MASTERARBEIT | MASTER'S THESIS

Titel | Title

Language attitudes and emotional tone: the effect of mood on speaker evaluations

verfasst von | submitted by

Nadine Pfandl BEd

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Education (MEd)

Wien | Vienna, 2025

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 199 507 519 02
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium Lehramt Sek (AB) Unterrichtsfach Englisch Unterrichtsfach Latein
Betreut von | Supervisor:	Univ.-Prof. Mag. Dr. Nikolaus Ritt

Acknowledgements

This project would not have been possible without the support of a few people; therefore, I would like to use this opportunity to express my sincere gratitude.

First and foremost, I want to thank my supervisor Univ-Prof. Mag. Dr. Nikolaus Ritt, whose expertise and support has guided me through this whole thesis. His feedback shaped the research design and the overall direction of the study.

Moreover, I am grateful for the help of my parents, who encouraged me not only during writing this thesis but throughout my whole studies. In addition, I would like to thank my partner Niklas and all of my friends for the continuous emotional support. Furthermore, I want to express my gratitude to my close friend Vanessa with whom I went from beginning to the end of the studies.

A special thank you also goes to the 180 voluntary participants as well as friends and family, who were willing to distribute the survey eagerly.

Table of content

List of Figures

List of Tables

List of abbreviations

AI	Artificial Intelligence
APT	Accent Prestige Theory
GA	General American
GB	Great Britain
GPPT	Gender and Prestige Preference Theory
L1	First Language
L2	Second Language
LX	Any foreign language beyond L1
MGT	Matched-guise Technique
RP	Received Pronunciation
SVLR	Scottish Vowel Length Rule
SSE	Standard Scottish English
TTS	Text-to-Speech
UK	United Kingdom
US	United States

1 Introduction

English has taken on the role of a global language, encompassing a great diversity of varieties which may evoke mental representations in the listener's mind - frequently even without a visual imagination of the speaker. As a result, various judgements are either consciously or unconsciously made, solely based on auditory features. These first impressions can center around multiple dimensions and might even lead to stereotype-based evaluations and therefore shape the overall perception towards a person. On the basis of certain linguistic or regional characteristics, potentially indicating a specific accent or dialect, interlocutors infer conclusions about social status, especially with regards to upper or lower class. Subsequently, these can affect further ratings of character traits, including qualities like warmth and dynamism. In addition, implications about someone's competences emerge, such as the perceived intelligence or educational background.

Although some judgements may initially seem of little consequence, they can still serve as impactful factors on further attitudes towards speakers, as presumptions can foster prejudiced thinking. The previously described instances belong to the broader field of language attitudes, a research domain which has gained remarkable interest in academia. To investigate perceptions towards different speakers, scholars have employed various approaches, with the matched-guise technique (MGT) being among the most frequently used. This method exclusively focuses on linguistic features and avoids potential influential factors, like gender or pitch, as the same speaker reads out a neutral text passage in different accents. People are then asked to describe their attitudes in terms of "social status, solidarity and dynamism" (Giles & Watson 2013: 28). Recent studies concluded that among English varieties, Received Pronunciation (RP) tends to be associated with the highest degree of prestige, while General American (GA) often receives better outcomes in solidarity traits (Carrie & McKenzie 2018: 324); however, in light of the social identity theory, evaluations may depend on someone's personally used English variety, as people commonly prefer accents that seem most familiar (Dalton-Puffer et al. 1997: 117).

To obtain results, this project selected three speakers, representing Received Pronunciation (RP), General American (GA) and Scottish Standard English (SSE), following the approach of the MGT. The aim was not only to explore attitudes towards the three varieties, but also to observe whether mood affects the perceptions towards the concerned voices; hence, the RP and GA sample occurred twice, each with a shift in emotional tone. Participants of the online survey were required to rate the included audio files in terms of different character traits, as well as their pronunciation clarity and perceived suitability as a news reporter. In addition, items asked

respondents about their willingness to meet the individual speakers for a coffee and the voices' possible regional provenances. Moreover, a rank ordering scheme was included on a scale from one to five stars.

The purpose of this thesis is to answer the following set of research questions. First, it examines the potential impact of speakers' own variety. Secondly, the project analyzes whether participants realize that two voices occur twice, differing in mood, and how emotional tone influences respondents' perceptions. As all samples were AI-generated, this project also seeks to investigate if the files are perceived as computer-produced. Moreover, influential factors, such as age, gender or mother tongue on speaker evaluations are assessed. Lastly, differences in recognition rates need to be scrutinized.

In terms of the structure of this thesis, chapters two to five elaborate on the most important concepts from the literature, including key definitions such as accent or dialect. Subsequently, the concerned language varieties, specifically Received Pronunciation (RP), General American (GA) and Standard Scottish English (SSE) with focus on their most important linguistic features, are described in greater detail. Theoretical frameworks regarding the domain of language attitudes as well as corresponding research approaches follow. The methodology part can be found in chapter six, describing research questions and hypotheses, the overall research, questionnaire design and a brief overview of AI in academia. Furthermore, the selection of the reading passage is explained along with crucial features of the chosen speakers. Finally, results of the conducted study are provided in section seven and afterwards compared to former findings from the literature.

As previously stated, initial impressions based on certain speech patterns may easily result in stereotypical judgements. Especially in times of increasing globalization, exposure to foreign languages is steadily increasing; thus, this specific research domain requires even deeper engagement.

2 Study of language

The first part of this chapter describes general implications of the concept of language with more detailed information on the differences between accents, dialects or varieties.

2.1 Definition of language

Undoubtedly, language is an indispensable part of life and therefore often taken for granted. This highlights the importance of observing, why it exists and how it can be defined, which proves to be highly complex (Rastall 2018: 3). Some may consider language as a separate instrument, which our mind uses for various purposes; however, both are parts of the "same cognitive activity" and therefore coexist (Rastall 2018: 5). Despite the difficulties in finding a concise definition, its functions are evident. We use the power of language to express thoughts or call for actions (Rastall 2018: 5); nevertheless, to put it in Rastall's words (2018: 5): "language, as acts of speaking, is **not** *for* communication: it *is* communication".

Language constitutes a crucial component of speech, which can be described with the following terms: "social, permanent and abstract" (Milewski 2019: 7, 8). The first descriptor, "social", refers to the ability to talk to those, who are part of the same language community based on grammar, syntax, vocabulary or norms to ensure mutual understanding (Milewski 2019: 8). Besides that, the aspect of permanence underlines the continuous concept of speech, as languages persist over a long period, although they have naturally undergone several changes (Milewski 2019: 8; Rastall 2018: 9). The abstractness describes relations to different concepts, which all address symbolic categories (Milewski 2019: 8).

2.2 Accent

An accent reflects a person's identity. As stated by Rangan (2023: 3), the *Oxford English Dictionary* describes it as "a way of pronouncing a language that is distinctive to a country, area, social class, or individual". Generally, an accent depends on regions or social class and can therefore serve as a marker of otherness, frequently leading to a highly comparative notion; nevertheless, grammar, syntax and vocabulary mostly comply with the standard language, differing primarily in pronunciation (Giles 1970: 213; Hughes et al. 2005: 2). Rangan (2023: 3) conceptualizes it with the following words: "global and local, racialized, gendered, ethnic and national, cosmopolitan and provincial, unconscious and performative, visual, audial, gestural and intertextual". These terms illustrate the various dimensions that the concept of accent covers. A lot of speakers instinctively associate an accent with a sound that is different from what is familiar and therefore, often negatively connotated (Rangan 2023: 4). As a result,

listeners may immediately draw conclusions about a person's competences (Laroche et al. 2022: 1210); however, it needs to be emphasized that everyone speaks with an accent. Speakers of a standard variety are in most cases not perceived as having a particular one, since it is more subtle than others (Yule 2016: 279).

2.3 Dialect

Whereas an accent primarily focuses on pronunciation, a dialect involves deviations from the norm in multiple areas, such as vocabulary or grammar (Ammon et al. 2004: 267, 268). It rarely appears in written forms, except for the purpose of representing the dialect itself. A dialect tends to occur in informal situations, although this may vary depending on the region. Ammon et al. (2004: 270) point out that dialects are frequently assumed "to lack communicative function […] and to function less well for trans-regional communication". This has been negated by Trudgill (Ammon et al. 2004: 270). Unquestionably, the norm is often seen as superior, but Trudgill considers them as equal, due to their same functionality (Ammon et al. 2004: 270).

2.4 Language variety

Variation counts as a key element of language existence. This term refers to all kinds of language variations without any negative or positive connotation. In contrast, accents or dialects are often associated with lower prestige, depending on the sociolinguistic context (Bieswanger & Becker 2017: 174). It is necessary to consider that rules and standards - similar to the concept of language itself - are not persistent but change over time. Certain variants might become outdated, new ones develop and others change their speaker community (Datska 2019: 229). This is due to the fact that language always underlies social principles.

Linguistic variation is not confined to a single field of language, but includes all areas, such as syntax, semantics, pragmatics, vocabulary or morphology (Bieswanger & Becker 2017: 173). Bieswanger and Becker (2017: 173) support their argument with a concrete example from British and American English. Even though they use different words, like *lift* and *elevator*, they address the same notion. There are three forms, namely sociolects, geographical and functional varieties. Moreover, a standard variety can be found in most communities but this is seen as a separate category and therefore, not included in the before mentioned types. As Yule (2016: 269) explains, most speakers view the norm as an "idealized variety", since it mostly occurs in standardized institutions, such as education and media.

2.5 Standard Language

A standard language is understood as a "uniform, minimally varying (homogenous) form of language" which always coexists with other varieties (Walsh 2021: 773). Speakers who adhere to a specific norm are often perceived as monolinguals. Research on language attitudes demonstrated that standard variants are associated with a higher status and hence, sometimes even regarded as an ideology (Walsh 2021: 774). Every language typically possesses a standard form, such as Standard British or Standard American. As Garrett (2010: 6) stresses, these norms considerably shape the domain of language attitudes. They are commonly associated with a high degree of correctness; thus, Garrett (2010: 7) assumes that this issue lays the foundation for their corresponding level of prestige. As a result, non-standard varieties are considered as inferior; however, it is important to note that standard languages are not inherently prestigious. Instead, speakers highly affect the development. If people of upper-class groups use the given variant, "the standard form carries overt prestige" (Hill 2011: 1). Frequently, this status is then linked to wealth and success, often leading to a desire of acquiring the concerned variety (Hill 2011: 1).

3 English language varieties

The following chapter gives an overview about English as a world language as well as three different varieties, more precisely General American (GA), Received Pronunciation (RP) and Scottish English, which form the basis for this whole thesis.

3.1 English as a world language

John Adams accurately predicted the future role of English in 1780 with the following statement: “English will be the most respectable language in the world and the most universally used read and spoken in the next century, if not before the close of this one” (as cited in Bieswanger & Becker 2017: 34). This reflects the role of English in the 21st century, which can be attributed to its tremendous power in economy or politics of Great Britian (GB) as well as the United States (US), but not so much to its linguistic qualities. In the three circles of English the number of English speakers is presented. The inner circle, which includes speakers from the US and UK, contains 320 to 380 million people, the outer circle (Singapore or India) 300 to 500 million and lastly, the expanding circle (Russia, China) compromises 500 to 1000 million (Bieswanger & Becker 2017: 35); hence, it became evident that non-native speakers form the biggest group; however, the definition of “speaker” frequently seems difficult to define, as proficiency levels sometimes do not set clear boundaries.

Despite the widespread use of English in diverse communities, its status as a lingua franca solidifies the role as a world language. English is particularly prominent in conversations, where interlocutors do not share a mother tongue. India plays a crucial role in this context, as both Hindi and English are used most frequently in communicating across the country. Moreover, it occurs in international business or academia and even in the field of travelling, such as air traffic control (Bieswanger & Becker 2017: 36).

In short, English, unlike other languages, has achieved a global status with a significant expansion not only in geographical but also in practical terms; nevertheless, it needs to be considered that this also results in a high exposure to potential changes and variations (Bieswanger & Becker 2017: 36).

3.2 General American (GA)

First and foremost, it is crucial to acknowledge the great variation of English within the United States. Consequently, the term General American (GA) does not represent its numerous

linguistic varieties (Hillenbrand 2003: 121); still, it is considered suitable for describing the specific phonological form, regarded as the standard language (Fodde 2013: 12).

This linguistic variation reflects the diversity within the American population, resulting from local differences. As a result, multiple varieties of the standard language are an indispensable part, particularly for expressing regional identity; nevertheless, a norm is necessary to ensure mutual intelligibility (Fodde 2013: 10, 11). Generally, the development of American English as an official language shows many parallels to British English, especially up until the 19th century (Fodde 2013: 12). To put it in Fodde's words, key factors were "the emerging of economically prestigious social strata of population […], alongside geographical distribution" (Fodde 2013: 12).

Furthermore, immigration had a major influence on the US. Consequently, vocabulary and pronunciation had to undergo several changes, caused by people's native languages (Fodde 2013: 12). Notably, these did not stop in the past but continue until present times. Following immigrant groups influenced the development of the American English language: Indian, German, Dutch, French, Italian, Yiddish, Hispanics as well as many others. First, the German people in the 17th or 18th century were leaving their home due to religious persecution, the majority being from Bavaria. As a result, several 'German cities' have emerged in the US. Through contact with the European refugees, Americans adopted certain words in their own language habits, such as *kindergarten*. Many of them relate to food and drink, for example *hamburger*, *pretzel* or *sauerkraut* (Kovecses 2000: 33). Even in the educational context, some German vocabulary is still present (*seminar* or *semester*) (Kovecses 2000: 33). Later, in the 19th and 20th century, the number of immigrants drastically increased and as Kovecses (2000: 34) puts it, US was known as "a country of the 'huddled masses'". Moreover, Italians also had a slight impact on GA; however, their influences are not particularly noteworthy in the linguistic domain, since they mostly affected words, like *pasta*, *spaghetti* or *espresso*; therefore, mostly relating to the culinary field (Kovecses 2000: 35).

Generally, hundreds of millions of people use GA daily. Linguists have even presented a growing number in the future (Kovecses 2000: 8).

3.2.1 Characteristics of GA

In comparison to RP, GA is associated with an economical character, reflected in its "use of shorter words […], shorter spellings […] and shorter sentences […]" (Kovecses 2000: 13). Moreover, a regular component is evident in the observation of specific past forms (*burned – burnt*) or the standardized use of *have* ("*Do you have" – "Have you"*) (Kovecses 2000: 14).

Besides that, GA appears to be very direct due to the removal of specific structures, such as *so do I*, which speakers of American English replace by *I do too* (Kovecses 2000: 14). Another key element of GA is the tendency to informality, as speakers address others on first-names basis or include "relationship markers in discourse" (Kovecses 2000: 14). Additionally, Standard American English incorporates many dynamic features, as evident in the preference of goal-oriented words (*to take a shower – to have a shower*) (Kovecses 2000: 15).

3.2.2 Sounds of GA

GA compromises segmental and suprasegmental elements and records 17 vowels as well as 24 consonants (Datska 2019: 232). A key element in standard GA is the free phonemic variation, which describes "a formal kind of variation between the phonemes in a word phonemic structure" (Datska 2019: 232). It is noteworthy that it does neither affect the meaning nor a word's lexical and grammatical category. Vowels play a crucial role in this context, since they are mainly responsible for leading to a different way of pronunciation (Datska 2019: 232). The most frequent sounds under consideration are: /ɑ:/, /e/, /ə/, /ɪ/, /oʊ/ and /ɔ:/, followed by /e/, /ə/ and /ɪ/ (Datska 2019: 233). Overall, a tendency to the center of the mouth is the most common reason for vowel variation. The vowels also in combination with other sounds are presented in Figure 1.

Regarding consonants, /Ø/, /s/, /z/, /j/, /t/, /tʃ/ are highly present in phonetic variation. The first symbol describes the omission of a certain sound in one variety, where others would include a consonant at the same position. A concrete example refers to the /Ø-j/ - variation, highly typical for Standard American English. Specifically, /j/ is usually avoided between /t/, /d/ and /n/ such as in *tune* (Datska 2019: 233). A common descriptor for GA is the term rhotic, meaning that the /r/ sound is prominent in all instances, either before or after a vowel, for instance in *bar* or *square* (Kovecses 2000: 25; Kuecker et al. 2015: 623). There are different options to produce this sound, but speakers of GA tend to slightly twist the tongue to the back, which is known as the retroflex /r/. As Hannisdal (2020: 256) presents, /t/ flapping is a typical phenomenon in American English and can be understood as "the realization of /t/ as a voiced alveolar flap" depending on the surrounding environment. It primarily occurs between two vowels as well as between /r/ and a vowel. Besides that, the precedence of /n/ before /t/ also causes a t-flapping (Hannisdal 2020: 258). Another common sound of GA is /æ/ as in *bathroom*, which is pronounced with unrounded lips and a slightly open mouth. It frequently occurs in words, where it precedes /f/, /s/ or /th/ (Kovecses 2000: 25). The phoneme /u/ in *new* in GA follows the model of a /j/-less pronunciation in comparison to RP. Lastly, the in RP frequently used short /ɒ/ vowel in, for instance *hot*, is produced similarly as /ɑ:/ (Kovecses 2000: 26).

beat	i	*boot*	u	*peer*	ɪɹ
pit	ɪ	*put*	ʊ	*pear*	ɛɹ
hate	ej	*boat*	ow	*part*	ɑɹ
pet	ɛ	*bought*	ɔ	*hurt*	əɹ
pat	æ	*pot*	ɑ	*cure, jury*	ʊɹ
path	æ	*soft*	ɔ	*four*	ɔɹ
but	ʌ	*palm*	ɑ	*baby*	i
bite	aj	*out*	aw	*runner*	əɹ
		choice	ɔj	*sofa*	ə

Figure 1. GA vowels (Rogers 2014: 120)

3.3 Received Pronunciation (RP)

3.3.1 Characteristics of RP

First and foremost, it must be noted that finding a concise definition of RP is highly complex and has caused multiple debates among researchers. As Wells (1982: 279) outlines, although the majority of British residents is frequently exposed to this variety, only a minor proportion actively uses it, which leads to mental representation even without a correct identification. Some scholars describe RP as both outdated and heavily shaped by social factors, as the accent is primarily spoken by the socio-economic elite, resulting in a high degree of class-oriented meanings (Mugglestone 2012: 1900). Mugglestone (2012: 1908) relates this issue to the phrase "ideological baggage". The persistence of the term has already been challenged by alternative descriptors such as "BBC Pronunciation" or "Reference Pronunciation", indicating a lack of consensus within academia.

In addition, it has been turned out difficult to assign RP to a specific regional provenance and was therefore even considered as a "non-regional pronunciation" by Collins and Mees in 2003 (as cited in Mugglestone 2012: 1900). Yet, discussions arose whether this can be regarded as an appropriate description, as most researchers would still attribute it to the South of England (Mugglestone 2012: 1908).

Briefly, the problematic nature of RP originates from three areas: firstly, the challenge in finding a definition without class bias; secondly, providing a clear identification of the accent itself and lastly, its embedding "in a temporal continuum", leading to associations of conservatism (Mugglestone 2012: 1909).

Despite these concerns, this thesis labels the accent of the speaker from England as RP to ensure terminological consistency. Furthermore, it was considered a suitable descriptor for representative matters in the provided diagrams. Following a critical analysis of the term RP,

the following section presents more detailed information on most important findings of the accent.

Received Pronunciation (RP) emerged in 18th century London (Fisher 2001: 71). The term was developed by John Walker, who described it as "generally adopted or approved" (Fisher 2001: 72) and "accepted in the most polite circles of society" (Hughes et al. 2005: 3). Even though times have changed, people of upper classes are still most common to use RP. Today, researchers estimate that only around three to five percent actively speak the given accent. This raises the question, why RP is frequently taught at schools compared to other accents. The authors attributed this to its high prestige and power; therefore, speakers often consider RP as most desirable to learn (Hughes et al. 2005: 3).

Moreover, due to its frequent use by news reporters, RP has easily spread across the country, leading to common understanding of the accent. Moreover, the majority of research has focused on RP, which is partly a result of the high demand among teachers and learners (Hughes et al. 2005: 3); nevertheless, some native speakers perceive it as unnatural, which often depends on geographical factors. In areas, where the local dialect differs from RP, locals consider it as even more artificial or snobbish (Hughes et al. 2005: 4). Consequently, an increasing number of especially young people, tend to adopt more casual ways of pronunciation, which even led to the emergence of new varieties, such as the Estuary English, combining features from RP and the variety of London's working class, Cockney (Hughes et al. 2005: 5).

