013. "The World System of Englishes. Accounting for the transnational importance of mobile and mediated vernaculars." *English World-Wide* 34, no. 3: 253–278. Accessed 21 February 2020. EBSCOHost.

Meecham, Marjory and Janie Rees-Miller. 2001. "Language in social contexts". In *Contemporary Linguistics: An Introduction*, edited by William O'Grady, John Archibald, Mark Aronoff and Janie Rees-Miller. 4th ed. Bedford: St. Martin's.

Mencken, Henry Louis. 1921. *The American language: An inquiry into the development of English in the United States.* 2nd edition. Accessed 18 April 2020. www.bartleby.com/185/.

Milroy, James. 2007. "Language Attitudes." In *The Routledge Companion to Sociolinguistics*, edited by Carmen Llamas, Louise Mullany, and Peter Stockwell. Abingdon: Routledge. 133–139.

Mollin, Sandra. 2006. *Euro-English: Assessing Variety Status*. Tübingen: Narr.

MOT Oxford Dictionary of English. 2005. Kielikone Oy. Accessed 17 April 2020. mot.kielikone.fi.

National Curriculum. 2014. *Perusopetuksen opetussuunnitelman perusteet.* [Bases for the curriculum of comprehensive education.] Accessed 26 April 2020. https://www.oph.fi/sites/default/files/documents/perusopetuksen_opetussuunnitelman_p erusteet_2014.pdf.

National Curriculum. 2019. *Lukio-opetuksen opetussuunnitelman perusteet.* [Bases for the curriculum of upper secondary education.] Accessed 26 April 2020. https://www.oph.fi/sites/default/files/documents/lukion_opetussuunnitelman_perusteet_ 2019.pdf

Nordquist, Richard. 2019a "Definition of English as a Second Language (ESL." *ThoughtCo*, 1 July 2019. Accessed 26 April 2020. https://www.thoughtco.com/english-as-a-second-language-esl-1690599

———. 2019b "Americanism in Language." *ThoughtCo*, 3 July 2019. Accessed 18 April 2020. https://www.thoughtco.com/what-is-americanism-words-1688984.

Osborne, Jason W. 2015. *Best Practices in Logistic Regression*. Los Angeles: SAGE. Accessed 25 April 2020. SAGE Research Methods.

Otwinowska, Agnieszka. 2015. *Cognate Vocabulary in Language Acquisition and Use: Attitudes, Awareness, Activation*. Blue Ridge Summit, PA: Multilingual Matters. Accessed 11 March 2020. De Gruyter.

Pearce, Michael. 2007. *The Routledge Dictionary of English Language Studies*. London: Routledge.

Pelanda, Brian Lee. 2011. "Declarations of Cultural Indeoendence: The Nationalistic Imperative behind the Passage of Early American Copyright Laws, 1783-1787." *Journal Of The Copyright Society Of The USA* 58, no. 2: 431–454. Accessed 20 April 2020. HeinOnline Law Journal Library.

Penhallurick, Rob. 2003. *Studying the English Language.* 2nd edition. Basingstoke: Palgrave Macmillan.

Quirk, R., Greenbaum, S., Leech, G. & Svartvik, J. 1972. *A Grammar of Contemporary English*. London: Longman.

Rickford, John R. and Penelope Eckert. 2001. *Style and Sociolinguistic Variation*. Cambridge: Cambridge University Press.

Rindal, Ulrikke. 2010. "Constructing Identity With L2: Pronunciation and Attitudes Among Norwegian Learners of English." *Journal of Sociolinguistics* 14, no. 2: 240–261. Accessed 25 April 2020. Wiley Online Library.

Rindal, Ulrikke, and Caroline Piercy. 2013. "Being 'Neutral'? English Pronunciation Among Norwegian learners.(Report)." *World Englishes* 32, no. 2: 211–229. Accessed 25 April 2020. Wiley Online Library.

Seidlhofer, Barbara. 2011. *Understanding English as a Lingua Franca*. Oxford: Oxford University press.

Sjöholm, Kaj. 2000. "Attitudes towards English and Two National Languages in Finland." In *Perspectives on Language and Communication in Multilingual Education*, edited by Kaj Sjöholm, and Anna-Lena Østern. Vasa: Åbo Akademi University. 121–143.

Statistics Finland. 2010. Accessed 21 April 2020. https://www.stat.fi/artikkelit/2010/art_2010-09-27_002.html?s=0

Statistics Finland. 2017. Accessed 31 March 2020. https://www.tilastokeskus.fi/tup/maahanmuutto/maahanmuuttajat-vaestossa/ulkomaan-kansalaiset.html.

Statistics Finland. 2018a. Accessed 31 March 2020. https://www.stat.fi/til/matk/2018/11/matk_2018_11_2019-01-10_tau_010_fi.html.

———. 2018b. Accessed 12 April 2020. https://www.stat.fi/tietotrendit/artikkelit/2018/alle-viidennes-opiskelijoista-opinnoissa-joissa-tasaisesti-naisia-ja-miehia-koulutusalojen-eriytyminen-jatkuu/.

Steinskog, Dag, Dag Tjøstheim, and Nils Kvamstø. 2007. "A Cautionary Note on the Use of the Kolmogorov-Smirnov Test for Normality." *Monthly Weather Review* 135, no. 3: 1151–1157. Accessed 12 March 2020. *https://doi.org/10.1175/MWR3326.1*

Tagliamonte, Sali. 2012. *Variationist Sociolinguistics: Change, Observation, Interpretation.* Malden, Mass: Wiley-Blackwell.

Tergujeff, Elina. 2012. "The English Pronunciation Teaching in Europe Survey: Finland." *Apples – Journal of Applied Language Studies* 6, no. 1: 29–45.

Thode, Henry J. 2002. *Testing for normality*. New York: Marcel Dekker.

Watt, Dominic. 2007. "Variation and the Variable." In *The Routledge Companion to Sociolinguistics*, edited by Carmen Llamas, Louise Mullany, and Peter Stockwell. Abingdon: Routledge.

Wardhaugh, Ronald. 2010. *An Introduction to Sociolinguistics*. 6. ed. Chichester ; Malden (Mass.): Wiley-Blackwell.

Webster, Noah. 1789. *Dissertations on the English Language.* Boston, MA. Accessed 20 April 2020. Gale Eighteenth Century Collections Online.

Wittgenstein, Ludwig. 1968. *Philosophische Untersuchungen: Philosophical Investigations*. 3rd edition. Oxford: Blackwell.