3.3.2 Sounds of RP

Generally, Received Pronunciation compromises 24 consonants and 20 vowels (Datska 2019: 231). RP is classified as a non-rhotic variety, which distinguishes it from GA and Scottish English; therefore, the /r/ sound is omitted especially in the syllable coda and only present in the beginning or middle of words, preceding vowels, which can also be seen in the transcriptions in Figure 2 (Kaminska 1995: 98). Unlike in other varieties, the /j/ sound is pronounced after alveolars, such as *student*. In addition, the /h/ in the beginning of words, like *half*, is fully articulated (Rogers 2014: 111). The glottal stop /ʔ/, a type of plosive, is widely used in RP. The most frequent position is at the end of a syllable before another consonant, which is then referred to as glottalization. Speakers of RP are likely "to use glottal stop as a realization of /t/", which might be due to media influence of the London variety (Hughes et al. 2005: 43).; nevertheless, highly attentive speakers avoid glottalization (Hughes et al. 2005: 43). RP includes the following nine fricatives: /f/, /θ/, /s/, /ʃ/, /v/, /ð/, /z/, /ʒ/, /h/. The distinction between the voiced (/v/, /ð/, /z/, /ʒ/) and voiceless (/f/, /θ/, /s/, /ʃ/, /h/) fricatives is not always

clear. People who use RP do not fully voice the given consonants in word final positions, such as *beige* or *bathe*. The sound /3/ only occurs in the middle of a word, like *pleasure*, except for French loanwords (*prestige*). The sound /h/ exclusively appears at the beginning of a syllable before a vowel; however, the subsequent vowel affects the realization of /h/. A concrete example of the variety in realization refers to *heat* or *heart*; hence, it can either be described as a "voiceless vowel as well as a fricative", explained by Ladefoged and Maddieson (as cited in Hughes et al. 2005: 44). Generally, /h/ is often weakened or even dropped in unstressed words like *he* or *his*, as in *help him* (Hughes et al. 2005: 44).

beat	i	*boot*	u	*peer*	ɪə
pit	ɪ	*put*	ʊ	*pear*	ɛə
hate	ej	*boat*	əw	*part*	ɑ
pet	ɛ	*bought*	ɔ	*hurt*	ɜ
pat	æ	*pot*	ɒ	*cure, jury*	ʊə
path	ɑ	*soft*	ɒ	*four*	ɔ
but	ʌ	*palm*	ɑ	*baby*	i
bite	aj	*out*	aw	*runner*	ə
		choice	ɔj	*sofa*	ə

Figure 2. RP vowels (Rogers 2014: 110)

3.4 Scottish English

3.4.1 Characteristics of SSE

Dialects in Scotland are separated by the following borders: "the North Sea, the Highland Line, the Mid/North Line and the Scottish/English border" (Kingstone 2015: 316). The term Scottish English is not easy to define, as it includes a great number of different varieties, which can be taken from Figure 3. To put it in Douglas' (2009: 33) words, it can be understood as a "bipolar linguistic continuum" and reaches from *Scots*, which contains the most Scottish features from the past and is mostly associated with rural dialects, to *Scottish Standard English*, including many features from RP (see Figure 3). In contemporary usage, *Scots* is considered a predominantly spoken language, due to the missing standardization of orthography (Douglas 2009: 37). Linked to its historical separation from England, it is obvious that their own Scottish variety emerged; nevertheless, RP is still prominent. In terms of grammar, a high degree of correspondence with Standard English is detectable (Rogers 2014: 113). Overall, the main difference to RP or other English varieties lies in the matter of pronunciation. There are also some minor deviations in terms of grammar, which the following examples illustrate: while *Standard Scottish English* uses: '*The car needs washed'*, the equivalent in the Standard English

variety would be '*The car needs washing'* (Douglas 2009: 34). Just like RP and GA, Standard Scottish English is classified as one of the varieties of *International Standard English* and therefore, prominent in most formal instances in Scotland (Douglas 2009: 38).

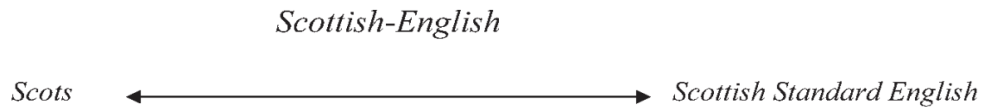


Figure 3.1 Scottish-English linguistic continuum

Figure 3. Range of linguistic Scottish English varieties (Douglas 2009: 33)

Generally, the status of *Scots* has been a controversial topic within the field of linguistics. The debate centers around the question, whether it should be viewed as an individual language or a common variety of English (Douglas 2009: 43). As stated by Kingstone (2015: 319), the answer frequently depends on the institution; hence, it becomes evident that politics play a crucial role, sometimes an even more important one than linguistics itself (Douglas 2009: 43). Undoubtedly, the interest in *Scots* has increased over the last decades. Studies by Murdoch in 1995 revealed that especially speakers of rural regions or "older, working-class manual workers" report a strong feeling of belonging to the given language, whereas younger upper-middle and middle-class language users usually identify themselves with the English language. In comparison, a research project from 2010 presented that 85% associate themselves as speakers of Scots, whereas around two thirds of the study population did not consider it as a separate language (Kingstone 2015: 320). Since 2013, *Scots* can be found in the Scottish census, marked as a separate language with around one third being able to properly speak it. Due to the increasing popularity, there were some considerations, whether *Scots* should be incorporated in schools or dictionaries, since there has hardly been an emphasis on developing or improving writing skills in Scots so far (Douglas 2009: 37; Kingstone 2015: 320); therefore, it is common to mostly use Standard English in a written form (Douglas 2009: 38).

3.4.2 Sounds of SSE

Similarly to GA, Scottish can be described as *rothic* (Rogers 2014: 114). The variety encompasses the following monophthongs: /i/, /e/, /ɛ/, /a/, /u/, /ʌ/, /ɑ/, /o/ (Wells 1982: 400). In contrast to RP, which uses two different vowels /ʊ/ and /u/, Scottish English creates a mixture of both, which is represented by /ʉ/, a high central vowel (Rogers 2014: 114) (see Figure 4). Another similar phenomenon can be detected in in the vowels of the RP pronunciation of *bath*

and *hat*. The Scottish variety merges these, creating one "low front vowel" (Rogers 2014: 114). While users of RP pronounce the sound /eɪ/, SSE speakers tend to favor /eː/. Instead of the schwa sound, Scots often report the unstressed /ʌ/ as in *sofa*. In addition, SSE does not distinguish between /o/ and /ɔː/, as speakers are more likely to pronounce /ɔ/ (Kaminska 1995: 43). At the beginning of a lexical item, /h/ remains present (Rogers 2014: 114). SSE lacks a differentiation between /æ/ and /ɑː/, using /a/ in both cases (Wells 1982: 403). As Rogers (2014: 114) states, "at the end of a morpheme, vowels, except /ɪ/ and /ʌ/, are long […] and […] vowels, except /ɪ/ and /ʌ/, are long before the consonants /v/, /ð/, /z/, /r/". A crucial component of Scottish Standard English is the Scottish Vowel Length Rule (SVLR) (Rathcke & Stuart-Smith 2015: 404). Overall, linguists classified the vowel duration in SSE as 'quasi-phonemic'. Whereas other English varieties do not put much emphasis on the duration of vowels for signaling the "dense/lax distinction", SSE provides several pairs, where the duration of the vowel makes a difference, such as "*brood* vs. *brewed*" (Rathcke & Stuart-Smith 2015: 405). The specific rule covers the most important principles of when vowels are produced as long sounds. Concrete examples are preceding voiced fricatives or /r/ and at the end of a morpheme. Consequently, short vowels occur in all other instances (Rathcke & Stuart-Smith 2015: 405). This phenomenon was first explained by Aitken in 1962 (Rathcke & Stuart-Smith 2015: 405).

beat	i	*boot*	ʉ	*peer*	ir
pit	ɪ	*put*	ʉ	*pear*	er
hate	e	*boat*	o	*part*	ar
pet	ɛ	*bought*	ɔ	*hurt*	ər
pat	a	*pot*	ɔ	*cure, jury*	ʉr
path	a	*soft*	ɔ	*four*	or
but	ʌ	*palm*	a	*baby*	e
bite	a͡e, ʌj	*out*	ʌw	*runner*	ər
		choice	ɒj	*sofa*	ʌ

Figure 4. Scottish vowels (Rogers 2014: 114)

4 Language Attitudes

Beyond visual appearances, such as age or gender, language plays a crucial role in making judgements about a speaker. Language is inherently complex and includes multiple dimensions like grammar, vocabulary, pronunciation and syntax. Variation in these fields are frequently due to geographical or social factors; hence, specific dialects or accents can convey more personal information. In the negative sense, certain impressions may result in prejudice or even discrimination, especially for domains like social class (Giles & Watson 2013: 1, 2, 28). According to Giles and Watson (2013: 28), speakers of a non-standard variety are often considered as unskilled or incompetent. These judgments are based on the three following aspects: "social status, solidarity and dynamism" (Giles & Watson 2013: 28). Studies revealed that users of non-standard varieties report lower results on all levels (Giles & Watson 2013: 28).

4.1 Definition

To put it in Baker's (1992: 9) words "attitudes is a term in common usage" and among the earliest interests in the domain of social psychology. Even though attitudes are highly essential elements of social psychology and used frequently by researchers in linguistics, finding a concise definition has proven to be difficult (Baker 1992: 11; Kircher & Zipp 2022: 2). In addition, the measurability of attitudes appears complex, primarily because of their origin, as they are never directly observable (Garrett et al. 2003: 2; Kircher & Zipp 2022: 2). This results in a wide range of definitions, differing in their detailedness (Garrett et al. 2003: 2).

Generally, an attitude can be understood as a reaction towards different objects accompanied by judgmental evaluations, as stated by McGuire (1985: 239). As presented by Garrett et al. (2003: 2), Henerson, Morris and Fitz-Gibbon define attitudes as items that people judge, shaped by feelings or values, while Oppenheim also elaborates on potential consequences of such evaluative behavior. Sarnoff (as cited in Garrett et al. 2003: 3) relates them to an either positive or negative reaction to certain items or objects. In the specific context of language attitudes, Peterson (2019: 8) provides the following definition: "beliefs or judgements people have about certain social styles of language, features of a language or varieties of a language". A parallel across all possible definitions is the issue of an "evaluative orientation" (Ahn 2017: 34). Two important descriptors for attitudes are *valence* and *intensity*, both equally important in the domain of attitude research. While the first term refers to the positive or negative judgements, the latter describes the firmness of the attitude (McKenzie & McNeil 2022: 15). Another similarity among researchers' findings is the difficulty of observing attitudes. As Kircher and Zipp (2022: 2) state, "attitudes are states of readiness, dispositions, or tendencies – and therefore

have no overt substance". This leads to challenges in choosing the most appropriate type of data to analyze attitudes (Kircher & Zipp 2022: 2).

As Hammine (2021: 380) explains, language attitudes "are often deep-rooted, subconscious and biased, often stemming from colonial domination". Generally, they have an impact on the persistence of a language and can therefore partially predict its development in the future (Kuznietsova 2024: 454). Attitudes convey judgements about various objects through evaluative descriptions, frequently shaped by emotions (Kuznietsova 2024: 454).

Overall, attitudes towards languages affect speakers' daily lives although people are often unaware of their effects (Hymes 1972) (as cited in Ahn 2017: 36). A common consequence of evaluative judgement are stereotypes, which might benefit some groups while discriminating others (Friedrich 2000) (as cited in Ahn 2017: 36).

4.2 Structure of Attitudes

A common model in research refers to the trilateral structure of attitudes introduced by Rosenberg and Hovland in 1960, dividing them into behavioral, cognitive and affective elements, (as cited in Garrett et al. 2003: 4). The cognitive part encompasses various beliefs or widely accepted norms. It stresses that attitudes are not an inherent part of people but rather acquired through personal experiences over time. Frequently, cognitive components of attitudes are equalized with an 'opinion' (Ahn 2017: 35).

An affective component refers to feelings and emotional reactions towards certain items. Even when speakers cannot clearly label a specific variety, they might still judge it either positively or negatively, which then influences the whole ongoing conversation. Linguists argue that attitudes tend to be based on extensive affective elements (Garrett et al. 2003: 4). A concrete example of this component is related to the social interaction. People may react, based on their impressions of a specific variety, either favorably or unfavorably to their interlocutor. (Ahn 2017: 35; Garrett et al. 2003: 4).

The behavioral element describes influences on speakers to perform certain actions (Garrett et al. 2003: 3); however, there have been some ongoing discussions on the genuine effect of attitudes on behavior. Studies by Bohner and Wanke in 2002 have proven that these can indeed considerably determine various actions (Ahn 2017: 35). A concrete example would be saving money to study English in English-speaking countries. (Ahn 2017: 35).

Even though attitudes are always analyzed based on these three elements, their interrelations still need further observation. Clore and Schnall claim that the components should not merely be understood descriptors but as "causes and triggers of attitudes" (as cited in Ahn 2017: 35).

4.3 Identity and Language

Identity is a multifaceted concept that relates to many multiple areas. Specifically, it refers to an individual and group belonging, containing feelings of both inclusion and exclusion, depending on whether someone is accepted as part of the in- or outgroup. Notably, a person can naturally belong to several groups simultaneously, depending on the context, such as family or work (Douglas 2009: 11). Literature claims that identities are not fixed. Instead, Douglas (2009: 13) compares it to a "strange country" that is re-explored on a daily basis and therefore described as dynamic and not static.

To draw conclusions between identities and language, it is important to consider the symbolic meaning of a language beyond its solely purpose of communication (Douglas 2009: 15). To put it in Douglas' words (2009: 15): "language can emphasize group identity, in that members of the group share the same language or language variety". In contrast, it serves as a marker of the outside for those, who lack the shared knowledge; nonetheless, when people have internalized the same linguistic features, it is frequently assumed that beliefs and norms overlap as well, thereby fostering the feeling of togetherness. A crucial concept in this context is the theory of dialogism, developed by Mikhail Bakhtin, a literary scholar (Douglas 2009: 16). He claims that there is a continuous "interaction between meanings"; thus, this dialogic element influences identity and creates a changing character. In addition, this notion highlights the importance of language within the creation of identity, especially regarding nationality. This is due to the link that languages draw between present and past days (Douglas 2009: 16).

4.4 Gender and Prestige

Women and men report different patterns in their language habits. Since the 1960s, researchers started to observe these differences in relation to social factors, such as prestige. As Trudgill (1974: 88) outlines, women aim at a more formal use of language than men, which, in the specific case of RP, becomes transparent through glottal stops. As mentioned before, highly sophisticated speakers tend to avoid it (Hughes et al. 2005: 43). This phenomenon is known as the *Gender and Prestige Preference Theory* (GPPT). In this context, prestige refers to upper classes (Angle & Hesse-Biber 1981: 449). All conducted studies revealed that social hierarchies are reflected in speech patterns. Hence, language of the wealthy population tends to be viewed

as more correct and sophisticated, whereas lower classes are said to deviate from the standard language in an incorrect way. Nevertheless, poor people show a higher level of empathy with others and can therefore, identify with them more easily. Moreover, they "may see considerable advantage in talking and identifying with the rich" (Angle & Hesse-Biber 1981: 450). Thus, one might assume that this results in an imitation of the language that is present in the domains of upper classes. With special regard to gender, the literature outlines that particularly women, lower in hierarchy, aim to adapt to the language habits of the wealthier people (Angle & Hesse-Biber 1981: 451).

The early stages of this theory go back to William Labov, who observed the standard American English together with a local Brooklyn dialect. Through his work, he pointed out that across all society classes, expect for the poorest population, women usually speak a variety closer to the norm. This phenomenon is based on the principles that females tend to show greater awareness to distinct linguistic elements than men (as cited in Angle & Hesse-Biber 1981: 451). These outcomes cannot only be applied on the American context, but also Peter Trudgill's findings in England revealed the same results, which he then linked to gender-specific ideals. While the standard language is associated with the female culture, varieties of lower classes are linked to the male population (Angle & Hesse-Biber 1981: 452; Trudgill 1974: 87). Trudgill (1974: 87) relates this motive to the "status-consciousness' that females are associated with. This also complies with the following outcomes: women assume to include more standard forms in their speech habits than they actually do, whereas the opposite pattern occurs for the males (Angle & Hesse-Biber 1981: 452).

As Fuertes et al. (2002: 349) claim, gender differences in attitudes towards varieties became evident through a study which analyzed perceptions of Australians towards English speakers with an Australian and a Greek accent. Results showed that Anglo Australian females reacted slightly more favorable towards the Australian than the Greek accent. A different pattern was recognizable for male participants, who significantly preferred the Australian variety. Women with a Greek accent considered the Greek examples as less suitable for formal settings. The cause lies in the female awareness of accent being a marker of status, increasing the willingness to belong to the associated upper class (Fuertes et al. 2002: 349).

4.5 Accent Prestige

To put it in the authors' words "accents are powerful interpersonal markers that influence evaluations of speakers" (Fuertes et al. 2002: 347). According to Giles (1970: 212) listeners base their judgements of accents on the following three principles: "aesthetic, communicative

and status". The first term centers around the degree, to which a speaker is perceived as pleasant or unpleasant. Communicative aspects evaluate the clarity while status reflects the assumed prestige of a particular accent (Giles 1970: 212). In addition, Kraut and Wulff (2013: 250) emphasize other influential elements, such as intelligibility, sex of the speaker, personal familiarity and attitudes towards a variety; thus, it becomes evident that rating an accent is shaped by several principles, resulting in a highly complex phenomenon (Giles 1970: 211, 212).

Stereotypes based on first impressions of certain varieties easily emerge especially when specific phonological segments are typical for various societal groups (Giles 1970: 211). Wilkinson proposed in 1965 that accent prestige follows a triangular hierarchy (as cited in Giles 1970: 212). Within Great Britain, RP and SSE are associated with the highest class, followed by "British regional accents" and lastly, "town and industrial accents" (Giles 1970: 212). As presented by Giles (1970: 223, 225), RP is generally associated with the highest degree of prestige and professionalism. Even though GA is regarded as less prestigious in this context, it can still achieve a status level. To get better insights into language attitudes, linguists frequently employ rating scales, ranging from high to low status (Giles 1970: 223); however, these findings are based on mean scores but miss to consider influential factors, such as age or social class and therefore, easily lead to generalization (Giles 1970: 224).

The accent prestige theory (APT) centers around two motives: status and solidarity (Fuertes et al. 2002: 347). The former focuses on perceptions of a speaker's education, intelligence or social class, whereas the latter refers to trustworthiness or politeness (Fuertes et al. 2002: 347). Even though a lot of research in this domain originated in Great Britain, studies have also been conducted in other countries, including the United States. Results from North America proved that standard-accent speakers were ranked highest in terms of status by most participants, aligning with findings from Great Britain; still, the solidarity dimension appeared to be more complex. While standard language users did not differentiate between standard and non-standard speakers, non-standard speakers reacted more favorably towards those with a non-standard accent (Fuertes et al. 2002: 348); thus, the feeling of belonging to a social community is a dominant factor. In particular, positive attitudes are more likely to appear for shared values and beliefs (Fuertes et al. 2002: 348).

4.6 Contact Hypothesis

Globalization has considerably influenced various domains of life, including the field of language attitudes, by increasing contact with foreign countries (Dörnyei & Csizer 2005: 327). As presented by Dörnyei and Csizer (2005: 328), the *Contact Hypothesis* influences perceptions

towards specific languages. It suggests that such contact shapes a person's attitude, not necessarily in a favorable way. For positive attitudes to develop, the following criteria need to be fulfilled: "equal group status, common goals, intergroup cooperation, authority support and friendship potential" (Dörnyei & Csizer 2005: 329). Researcher put particular emphasis on the last category, since settings that lack personal matters and therefore often not of great importance, are unlikely to positively influence attitudes (Dörnyei & Csizer 2005: 329).

Tourism is a crucial component of intercultural communication; however, with regards to the *Contact Hypothesis*, the chance of favorable influences may be limited. Conversations between tourists and locals are often characterized by triviality and consequently do not reach a personal level. As presented in the hypothesis, equal group status is essential but this is also frequently not achieved due to travelers' potential higher economic class (Dörnyei & Csizer 2005: 329, 330). Based on these arguments, such types of cross-cultural interactions do not automatically need to lead to positive foreign language attitudes (Dörnyei & Csizer 2005: 330); nonetheless, other factors can positively influence these contacts, particularly "self-interest, tourist densities and the socioeconomic status of the tourists" in comparison to the locals (Ward et al. 2001: 141).

5 Language Attitudes Research

The following section provides the most important information on current attitude research, including common approaches.

5.1 Accents recognition

A wide range of contemporary studies examined social and linguistic judgements based on accent recognition. This process can be particularly challenging for non-native speakers which therefore often leads to misinterpretations. To avoid potential confusion, the recognition of "segmental accent features", such as rhoticity in GA is crucial (Carrie & McKenzie 2018: 313, 314). Previous research revealed that RP and GA were easily recognized by German native speakers. This is due to the educational context, where these two varieties are most commonly taught (Carrie & McKenzie 2018: 314); hence, it becomes evident that prior language experiences as well as physical propinquity positively influence the identification process. Besides geographical influences, media, such as social networks play an essential role (Yook & Lindemann 2013: 290).

Most studies employ a forced-choice recognition task, in which participants are required to identify a specific accent out of a fixed selection; however, many benefits have appeared for more open formats, particularly the avoidance of random guessing (Carrie & McKenzie 2018: 314). Generally, most of the research is based on the matched-guise technique (MGT), which will be described in further course of this thesis; notably, these studies mostly do not evaluate whether the accent was correctly labeled. Instead, they observe respondents' individual answers and the correlating judgements (Yook & Lindemann 2013: 279).

Non-native English speakers tend to associate specific accents with RP or GA and overlook other common varieties, such as Australian in their answers (Yook & Lindemann 2013: 281). Moreover, it is noteworthy that native-like sounding makes a huge difference. Studies revealed that participants ranked Austrian English speakers, whom they identified as Austrians and consequently, as L2 speakers of English, lower than those, being considered as L1 users, even though they both stem from Austria. Similar findings were observable in other countries, such as Japan (Yook & Lindemann 2013: 282). Yook and Lindemann (2013: 282) explain that in these specific types of studies, respondents usually base their evaluations on voices. Doubtlessly, visual appearances influence language attitudes as well as the identification progress and might therefore lead to different results (Yook & Lindemann 2013: 282).

5.2 Approaches for studying language attitudes

As Baker (1992: 18) states, attitude measurement does not always reflect accurate insights due to the following reasons: first, people sometimes tend to answer questions, aligning with socially favorable norms, known as the halo effect (Baker 1992: 19); however, this does not necessarily happen consciously. Secondly, characteristics of the researcher such as age, gender or perceived prestige, as well as the overall research design or setting can significantly impact participants' attitudes (Baker 1992: 19).

Researchers present three main approaches to study language attitudes, namely the observation of the societal treatment of a language variety as well as direct and indirect (matched-guise technique) measures. Every method has its advantages and disadvantages; hence, the most suitable one needs to be chosen, depending on the purpose or the context of the study (Garrett et al. 2003: 15).

5.2.1 Societal treatment approach

This method focuses on how certain varieties and their users are treated within society. Crucial elements are "participant observation and ethnographic studies, or the analysis of a host of sources in the public domain" (Garrett et al. 2003: 15). A core element of the societal treatment approach is that respondents are not directly asked about their opinion (Kircher & Zipp 2022: 14; Ryan et al. 1988: 1068). Most studies following the given approach are analyzed quantitatively, addressing fields of print media, like dialects used in books (Garrett et al. 2003: 15; Kircher & Zipp 2022: 14). Only a minority of researcher apply this method, due to its assumed informal character as it is therefore considered suitable for research in social psychology (Garrett et al. 2003: 16; van Hout & Knops 1988: 7).