Ylönen, Hannele. 2007. "Kansainvälisyyskasvatuksen laajuus ja syvyys suomalaisessa englannin kielen opetuksessa." *3.2.4 Kehittämishanke*. Jyväskylän Ammattikorkeakoulu. Accessed 13 September 2019. https://www.theseus.fi/bitstream/handle/10024/19624/TMP.objres.494.pdf?sequence=1&isAllowed=y

Zahn, Christopher J, and Robert Hopper. 1985. "Measuring Language Attitudes: The Speech Evaluation Instrument." *Journal of Language and Social Psychology* 4, no. 2: 113–123. Accessed 25 April 2020. SAGE Journals.

Appendix 1 Test / Questionnaire Instructions

[The test was administered in Finnish. The English translations in bold are mine.]

[Blank lines indicate change of PowerPoint slides.]

Tällä testillä tutkitaan tiettyjä edistyneiden englannin opiskelijoiden kielenkäytön piirteitä.
This test is for investigating certain aspects of the language use of advanced English students.

- Ruudulle tulee 10 lyhyttä tekstiä yksi kerrallaan.
  There will 10 short texts on the screen, one at a time.
- Sinun tehtäväsi on kääntää tekstit suomesta englantiin.
  Your task is to translate the texts from Finnish into English.
- Kirjoita käännöksesi lomakkeessa tekstiä vastaavan numeron kohdalle.
  Write down your translation next to the number corresponding to the text.

- Jokaista tekstiä varten on aikaa 1 minuutti.
  There is 1 minute for each text.
- Ilmoitan kun ajasta on kulunut puolet ja uudestaan kun aikaa on jäljellä 10 sekuntia.
  I will notify you when half of the time has passed, and again when there is 10 seconds remaining.
- Kirjoita alas ensimmäisenä mieleen tuleva käännös.
  Write down the first translation that occurs to you.
- Muista, että testi on anonyymi, ja mahdollisista virheistä on siis turha stressata.
  Remember that the test is completely anonymous, and there is no reason to stress over possible mistakes.

1. Se oli ainoa ravintola pääkadulla, ja jono ylsi jalkakäytävälle asti.
    It was the only restaurant on the (main [AmE] / high {BrE}) street and the (line [AmE] / queue [BrE]) reached the (sidewalk [AmE] / pavement [BrE]).

2. Ravintolan johtaja oli kadonnut viikonloppuna.
    The restaurant manager had gone missing over the weekend.

3. Komisario Hunt kiersi ihmismassaa vastapäivään tiskille asti.
    Officer Hunt circled the mass of people (counterclockwise [AmE] / anticlockwise [BrE]) up to the counter.

4. Päätarjoilija näytti sairaalta ja käytti liian suuria housuja.
    The main waiter looked (sick [AmE] / ill [BrE]) and was wearing (pants [AmE] / trousers [BrE]) that were too big for him.

5. Hän osoitti Huntille sotkuisen taukotilan haastatteluja varten.
    He appointed a messy breakroom to Hunt for interviews.

6. Henkilökunta oli hermostunutta. Oliko heillä jotakin salattavaa?
    The staff was nervous. Did they have something to hide?

7. Yksi sanoi johtajan lähteneen syyslomalle, toinen arveli tämän olevan kotona.
One said the manager had gone on a (fall [AmE] / autumn [BrE]) (vacation [AmE] / holiday [BrE]), another thought him to be at home.

8. Johtajan asunto oli samassa rakennuksessa kaksi kerrosta ylempänä.
The manager's (apartment [AmE] / flat [BrE]) was in the same building, two storeys up.

9. Hunt oli jo odottamassa hissiä, kun lihamyllystä kuului räjähdys.
Hunt was already waiting for the (elevator [AmE] / lift [BrE]) when an explosion was heard from the meatgrinder.

10. Kukaan henkilökunnasta ei ollut tiennyt, että johtajalla oli sydämentahdistin.
Nobody in the staff had known that the manager had a pacemaker.

Some background questions:
11. Age
12. Gender
13. I would like to sound American
14. I would like to sound British

American English sounds
15. Pleasant
16. Sophisticated
17. Cool
British English sounds
18. Pleasant
19. Sophisticated
20. Cool

21. Have you spent more than a month in an English-speaking country?
1) No
2) Yes – Which country and for how long? (estimate)
If there are more than one such period, please mark them all.
22. In your everyday life, do you regularly talk with native British or American speaker(s)?
(At least once a week)
A) No
B) Yes, American
C) Yes, British
D) Yes, both

Write your e-mail at the end of the form, if you want to participate in a more detailed at a later date.
Thanks a lot for your answers!

Appendix 2 Test / Questionnaire Form

1. _____

   _____

2. _____

   _____

3. _____

   _____

4. _____

   _____

5. _____

   _____

6. _____

   _____

7. _____

   _____

8. _____

   _____

9. _____

   _____

10. _____

   _____

11. Age _____

12. Gender
- ☐ Female
- ☐ Male
- ☐ Other

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree
13.					
14.					
15.					
16.					
17.					
18.					
19.					
20.					

21.
- ☐ No
- ☐ Yes _____

_____

_____

22.
- ☐ A
- ☐ B
- ☐ C
- ☐ D

E-mail (voluntary) _____

My e-mail can be used to

☐ send me my test results

☐ contact me for further study

Appendix 3 Questionnaire Results

#	Age	Gender	AmE target	BrE target	Con-trasted target	AmE plea-sant	BrE plea-sant	Con-trasted plea-sant	AmE sophis-ticated	BrE sophis-ticated	Con-trasted sophis-ticated	AmE cool-ness	BrE cool-ness	Con-trasted cool-ness	Regu-lar contact
1	19	1	-4	3	-1	-4	4	0	-2	4	2	-3	3	0	4
2	19	1	-4	3	-1	-4	4	0	-1	5	4	-4	2	-2	4
3	19	2	-2	4	2	-4	4	0	-2	4	2	-4	2	-2	2
4	20	1	-3	3	0	-4	2	-2	-1	4	3	-4	3	-1	4
5	19	1	-4	3	-1	-4	4	0	-3	5	2	-4	4	0	3
6	21	2	-4	2	-2	-4	4	0	-3	3	0	-4	4	0	4
7	22	1	-4	2	-2	-4	4	0	-2	4	2	-4	2	-2	4
9	22	2	-3	4	1	-3	4	1	-2	4	2	-3	4	1	4
10	28	2	-4	2	-2	-3	2	-1	-2	4	2	-4	2	-2	4
11	23	2	-3	2	-1	-3	1	-2	-2	3	1	-4	2	-2	4
12	30	1	-3	4	1	-4	4	0	-3	4	1	-3	3	0	4
13	20	1	-3	4	1	-4	5	1	-2	5	3	-4	3	-1	4
15	19	1	-2	5	3	-2	5	3	-4	5	1	-2	5	3	4
16	20	1	-4	2	-2	-4	2	-2	-3	2	-1	-4	3	-1	1
17	21	2	-5	4	-1	-4	5	1	-4	5	1	-4	4	0	1
18	21	2	-2	4	2	-2	2	0	-3	4	1	-4	3	-1	4
19	23	1	-3	4	1	-4	4	0	-3	3	0	-3	3	0	1
20	27	1	-2	4	2	-4	5	1	-3	5	2	-4	4	0	4
21	26	1	-2	4	2	-2	4	2	-2	4	2	-2	4	2	4
22	22	1	-5	2	-3	-5	5	0	-3	5	2	-5	5	0	4
23	21	2	-1	5	4	-3	4	1	-2	4	2	-2	3	1	4
24	21	1	-2	4	2	-4	4	0	-1	5	4	-4	4	0	4
25	20	1	-5	3	-2	-4	4	0	-2	4	2	-4	3	-1	4
27	34	1	-4	4	0	-5	3	-2	-2	4	2	-3	3	0	1
29	19	1	-2	5	3	-3	5	2	-1	4	3	-3	4	1	4
30	20	1	-1	5	4	-2	5	3	-2	5	3	-3	3	0	4
31	20	1	-2	4	2	-4	4	0	-1	4	3	-4	3	-1	4
33	21	1	-5	1	-4	-5	4	-1	-3	5	2	-5	1	-4	4
34	21	1	-5	2	-3	-5	4	-1	-3	4	1	-5	3	-2	4
35	19	1	-2	4	2	-5	5	0	-3	4	1	-3	3	0	4
36	21	1	-4	3	-1	-4	4	0	-3	4	1	-4	4	0	4
38	21	1	-4	4	0	-3	4	1	-1	5	4	-5	3	-2	1
40	21	1	-4	2	-2	-5	5	0	-2	4	2	-5	5	0	4
42	20	1	-2	4	2	-4	4	0	-2	4	2	-3	3	0	4
43	20	2	-2	5	3	-3	4	1	-2	4	2	-4	4	0	4
45	20	1	-1	4	3	-2	4	2	-1	4	3	-4	3	-1	4
46	22	3	-3	3	0	-3	4	1	-1	4	3	-3	4	1	4
47	20	1	-3	3	0	-4	4	0	-1	4	3	-4	1	-3	1
48	26	1	-2	4	2	-2	4	2	-2	4	2	-3	4	1	4
49	24	2	-4	3	-1	-5	2	-3	-3	4	1	-3	3	0	1
Mean	21,80		-3,10	3,43	0,33	-3,68	3,88	0,20	-2,20	4,15	1,95	-3,68	3,23	-0,45	
Median	21		-3	4	0	-4	4	0	-2	4	2	-4	3	0	
Mode	20		-4	4	2	-4	4	0	-2	4	2	-4	3	0	
SD	3,28		1,19	1,03	2,08	0,94	0,99	1,32	0,85	0,66	1,08	0,80	0,95	1,34	

Notes: Categorical variables *gender* and *regular contact* are assigned numerical values in the order they appear in the questionnaire (Appendix 2).

Likert scale evaluations are coded from 1 (strongly disagree) to 5 (strongly agree)

SD: Standard deviation

Appendix 4 Corpus Search Sequences and Values of Distinctiveness

All searches conducted with the corpus collection and search tools of Brigham Young University (https://corpus.byu.edu/).

COCA: Corpus of Contemporary American English
BNC: British National Corpus

main street vs. high street
COCA "along the high street": 3; " along the main street": 51 (94.4% 'main')
BNC "along the high street": 17; " along the main street": 6 (73.9% 'high')
Value of distinctiveness = 0.683

line vs. queue
COCA "waiting in line": 481; "queueing": 25 (95.0% 'line')
BNC "waiting in line": 6; "queueing": 90 (93.8% 'queue')
Value of distinctiveness = 0.888

sidewalk vs. pavement
COCA "sidewalk cafe/caf": 67; "pavement cafe/caf": 0 (100.0% 'sidewalk')
BNC "sidewalk cafe/caf": 3; "pavement cafe/caf": 23 (88.4% 'pavement')
Value of distinctiveness = 0.884

counterclockwise vs. anticlockwise
COCA " anticlockwise": 13; "counterclockwise": 472 (97.3% 'counterclockwise')
BNC "anticlockwise": 48; "counterclockwise": 2 (96.0% 'anticlockwise')
Value of distinctiveness = 0.933

sick vs. ill
COCA "ill-looking": 3; "sick-looking": 12 (80.0% 'sick')
BNC "ill-looking" : 8; " sick-looking": 1 (88.9% 'ill')
Value of distinctiveness = 0.689

pants vs. trousers
COCA "his trousers": 643; "his pants": 2637 (80.4% 'pants')
BNC "his trousers": 324; "his pants": 40 (89.0% 'trousers')
Value of distinctiveness = 0.694

fall vs. autumn
COCA "in (the) fall": 8047; "in (the) autumn": 1017 (88.8% 'fall')
BNC "in (the) fall": 56; "in (the) autumn": 1130 (95.3% 'autumn')
Value of distinctiveness = 0.841

vacation vs. holiday
COCA " summer vacation": 674; "summer holiday": 59 (92.0% 'vacation')
BNC "summer vacation": 34; "summer holiday": 103 (75.2% 'holiday')
Value of distinctiveness = 0.672

apartment vs. flat
COCA "my apartment": 1831; "my flat": 115 (94.1% 'apartment')
BNC "my apartment": 34; "my flat": 164 (82.8% 'flat')
Value of distinctiveness = 0.769

elevator vs. lift
COCA "the lift": 1132; "the elevator": 4891 (81.2% 'elevator')
BNC "the lift": 737; "the elevator": 104 (87.6% 'lift')
Value of distinctiveness = 0.688