5.2.2 Direct approach

In contrast to the societal treatment approach, the direct method is straightforward and based on elicitation. Researchers typically use this method in interviews or surveys, where participants need to answer questions about language preferences, for instance. Through these responses, attitudes towards specific objects become evident. Thus, beliefs are directly provided by the interviewees, instead of being derived by researchers (Garrett et al. 2003: 16). A famous study was conducted by William Labov, in which he asked the population of New York City about two different ways of pronunciation (Garrett et al. 2003: 25). The first question addressed speakers' personal actual use while the second one to the form they thought they should use (Garrett et al. 2003: 25). Since then, the direct approach has been applied in multiple settings, for instance „to predict second-language learning and relative language use, to examine policy

issues such as bilingual education" and to investigate the existence or development of languages in general, like Gaelic (Garrett et al. 2003: 25).

Within the direct approach, researchers distinguish between "word of mouth" or "written response" techniques. The first term refers to interviews or surveys. The difference between these lies in the obligatory presence of the participants. Whereas interviews need to be held face-to-face, surveys can also be administered online or via telephone (Garrett et al. 2003: 26). "Written response" techniques include questionnaires with short or long text answers as well as rating scales. In both designs, difficulties can appear, relating to the following areas: "hypothetical questions, strongly slanted questions, multiple questions, social-desirability bias, acquiescence bias, characteristics of the researcher, the language employed in the process of data collection and group polarization" (Garrett et al. 2003: 27; Kircher & Zipp 2022: 14). Some of these are avoided by phrasing the questions appropriately, while others consider popular participants' trends. In order to counteract this issue, researchers have to adapt their study design (Garrett et al. 2003: 27).

The following section should now give a more detailed insight into the most popular approaches of the "word of mouth" and the "written response" technique, namely interviews and questionnaires.

As Kircher and Zipp (2022: 99) outline, interviews predominantly follow a qualitative research approach since the 1950s. The first linguist, who incorporated them in his research, was William Labov in 1966, where he observed linguistic variables of English speakers in New York (Karatsareas 2022: 99). These were analyzed regarding age, status or gender. Interviews are commonly conducted through a "one-to-one conversation", aiming to evoke information from the interviewee. The literature distinguishes between structured and unstructured formats. The former consists of multiple-choice or yes/no questions. A key feature is that every participant gets the same questions in the same order. The latter contains open-ended formats without given limitations in the depth of the answers (Karatsareas 2022: 99). The third type, semi-structured interviews are a mixture of the forementioned approaches. Researchers focus on including all the necessary topics but without adhering to a fixed order. Open-ended questions, enable participants to share their own beliefs and experiences in greater detail (Karatsareas 2022: 99).

Interviews offer numerous advantages, particularly because of their great variety in research. They can be an effective tool for collecting feedback after conducting a study and provide concrete insights into interviewees' "affect and cognition"; however, they are limited in analyzing real-life habits and behavior; still, interviews offer the possibility to draw conclusions

across various respondents on a more detailed level. Participants can talk freely about their values, beliefs or personal experiences with minimal restrictions (Karatsareas 2022: 101); nevertheless, researchers need to consider some constraints. Karatsareas (2022: 101) explains that interviews cannot be viewed as neutral events, as they are affected by the "spatial and temporal context". In addition, the degree of familiarity between the interviewer and the interviewee might have an impact on the outcomes. Moreover, the *acquiescence bias,* a general trend to avoid disagreeing, or the *social desirability bias*, an inclination to present oneself as positively as possible, play a crucial role and are more present in personal settings (Karatsareas 2022: 102).

As pointed out by Kircher (2022: 129), questionnaires have always been a popular tool in linguistics and are now an indispensable component in the field of language attitudes. They belong to the domain of direct approaches, as questionnaires explicitly ask participants about their beliefs, values or language habits. These can take the form of two types: open-ended or closed questions. The biggest difference lies in the format of the answers. While open-ended questions allow respondents to express their thoughts individually, closed questions present a set of fixed options (Kircher 2022: 129). Compared to interviews, questionnaires enable researchers to collect a large volume of data in a shorter period. In particular, the analysis and conclusion or comparison across respondents is faster, especially for closed questions, thereby making generalizations more feasible (Kircher 2022: 130); nevertheless, some limitations become evident. Firstly, misunderstandings cannot be clarified and therefore these might influence the research outcomes (Kircher 2022: 130). As with interviews, the *acquiescence* and *social desirability* bias count as potential restraints; nonetheless, this can be avoided by ensuring anonymity. Lastly, questionnaires often lack a high degree of individuality and hence frequently fail to observe more in-depth fields (Kircher 2022: 130).

It is important to note that not every method is suitable for every context (Garrett et al. 2003: 31). For instance, questionnaires require a certain level of language competency and thus, particularly children may need to be interviewed; however, this might also cause several difficulties.

5.2.3 Indirect approach

In contrast to the direct approach, the indirect method is commonly known for its subtle character and often used synonymously with the matched-guise technique (MGT), developed by Wallace E. Lambert (Loureiro-Rodriguez & Acar 2022: 185). He argued that participants' answers, collected through directly and precisely asking questions, frequently fail to represent

actual language habits or beliefs (Garrett et al. 2003: 51). Its core element involves the inclusion of audio files, which respondents are required to listen to and afterwards state their opinion regarding various topics, such as personality traits of the given speakers or the presumed intelligence, typically using rating scales (Garrett et al. 2003: 52).

As Loureiro-Rodriguez and Acar (2022: 185) outline, the MGT typically includes multilingual people, reading out a neutral text in a different accent and dialect. As Garrett et al. (2003: 52) state, "the technique attempts to control out all but the manipulated independent variable". In the context of accents, this means that the same person delivers the passage in different varieties without any variation in pace or pitch to solely focus on linguistic features (Garrett et al. 2003: 52). To minimize the risk that respondents recognize the previous speaker, additional audio files, known as distractors, are included (Loureiro-Rodriguez & Acar 2022: 186). After each recording, the audio file pauses. This, on the one hand, serves the purpose to let participants finish the given questions and on the other hand, also reduces the chance of speaker recognition (Garrett et al. 2003: 52).

A key element of questionnaires used with the MGT is the evaluation of various personality traits, which participants are asked to rank regarding the concerned speaker. Typical fields are "prestige, social attractiveness and dynamism" (Garrett et al. 2003: 53). In addition, researchers frequently include a prejudice scale to gain deeper insights into potential positive or negative attitudes, followed by questions about respondents' personal preferences (Loureiro-Rodriguez & Acar 2022: 186). An advantage of this technique lies in the high degree of honesty, as participants are not directly asked by interviewers, but indirectly and privately; thus, interviewees may not necessarily feel the urge to conform with socially accepted values (Loureiro-Rodriguez & Acar 2022: 188). As the method has been employed in multiple areas, researchers have the possibility to analyze similarities or differences across studies. Garrett et al. (2003: 57) explain that the MGT laid the foundation for further work in several research areas, including "social psychology of language and sociolinguistics" (Garrett et al. 2003: 57).

Despite all the positive effects that the approach has, limitations still need to be acknowledged. First, Garrett et al. (2003: 58) outline that participants are provided with a large number of recordings. This may lead to an augmentation of the language effects, that might not appear in real-life contexts. Moreover, respondents may misinterpret different accents and varieties and therefore associate them with incorrect grammar (Garrett et al. 2003: 58). Another frequently raised criticism concerns the reading passages. Since the speakers prepared the text in advance, spontaneous elements are eliminated compared to natural conversations (Loureiro-Rodriguez

& Acar 2022: 188). A further challenge lies in selecting a person, capable of presenting all required accents; hence artificial intelligence (AI) has been used in this thesis to overcome this obstacle. As previously mentioned, the content of the reading passages must remain neutral to avoid emotional reactions, which could influence the perception of the given language variety (Loureiro-Rodriguez & Acar 2022: 189); nevertheless, it is noteworthy that different interpretations can still occur due to various demographic factors and therefore affect the judgements of the speakers. As the authors (Loureiro-Rodriguez & Acar 2022: 189) outline, the most appropriate method for maximizing validity is a mixture between MGT and direct approaches, like questionnaires.

6 Research Design

The following chapter gives detailed information about the research design as well as its data collection process.

6.1 Aims and hypotheses

The study aims to analyze different attitudes towards three English varieties in five recordings, more precisely RP, GA und Scottish; however, the Scottish example should only serve as a distractor and is therefore, not observed in greater detail. Two examples are repeated but in a different mood. Based on the given goal and theoretical underpinnings, the following research questions emerged.

I. To which extent does participants' own spoken English variety influence their judgements of the provided examples?

II. Do participants realize that two varieties are spoken by the same speaker, differing in mood (conversational versus sad) and how does emotional tone affect the perception of the language variety?

III. Are participants able to detect that the recordings are AI-generated and therefore describe the speakers' way of talking accordingly?

IV. Which sociolinguistic variables influence participants' evaluations?

V. What factors influence respondents' rank ordering and recognition rates of the given varieties?

Derived from the forementioned questions, the subsequent hypotheses were developed.

H_1-I: Participants base their judgements of the given character traits on their self-identified English.

H_1-II: Respondents evaluate the speakers as separate examples regardless of the mood change.

H_1-III: In comparison to a happy mood, sadness is more negatively evaluated in all dimensions.

H_1-IV: Due to the current progress in AI, participants do not assume that the recordings are computer-generated.

H_1-V: Former residences in Anglophone areas evoke certain emotions and therefore have an impact on the character trait level.

H_1-VI: Participants' age, mother tongue and self-identified English influence the rank ordering results of the character traits.

6.2 Research methodology

As discussed in chapter 5.2, various approaches exist for investigating language attitudes; the matched-guise technique (MGT) is among the most popular ones and has laid the foundation for numerous research projects (Garrett et al. 2003: 57). Due to its many advantages, the technique was also selected for the purpose of this thesis. Whereas other methods directly elicit participants' opinion, the MGT provides an indirect manner and therefore, can increase honesty (Loureiro-Rodriguez & Acar 2022: 188).

In the MGT, people are required to listen to short recordings and evaluate various personality traits or the linguistic competence of the speakers, typically based on first impressions (Garrett et al. 2003: 52). As previously mentioned, only one person reads out the passages to reduce potential effects of voice or age (Garrett et al. 2003: 52). If any discrepancies occur across the various samples, it can be concluded that these stem from the language variety itself rather than speaker-related factors (Soukup 2009: 86). It needs to be stressed that this project did not use one single voice to record all examples. Instead, two people appeared twice, resulting in a total of three female speakers for five recordings. Another essential component of the MGT is the inclusion of a distractor, which was implemented in this study through the third example (Loureiro-Rodriguez & Acar 2022: 186).

Participants were not informed about the specific research design and therefore believed that the study included five different people. Moreover, it was not disclosed that the recordings were AI-generated. As stated by Soukup (2009: 87), finding volunteers, capable of speaking all varieties needed, might seem nearly impossible; thus, artificial intelligence proved to be a practical alternative. Furthermore, respondents should be distracted from the fact that they are evaluating different English varieties. To achieve this, the description of the survey used the context of reporting the news and participants should evaluate the voices accordingly.

6.3 Artificial Intelligence

Artificial Intelligence (AI) is a common tool in modern life and understood as "any computational system producing intelligence behavior" (Rochel 2022: 39). Its primary goal is to create an end product that simulates human behavior (Chan & Hu 2023: 2). Within the frame of this thesis, an AI website, namely Murf AI, was used to generate audio files with several English varieties and different emotions. If researchers plan to integrate AI in their projects,

multiple aspects require careful consideration, especially ethical regulations of AI guidelines focusing on explainability and transparency (Balasubramaniam et al. 2023: 1; Rochel 2022: 38).

Within the scope of this research project, text-to-speech (TTS) technologies were employed, which have gained great interest, especially since the rapid improvement in AI technologies. Generally, these programs convert written texts into authentic audio files and therefore imitate a real-life setting (Amin 2024: 888; Khanam et al. 2022: 398). It is noteworthy that pitch, pace and even other linguistic features, like pronunciation or even specific language varieties can be easily adapted (Fitria 2023: 3). A TTS software was originally developed to assist people with disabilities; nevertheless, its use has increased over the last years and is now implemented in multiple fields of everyday lives. There are no limitations and all types of texts, such as books, articles or websites can be converted. Depending on the platform used, the generated audios can sound very natural and realistic (Fitria 2023: 3).

Due the improvement and development of various AI platforms, these can benefit academia in many ways (Ali 2020: 1). A key advantage is the increase of "productivity and efficiency", thereby fastening the research process (Dugošija 2024: 277). This has also become evident through the design of this thesis, as it proved to be less time consuming to use the specific website for imitating human voices than working with real people. Besides that, flexibility is enhanced. If troubles with a particular accent had arisen in this specific research design, a more suitable one could have easily been selected (Köchling & Wehner 2023: 46). There are no local constraints as websites allow quick access (Dugošija 2024: 275). Moreover, depending on the website, AI produces authentic samples, frequently indistinguishable from real human designs (Chan & Hu 2023: 2).

Even though AI has positively influenced linguistic research, its challenges and limitations require careful consideration. Especially in the field of language attitudes, various platforms could be responsible for minimizing real-life communication with native speakers in the research progress.; hence, it might lead to an overdependence on technology (Dugošija 2024: 278).

6.4 Questionnaire Design

The study follows a quantitative approach, using a questionnaire created with Jotform. Its design is based on former research conducted in the domain of language attitudes (Soukup 2009: 86). As Wilson and Dewaele (2010: 103, 108) state, online surveys address a great

number of people and therefore contribute to a high degree of validity. It starts by collecting sociodemographic data, like age, gender, country of birth and mother tongue. Besides that, the English proficiency level, field of study or work, former travel experiences to Anglophone countries as well as the individual spoken variety of the respondents are gathered. Subsequently, the five audio recordings with the corresponding questions appeared. At the beginning of each set, a four-point Likert scale, ranging from "very" to "not", was provided to observe participants' evaluations in terms of eleven adjectives regarding speakers' accents, adapted from former research (Soukup 2009: 86). Zahn and Hopper (1985) (as cited in Soukup 2009: 107) grouped the concerned adjectives into following categories: 'superiority' (education or intelligence), 'attractiveness' (friendliness, honesty) and 'dynamism' (laziness, enthusiasm); however, it needs to be considered that these are not 'universal', as their importance may vary depending on various circumstances and context (Soukup 2009: 107). Solely positive character traits, except *snobbish*, were presented on the left side of the item to ensure consistency (Soukup 2009: 108). An even range of numbers was chosen to avoid central tendency, especially in the context of a larger set of elements to rate.

Similar principles were applied to the following two items, which asked respondents for their perception of the perceived clarity of the speakers' pronunciation and the suitability for presenting the news. Unlike question one, items two and three employed a five-point rating scale (1= "unclear" 5= "very clear"). While the forementioned elements worked with fixed numbers, the following section asked participants to describe their perception of the presentation as a news reporter in one word to provide variety within the survey and avoid fatigue effects. No constraints were stated to collect the most personal insights into respondents' attitudes.

Lastly, people had to decide, based on a yes/no question, whether they would like to meet the respective speaker for a coffee to observe patterns in solidarity (Giles & Watson 2013: 28). After completing the question set for all five recordings, respondents should rank the accents according to their preferences from 1 (least favorite) to 5 stars (most favorite). Next, participants had to describe their assumptions about the speakers' country of origin. No specific guidelines were included regarding the explicitness of their answers. An open format was employed to avoid random guesses (Carrie & McKenzie 2018: 314).

Respondents were asked to listen to each recording once while completing the questions; however, the survey was not separated into different parts, allowing participants to go back to the audio files for the final part of the questionnaire.

The survey started with the GA sample, followed by the RP tape. Both speakers read the text passage in a conversational style. Next, the SSE voice occurred, once again delivering a happy feeling, which was solely included for the purpose of a distractor. Afterwards, a shift in mood appeared. The sad emotion was then introduced with the same GA variety as in the beginning to provide sufficient time between its first appearance to minimize the chance of speaker recognition. Lastly, respondents encountered the RP voice in a melancholic manner.

6.5 Data collection

The data collection process started in September 2024 and continued until January 2025. Before that, a pilot study was conducted to detect potential limitations of the research design. For this purpose, the questionnaire was sent to ten students from the Department of English and American Studies but no changes were made after this step.

In total, 180 completed samples could be collected. The study did not aim at a specific target group, but everyone with an intermediate level of English was invited to participate. The survey was for instance distributed across various schools and institutions in Austria, such as the University of Vienna.

As stated at the beginning of the survey, participation was entirely voluntary, and people could withdraw without any negative consequences. All answers were anonymous and treated confidentially; thus, respondents' privacy was ensured throughout the whole study. The gathered data was only used for the purpose of this thesis.

6.6 Speakers

The MGT is among the most frequently used approaches for exploring language attitudes, primarily due to its research design, focusing solely on linguistic features and thereby avoiding personal influences, such as pitch or gender (Yook & Lindemann 2013: 279; Garrett et al. 2003: 52); still, according to Soukup (2009: 87), a possible limitation of the technique is finding someone, who is capable of speaking all required language varieties naturally, as the artificial sounding of a speaker's accent might distort the outcomes (Soukup 2009: 87). The use of AI aims to decrease this possibility. Similar to other studies, it proved to be impossible to find samples authentically conveying all three varieties, even with the AI assistance; therefore, three different voices were chosen, all represented by young females. Since they are computer-generated demographic information is limited to sex and an approximate age group.

The characteristics of the respective varieties have already been described in chapter 3; therefore, this section aims to relate these findings to the specific audio recordings, included in this project.

6.6.1 GA recording

The recording represents several crucial features of the GA variety. First, rhoticity is strongly present, occurring in both pre- or postvocalic positions. This becomes evident through the given audio file in words such as *traveler (*/ˈtrævələr/), *desire* (/dəˈzaɪər/), *nature* (/ˈneɪʧər/) or *arts* (/ɑrts/) (Kovecses 2000: 25; Kuecker et al. 2015: 623). Furthermore, t-flapping is a common characteristic of the American English variety, which appears intervocalically or between /r/ and a following vowel (Hannisdal 2020: 258). In this recording, the speaker articulates the sound in *motivated* (/ˈmoʊtəˌveɪtəd/) or *communities* (/kəˈmjunətiz/), for instance. GA tends to pronounce the short British /ɒ/ vowel longer, turning into the long /ɑ/. A concrete example is the word *possible* (/ˈpɑsəbəl/) (Kovecses 2000: 26). In contrast to RP, GA speakers voice fricatives at the end of words as in *attractions* (/əˈtrækʃənz/) (Hughes et al. 2005: 44). Additionally, the /æ/ sound commonly appears in GA where other varieties would use the long /ɑː/ (Kovecses 2000: 25). This phenomenon can be detected in *traveler* (/ˈtrævələr/). Notably, the /æ/ sound is similar for all varieties; however, the pronunciation slightly differs. Speakers of GA do not round their lips but slightly open the mouth, which also appears in the recording of speaker 1 (Kovecses 2000: 25). Moreover, the diphthong /oʊ/ is more common in GA than RP variant /əʊ/; thus, it starts with an /o/ and glides to /ʊ/. This sound is observed in *motivated* (/ˈmoʊtəˌveɪtəd/) or *location* (/loʊˈkeɪʃən/).

6.6.2 RP recording

A common descriptor for RP is non-rhotic, meaning that speakers do not pronounce the /r/ sound at the end of a syllable and in post-vocalic positions before a consonant; thus, an active realization of /r/ solely appears at the beginning of a word or when it precedes vowels (Kaminska 1995: 98). This pattern also becomes evident in the given RP recording. The omission of the post-vocalic /r/ occurs in words, such as *learn* (lɜːn), *concerned* (kənˈsɜːnd) or *forms* (fɔːmz/). The so-called linking /r/ is a frequent phenomenon in RP, which describes the active pronunciation of /r/ between two words, appearing between two vowels (Brown 1988: 144). In the RP audio file, this can be heard in the phrase *traveler is* (/ˈtrævələrɪz/). In the coda of a syllable, the /r/ is not actively realized, as in *desire* (/dɪˈzaɪə/) or *nature* (ˈneɪʧə). In addition, the chosen voice differentiates between dark /ɬ/ and light /l/. The first is found in words such as *local* (/ˈləʊkəl/) or *language* (/ˈlæŋgwɪʤ/), whereas the latter occurs in *possible* (/pɒsəbɬ/). H-dropping is not prominent in the RP recording, since attentive speakers aim to avoid it (Hughes

et al. 2005: 44). Instead, full realizations of /h/ occur in *history* (/ˈhɪstəri/) or *heritage* (/ˈhɛrɪtɪʤ/). Unlike GA, RP tends to use the short /ɒ/ vowel in *possible* (/ˈpɒsɪbəl/), for instance (Kovecses 2000: 26).

6.6.3 SSE recording

A common linguistic feature of SSE is its rhoticity, which the concerned speaker in the survey realizes through words, such as *desire* (/dɪˈzaɪɹ/), *nature* (/ˈneːtʃəɹ/) or *forms* (/fɔɹmz/). As stated by Kaminska (1995: 43), in contrast to RP, SSE does not distinguish between /o/ and /ɔː/. Instead, they use /ɔ/, which is evident in the pronunciation of *forms* (/fɔɹmz/) in the given recording. Moreover, Scottish Standard English combines the vowels /æ/ and /ɑː/, resulting in a realization of /a/ for all cases, which becomes evident through *traveler* (/ˈtɹavələɹ/) (Rogers 2014: 114). Rogers (2014: 114) elaborates on the realization of the vowels /ʊ/ and /u/, that are frequently merged to /ʉ/, a high central vowel. This characteristic of SSE can be found in *too* /tʉː/.

6.7 Reading passage

The chosen text focuses on the topic of cultural tourism. To align with the aims and context of this study, several criteria need to be fulfilled. First, the reading passage should not evoke any emotions while listening and must avoid political and controversial issues, as the content should not distract participants from the speakers' linguistic competence. Moreover, it had to be universally understandable and not tied to a specific cultural context. Additionally, the use of language was formal but accessible to ensure proper understanding for intermediate English proficiency levels. Besides that, simple and easy to follow sentences are preferred over a complex syntactic structure. Since respondents believed that they assess the recordings for their suitability as a news reporter, the text also had to be appropriate for this specific genre. Lastly, the length of the audio file was limited to around one minute, depending on the language variety and the specific mood (conversational versus sad). The maximum duration was 1:03 minutes (speaker E), whereas the shortest file lasted 46 seconds (speaker A). This variation of 17 seconds is due to the slower pace of the melancholy RP speaker in comparison to the happy GA file.

The text was directly taken from the Revfine website, a platform elaborating on topics around travelling and tourist industry. The specific subject of the respective section refers to different sectors of tourism. No adaptions were made since the passage has already seemed appropriate for the purpose of this research design.

Cultural Tourism

Cultural tourism is a type of tourism where the traveler is motivated by a desire to learn about and engage with the history, traditions, practices and cultural attractions of a location. It is often concerned with experiencing other ways of life and engaging with people, but cultural tourism can also help to fund conservation efforts too. By its very nature, cultural tourism can take many different forms, with relevant activities including visiting monuments, museums and religious sites, attending festivals, engaging with the arts or learning the local language. It can also provide communities with a means of monetizing their cultural heritage, although efforts do need to be taken to limit some of the possible negative side effects of a booming tourism industry.

(https://www.revfine.com/tourism-industry/#cultural-tourism , Barten n.d.)

6.8 Participants

Online surveys address a great variety of people regardless of their geographical location (Wilson & Dewaele 2010: 103, 108). This study did not focus on a specific target group but aimed to collect as many different samples as possible to observe potential patterns related to ages, gender, English proficiency level or former residences in Anglophone countries; thus, the questionnaire was shared on social media sites and sent to friends, colleagues and different departments at the University of Vienna.

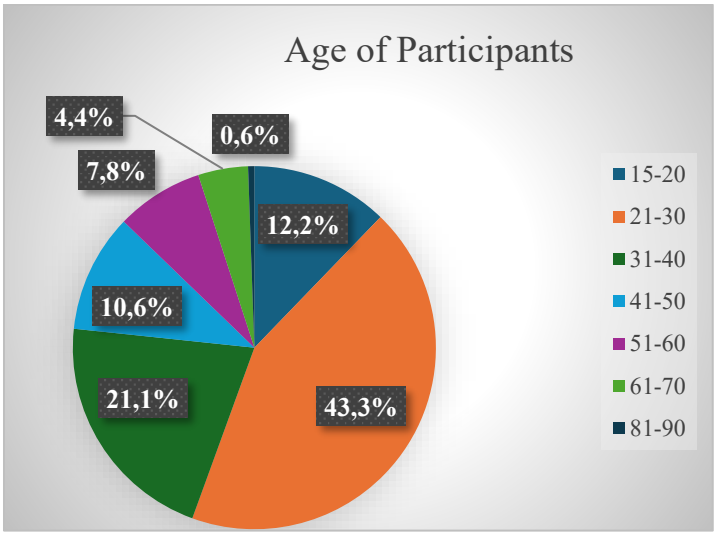


Figure 5. Age

The age of the participants ranged from 15 to 89 years (see Figure 5). The largest group consisted of respondents between 20 and 30 years. This might be due to the frequent distribution among university students. Followed by that, 21,1% were aged 31 to 40. One completed survey was submitted by a person of 89 years.

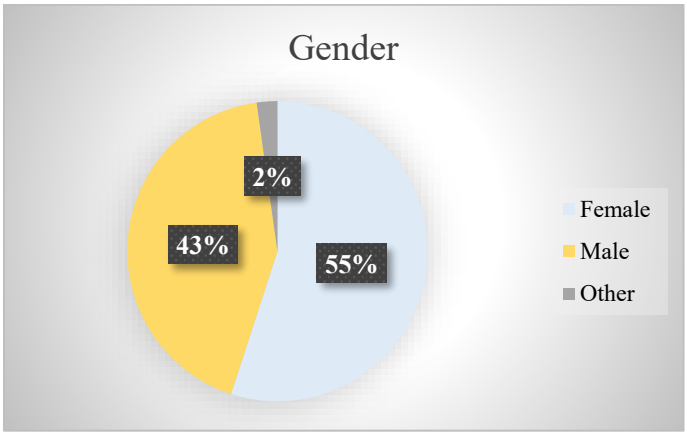


Figure 6. Gender

The majority of respondents is female (n=99), which makes around 55% in total. 77 participants reported to be male (43%), while 4 people identified as none of the two (2%) (see Figure 6).

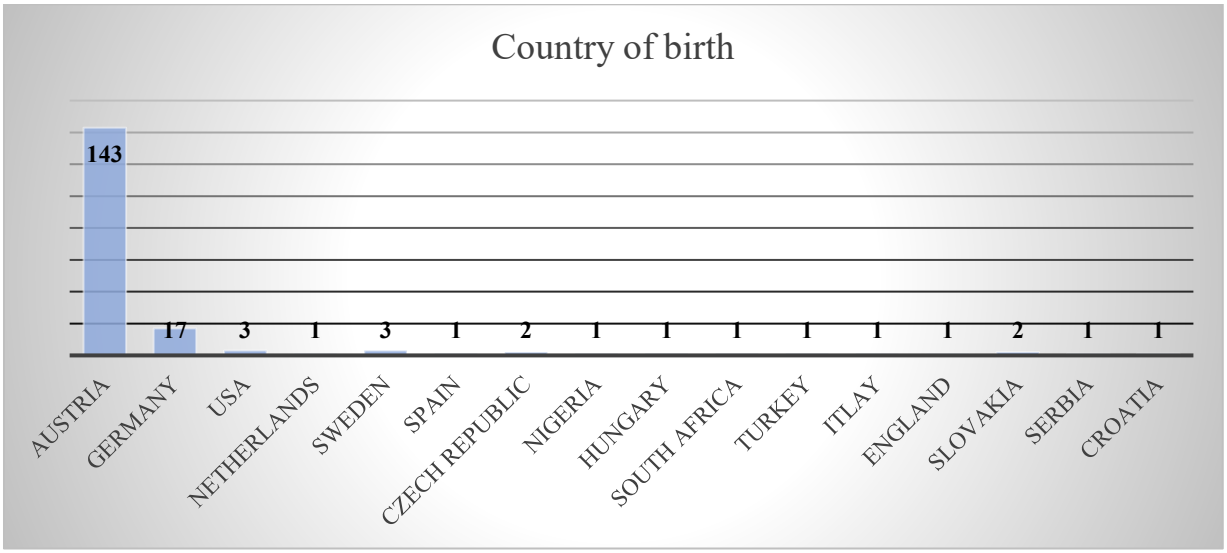


Figure 7. Country of birth

Regarding participants' country of origin, a total of 16 different places were reported. As can be taken from Figure 7 above, the majority was born in Austria (n=143), amounting to 79,4%. Seventeen respondents mentioned Germany as their country of birth. Apart from that, the United States and Sweden appeared three times, while Czech Republic as well as Slovakia were

each named twice. The remaining countries, Netherlands, Spain, Nigeria, Hungary, South Africa, Turkey, Italy, England, Serbia and Croatia occurred once.

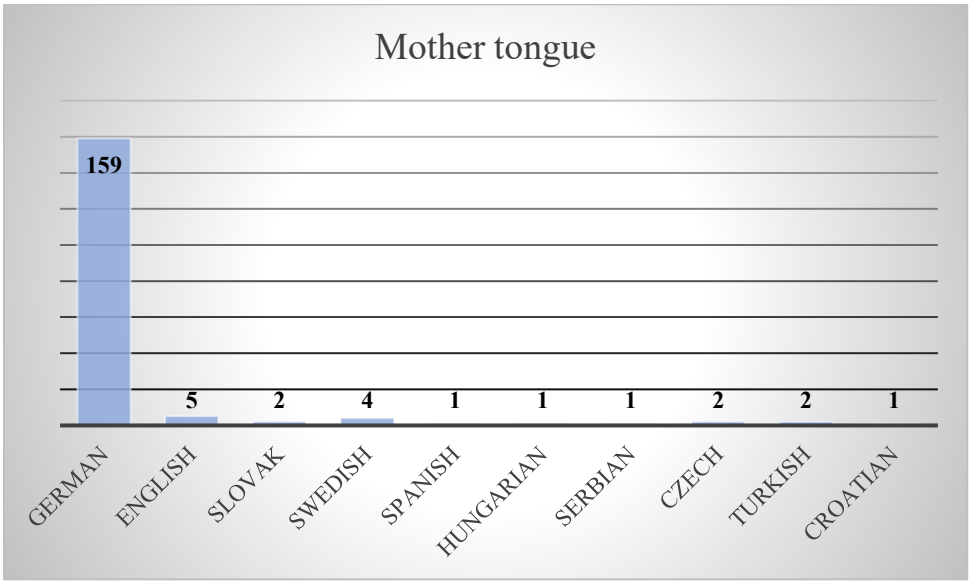


Figure 8. Mother tongue

Item 4 of the questionnaire revealed that respondents reported ten different languages. Overall, the distribution of participants' mother tongue presented a similar pattern as the results for country of birth. As visible in Figure 8, the majority, particularly 88,3% (n=159) stated German as their first language, significantly outnumbering the five English native speakers. Four people reported Swedish as their first language, two Slovak, one Spanish, one Hungarian, one Serbian, two Czech, two Turkish and one Croatian.

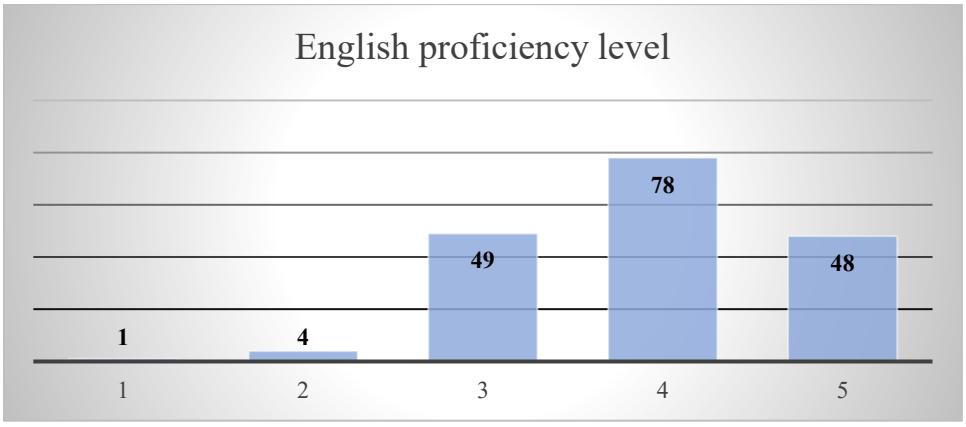


Figure 9. English proficiency level

Since the survey was distributed within the Department of English and American Studies and completed by numerous English teachers, it becomes evident that 70% (n=126) of the

participants described their English proficiency level as either good (=4) or very good (=5) (see Figure 9). Only one person reports to have insufficient English language skills. In contrast, 27,2% (n=49) classified their knowledge as average.

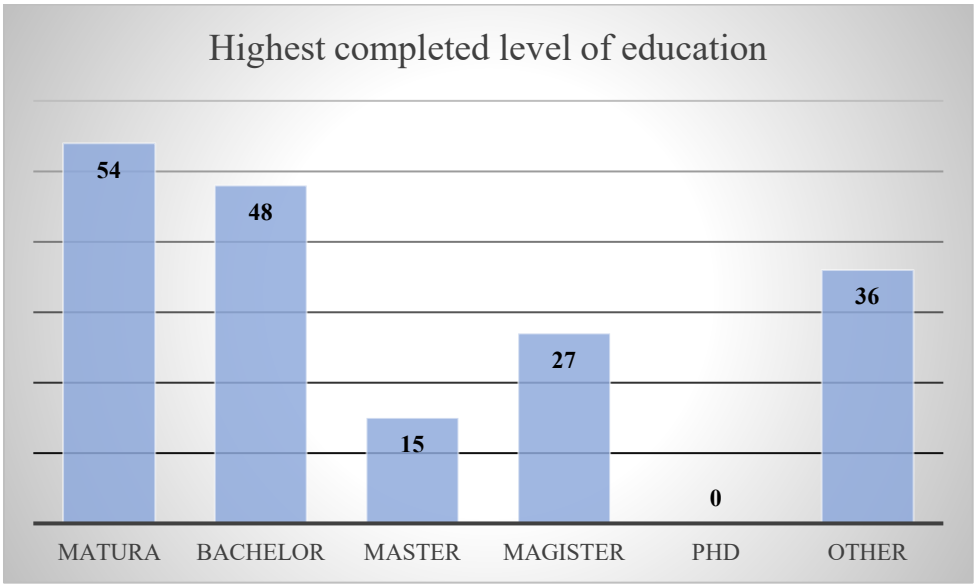


Figure 10. Highest completed level of education

Regarding respondents' educational background, the following results can be obtained from Figure 10: first, most participants (n=54) reported the Matura, the Austrian exam for leaving secondary schools, as their highest completed level of education. Closely followed by that, 26,7% (n=48) successfully completed a bachelor's degree. Next, 15 respondents finished studies in a Master program and 27 held a Magister degree. Thirty-six participants selected the *Other* option.

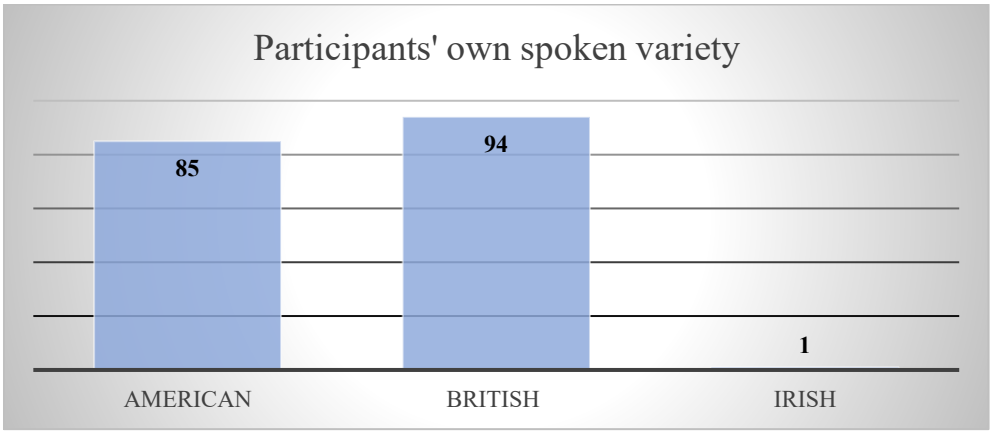


Figure 11. Participants' own spoken variety

Doubtlessly, RP and GA are among the most widely spoken varieties, which is also mirrored through this survey and presented in Figure 11 (Kovecses 2000: 8). Ninety-four participants

outlined standard RP as their own spoken variety of English; nevertheless, the distribution between these two was relatively balanced as 85 respondents claimed to primarily follow GA principles; therefore, no substantial discrepancies could be observed in this pattern, although one statistical outlier appeared by selecting Irish.

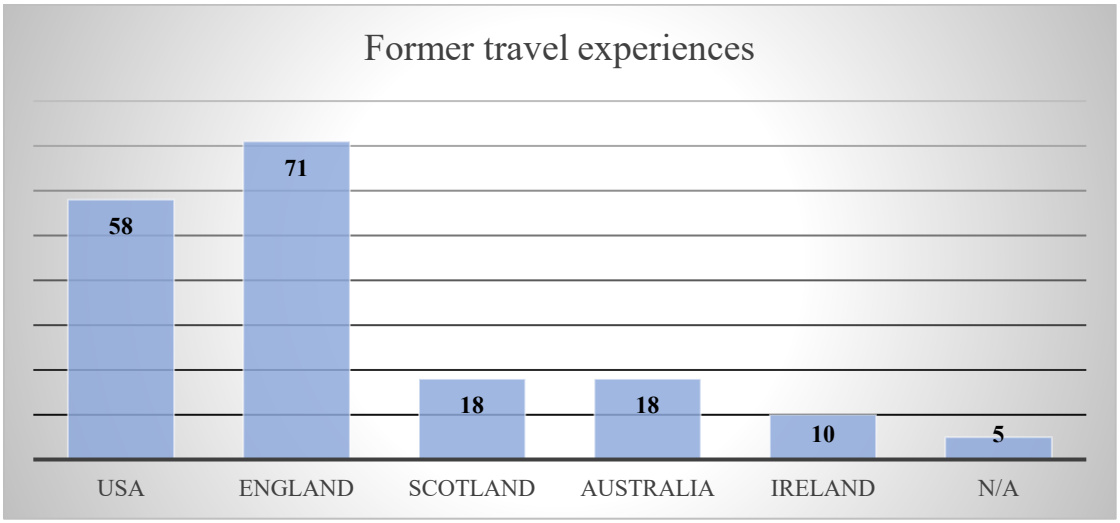


Figure 12. Former travel experiences

Corresponding with participants' self-identified English, the USA and England were among the most popular travel destinations (see Figure 12). Five respondents did not select any of the provided options; hence, it can be assumed that they have not visited any of these countries, reducing the sample size from 180 to 175 answers. Seventy-one people reported England, adding up to a total of 40,6%, followed by 58 votes for the USA, amounting to 33,1%. An equal number of participants ranked either Australia or Scotland as their favorite destination. Lastly, Ireland appeared ten times.

7 Results

This section describes the analysis of the data collection in accordance with the hypotheses and research questions stated in 6.1.

7.1 Analysis of mean scores regarding character traits

To get an initial overview of the entire data set, average scores for all character traits, consisting of one negatively and ten positively connoted adjectives are provided. To convert the given four-point Likert scale into numerical data, values ranging from 0 (=*not*) to a maximum of 3 points (=*very*) were assigned; thus, 0 points symbolize that participants do not associate the characteristic with the concerned speaker, while 3 points indicate full agreement with the adjectives.

As presented in Table 1, significant differences were observed between the conversational and the sad recordings, indicating that mood served as an influential factor. To obtain data, mean scores of each character traits were calculated by adding up the received evaluations and subsequently dividing them by the number of respondents.

Table 1. Mean scores of character traits for all recordings

Variable	GA happy	RP happy	SSE happy	GA sad	RP sad
trustworthy	2.61	2.64	1.98	1.05	0.65
friendly	2.63	2.63	2.06	1.02	0.72
polite	2.70	2.60	1.94	1.08	0.65
intelligent	2.56	2.84	2.34	1.04	1.02
hardworking	2.49	2.61	2.20	1.01	0.83
outgoing	2.55	2.62	2.18	0.78	0.47
snobbish	0.66	1.58	1.82	0.66	1.02
reliable	2.62	2.66	2.17	1.02	0.77
wealthy	2.40	2.70	2.11	0.93	0.93
funny	2.35	2.21	1.71	0.57	0.37
confident	2.70	2.73	2.11	0.81	0.50

Overall, the happy RP recording achieved the highest mean score in seven out of eleven categories, namely *trustworthy* (M=2.64), *intelligent* (M=2.84), *hardworking* (M=2.61), *outgoing* (M=2.62), *reliable* (M=2.66), *wealthy* (M=2.70) and *confident* (M=2.73). It received a score over M=2 in ten descriptors, except for the perceived snobbishness, where the variety obtained on average M=1.58 points, closely behind SSE.

GA conversational revealed consistently positive results with values from M=0.66 for *snobbish,* to M=2.70 for *polite* and *confident*. In some categories, such as *friendly, reliable, trustworthy* and *confident*, RP and GA attained a maximum difference of only 0.04. Interestingly, the American variety displayed the only instance where the same score was obtained for a single character trait (*snobbish*) regardless of the change of emotional tone (M=0.66).

Participants perceived SSE as the most snobbish variety among all included samples (M=1.82). Apart from that, Scottish Standard English neither reached the highest nor lowest score in any other character traits. It can be noted that respondents reacted more favorably to status related elements, like *intelligent* (M=2.34) and *hardworking* (M=2.20). In comparison to GA and RP, SSE was rated more negatively in the field of warmth, including characteristics such as *friendly* (M=2.06) and *polite* (M=1.94).

A different trend from the first three recordings to speaker 4 and 5 was detectable. While the conversational audio files hardly reached values below M=2.0, the sad voices rarely exceeded M=1.0. The GA recording, for instance, attained the lowest average score of M=0.57 (*funny*). Moreover, traits related to dynamism were particularly negatively reviewed, including *outgoing* (M=0.78) and *confident* (M=0.81). In contrast, the highest number became evident in the evaluation of *polite* (M=1.08).

Similar patterns appeared in the analysis of the melancholic RP recording; however, it is noteworthy that only two adjectives, namely *intelligent* (M=1.02) and *snobbish* (M=1.02), obtained a score over M=1.0. A resemblance to the conversational RP voice was observed in the relatively high rating of the assumed intelligence of the speaker. The lowest average score (M=0.37) was once again yielded in a dynamic dimension (*funny*), marking the worst rating throughout all speakers. A closer comparison of the sad audio files revealed that participants ranked the GA voice higher or equal in almost all dimensions except *snobbish* (M_{GA}=0.66; M_{RP}=1.02).

As stated in chapter 4, people often base their evaluations of speakers with different language varieties on the following dimensions: “social status, solidarity and dynamism” (Giles &

Watson 2013: 28). This paper employs similar dimensions with slight adaptations, substituting solidarity with warmth.

The following part describes the findings for the individual fields based on the calculation of mean scores. To obtain results, the average values of each category were summed and consequently divided by the total number of adjectives assigned to that group. Table 2 presents the given dimensions with the respective adjectives included in the survey. The adjective *snobbish* was excluded from the table due to its negative connotation in contrast to the other positively connoted adjectives in the field of social status; therefore, its mean scores will not be compared with the other character traits, but a closer analysis of the perceived snobbishness will follow in further course of this paper.