Appendix 5 Lexical Choices Weighted for Distinctiveness

#	main street	high street	line	queue	sidewalk	pavement	counter-clockwise	anti-clockwise	sick	ill
1	-.683	0	0	.888	-.884	0	0	0	-.689	0
2	-.683	0	-.888	0	-.884	0	0	0	0	.689
3	-.683	0	0	.888	-.884	0	-.933	0	-.689	0
4	-.683	0	-.888	0	-.884	0	0	0	-.689	0
5	-.683	0	0	.888	0	.884	-.933	0	0	.689
6	-.683	0	-.888	0	-.884	0	0	0	0	.689
7	-.683	0	-.888	0	0	.884	-.933	0	0	.689
9	-.683	0	-.888	0	-.884	0	-.933	0	0	.689
10	-.683	0	-.888	0	-.884	0	-.933	0	-.689	0
11	-.683	0	-.888	0	-.884	0	-.933	0	0	.689
12	-.683	0	0	.888	0	0	0	0	-.689	0
13	-.683	0	-.888	0	-.884	0	-.933	0	0	.689
15	-.683	0	0	.888	0	0	0	0	0	.689
16	-.683	0	-.888	0	-.884	0	-.933	0	-.689	0
17	-.683	0	-.888	0	0	0	0	0	0	.689
18	-.683	0	0	.888	0	.884	-.933	0	-.689	0
19	-.683	0	0	.888	0	0	0	0	-.689	0
20	-.683	0	-.888	0	0	.884	-.933	0	0	.689
21	-.683	0	0	.888	0	.884	0	0	-.689	0
22	-.683	0	-.888	0	0	0	0	0	0	.689
23	-.683	0	-.888	0	0	0	-.933	0	-.689	0
24	-.683	0	0	.888	0	0	0	0	-.689	0
25	-.683	0	-.888	0	-.884	0	0	0	-.689	0
27	-.683	0	0	.888	0	.884	0	0	0	0
29	-.683	0	0	.888	0	.884	-.933	0	0	.689
30	-.683	0	-.888	0	0	.884	-.933	0	-.689	0
31	-.683	0	-.888	0	-.884	0	-.933	0	-.689	0
33	-.683	0	0	.888	-.884	0	-.933	0	-.689	0
34	-.683	0	-.888	0	-.884	0	-.933	0	-.689	0
35	-.683	0	-.888	0	-.884	0	-.933	0	0	.689
36	-.683	0	0	.888	-.884	0	0	0	-.689	0
38	-.683	0	0	.888	0	.884	0	.933		.689
40	-.683	0	0	.888	0	.884	-.933	0	0	.689
42	-.683	0	0	.888	-.884	0	0	.933	0	.689
43	-.683	0	-.888	0	-.884	0	-.933	0	0	.689
45	-.683	0	-.888	0	0	.884	0	0	0	.689
46	-.683	0	0	.888	0	.884	-.933	0	-.689	0
47	-.683	0	0	.888	-.884	0	-.933	0	0	.689
48	-.683	0	0	.888	-.884	0	0	0	0	.689
49	-.683	0	0	.888	-.884	0	-.933	0	0	.689

Notes: Values in cells are values of distinctiveness (cf. Appendix 4)

Red cells indicate omitted word pairs.

#	pants	trousers	fall	autumn	vaca-tion	holiday	apart-ment	flat	elevator	lift	AmE%	BrE%
1	0	.694	0	0	0	0	-.769	0	-.688	0	70.1 %	29.9 %
2	-.694	0	0	.672	-.841	0	-.769	0	-.688	0	80.0 %	20.0 %
3	-.694	0	0	.672	0	0	0	.769	-.688	0	69.9 %	30.1 %
4	-.694	0	-.672	0	0	0	-.769	0	-.688	0	100.0 %	0.0 %
5	-.694	0	0	.672	-.841	0	-.769	0	-.688	0	34.1 %	65.9 %
6	-.694	0	-.672	0	-.841	0	-.769	0	-.688	0	89.9 %	10.1 %
7	-.694	0	-.672	0	-.841	0		0	-.688	0	79.7 %	20.3 %
9	-.694	0	0	.672	0	.841	-.769	0	-.688	0	71.6 %	28.4 %
10	-.694	0	0	0	-.841	0	-.769	0	-.688	0	100.0 %	0.0 %
11	-.694	0	0	0	-.841	0	-.769	0	0	0	89.2 %	10.8 %
12	-.694	0	0	.672	0	.841	-.769	0	-.688	0	60.4 %	39.6 %
13	-.694	0	0	0	0	0	0	.769	-.688	0	76.6 %	23.4 %
15	0	.694	-.672	0	0	.841	-.769	0	-.688	0	48.7 %	51.3 %
16	-.694	0	0	.672	-.841	0	-.769	0	-.688	0	91.3 %	8.7 %
17	0	.694	0	.672	0	.841	-.769	0	-.688	0	52.3 %	47.7 %
18	-.694	0	0	.672	0	.841	-.769	0	0	.688	48.7 %	51.3 %
19	0	.694	0	0	-.841	0	-.769	0	-.688	0	70.7 %	29.3 %
20	-.694	0	0	.672	0	0	-.769	0	0	0	63.9 %	36.1 %
21	-.694	0	0	0	0	.841	-.769	0	-.688	0	57.4 %	42.6 %
22	-.694	0	0	0	0	.841	-.769	0	-.688	0	71.7 %	28.3 %
23	-.694	0	0	.672	-.841	0	-.769	0	-.688	0	90.4 %	9.6 %
24	0	.694	0	.672	0	0	-.769	0	0	.688	43.7 %	56.3 %
25	-.694	0	0	.672	-.841	0	-.769	0	-.688	0	90.1 %	9.9 %
27	-.694	0	0	0	0	.841	-.769	0	-.688	0	52.0 %	48.0 %
29	0	.694	0	.672	0	0	0	.769	0	.688	23.4 %	76.6 %
30	0	.694	0	0	-.841	0	-.769	0	-.688	0	77.7 %	22.3 %
31	-.694	0	0	0	-.841	0	-.769	0	-.688	0	100.0 %	0.0 %
33	-.694	0	-.672	0	-.841	0	-.769	0	-.688	0	88.5 %	11.5 %
34	-.694	0	-.672	0	0	.841	-.769	0	-.688	0	89.1 %	10.9 %
35	-.694	0	0	0	0	.841	0	.769	-.688	0	67.5 %	32.5 %
36	-.694	0	-.672	0	0	.841	-.769	0	0	.688	64.5 %	35.5 %
38	-.694	0	-.672	0	0	.841	-.769	0	-.688	0	45.3 %	54.7 %
40	-.694	0	0	.672	0	.841	-.769	0	-.688	0	48.7 %	51.3 %
42	-.694	0	-.672	0	0	0	-.769	0	-.688	0	63.6 %	36.4 %
43	-.694	0	0	.672	0	0	-.769	0	0	.688	70.3 %	29.7 %
45	-.694	0	0	.672	0	0	0	.769	-.688	0	49.5 %	50.5 %
46	-.694	0	0	0	0	.841	-.769	0	-.688	0	63.0 %	37.0 %
47	0	.694	0	.672	0	0	-.769	0	-.688	0	57.3 %	42.7 %
48	0	.694	0	0	-.841	0	-.769	0	-.688	0	63.0 %	37.0 %
49	-.694	0	0	0	-.841	0	-.769	0	-.688	0	77.7 %	22.3 %

Abbreviations: AmE%: American English lexical choices weighted for distinctiveness
BrE%: British English lexical choices weighted for distinctiveness
SD: Standard deviation

Mean	68.7	31.3
Median	70.0	30.0
SD	18.6	18.6

Appendix 6 Interviewed Participants

Participant #	% of Lexical Choices AmE	Target Preference*	Evaluation of Sophistication*	Evaluation of Coolness*	Evaluation of Pleasantness*
15	47.5	3	1	3	3
25	89.9	-2	2	-1	0
34	88.9	-3	1	-2	-1
45	50.7	3	3	-1	2
49	77.7	-1	1	0	-3
Range among the whole sample	23.9 – 100.0	(-4) – 4	(-1) – 4	(-4) – 3	(-3) – 3

* The variable is based on a 9-point scale.

(−4) represents the strongest possible evaluation of American English over British English.

4 represents the strongest possible evaluation of British English over American English.

Appendix 7 Interview Framework

[Interviews were conducted in Finnish. The English translations in bold are mine.]

1. Mitä sanoja, jotka eroavat amerikanenglannissa ja brittienglannissa, tulee ensimmäisenä mieleen?
   Which words that differ between AmE and BrE come to mind first?

2. Tiedätkö, mitkä seuraavista sanoista ovat amerikanenglantia ja mitkä brittienglantia?
   Do you know, which of the following words are AmE and which ones are BrE?

Main street	High street
Ill	Sick
Pants	Trousers
Lift	Elevator
Flat	Apartment
Vacation	Holiday
Autumn	Fall
Queue	Line
Anticlockwise	Counterclockwise
Sidewalk	Pavement

3. Alleviivaa sanat, joita itse käytät. Voit alleviivata molemmat vaihtoehdot.
   Underline the words that you use yourself. You can underline both alternatives.

4. Kuinka hyvin seuraavat adjektiivit vastaavat mielikuvaasi amerikanenglannista?
   How well do the following adjectives represent your view of AmE?

American English sounds	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
Competent					
Likeable					

5. Kuinka hyvin seuraavat adjektiivit vastaavat mielikuvaasi brittienglannista?
   How well do the following adjectives represent your view of BrE?

British English sounds	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
Competent					
Likeable					

6. Kuvaile brittienglantia omin sanoin.
   Describe BrE in your own words

7. Kuvaile amerikanenglantia omin sanoin.
   Describe BrE in your own words

8. Mitä amerikanenglannin muotoa arvioit edellisissä kysymyksissä?
   Which variety of AmE did you evaluate in the previous questions?

   - Apukysymyksiä, mikäli arvioidun variantin määrittäminen on vaikeaa:
   - **Auxiliary questions, in case it is difficult to determine the variety evaluated:**
   - Minkälainen ihminen sitä puhuu?
   - **What kind of a person speaks it?**
   - Esimerkki julkisuuden henkilöstä, joka puhuu sitä?
   - **An example of a famous person who speaks it?**
   - Missä sitä puhutaan?
   - **Where is it spoken?**

9. Mitä brittienglannin muotoa arvioit edellisissä kysymyksissä?
   Which variety of BrE did you evaluate in the previous questions?

   Apukysymyksiä, mikäli arvioidun variantin määrittäminen on vaikeaa:
   Auxiliary questions, in case it is difficult to determine the variety evaluated:
   - Minkälainen ihminen sitä puhuu?
   - **What kind of a person speaks it?**
   - Esimerkki julkisuuden henkilöstä, joka puhuu sitä?
   - **An example of a famous person who speaks it?**
   - Missä sitä puhutaan?
   - **Where is it spoken?**

10. Herättääkö vahvasti amerikanenglantia puhuva suomalainen erilaisen reaktion kuin natiivi?
   Does a non-native speaker with a clearly American variety of English prompt a different reaction than the natives?
   - Millä tavalla?
   - **In what way?**
   - Miksi?
   - **Why?**

11. Herättääkö vahvasti brittienglantia puhuva ei-natiivi erilaisen reaktion kuin natiivi?
   Does a non-native speaker with a clearly British variety of English prompt a different reaction than the natives?
   - Millä tavalla?
   - **In what way?**
   - Miksi?
   - **Why?**

12. Mitkä kielen osa-alueet vaikuttavat arvioosi eri kielimuodoista?
   Which subsystems of language affect your evaluations of different varieties?
   - Kielioppi
   - **Grammar**
   - Ääntäminen
   - **Pronunciation**
   - Sanasto
   - **Vocabulary**

13. Minkälaisissa tilanteissa käytät englantia?
   In what kind of situations you use English?

14. Pyritkö omassa englannissasi johonkin natiivikielimuotoon? Mihin? Miksi?
   Do you strive for some native variety in your own English? Which one? Why?

15. Jos kyllä, mihin kielen osa-alueisiin kiinnität huomiota?
   If yes, to which linguistic subsystems you pay attention?
   - Kielioppi
   - **Grammar**
   - Ääntäminen
   - **Pronunciation**
   - Sanasto
   - **Vocabulary**

16. Onko tärkeää kuulostaa natiivilta? Mitä etua siitä on?
   Is it important to sound like a native? What is the advantage?

17. Sopiiko amerikanenglanti tai brittienglanti paremmin tiettyihin tilanteisiin?
    Is AmE or BrE better suited for certain situations?
    esim. yliopisto, työhaastattelu, kaveripiiri, harrastukset
    e.g. university, job interview, circle of friends, hobbies

18. Huomaatko muuttavasi puhetapaasi eri tilanteissa?
    Do you notice changing the way you speak in different situations?
    Miten? Painottuvatko AmE tai BrE piirteet joissakin tilanteissa?
    How? Are features of AmE or BrE emphasised in some situations?

19. Huomaatko muiden ei-natiivien muuttavan puhetapaansa eri tilanteissa?
    Do you notice other non-natives changing the way they speak in different situations?
    Miten? Painottuvatko AmE tai BrE piirteet joissakin tilanteissa?
    How? Are features of AmE or BrE emphasised in some situations?

20. Onko tärkeää pitäytyä yhdessä kielimuodossa? Miksi? Miksi ei?
    Is it important to stick to one variety? Why? Why not?

21. Mitä englannin muotoa kuulet eniten yliopistolla?
    Which variety of English do you hear the most at the university?

22. Oletko mielestäsi oppinut englantia enemmän koulussa vai vapaa-ajalla?
    Do you think you have learned more English at school or in your spare time?