Table 2. Character trait dimensions

Social Status	Warmth	Dynamism
intelligent	trustworthy	outgoing
hardworking	friendly	funny
wealthy	polite	confident
	reliable	

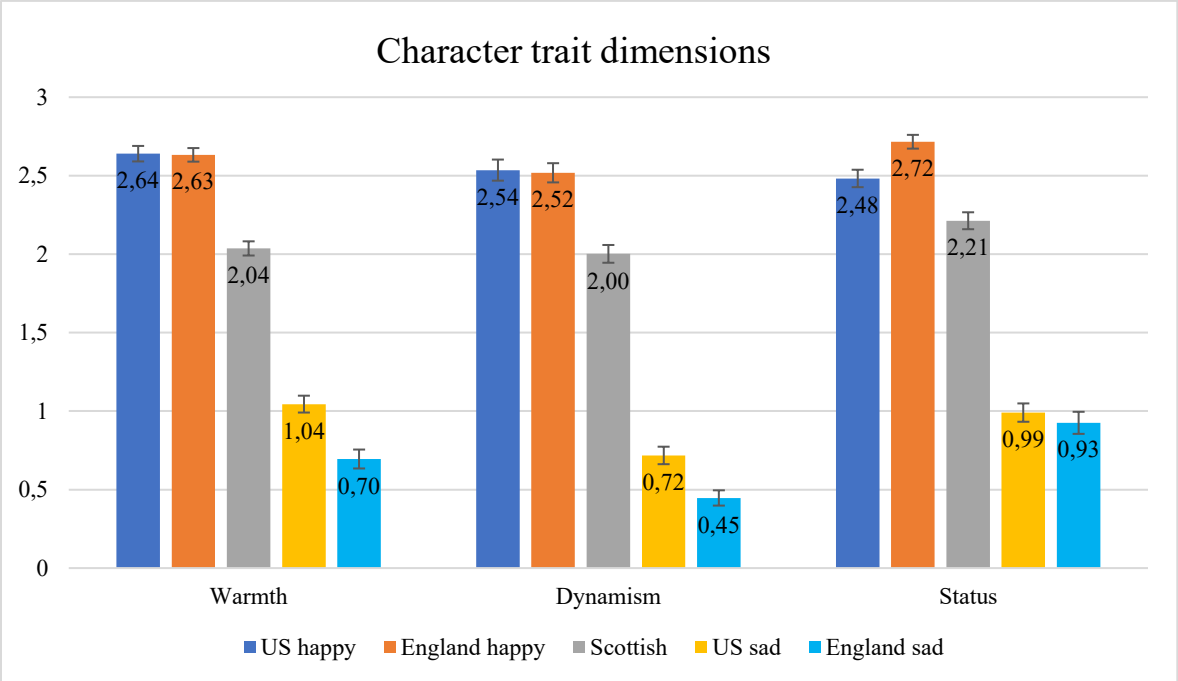


Figure 13. Comparison of character trait dimensions

As illustrated in Figure 13, the conversational GA and RP files did not reveal great variance in the dimensions of warmth and dynamism. Whereas GA reached a score of M=2.64 in the

forementioned category, RP gained M=2.63 points, resulting in a slight difference of 0.01. Another decent discrepancy occurred in the analysis of dynamic adjectives. The English variety scored a mean of M=2.52 and the American equivalent reached M=2.54 points, differing by only 0.02. The overall highest score was recorded by the conversational RP speaker in the field of social status with an average number of M=2.72 out of a maximum of 3.0; thus, the analysis of social status revealed greater discrepancies between the GA (M=2.48) and RP (M=2.72) recording.

The SSE conversational recording was rated lower than the happy GA and RP in all of the three categories. First, warmth with a mean rating of M=2.04 was significantly below the average of the varieties presented before. Secondly, the dynamic aspect of the concerned speaker outlined only marginal discrepancies to the characteristics in the field of solidarity with a mean value of M=2.0. The highest number of points were gained regarding social status (M=2.21).

As can be taken from Figure 13, none of the conversational audio files went below an average of M=2.0. In contrast, an opposite pattern occurred for the sad recordings. The highest score was attained in the field of warmth with a mean value of M=1.04; thus, this reflects the conversational American English recording, where warmth was also ranked highest but with a major difference of 1.60. While the aspect of social status reported the lowest score (M=2.48) of the conversational GA speaker, the sad GA voice attained the second highest number of points out of the three respective categories with a mean of M=0.99. The lowest number of points was recorded for dynamic adjectives (M=0.72) which simultaneously showed the biggest difference within the variety of 1.82 points.

A comparison with the conversationally read RP sample confirms that social status remained the highest ranked category, outlining a mean score of M=0.93; however, in terms of warmth and dynamism major discrepancies emerged between the two RP recordings, effected by a mood change. The first mentioned field obtained a score of M=0.70 and therefore created a discrepancy of 1.93 points, while the latter was ranked lowest out of all dimensions (M=0.45) and hence, showed the largest difference to the conversational RP style with 2.07 points in total.

To sum up, the highest values were achieved by the happy RP voice in the dimension of social status (M=2.72), closely preceding the scores in the field of warmth of the GA speaker (M=2.64).

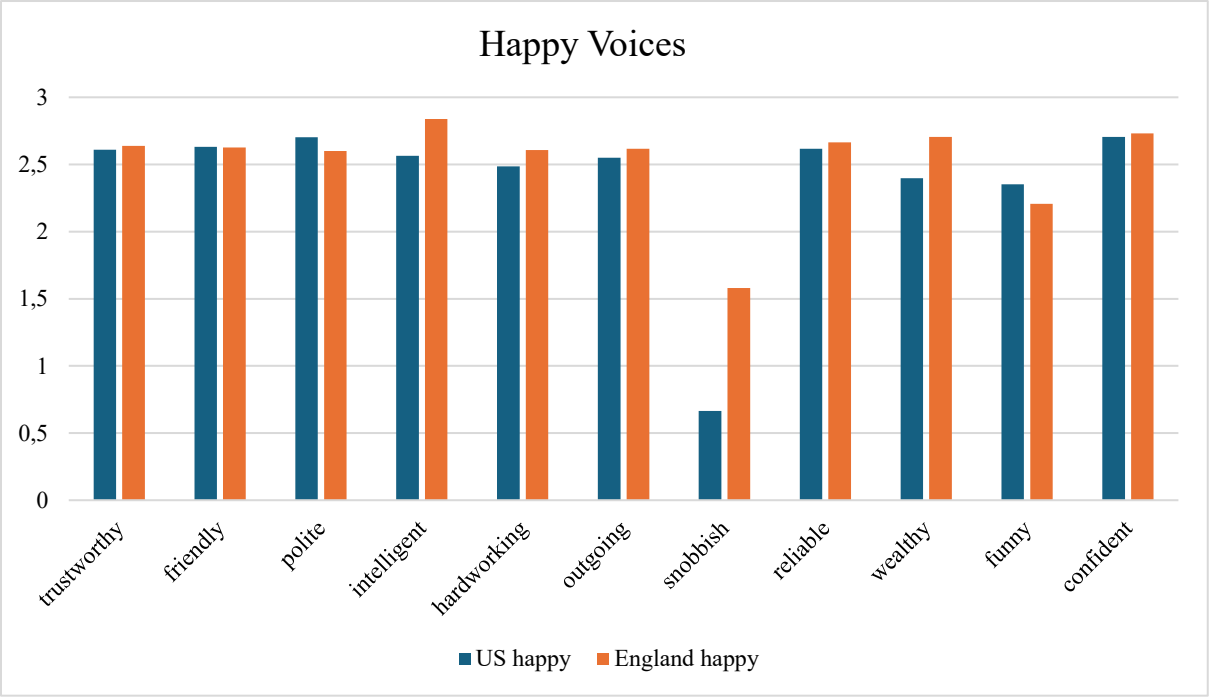


Figure 14. Happy voices all respondents

As visible in Figure 14, participants expressed similar attitudes towards the happy GA and RP variety regarding the given character traits. The most prominent discrepancy appeared in the evaluation of snobbish (M_{GA}=0.66; M_{RP}=1.58), which is discussed in chapter 7.11 in greater detail. In the field of warmth, including *trustworthy*, *friendly*, *polite* and *reliable*, the distribution of maximum values was balanced. While GA received slightly higher or equal points in the evaluation of *polite* (M=2.70) and *friendly* (M=2.63), RP achieved better results in *trustworthy* (M=2.64) and *reliable* (M=2.66); however, differences were only marginal with a maximum of 0.3 out of all adjectives. Interestingly, significant divergences emerged in the evaluation of social status. A closer look at the ratings belonging to social status revealed that the maximum score was attained for the perceived intelligence of the England speaker with a mean of M=2.84 points, which was also the highest value out of all audio files. In contrast, the GA voice garnered an average of M=2.56, adding up to a difference of 0.28. Another concrete example refers to the analysis of the characteristic *wealthy*. Whereas the American speaker obtained a mean value of M=2.40, the England counterpart received 2.70 points. The same trend appeared in the evaluation of *hardworking* (M_{GA}=2.49; M_{RP}=2.61). Both varieties attained their lowest scores in different dimensions.

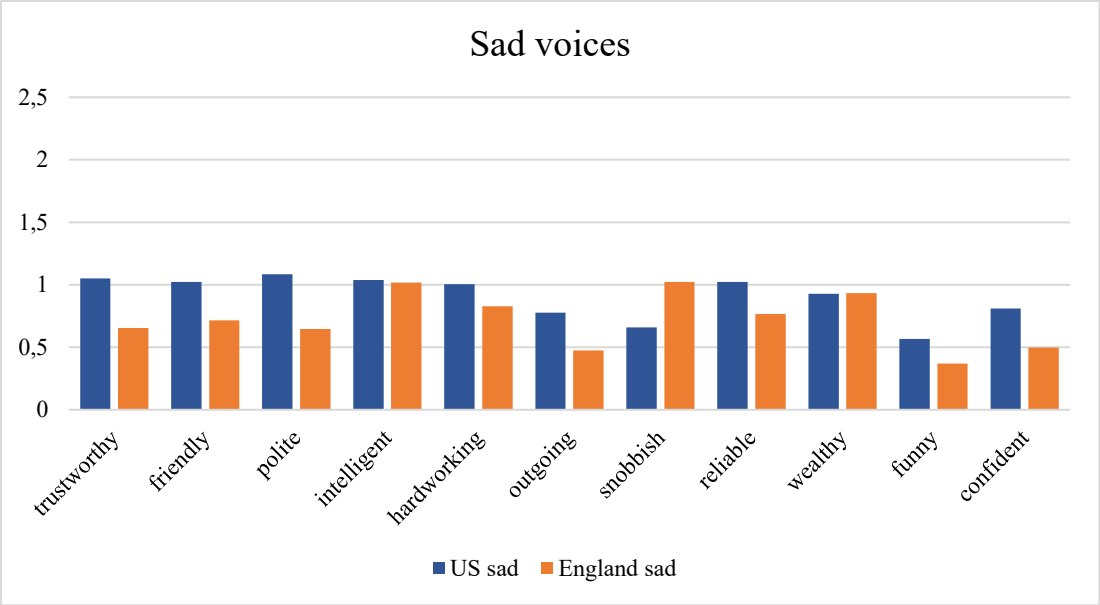


Figure 15. Sad voices all respondents

Figure 15 illustrates the distribution of mean scores for the GA and RP variety, each presented in a sad tone. Overall, the American speaker received higher points across all character traits except *snobbish* and *wealthy*; however, due to the negative connotation of *snobbish*, a lower score indicates more positive attitudes. The GA voice outperformed the England file in the category of warmth and also reported the biggest differences out of all forementioned categories, including the adjectives *trustworthy* (M_{GA}=1.05; M_{RP}=0.65), *friendly* (M_{GA}=1.02; M_{RP}=0.72), *polite* (M_{GA}=1.08; M_{RP}=0.65), *reliable* (M_{GA}=1.02; M_{RP}=0.77) with discrepancies ranging from 0.25 to 0.43 points. In contrast, an almost identical mean was obtained for the category of *wealthy*. Furthermore, in the field of social status, the term *intelligent* revealed resemblances in the rating of the GA and RP audio file, outlining a difference of 0.02. Character traits of dynamism generally yielded the lowest mean values. Specific examples refer to *outgoing* (M_{GA}=0.78; M_{RP}=0.47), *funny* (M_{GA}=0.57; M_{RP}=0.37) and *confident* (M_{GA}=0.81; M_{RP}=0.50). It is noteworthy that the sad RP voice reached lower scores than the American speaker in all dynamic characteristics; however, a different pattern occurred in the evaluation of the happy RP counterpart, exceeding the results of the GA file.

A comparison between the conversational and sad recordings revealed opposite patterns (see Figure 16). While the distribution of the maximum scores for the happy RP and GA variety was balanced throughout all character traits, the melancholy American speaker stood out with its higher means, except in the evaluation of *snobbish* and *wealthy*. Both English files scored their maximum values in the domain of the perceived intelligence, but with a remarkable difference of 1.82 points. The conversational and sad GA speaker reached the maximum number of points

in different categories. The first mentioned was rated high regarding confidence (M=2.70), while the latter reached the greatest value in the evaluation of *polite* (M=1.08).

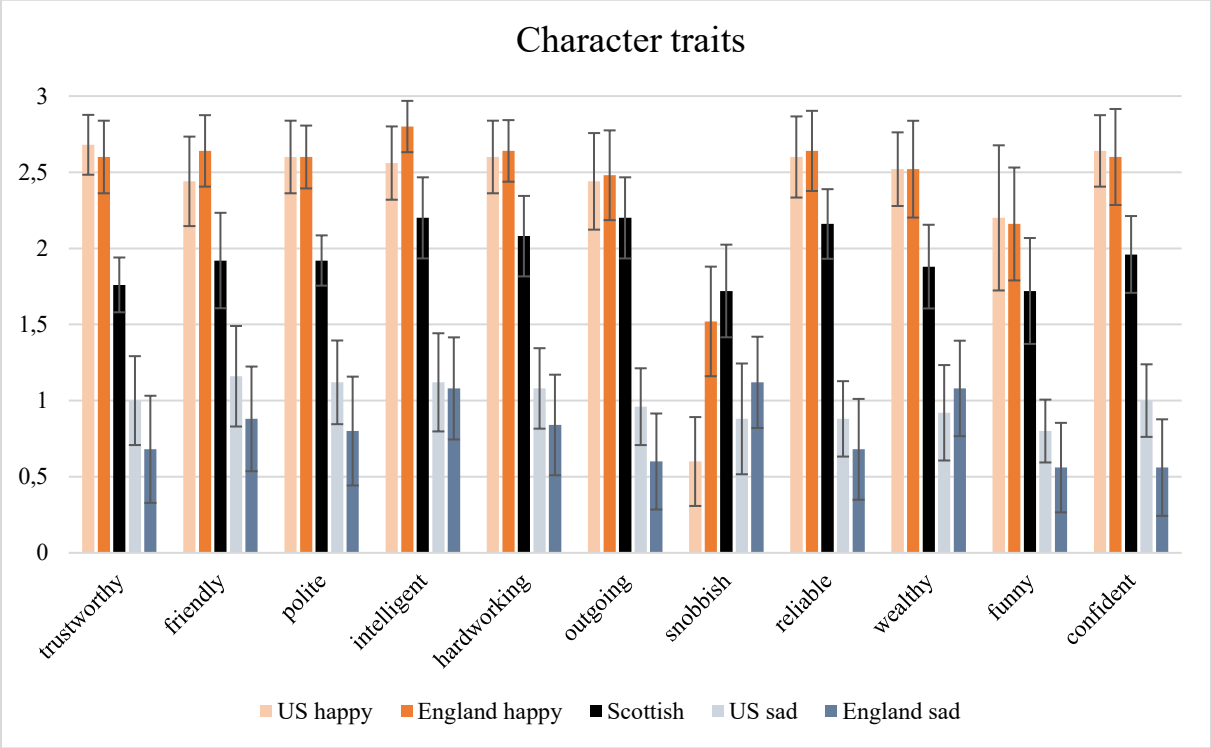


Figure 16 .Character traits all respondents

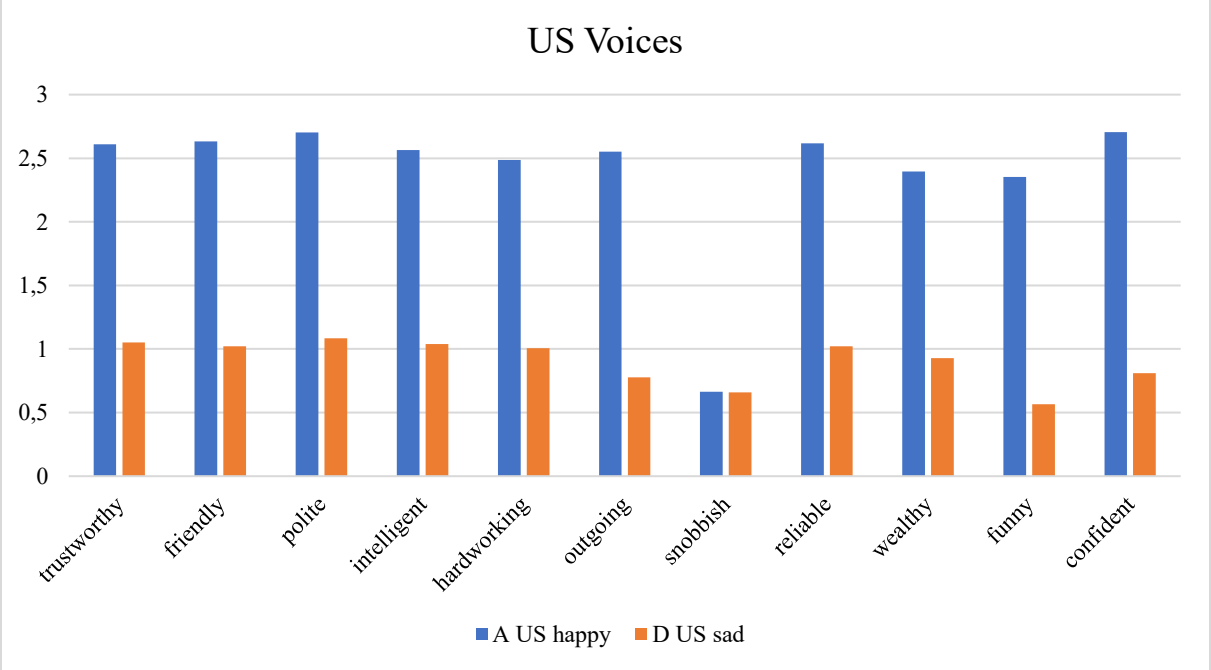


Figure 17. US voices all respondents

Figure 17 above as well as the graphical illustration below (see Figure 18) present mean scores regarding different character traits as described in Table 1. Both bar charts aim to provide a

concise comparison between the happy (A, B) and sad speakers (D, E) to emphasize the remarkable influence of emotional tone on participants' evaluations within one English variety.

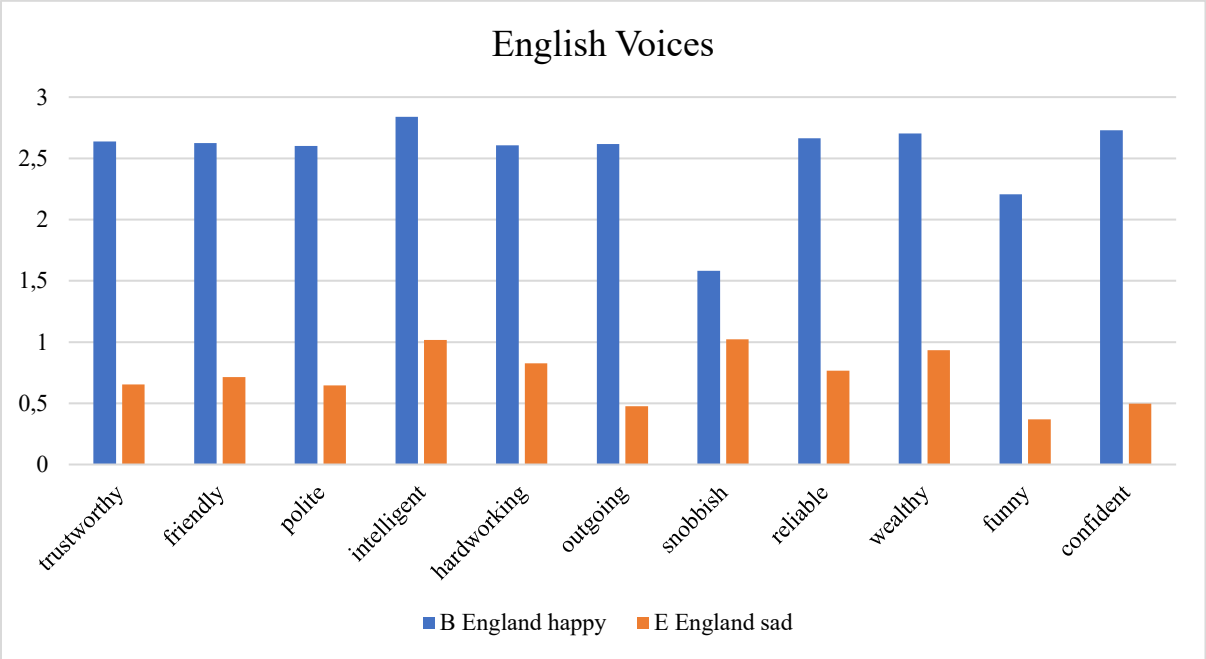


Figure 18. English voices all respondents

7.2 Evaluations of pronunciation clarity

To assess the clarity of speakers' pronunciation (see Figure 19), average values of all samples were calculated with a minimum of 1 (unclear) to a maximum of 5 points (very clear). For calculating mean scores, all points were added and divided by the number of respondents. A closer look at Figure 19 revealed consensus among respondents' rating of the individual samples. A notable decreasing trend became transparent from the conversational to the sad recordings. The first RP speaker received the highest mean score (M=4.85), closely followed by the conversational GA voice (M=4.80) with a gap of 0.05. The SSE speaker was set in the middle with a mean of M=4.16. A difference of 1.32 separated the third place and the fourth-ranked sad GA sample (M=2.84), indicating that the sad recordings were perceived considerably inferior in terms of pronunciation compared to the conversationally read examples. The emotional RP variety scored the lowest mean value with M=2.51 out of 5 points. While the happy RP speaker obtained a higher mean evaluation (M=4.85) than the GA equivalent (M=4.8), an opposite pattern emerged for the melancholy audio files. Overall, a discrepancy of 2.34 points appeared from the first (conversational RP) to the last place (RP sad).