23. Mitä englannin muotoa seuraamissasi medioissa enimmäkseen kuulee?
    Which variety of English do you hear mostly in the media you follow?

24. Oliko sinulle koulussa tärkeää saada englannista hyviä arvosanoja?
    Was it important for you to get good grades in English at school?

25. Puhuiko viimeisin englanninopettajasi AmE, BrE vai jotakin muuta?
    Did your last teacher in school speak AmE, BrE or something else?
    - Entä aiemmat opettajat?
    - **How about earlier teachers?**

26. Piditkö englannin tunneista ja opettajista?
    Did you like the English classes and teachers?

Appendix 8 Participants Excluded on the Basis of Residence in an English-speaking

Residence (> 1 month)	% of American Lexical Choices (Mean)
No residence in an English-speaking country (N=40)	68.7
Residence in the USA (N=5)	80.9
Residence in the UK (N=1)	23.4
Residence in Australia (N=1)	24.2
Residence in New Zealand (N=1)	28.6

Appendix 9 Finnish Summary

Tämän tutkimuksen pyrkimyksenä oli selvittää, korreloiko suomalaisten yliopistotason englannin opiskelijoiden (= 40) pyrkimys kuulostaa amerikkalaiselta tai britiltä (jatkossa *kohdepreferenssi*) kyseiselle varieteetille tyypillisten sanojen käytön kanssa (jatkossa *leksikaaliset valinnat*). Edelleen tutkin, korreloivatko kohdepreferenssi sekä amerikan- ja brittienglannin *statuksesta ja kompetenssista* sekä *sosiaalisesta vetovoimasta* tehdyt arviot (jatkossa *kieliasenteet*).

Tulevina englannin kielen ammattilaisina osallistujien kielelliset tendenssit voivat valaista englannin kielen tulevaisuutta Suomessa. Onko suunta kohti amerikanenglantia (AmE) vai brittienglantia (BrE)? Lisäksi mahdolliset korrelaatiot mitattujen muuttujien välillä kertoisivat opiskelijoiden kyvystä tehdä sosiaalisesti motivoituja sanavalintoja ja käyttää englannin natiivivarieteetteja identiteettinsä muodostamisen välineenä. Mikäli tällainen havainto tehtäisiin, asettaisi se paineita opetuksessa käytettävien mallien omavalintaisuudelle, sillä tutkimuksen osallistujien voidaan katsoa toimivan esimerkkinä edistyneistä englannin oppijoista Suomessa. Edelleen laajakantoisemmin osallistujien voidaan katsoa edustavan englannin vieraskielisistä käyttäjiä, joiden asema englanninkielisessä maailmassa on edelleen kiivaan keskustelun aiheena. Mikäli vieraskielisillä käyttäjillä havaitaan kyky käyttää englannin varieteetteja sosiaalisen identiteettinsä ilmaisemiseen, tämä tukisi heidän asemaansa tasavertaisina englannin käyttäjinä.

Aineisto kerättiin eräässä suomalaisessa yliopistossa ennen suositun perusopintojen kurssin alkua, jolloin saatiin melko kattava otos opintojensa alkupuolella olevista opiskelijoista. Ajatuksena oli, että osallistujat eivät ole vielä akateemisesti kovin pitkälle koulutettuja, ja näin tulokset edustaisivat paremmin edistyneitä englannin käyttäjiä yleensä. Toki osallistujien ei silti voida katsoa olevan kattava otos edistyneistä englannin käyttäjistä. Osallistujat, jotka olivat viettäneet ainakin yhden yli kuukauden jakson englanninkielisessä maassa, jätettiin lopullisen analyysin ulkopuolelle, sillä tutkimuksen keskiössä oli englanti vieraana kielenä (EFL), ja pitkän natiivikontaktin todettiin vaikuttavan tuloksiin.

Aluksi osallistujat tekivät käännöstehtävän, johon oli sisällytetty kaikkiaan kymmenen sanaa, joiden kääntäminen edellytti valintaa AmE:n ja BrE:n varianttien välillä. Sanavalintojen tietoisen monitoroinnin välttämiseksi osallistujille ei kerrottu testin tarkoitusta. Käännöstehtävän jälkeen osallistujat arvioivat 5-portaisella Likert-asteikolla haluaan kuulostaa amerikkalaiselta tai brittiläiseltä sekä termien *sophisticated* (sofistikoitunut), *cool* (cool) ja *pleasant* (miellyttävä) sopivuutta kuvaamaan amerikanenglantia ja brittienglantia. Termi *sofistikoitunut* edusti statusta ja kompetenssia, *cool* ja *miellyttävä* puolestaan sosiaalista vetovoimaa

Aiemman tutkimuksen valossa oli syytä odottaa sanaston olevan pääosin amerikkalaista. Gonçalves et al. (2018) havaitsivat Twitter-viestejä tutkimalla englannin amerikkalaistumisen olevan

maailmanlaajuinen ilmiö, joka on erityisen voimakasta maissa, joissa englannilla ei ole virallista asemaa. Suomessa käytetty sanasto oli näidenkin maiden joukossa yksi amerikkalaistuneimmista (2018, 11). Myös Forsmanin (2000) aiempi tutkimus lukio-opiskelijoiden parissa tukee amerikanenglantilaisen sanaston suosiota Suomessa.

Jotta käännöstesti antaisi mahdollisimman luotettavan kuvan osallistujien sanaston amerikkalaisuudesta/brittiläisyydestä, jokaisen tutkittavan sanaparin merkitsevyys painotettiin sen mukaan, kuinka eksklusiivisia sanat olivat varieteeteilleen. Tämä perustui sanojen esiintymistaajuuteen suhteessa vastineisiinsa sekä amerikan- (COCA; Davis 2008) että brittienglannin (BNC; 2004) korpuksissa. Lopputulos oli, että osallistujien **sanasto oli keskimäärin 68,7% amerikkalaista**.

Muualla Euroopassa tehdyissä tutkimuksissa yliopistotason englanninopiskelijoiden BrE oli selvästi suositumpi englannin mallina (Ladegaard Sachdev 2006, 101; Waniek-Klimczak and Klimczak 2005, 242). Tässä tutkimuksessa 19 osallistujaa halusi mieluummin kuulostaa britiltä ja 16 amerikkalaiselta. Likert-asteikolla varieteettien mediaanien erotus oli 1 piste BrE:n hyväksi. **BrE oli siis hieman suositumpi malli, mutta ero ei ollut tilastollisesti merkittävä** (Van der Waerden testi, $p = 0,329$). Mainitut aiemmat tutkimukset ovat jo kohtalaisen vanhoja, ja on epäselvää, missä määrin AmE:n suurempi suosio niihin verrattuna johtuu AmE:n suosion noususta ja missä määrin maakohtaisista eroista.

Mitä tulee kieliasenteisiin, aiemman tutkimuksen nojalla oli syytä hypotetisoida, että brittienglanti yhdistettäisiin selvästi enemmän statukseen ja kompetenssiin ja amerikanenglanti vastaavasti hieman enemmän sosiaaliseen vetovoimaan. Tämä hypoteesi vahvistui osittain, sillä BrE arvioitiin 2 pistettä (mediaanien erotus) korkeammalle *sofistikoituneisuuden* asteikolla. Toisin sanoen **BrE:n arvioitiin omaavan selvästi enemmän statusta ja kompetenssia.** Toisaalta **AmE:n ja BrE:n sosiaalinen vetovoima arvioitiin lähes yhtäläisesti** sekä *cooliuden* että *miellyttävyyden* suhteen, sillä arvioiden mediaanierotus oli 0.

Sanavalintojen ja kohdepreferenssin korreloimiseksi arviot kohdepreferenssistä rinnastettiin laskemalla erotus arvioille AmE:sta ja BrE:sta henkilökohtaisina englannin malleina. AmE:n arvot vähennettiin BrE:n arvoista, jolloin saatiin 9-portainen asteikko suurimmasta mahdollisesta AmE-preferenssistä (–4) suurimpaan mahdolliseen BrE-preferenssiin (4). Toisen muuttujan ollessa asteikollinen käytettiin metodina ei-parametrista Spearmanin järjestyskorrelaatiota. Tuloksena oli melko heikko (0,347), mutta tilastollisesti merkittävä ($p = 0,28$) korrelaatio, joten **kohdepreferenssi ja leksikaaliset valinnat ovat suhteessa toisiinsa suomalaisilla yliopistotason englannin opiskelijoilla**.

Lopuksi tutkin vielä kohdepreferenssin ja kieliasenteiden suhdetta toisiinsa. Metodina käytin ordinaalista regressioanalyysia, jossa selitettävänä muuttujana oli kohdepreferenssi ja selittävinä muuttujina sofistikoituneisuus, coolius ja miellyttävyys. Kaikki muuttujat kontrastoitiin AmE:n ja BrE:n suhteen kuten edellä, jolloin saatiin neljä 9-portaista asteikkoa. Coolius osoittautui parhaiten kohdepreferenssiä ennakoivaksi muuttujaksi. Sen regressiokerroin oli 0,678 ($p = 0,017$), eli yhden pisteen muutos arviossa varieteetin cooliudesta ennakoi 0,678 pisteen samansuuntaista muutosta arviossa kohdepreferenssistä. Miellyttävyyden regressiokerroin oli 0,594 ($p = 0,043$) ja sofistikoituneisuuden 0,417, mutta tämä jälkimmäinen ei ollut tilastollisesti merkittävä tulos ($p = 0,125$). Tulosten perusteella siis **sosiaalinen vetovoima ennakoi kohdepreferenssiä voimakkaammin ja merkittävämmin kuin status ja kompetenssi**.

Intuitiivisin johtopäätös tuloksista olisi tiivistetysti, että opiskelijat valitsevat englannin mallin pääosin sen sosiaalisen vetovoiman perusteella, mutta ovat jakautuneita sen suhteen kumpaan varieteettiin sen yhdistävät. Mallin valinta puolestaan vaikuttaisi jossakin määrin ohjaavan sanavalintoja, mutta AmE:n vaikutus näkyy kuitenkin voimakkaasti myös BrE:hen tähtäävillä opiskelijoilla. Korrelaatioiden syy-seuraussuhteet ovat kuitenkin epävarmoja. On esimerkiksi mahdollista, että englannin malli valitaan pragmaattisin perustein, ja tämän jälkeen valittu malli aletaan kokea myös sosiaalisesti houkuttelevammaksi. Korreloivien muuttujien taustalla voi myös olla molempiin vaikuttava muuttuja. Esimerkiksi altistuminen pääosin tietylle varieteetille voisi vaikuttaa sekä sanavalintoihin että kohdepreferenssiin.

Sosiaalisissa tieteissä ilmiöt ovat kuitenkin harvoin yksioikoisia. Tässäkin tapauksessa yllä mainitut vuorovaikutussuhteet, sekä mahdollisesti useat muut, lienevät kaikki osittain tosia. Kaikkien tutkittujen muuttujien suhteiden syväluotaavaa analyysi vaatisikin huomattavasti enemmän dataa sekä mahdollisesti dynaamisten järjestelmien teoriassa käytettyjä metodeja, joihin tämän tutkimuksen resurssit eivät olleet riittäviä. Joitakin voimakkaimpia tendenssejä on kuitenkin mahdollista selvittää pienemmässäkin mittakaavassa. Tätä tarkoitusta varten haastattelin viittä osallistujaa kysyäkseni heidän omaa näkemystään sanastoonsa ja asenteisiinsa vaikuttavista tekijöistä. Jossain määrin pystyin myös nojaamaan aiempiin eurooppalaisiin tutkimuksiin kieliasenteista ja niiden vaikutuksesta englannin mallin valintaan.

Yliopistotason englannin opiskelijat eri puolilla Eurooppaa ovat useimmin antaneet affektiivisia syitä englannin mallin valinnalle (Janicka, Kul and Weckwerth 2005; Ladegaard and Sachdev 2006; Carrie 2017). Osalla opiskelijoista taustalla ovat kuitenkin pragmaattisemmat syyt kuten hyöty- ja helppous näkökohdat. Janicka, Kul and Weckwerth (2005, 254) totesivat, että puolalaisilla opiskelijoilla affektiiviset syyt painottuivat BrE:hen tähtäävillä opiskelijoilla ja pragmaattiset syyt AmE:hen tähtäävillä opiskelijoilla. Tämän tutkimuksen haastatelluilla

osallistujilla vaikutti olevan vastaava tendenssi. Molemmat BrE:hen tähtäävät haastateltavat antoivat ainoastaan affektiivisia perusteluja valinnalleen: "[…] kivemman kuulosta, hienompaa. Se kuulostaa vaan niin paljon kauniimmalta" (Osallistuja #15), "[…] se on silleen hienompaa, ja tulee mieleen sillain ylemmässä luokassa oleva henkilö" (Osallistuja #45). Kaikki AmE:hen tähtäävät osallistujat perustelivat valintaansa AmE:n tuttuudella ja helppoudella. Osallistuja #25 esitti tosin myös affektiivisen arvion ja sanoi AmE:n olevan "rennompaa".