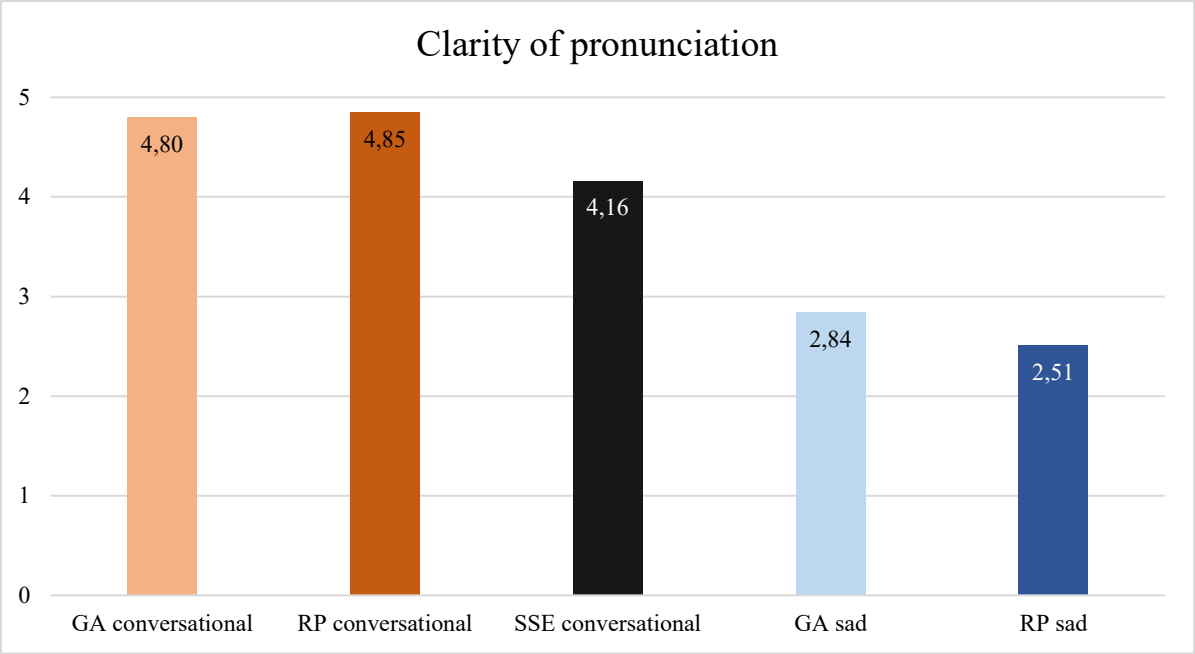


Figure 19. Mean scores of pronunciation clarity by all respondents

7.3 Mean scores regarding suitability as a news reporter

The calculation of the mean scores regarding the suitability as a news reporter (see Figure 20) followed the same principles as the analysis of the pronunciation clarity. Participants were required to rank the individual speakers with a minimum of 1 to a maximum of 5 points. The total score was then divided by the number of received answers to determine average values. The conversational RP speaker was ranked in first place with a mean of M=4.79 and therefore considered to be most suitable for reading the news. The first GA recording received the second highest score (M=4.64) with a difference of 0.15. Then the SSE voice followed with M=3.77, showing a gap of 0.87 to the happy GA equivalent in second place. Similar to the review of the pronunciation, the sad speakers obtained remarkably lower scores than the first three examples. The emotional GA variety attained M=1.96 points while the melancholic RP sample was regarded as the least favorite in terms of news reporter suitable with a means of M=1.76 and therefore presented a gap of 3.03 to the happy equivalent. These findings align with analysis of the pronunciation aspect. In both dimensions, the conversational RP voice weas ranked in first place, while the sad equivalent reached the lowest score and was therefore not considered as a good news reporter.

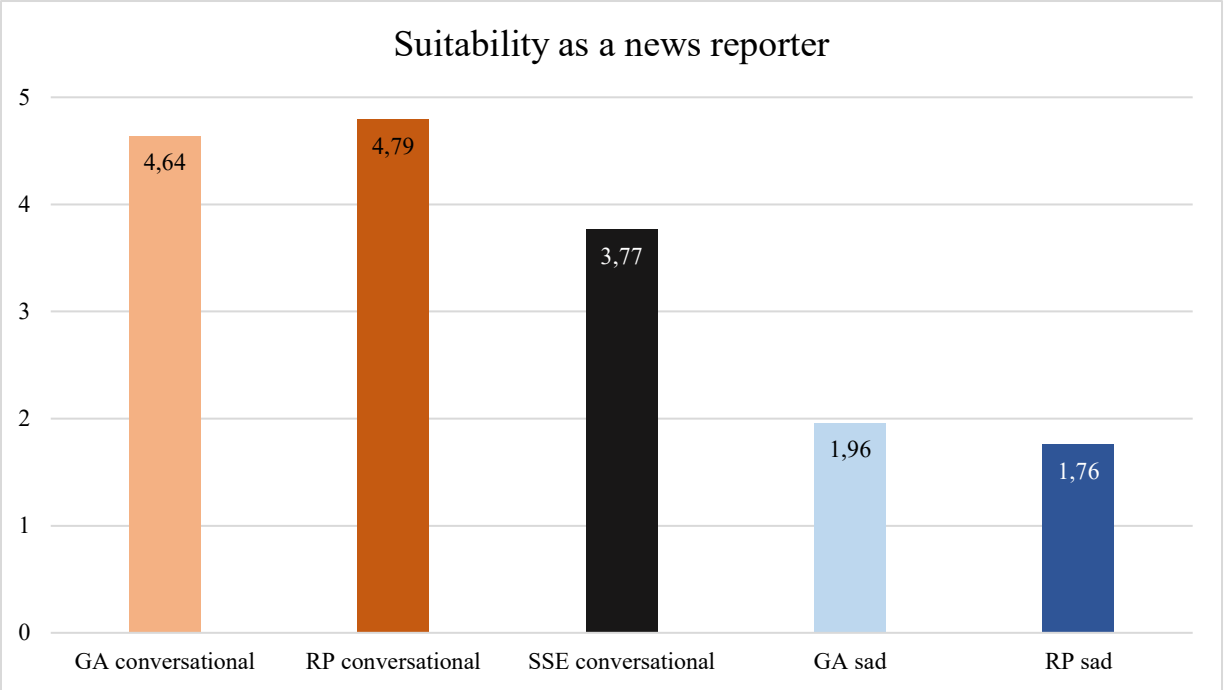


Figure 20. Mean scores of suitability as a news reporter by all respondents

7.4 Evaluations of participants' willingness to meet the speakers for a coffee

To convert the yes/no question regarding participants' willingness to meet the individual speakers for a coffee into numerical data, a scale from 0 (*no*) to 1 (*yes*) was applied. Afterwards, all scores were added and divided through the number of respondents. The y-axis in Figure 21 indicates the reached values. The illustration below shows that most respondents (n=163) were willing to meet Speaker 2 (RP conversational) for a coffee, garnering M=0.91 out of a maximum of 1. The second highest evaluation was achieved by the happy American voice, attaining 152 votes in favor of meeting the respective speaker, equaling 0.84 points. The SSE audio file was voted into third place, as 104 participants were willing to have a coffee with the Scottish speaker, resulting in a value of M=0.58. A remarkable divergence became evident in the analysis of the two sad voices. Speaker 4 (GA sad) in fourth position, scored 0.47 points less than the SSE sample, leading to only 0.11 achieved points. Similar to the happy voice from the USA and England, the difference between the sad samples remained slight with 0.06. While 152 respondents wanted to meet the conversational GA speaker, only 19 votes, in favor of having coffee, became evident for the sad counterpart. A similar trend occurred for both RP recordings. Whereas the conversational voice from England reached the maximum score of M=0.91 (n=163), the melancholy equivalent only yielded an average M=0.05 (n=9); thus, as visible in Figure 21, the maximum and minimum values were obtained by an RP variety.

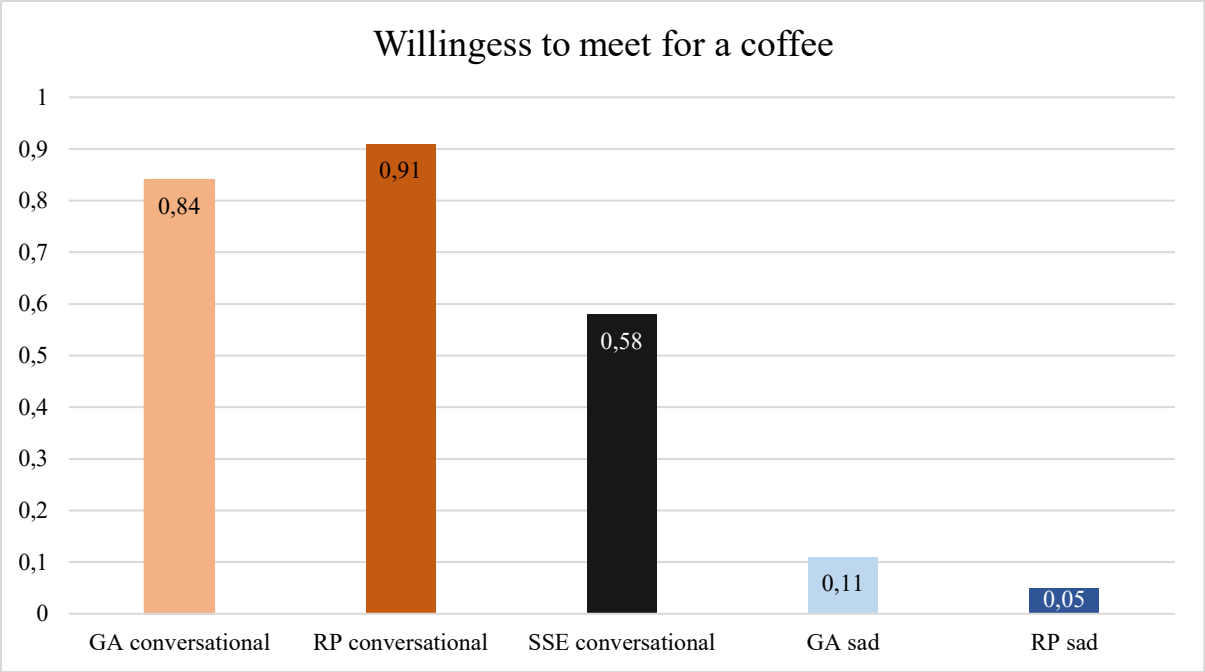


Figure 21. Willingness to meet for a coffee all respondents

7.5 Respondents' descriptions of speakers' presentation style

The questionnaire asked participants to describe each speaker's presentation style in one word. This section analyses respondents' answers and subsequently groups similar descriptors. The aforementioned dimensions (warmth, social status, dynamism) were employed with an addition of the category of presentation style. The Scottish variety was excluded to primarily focus on the comparison between the RP and GA counterparts with special emphasis on the effect of mood change.

Table 3 presents the descriptors submitted for the conversational GA speaker. The majority of respondents focused their evaluations on the field of warmth with *nice* (n=27) as the most frequent association rather than assumed competencies. In terms of presentation style, positive adjectives, such as *clear*, *bright* (n=14) or *excellent* or *good* (n=16) prevailed; nevertheless, a minority (n=2) noted that the pace was too fast and therefore, hard to follow. Additionally, some participants (n=7) considered the speaker's presentation manner of the news as *boring* and *monotonous*. In contrast, others perceived it as *engaging*, *lively* (n=17) and *natural* (n=3). Further descriptive terms ranged from *convincing* (n=1), *informative* (n=4) to *fluent* (n=3), reflecting consistent positive attitudes towards the respective audio recording. For the dimension of warmth, only optimistic connotations occurred. Concrete examples were *nice* (n=27), *friendly* (n=17), *positive* (n=16) and *warm* (n=15). Similar patterns emerged in the other fields of social status and dynamism. The first mentioned included the terms *professional* (n=9),

sophisticated (n=2) and *competent* (n=1), whereas the latter encompassed *confident* (n=2) and *motivated* (n=1).

Table 3. Descriptors for happy GA speaker all respondents

Warmth	Presentation style	Social status	Dynamism
nice (27)	**excellent/super/ good/cool** (16)	professional (9)	**engaging/lively** (17)
friendly (17)	**clear/bright** (14)	sophisticated (2)	average (3)
positive (16)	automated/monotonous/bored (7)	competent (1)	confident (2)
warm (15)	informative/factual (4)		motivated (1)
welcoming/inviting (8)	natural (3)		automated (1)
open minded (4)	fluent/direct/concise (3)		motivated (1)
reliable (1)	calm/relaxed (3)		
	fast/difficult to understand (2)		
	thoughtful/structured (2)		
	convincing (1)		
	smooth (1)		

Table 4 illustrates an opposite pattern as Table 3 presented above. Overall, respondents tended to focus on professionalism in terms of presentation style rather than on social aspects. The overall most frequent association was *professional* (n=46). A closer look at the dimension of presentation style revealed that respondents were highly convinced of the happy RP speaker, as indicated through descriptors like *good* (n=30), *academic* (n=1), *prepared* (n=14) or *comprehensible* (n=7). Only a minority submitted negatively connoted descriptors such as *monotonous* (n=2). In the category of warmth, similar descriptors were detectable as in the analysis of the GA equivalent but the number of responses significantly varied. Whereas the happy American speaker received 27 votes for *nice*, the counterpart from England only obtained

one answer regarding this characteristic. The same applied for *friendly*. Participants associated the RP representative with a high social status, underlined by adjectives such as *posh* (n=4), *rich* (n=5) and *upper class* (n=1). Interestingly, none of these terms appeared in the description of the GA sample; nonetheless, the dimension of dynamism proved similar trends as the analysis of the American voice. Among the most frequently mentioned terms were *confident* (n=2) and *engaging* (n=3); however, the GA speaker tended to be perceived as more dynamic based on the number of received responses.

Table 4. Descriptors for happy RP speaker all respondents

Warmth	Presentation style	Social status	Dynamism	Other
friendly (1)	**good** (30)	**professional** (46)	engaging (3)	British (1)
nice (1)	**clear/bright** (16)	**skilled** (18)	confident (2)	conceited (1)
polite (1)	prepared (14)	rich (5)	motivated (2)	familiar (1)
sweet (1)	comprehensible (7)	posh (4)	dynamic (1)	metro (1)
trustworthy (1)	adequate (4)	competent (1)	outgoing (1)	snobbish (1)
positive (1)	informative (4)	upper class (1)		
	natural (4)	classy (1)		
	serious (2)			
	monotonous (2)			
	academic (1)			

As visible in Table 5, emotional tone had a significant influence on participants' personal evaluations. The most frequently mentioned descriptor was *sad* with a total number of 54 occurrences, followed by *insecure* (n=32). Overall, negatively connotated terms outnumbered positive adjectives in all dimensions. Only three optimistic expressions, specifically *friendly* (n=1), *professional* (n=11) and *competent* (n=1) were observable. Especially the category of dynamism demonstrated substantial discrepancies between the conversational and sad GA recording. While the happy file consisted of highly dynamic terms, like *lively* or *engaging*, the latter presented an opposite pattern with stagnant descriptors (*unmotivated* (n=6); *depressed*

(n=1)). Another difference concerned the descriptions in the field of warmth. Whereas the happy American speaker received numerous mentions of *nice* (n=27) and *friendly* (n=17), the melancholy counterpart only garnered one instance of *friendly*; however, both GA samples showed similarities in their style of responses, as participants focused on the evaluations of the suitability as a news reporter rather than social characteristics. Surprisingly, a closer look at the dimension of social status revealed that more answers referring to professionalism were obtained for the sad speaker (n=11), contradicting the numerical evaluations presented in chapter 7.1.

Table 5. Descriptors for sad GA speaker all respondents

Warmth	Presentation style	Social status	Dynamism	Other
friendly (1)	**sad** (54)	professional (11)	unmotivated (6)	artificial (6)
	insecure (32)	competent (1)	depressed (1)	arrogant (3)
	monotonous/ bored (23)			average (1)
	dramatic (11)			
	annoyed (10)			
	inappropriate (7)			
	negative/ cold (7)			
	slow (3)			
	awkward (2)			
	nasal (1)			

As immediately evident in Table 6, speaker 5 (RP sad) did not receive any positively connoted adjectives in the domain of warmth, besides *empathetic* (n=1). In contrast, *angry* (n=2) or *unfriendly* (n=2) were noted. A similar trend occurred in the other dimensions, where predominantly unfavorable attitudes dominated. A parallel in the analysis of the sad GA and RP file consisted of *sad* as the most common descriptor with 59 submissions for the speaker from England. While the American voice was perceived as *insecure* with a total number of 32 responses, the second most frequent evaluations of the RP sample highlighted its dramatic presentation (n=26), closely followed by the perceived monotonous reading (n=23).

Considerable differences were observable for the dimension of social status. While the conversational RP speaker received many positive terms, such as *professional* (n=46) or *skilled* (n=18), the sad counterpart only garnered one term (*rich*). Similar applied the field of dynamism, where the conversational voice from England achieved highly favorable results, whereas the melancholic equivalent was described as *unmotivated* (n=2) or even *depressed* (n=1).

Table 6. Descriptors for sad RP speaker all respondents

Warmth	Presentation style	Social status	Dynamism	Other
angry (2)	**sad** (59)	rich (1)	unmotivated (2)	arrogant (7)
unfriendly (2)	**dramatic** (26)		depressed (1)	artificial (3)
empathetic (1)	**monotonous/bored** (23)		relaxed (1)	snobbish (1)
	inappropriate (22)			average (1)
	insecure (13)			
	annoyed (11)			
	good/flowing (2)			
	informative (2)			

A comparison of all descriptors showed that emotional tone was highly influential on participants' attitudes. Considerable differences could be detected. While positively connoted terms dominated in the evaluations for the happy speakers, predominantly negative or pessimistic adjectives emerged for the sad counterparts. These either solely referred to their professionalism and linguistic aptitude or to social aspects; thus, it can be inferred that respondents did not realize that the same voices appeared for the happy and sad samples.

7.6 Effects of gender on participants' evaluations

The following chapter compares the mean scores for different character traits based on participants' self-reported gender, outlining differences as well as parallels. Since most respondents identified as either male (n=77) or female (n=98), these two groups were analyzed in greater detail to ensure statistical reliability. The same procedure as in chapter 7.1 was employed to obtain numerical data. Each character trait received a certain number of points (*very*=3; *not* =0) which were subsequently divided by the number of responses. In this specific case, data had to be filtered according to gender.

Table 7. Mean scores character traits by female participants

Variables	GA happy	RP happy	SSE happy	GA sad	RP sad
trustworthy	2.58	2.69	2.03	1.02	0.69
friendly	2.54	2.60	2.05	0.97	0.73
polite	2.64	2.61	1.97	1.02	0.74
intelligent	2.54	2.82	2.28	1.04	1.03
hardworking	2.44	2.60	2.21	1.02	0.90
outgoing	2.51	2.56	2.15	0.72	0.48
snobbish	0.69	1.62	1.70	0.67	1.10
reliable	2.60	2.63	2.16	1.06	0.82
wealthy	2.34	2.69	2.06	0.91	0.95
funny	2.23	2.05	1.68	0.52	0.37
confident	2.68	2.79	2.09	0.79	0.52

Table 8. Mean scores character traits by male participants

Variables	GA happy	RP happy	SSE happy	GA sad	RP sad
trustworthy	2.65	2.57	1.92	1.09	0.61
friendly	2.75	2.65	2.04	1.09	0.70
polite	2.78	2.59	1.90	1.17	0.54
intelligent	2.61	2.86	2.42	1.03	1.01
hardworking	2.55	2.61	2.17	1.00	0.75
outgoing	2.61	2.67	2.23	0.82	0.48
snobbish	0.64	1.51	1.94	0.66	0.94

reliable	2.66	2.70	2.18	0.97	0.70
wealthy	2.47	2.71	2.14	0.94	0.92
funny	2.51	2.38	1.73	0.64	0.38
confident	2.74	2.66	2.13	0.83	0.48

The analysis of gender-based differences showed that overall, women and men did not report substantial discrepancies in evaluating the individual character traits (see Tables 7 and 8); nevertheless, the results of the happy American speaker indicated slightly more positive perceptions among male participants across all dimensions except *snobbish*; however, distinctions remained marginal with a maximum gap of 0.28 points, observed in the category of *funny*. The conversational RP voice showed similar trends. While higher values from female participants emerged in the domains of *trustworthy*, *polite*, *snobbish* and *confident,* men assigned more points to the seven remaining characteristics. Regarding the SSE speaker, women attributed more favorable attitudes towards trustworthiness, friendliness, politeness and the adjective *hardworking* of the concerned audio file; nevertheless, discrepancy remained minor with a maximum of 0.14 points. Furthermore, mood did not prove to be a significant influential factor in terms of gender differences, since no remarkable variations could be detected. While men generally tended to record more optimistic positions towards the sad American speaker, an opposite trend arose for the melancholic voice from England, as female participants ranked the concerned variety higher in eight categories (*trustworthy, friendly, polite, intelligent, hardworking, reliable, wealthy* and *confident*).

In summary, gender did not have a significant impact on participants' attitudes towards the five varieties. Differences remained marginal with a maximum gap of 0.28 points, observed in the evaluation of the conversational GA speaker for the term *funny*. Overall, men tended to prefer the American voice, regardless of the change in tone, as well as the happy RP audio file, while women expressed more positive perspectives towards the sad RP voice. The conversational SSE sample presented a balanced pattern with no clear gender-based preferences.

As immediately visible in Figure 22, the calculation of mean scores regarding the perceived suitability as a news reporter also did not reveal any significant gender-based discrepancies. While the x- axis symbolizes the concerned English varieties, the y-axis indicates the number of points ranging from zero to five. Generally, women rated all varieties higher except for the happy American speaker. The overall maximum score was achieved by the conversational RP sample with M=4.84 points. The highest score among males was achieved by the happy

American file (M=4.71). As with the analysis of the individual character traits, differences remained marginal, ranging from 0.03 to 0.29 points.