Yritin selvittää haastatteluilla myös sanavalintojen ja kohdepreferenssin välisen korrelaation syytä. Kysyttäessä, mihin kielen osa-alueeseen haastateltavat kiinnittävät huomiota pyrkiessään preferoimaansa varieteettiin, kaikki vastasivat keskittyvänsä ääntämiseen. Ainoastaan osallistuja #8 sanoi kiinnittävänsä jonkin verran huomiota myös sanastoon. Koska tietyille varieteeteille tyypilliset sanat eivät vaikuta olevan tietoisen monitoroinnin kohteena, voisi perustellusti päätellä, että sanaston vaihtelu on alttiimpaa kielisyötteen vaikutukselle. Koska paljon aikaa englanninkielisissä maissa viettäneet vastaajat rajattiin tutkimuksen ulkopuolelle, ja harva osallistuja ilmoitti olevansa säännöllisessä kontaktissa natiivipuhujiin, amerikkalainen ja brittiläinen kielisyöte rajoittuu lähinnä mediaan.

Haastattelujen perusteella media näyttäisi olevan yhteydessä sanavalintojen lisäksi myös kohdepreferenssiin. Ainoastaan BrE:hen pyrkivät haastateltavat sanoivat seuraavansa brittiläistä mediaa. He myös tuottivat huomattavasti enemmän BrE sanoja kuin muut haastatellut sekä osallistujat keskimäärin. Median vaikutus selittäisi myös AmE:n sanaston yleisyyden kaikkien osallistujien keskuudessa, sillä 4/5 haastelluista osallistujista arvioi oppineensa englantia enemmän [AmE-vaikutteisesta] mediasta kuin [BrE-vaikutteisesta] muodollisesta opetuksesta. Jopa ainoa vastaaja, joka koki oppineensa englantia yhtä paljon koulussa kuin vapaa-ajallakin totesi sanaston tulevan pääosin mediasta: "[…] koulussa se sanasto on lähinnä niitä jotain kappaleita siinä tekstikirjassa, niin ei siitä opi läheskään tarpeeksi sanastoa" (osallistuja #15).

Kuten todettua muuttujien välisen vuorovaikutuksen tarkat mekanismit eivät kuitenkaan näin suppealla tutkimuksella selviä. Toisaalta tutkimuksen kenties tärkeimmän päätelmän kannalta ne ovat toissijaisia. Kohdepreferenssin ja kieliasenteiden välinen korrelaatio itsessään osoittaa, että oman englannin idiolektin halutaan välittävän tiettyjä sosiaalisia signaaleja. Se, tapahtuuko tämä muuttamalla kohdepreferenssiä, siihen yhdistettyjä asenteita, tai molempia, on mielenkiintoinen kysymys, mutta ei vaikuta siihen, että englannin varieteetti on selvästi väline henkilökohtaisen identiteetin muodostamisessa.

Korrelaatio kohdepreferenssin ja sanavalintojen välillä puolestaan kertoo, että osallistujat käyttävät englannin varieteetteja onnistuneesti sosiaalisen identiteettinsä ilmaisemiseen, eli lähentävät todellista minäänsä ideaaliminäänsä (Dörnyei ja Chan 2013, 438). Taustalla olevilla syy-

seuraussuhteilla on merkitystä toisen kielen oppimisen tutkimukselle, ja niitä tulisi tutkia enemmän, mutta joka tapauksessa on selvää, että myös ei-natiivit voivat tehdä sosiaalisesti motivoituja sanavalintoja. Tämä vahvistaa englantia vieraana kielenä käyttävien asemaa legitiimeinä englannin käyttäjinä, joilla on omaleimaisia tapoja käyttää englantia sosiaalisten merkitysten luomiseen. Tätä tukevat myös havainnot siitä, että norjalaiset lukio-opiskelijat eivät aina pyri mihinkään natiivivarieteettiin (Rindal ja Piercy 2013) ja toisaalta hyödyntävät natiivivarieteetteja eri rekistereinä (Rindal 2010).

Se, että ainakin edistyneemmät englanninopiskelijat käyttävät englannin eri varieteetteja identiteettinsä rakentamisen välineinä, asettaa myös paineita muodolliselle opetukselle omalähtöisen englannin mallin tavoittelun tukemiselle. Haastattelujen perusteella tämä on erityisen tärkeää suhteessa ääntämiseen. Toisaalta sanaston merkitys kielelliselle identiteetille ei vaikuta olevan kovin suuri, vaikka osallistujien sanasto korreloikin kohdepreferenssin kanssa tilastollisesti merkittävällä tavalla. Asiaa tulisi toki tutkia laajemmalla otoksella sekä alemmilla koulutusasteilla, mutta tämän tutkimuksen perusteella voisi olla perusteltua painottaa sanastonopetuksessa ymmärrettävyyttä enemmän kuin omavalintaista kielellistä identiteettiä. Nähdäkseni tähän on kolme mahdollista näkökulmaa:

1) Amerikkalaisen sanaston suosiminen: Tämä tutkimus osaltaan vahvistaa sanaston amerikkalaistumisen globaalilla tasolla. Voisi olla perusteltua myötäillä tätä kehitystä myös muodollisessa opetuksessa.

2) Brittiläisen sanaston suosiminen: On viitteitä siitä, että koulun ulkopuolella altistutaan pääasiassa amerikanenglannille, mikä on syynä sanaston yleiselle amerikkalaistumiselle. Monipuolisen sanaston oppimisen kannalta voisi siis olla järkevää edelleen opettaa pääosin brittiläistä sanastoa kouluissa.

3) Kansainvälisesti yleisimmän sanaston suosiminen: Suurin osa englanninkielisestä kommunikaatiosta käydään ei-natiivien välillä. Ymmärrettävyyden kannalta voisi olla tehokkainta selvittää kansainvälisesti käytetyimmät muodot ja käyttää ensisijaisesti niitä muodollisessa opetuksessa.

Suomalaisten opiskelijoiden englannin kielen käyttö identiteetin osana tukee myös näkemystä ei-natiivien englannin käyttäjien legitimiteetistä. Englantia vieraana kielenä käyttävien ihmisten tekemät sosiaalisesti motivoidut kielelliset valinnat asettavat paineita sosiolingvistisen kompetenssin käsitteen uudistamiselle. Toistaiseksi sosiolingvistisen kompetenssin on katsottu edustavan ainoastaan kykyä emuloida natiivipuhujien sosiolingvististä variaatiota (Durham 2014, 14). Olisi

perusteltua muuttaa määritelmä huomioimaan ei-natiivien omaleimaiset tavat luoda sosiaalisia merkityksiä englannin kielen avulla.