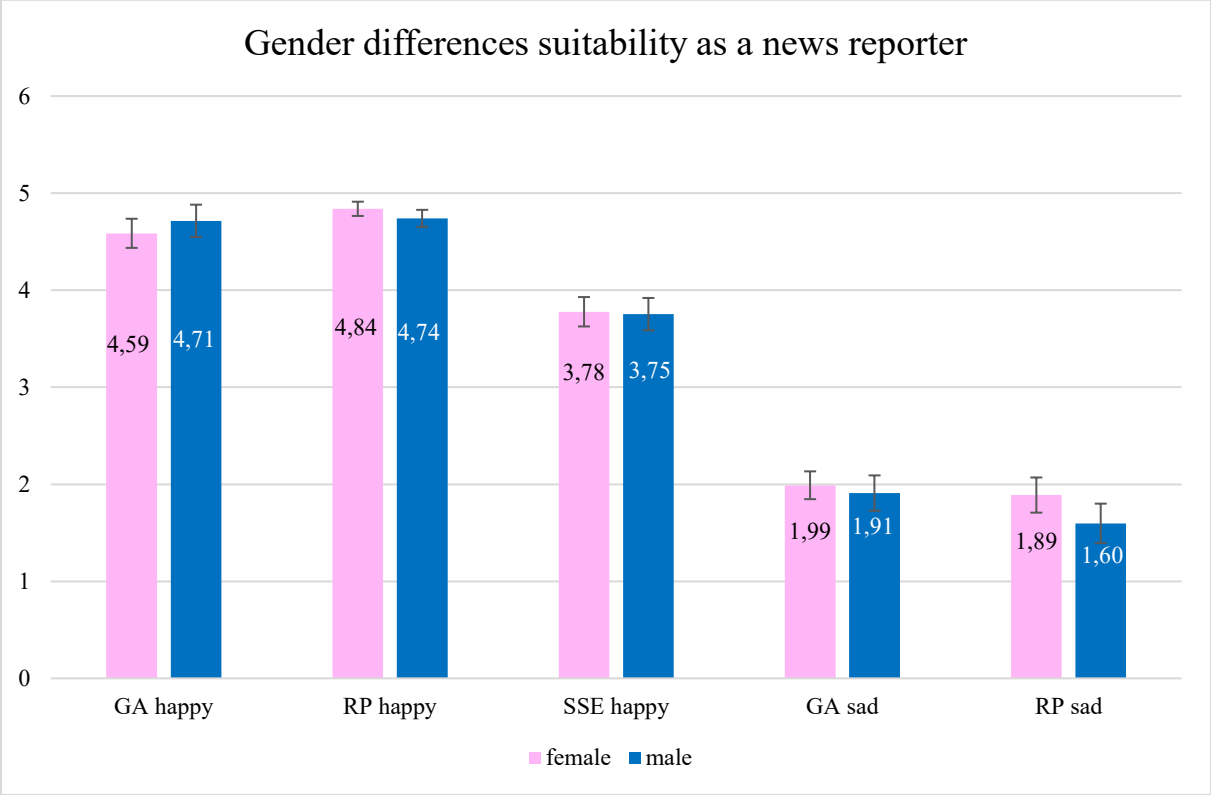


Figure 22. Gender differences suitability as a news reporter

The graphical illustration below (see Figure 23) outlines mean scores concerning the reported clarity of pronunciation. The pink bars represent the responses from female participants, while the ones, colored in blue, portray the average values of the male group. Similar trends could be observed as for the perceived suitability as a news reporter. The SSE file received a mean of M=4.16 from both groups, while the largest gender-based discrepancy appeared in the analysis of the sad RP speaker, showing a gap of 0.35. A closer analysis revealed that female participants reported higher values for all audio recordings. In terms of the conversational voices, mean scores consistently remained over 4 points, while an opposite pattern again emerged for the sad voices, with values not exceeding a mean of M=2.9. Overall, the melancholic speaker from England was ranked last, whereas the happy equivalent received the greatest mean scores with M=4.88 and M=4.82; thus, another resemblance to the perceived suitability as a news reporter appeared. The happy American file reached slightly lower values than the speaker from England with M=4.82 and M=4.77 points.

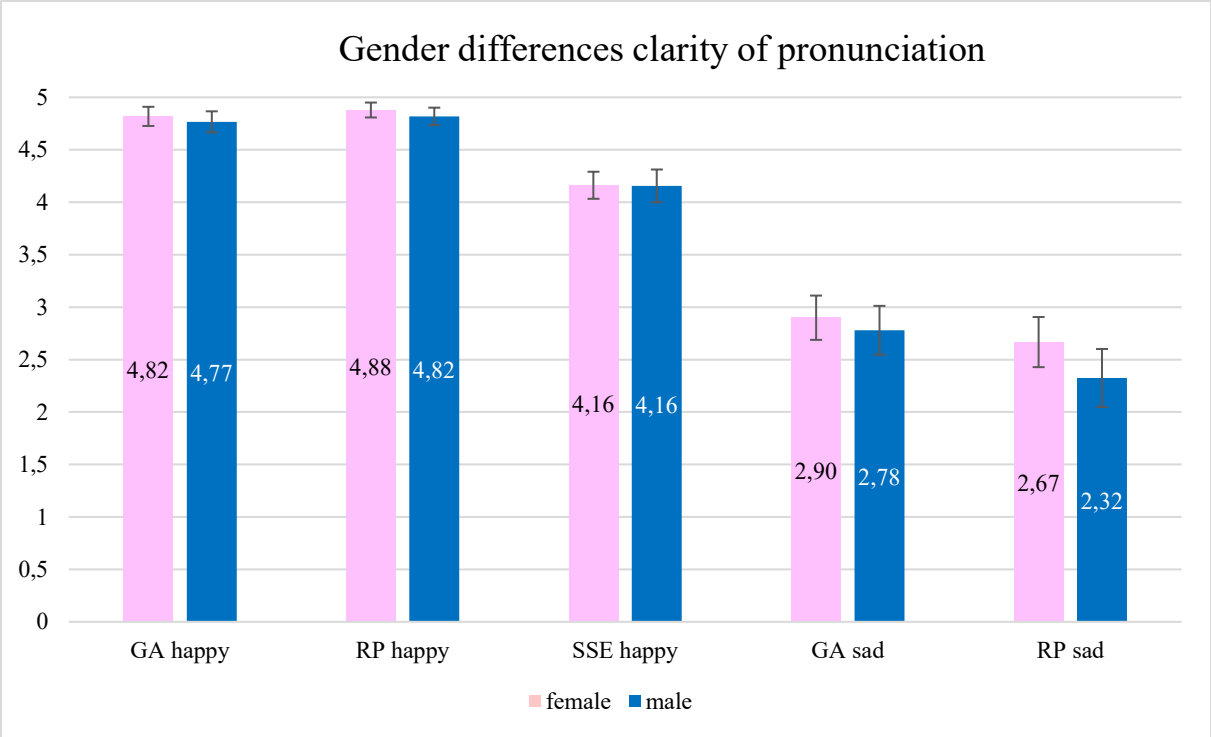


Figure 23. Gender differences clarity of pronunciation

Generally, confidence intervals remained consistent throughout both dimensions, indicating that the outcomes were comparable across both groups. In addition, female and male participants reported resemblances in their attitudes.

Figure 24 presents mean values for participants' willingness to meet the concerned speakers for a coffee. The yes/no question was converted into numerical data by assigning 1 point to *yes* and 0 points to *no*. Scores were then summed and divided by the number of responses. The y-axis illustrates the obtained values, ranging from 0 to 1, while the individual speakers are presented on the x-axis. As immediately visible, participants expressed less favorable attitudes towards the sad recordings; however, gender did not appear as a significant influential factor, as no remarkable discrepancies could be observed. The largest differences were recorded for the conversational GA voice with a value of 0.12. Most consensus between women and men occurred in the evaluation of the SSE sample. The three remaining varieties (RP happy, GA and RP sad) outlined similar discrepancies ranging from 0.03 to 0.04 points.

To conclude, findings revealed that the emotional tone had a remarkable impact on respondents' perceptions of the individual voices, regardless of their gender. In contrast, gender itself did not considerably influence participants' judgements. The calculation of average values and confidence intervals showed only minor distinctions in all dimensions.

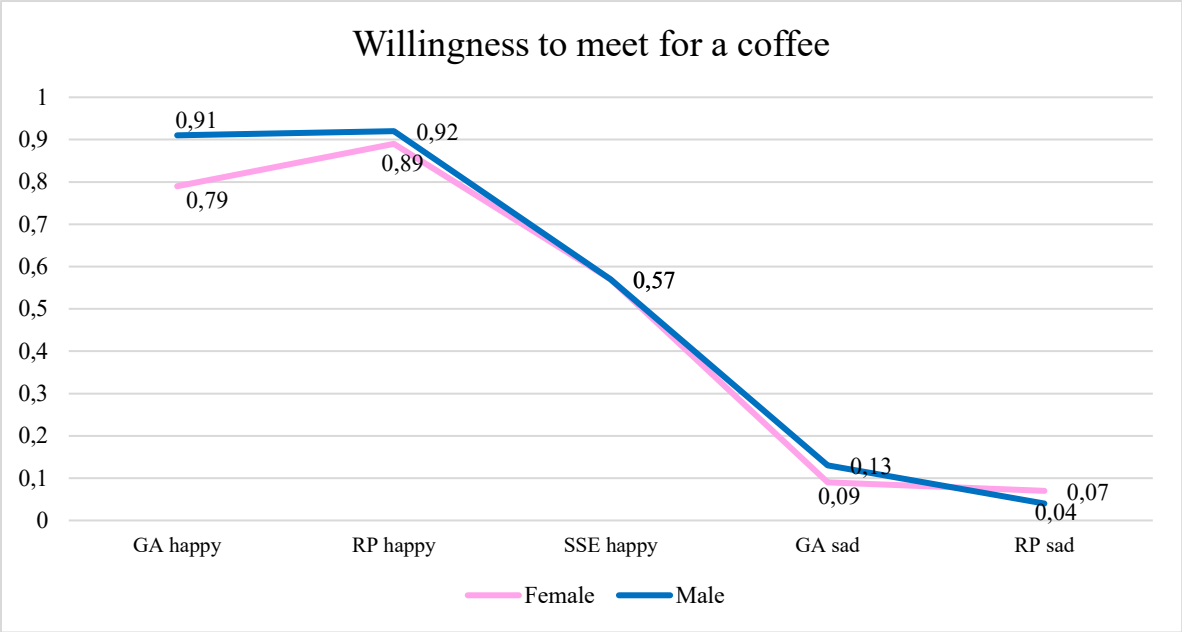


Figure 24. Gender differences willingness to meet for a coffee

7.7 Influence of prior residences in Anglophone areas

This section examines the influence of prior residences in Anglophone countries on respondents' attitudes towards the five chosen varieties. It is important to note that these results specifically refer to living in foreign countries, rather than short travel experiences. Mean scores are calculated to outline similarities and differences between participants with and without experience. In total, 28 respondents reported having lived in areas, where English is considered an official language. In comparison, 152 people without this experience participated in the given online survey; however, it needs to be noted that two out of these 28 have resided in two different English-speaking countries, particularly England or UK and USA. The pie chart below (see Figure 25) provides a more detailed overview of the destinations mentioned. The largest group consisted of people, who have resided in Great Britain (n=17), followed by the US (n=9). Besides that, other responses referred to Australia (n=2) or South Africa (n=2).

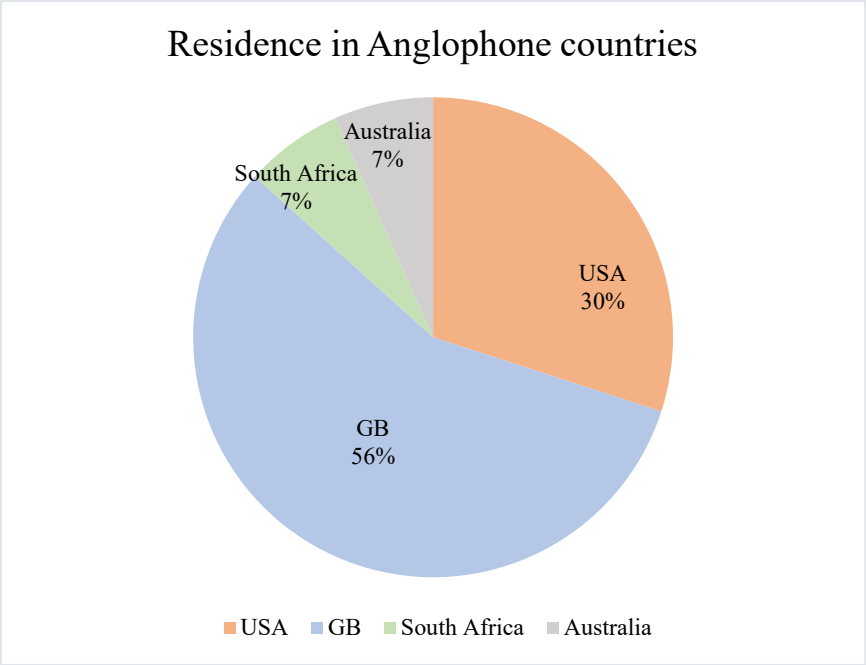


Figure 25. Prior residences in Anglophone countries

Table 9. Mean scores character traits by participants with former residence in Anglophone areas

Variables	GA happy	RP happy	SSE happy	GA sad	RP sad
trustworthy	2.5	2.82	1.96	0.96	0.89
friendly	2.64	2.61	1.96	0.89	0.79
polite	2.68	2.75	2.07	1	0.89
intelligent	2.39	2.86	2.36	0.96	1.07
hardworking	2.29	2.70	2.29	0.96	1.11
outgoing	2.5	2.68	2.25	0.46	0.46
snobbish	0.64	1.46	1.71	0.57	1.14
reliable	2.39	2.68	2.29	0.79	0.93
wealthy	2.25	2.68	2.21	0.75	1.07
funny	2.29	2.32	1.75	0.29	0.43
confident	2.61	2.75	2.07	0.68	0.61

Table 10. Mean scores character traits by participants without former residence in Anglophone areas

Variables	GA happy	RP happy	SSE happy	GA sad	RP sad
trustworthy	2.63	2.61	1.99	1.07	0.61
friendly	2.63	2.63	2.07	1.05	0.70
polite	2.71	2.57	1.91	1.1	0.60
intelligent	2.60	2.83	2.33	1.05	1.01
hardworking	2.52	2.59	2.18	1.01	0.77
outgoing	2.56	2.61	2.17	0.83	0.48
snobbish	0.67	1.60	1.83	0.68	1
reliable	2.66	2.66	2.15	1.07	0.74
wealthy	2.42	2.71	2.09	0.96	0.91
funny	2.36	2.19	1.70	0.62	0.36
confident	2.72	2.73	2.12	0.83	0.48

Tables 9 and 10 above present differences in mean scores between participants with and without former residence in Anglophone areas. Overall, some discrepancies were observable, indicating that having lived in English-speaking countries partially influences respondents' attitudes towards the individual speakers. The GA voice revealed higher values by participants who have not resided in regions with English as an official language in all domains except *friendly*. The largest divergence of 0.27 points emerged in the perceived reliability of the American sample. Interestingly, the two groups did not assign their highest and lowest values to the same character traits. While those who have lived abroad, reported the lowest ratings for the speaker's wealthiness (M=2.25) and their highest for the perceived politeness (M=2.68), respondents without this experience gave the greatest score for the voice's confidence (M=2.72) and their lowest for the funniness (M=2.36).

A contrasting trend appeared in the analysis of the happy recording from England. The group with a former residence in Anglophone areas outlined higher mean values in nearly all categories, except for *friendly* and *wealthy*; however, discrepancies in both characteristics amounted to a maximum of only 0.03 points. The remaining adjectives outlined a higher variation from 0.02 to 0.21 points. Interesting results can be obtained by a closer look at the calculations of the sad GA voice. As shown in previous evaluations, mood had a remarkable impact on participants' attitudes. The sad version received significantly lower ratings than the happy counterpart. In addition, it became evident that respondents without a former residence

in Anglophone countries assigned higher means for all categories than those with such experience, revealing differences from 0.05 to 0.37. One similarity in both data sets was the highest rating for *polite*. Respondents, who lived in Anglophone regions, reported a mean of M=1, whereas participants without a former residence in countries with English as an official language showed an average of M=1.1. The lowest points emerged in the analysis of *funny*, including values M=0.29 and M=0.62. An opposite pattern emerged for the sad voice from England. Participants with a former residence in an Anglophone area reported higher average points than the other group except for *outgoing*. The largest discrepancies were found in the evaluations of *hardworking*. While internationally experienced respondents ascribed a mean of M=1.11, participants without this experience outlined a total of M=0.77, resulting in a difference of 0.34. Most consensus between the two groups occurred in the assessment of the speaker's perceived intelligence with a marginal discrepancy of 0.06.

Figure 26 below illustrates both differences and parallels in participants' ratings of the perceived suitability of the individual speakers, based on whether they have lived in Anglophone countries. The calculation of mean scores and the presentation of the chart follows the same principles as the evaluations in chapter 7.6. The x-axis represents the respective English varieties, while the y-axis symbolizes the received mean values on a scale from one to five possible points.

A first glimpse already reveals that the overall highest score was obtained by the conversational RP voice and was therefore considered most suitable for reporting the news by both groups with means of M=4.86 and M=4.78 points. Respondents with travel experiences expressed slightly more positive perceptions.

A different pattern emerged for the happy GA speaker, where participants, without extended residence in Anglophone areas recorded a higher suitability (M=4.66); however, no consistent tendencies were observable in the comparison of the two mentioned groups. While the happy England and SSE sample received greater values by respondents with such international experience, a dissimilar trend accounted for the three remaining samples, namely happy and sad GA as well as melancholic RP. Specifically, the Scottish variety outlined means of M=3.82 and M=3.76 points, while both sad speakers again reached remarkable lower scores. Average scores of M=1.82 and M=1.98 put the sad GA sample in fourth position, outlining the biggest difference between the two data groups with 0.16 points. Respondents considered the melancholic RP voice as least appropriate for reading the news with a mean of M=1.64 (with experience), and M=1.78 (without experience).

Overall, both data sets described similar attitudes towards the individual speakers with minor discrepancies ranging from 0.06 to 0.16 points.

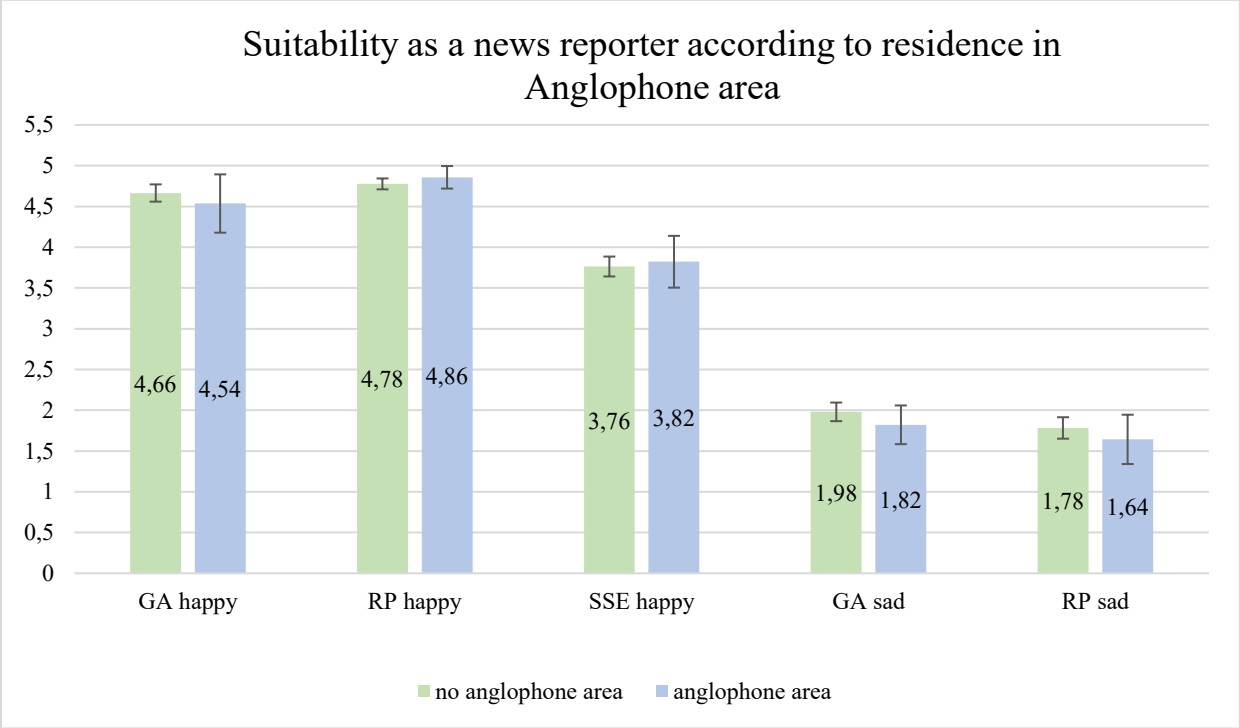


Figure 26. Comparison perceived suitability as a news reporter based on prior residences in Anglophone areas

Figure 27 presented below outlines comparisons of the pronunciation clarity according to prior travel experiences to English speaking countries. While the blue bars indicate mean scores, assigned by people, who have lived in Anglophone areas for a specific time, the ones colored in green symbolize those without this experience. An interesting pattern is revealed by the first glimpse at the findings presented in the illustration below. The group without this international experience reported lower values for all speakers except the sad GA voice; however, it needs to be stressed that discrepancies remained marginal with a maximum of 0.15. In general, a similar trend could be observed as in the previous evaluations with a remarkable decrease in mean scores for the sad speakers. The overall highest values were garnered by the happy English voice with M=4.93, evaluated by participants with Anglophone residence experiences, and a mean of M=4.84 out of five possible points, distributed by those, who claimed to never have lived in an English-speaking country. The happy GA sample was put in second position by both groups, outlining similar means. As visible in Figure 27, the SSE reported a difference of 0.11 points, while the sad GA voice also recorded a discrepancy of 0.11. The overall lowest values were obtained by the sad RP file, regardless of the international experience.

In short, findings revealed no significant impact of prior residences in Anglophone countries on the judgements of pronunciation clarity. Wider confidence intervals appeared for the internationally experienced group, which might be due to smaller sample size, simultaneously suggesting less variability.

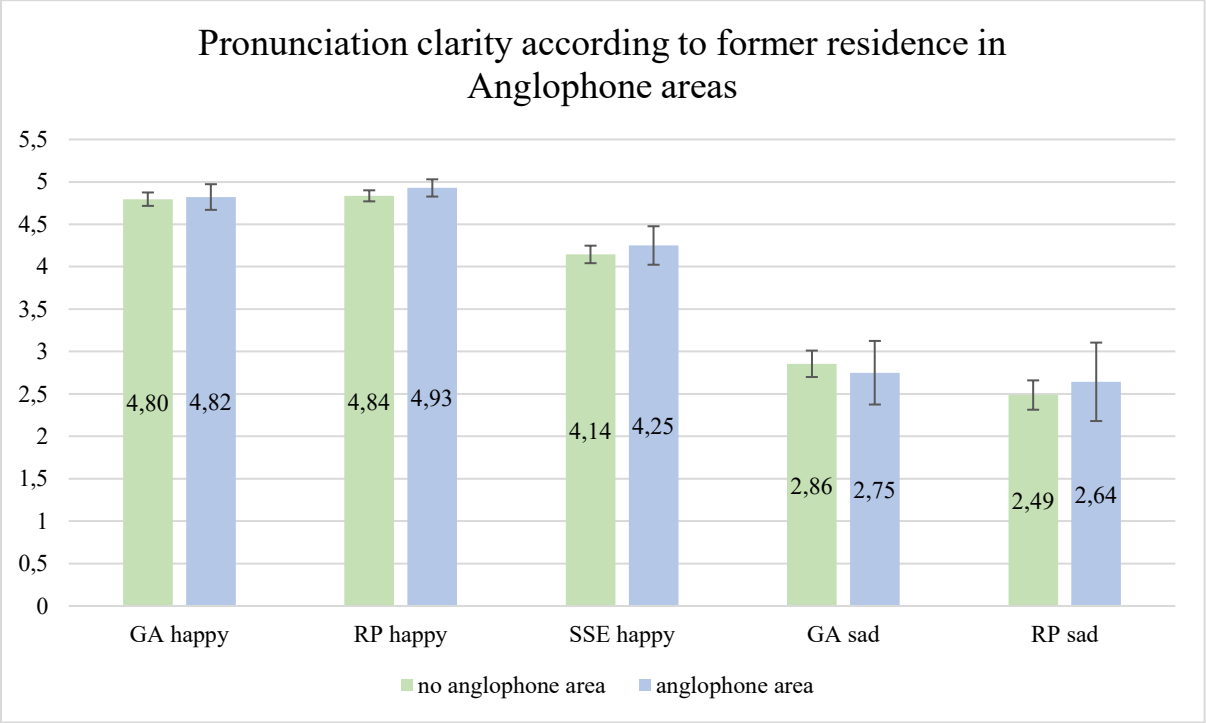


Figure 27. Comparison pronunciation clarity according to former residences in Anglophone countries

The graphical illustration (see Figure 28) below shows the distribution of mean score regarding participants' willingness to meet the individual speakers for a coffee. The two bars, colored in green and blue, outline differences according to prior travel experiences to English-speaking areas. To obtain numerical data, values were assigned on a scale from 0 (*no*) to 1 (*yes*) and afterwards divided through the number of answers received. Through a closer look at Figure 28, nearly identical mean scores were discovered for the conversational GA speaker, differing by only 0.03 points. Additionally, respondents of both groups reported highly similar attitudes towards the sad RP sample with a discrepancy of 0.02; nevertheless, the sad recordings were ranked remarkably lower in comparison to the conversational counterparts, regardless of previous travel experiences. The highest value out of all dimensions occurred in the analysis of the happy RP voice with a mean of 0.93, assigned by participants, who have lived in an Anglophone country. Respondents without this experience described less favorable attitudes (M=0.9). The most significant variance among both groups emerged for the SSE file, presenting values from 0.75 to 0.55; thus, people, having lived in an English-speaking country, tend to be

more willing to meet the Scottish speaker for a coffee; however, even the higher score of 0.75 only sufficed for the third place.

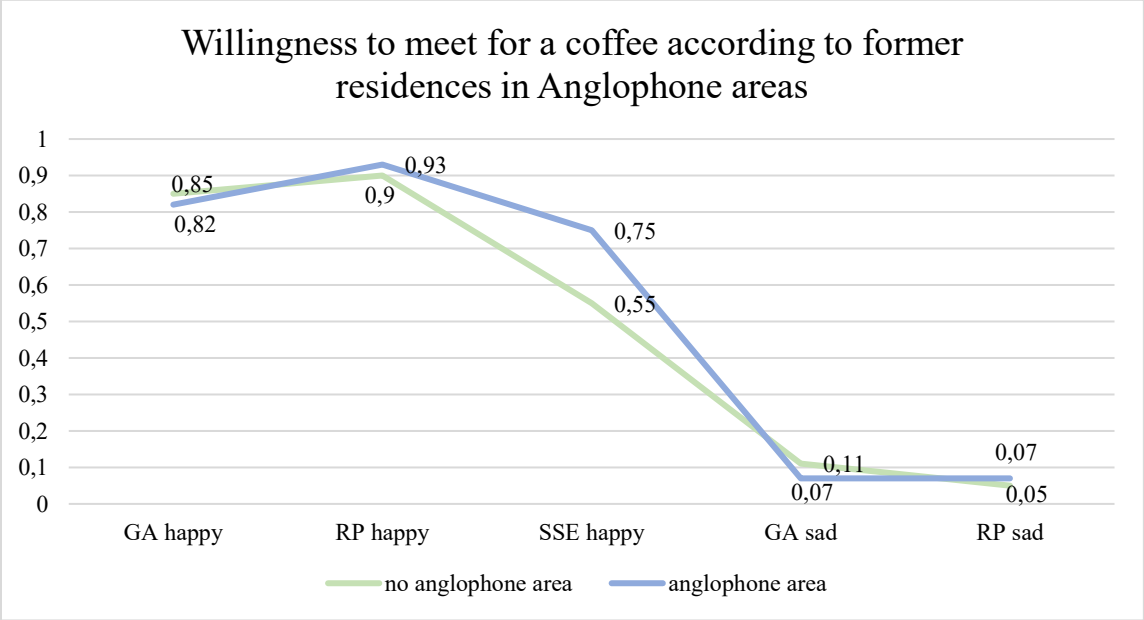


Figure 28. Comparison willingness to meet for a coffee according to residences in Anglophone areas

To conclude, similar to the dimensions of clarity and perceived suitability as a news reporter, participants' willingness to meet for a coffee was not influenced by prior travel experiences in Anglophone regions. Three varieties, specifically GA and RP happy as well as RP sad, were assessed with differences from only 0.02 to 0.03. Likewise, the melancholic American speaker outlined a difference of just 0.04 points. Overall, discrepancies did not exceed 0.2 points, as presented in the bars of the SSE voice.

The sample size of participants with a former residence in Anglophone (n=28) areas was considerably smaller than the group without such experience (n=152). As a result, confidence intervals were wider for the internationally experienced respondents. This suggests that findings are less reliable for respondents, who spent a longer period in regions with English as an official language.

7.8 Evaluations dependent on mother tongue

Figure 29 presents the number of English and German native speakers. As immediately visible, the majority uses German as their mother tongue (n=159), while only a minor part reported English as their first language (n=5). Of the five respondents, four claimed to speak American

English. In total, a diverse range of L1s was reported, such as Swedish (n=4), Slovak (n=2) and Czech (n=2) or Turkish (n=2); however, this chapter focuses exclusively on the comparison between English and German native speakers and their potential influence on the evaluated dimensions. It should be noted that the sample size of English native speakers was significantly smaller compared to participants with German as their mother tongue.

Figure 29. Mother tongue German and English

Table 11. Mean scores character traits by German native speakers

Variables	GA happy	RP happy	SSE happy	GA sad	RP sad
trustworthy	2.61	2.62	1.99	1.04	0.63
friendly	2.63	2.64	2.08	1.02	0.73
polite	2.69	2.60	1.96	1.08	0.63
intelligent	2.56	2.84	2.33	1.05	1.03
hardworking	2.51	2.59	2.18	0.99	0.82
outgoing	2.54	2.62	2.21	0.79	0.49
snobbish	0.67	1.6	1.82	0.66	0.99
reliable	2.61	2.68	2.17	1.02	0.78
wealthy	2.41	2.72	2.1	0.92	0.94
funny	2.35	2.22	1.73	0.57	0.40
confident	2.70	2.74	2.13	0.83	0.51

Table 12. Mean scores character traits by English native speakers

Variables	GA happy	RP happy	SSE happy	GA sad	RP sad
trustworthy	2.6	2.8	1.8	1.8	1.2
friendly	2.8	2.6	1.8	1.6	0.4
polite	2.8	2.8	1.8	1.6	1
intelligent	2.4	3	2.4	1.4	0.8
hardworking	2.2	2.6	2.4	2	1.2
outgoing	2.8	2.6	2	0.8	0.2
snobbish	0.6	1.6	2	0.6	1.6
reliable	2.4	2.6	2	1.4	1
wealthy	2.2	2.6	2.2	1.2	0.8
funny	2.4	2	1.6	0.4	0
confident	3	2.8	1.8	1	0.6

Tables 11 and 12 above outline differences in mean scores for the given character traits according to participants' mother tongue. The happy GA voice did not reveal a clear pattern regarding the influence of respondents' L1 on their language attitudes. German native speakers reported higher scores for five out of ten dimensions, indicating that the distribution of greater mean scores was balanced between the two groups. Moreover, participants, regardless of their mother tongue, assigned their greatest values to the same character trait (*confident*). German native speaker assigned a mean of M=2.7, whereas L1 users of English recorded an average score of M=3; however, the lowest points differed in the two groups. Respondents with German as their mother tongue ascribed their minimum score to the dimension of dynamism, more precisely *funny*, with a mean of M=2.35. In contrast, English natives gave their lowest values to the field of social status, outlining an average score of M=2.2 for *hardworking* and *wealthy*.

A similar trend continued for the happy equivalent from England. English L1 users showed greater values for *trustworthy, polite, intelligent, hardworking* and *confident*, while German natives reported higher points for the remaining five adjectives; hence, no clear influence of a person's mother tongue could be detected regarding the happy speaker from England. Generally, differences remained smaller compared to the American voice (see Table 11 and 12). Both groups awarded their maximum scores to the category of the perceived intelligence of the speaker. Likewise, the two data sets distributed the minimum values to the same character trait (*funny*), outlining means from M=2, from English natives, to M=2.22 by L1 users of German.

Interesting results can be obtained by a closer look at the columns, when observing mean scores for the SSE voice. German native speakers assigned lower points for all characteristics of the social status dimension. Overall, discrepancies ranged from 0.07 to 0.33. The calculations of the sad US sample showed that German native speakers reported less favorable attitudes towards all character traits except *funny*. Great discrepancies were observable, especially for *trustworthy.* While L1 users of German allocated a mean of M=1.04, English natives awarded a total of M=1.8, resulting in a remarkable difference of 0.76. Character traits, belonging to the field of warmth, revealed considerable discrepancies as well; however, this pattern did not apply to the equivalent from England. Respondents with German as their mother tongue recorded greater mean scores for *friendly*, *intelligent*, *outgoing*, *wealthy* and *funny*, whereas the English natives reported more positive attitudes towards the other five characteristics; hence, no clear pattern regarding the influence of participants' mother tongue was recognizable for the sad speaker from England.

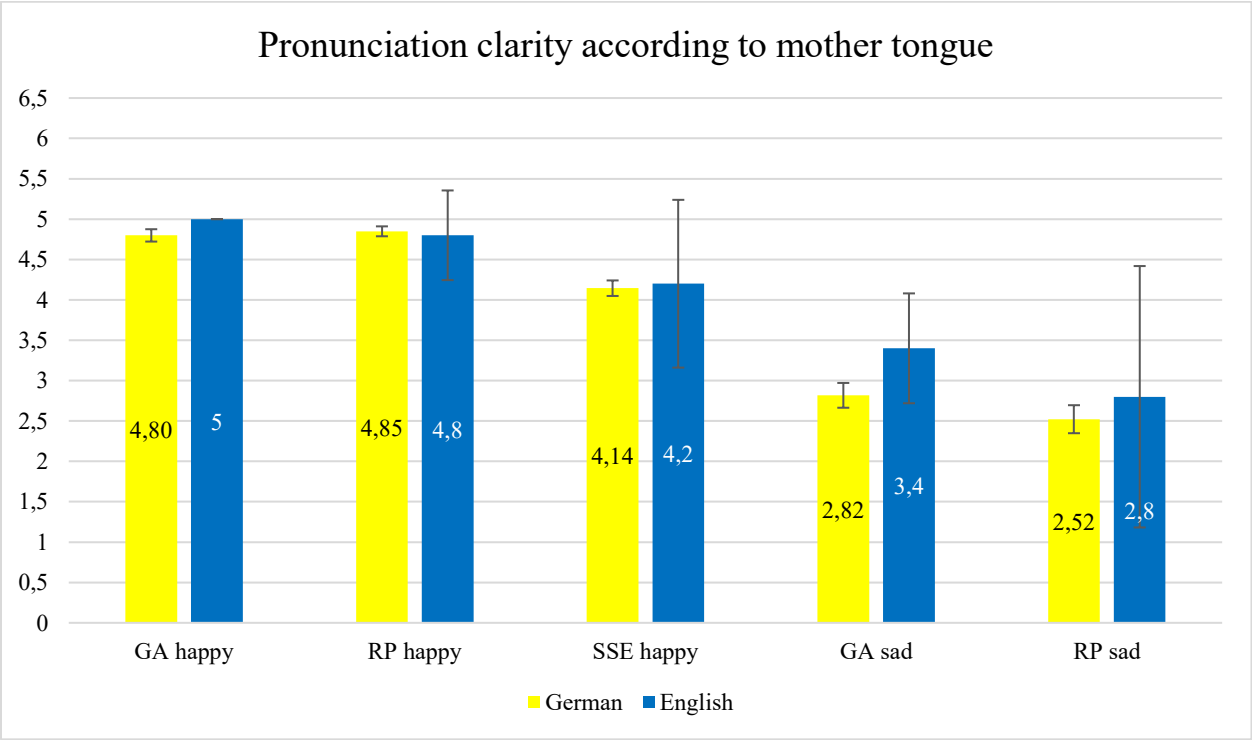


Figure 30. Comparison pronunciation clarity according to mother tongue

The illustration above (see Figure 30) shows the distribution of mean scores for pronunciation clarity based on respondents' mother tongue. The blue bars represent L1 users of English, while the yellow ones symbolize German native speakers. As visible in Figure 30, the maximum score, 5 out of 5 points, was achieved by the conversational GA recording, as assessed by the English native speakers. In comparison, L1 users of German reported slightly more negative

attitudes with a still high average of M=4.8, adding up to a difference of 0.2. An analysis of the happy RP equivalent showed that participants with German as their mother tongue assigned marginally higher values (M=4.85) than the English native group (M=4.8). The SSE speaker was overall ranked in third place. L1 users of English described the pronunciation of the respective speaker as clearer (M=4.2) compared to German native speakers (M=4.14). The sad GA sample revealed the largest discrepancy in means of 0.58 points with higher values distributed by English natives (M=3.4). Both groups assigned their lowest values to the sad RP voice; nevertheless, participants, who reported English as their L1, considered the speaker's pronunciation as more intelligible (M=2.8) than the German data set (M=2.52).

To conclude, the calculation of average scores revealed that participants with L1 English tended to outline greater values for every variety except the conversational GA voice.

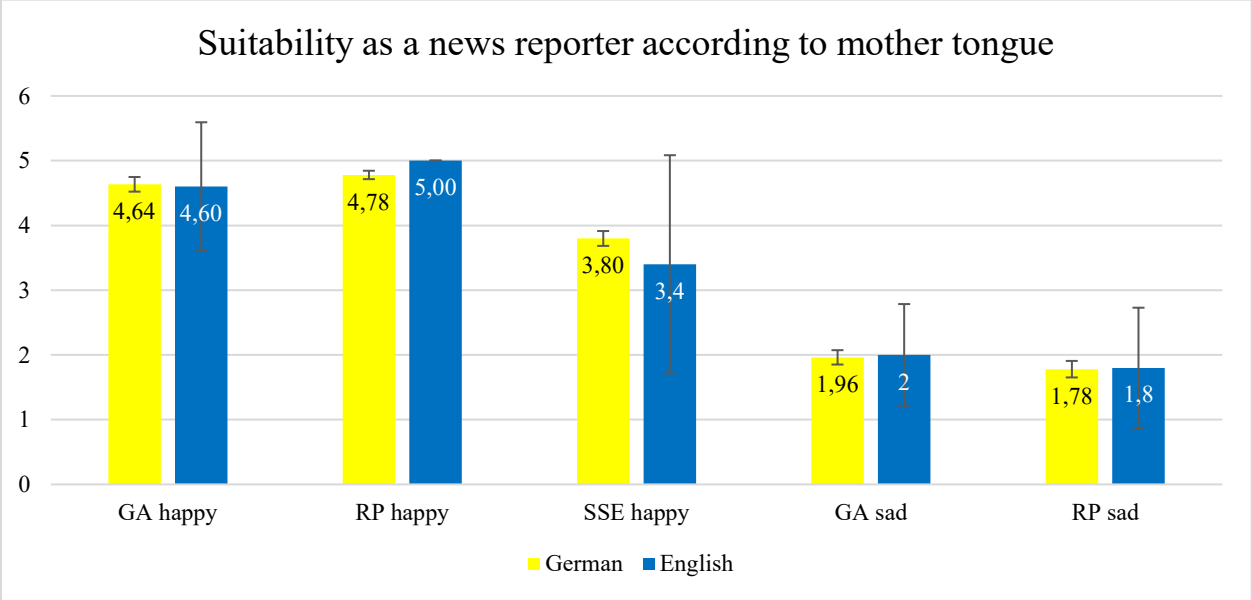


Figure 31. Comparison suitability as a news reporter according to mother tongue

Figure 31 illustrates mean scores regarding the perceived suitability as a news reporter, focusing on the influence of participants' mother tongue. The happy RP voice was ranked first, with scores ranging from M=5.0, assigned by native English speakers, to M=4.78, reported by respondents with German as their L1, resulting in a difference of 0.22. The second highest values were obtained by the conversationally read GA sample. Once again, German natives described the concerned example as more appropriate for reading the news (M=4.64) than the English L1 data set (M=4.6); however, a discrepancy of 0.04 points in mean scores remained marginal. Participants put the SSE recording in third place. While German native speakers distributed a mean of M=3.80 out of 5 possible points, L1 users of English ascribed a score of

M=3.4; hence, it became evident that respondents with German as their mother tongue tended to provide higher values for the happy voices; nevertheless, this trend changed in the evaluation of the sad audio files. A first glimpse at the bars of the melancholic samples revealed that slightly greater scores were ascribed by the English native group. Differences ranged from 0.02 to 0.04 points. Whereas the sad GA speaker reached a maximum of 2.0 points in the calculations of L1 English users, German natives assigned an average value of M=1.96. While the happy voice from England was considered most suitable for reporting the news, the sad counterpart with scores of M=1.78 and M=1.8 ended in last place, suggesting unfavorable perceptions by both data sets.

In short, results proved that participants' mother tongue influenced the evaluation of a good news reporter. English natives ascribed higher scores for all varieties except the happy GA and SEE sample; nonetheless, both groups distributed their greatest to the conversational RP sample, even though the majority of L1 users of English reported to speak American English.

As stated at the beginning of this section, remarkably fewer respondents reported English as their L1 (n=5), while the majority belonged to the German native group (n=159); thus, remarkable differences emerged between the groups' confidence intervals. Wide intervals indicate a considerable uncertainty, thereby making the discrepancies not statistically significant.

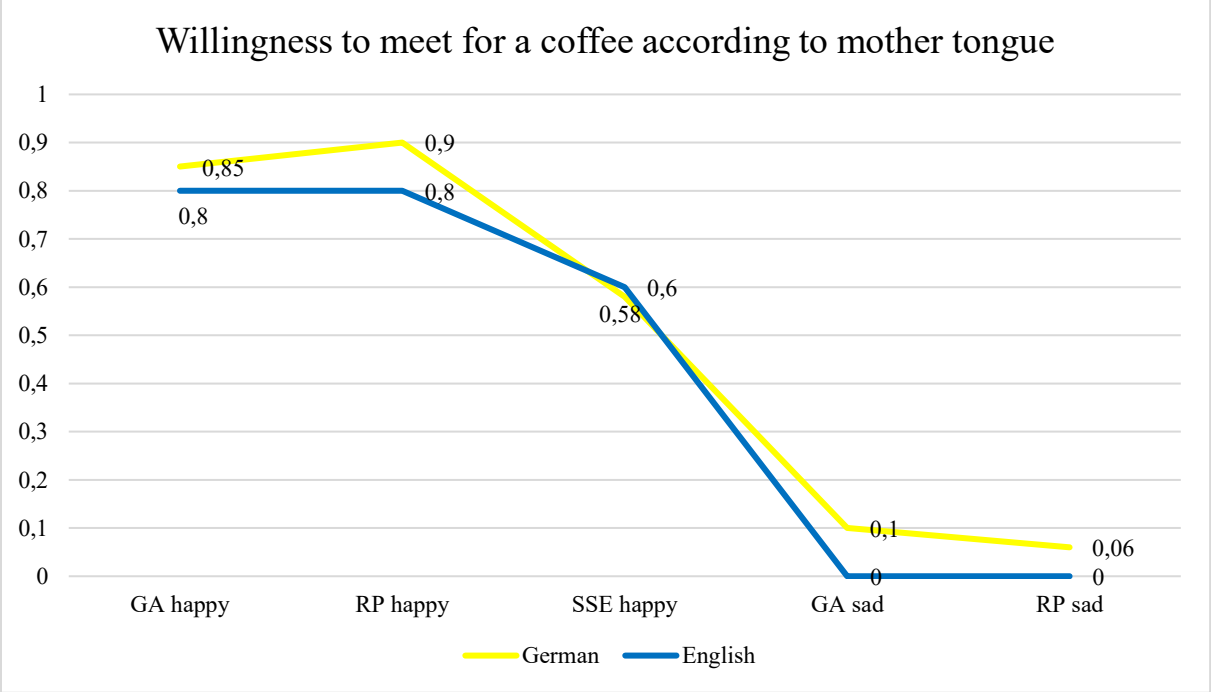


Figure 32. Comparison willingness to meet for a coffee according to mother tongue

The illustration above (see Figure 32) presents mean scores of respondents' willingness to meet the respective speakers for a coffee according to their mother tongue. Similar principles were applied to gather numerical data as discussed in chapters 7.6 and 7.7. A closer examination revealed that L1 users of German, represented by the bars in yellow, showed a higher willingness to meet the individual voices than the English native speakers besides the SSE recording. Overall, the greatest value was obtained by the happy RP sample with a mean of M=0.9 out of 1 possible points, assessed by L1 users of German. In comparison, the English group showed an average of 0.8, resulting in a discrepancy of 0.1. A similar pattern emerged for the conversational GA sample, with scores from 0.85 to 0.8, indicating an even more marginal difference of 0.05. Generally, respondents were slightly more willing to have coffee with the English than the American speaker. As previously stated, the SSE sample remained the only instance, where English natives outlined more positive attitudes than the German data set; however, distinctions stayed marginal with 0.02 points. Interesting insights can be gained by a more detailed observation of the sad voices. As Figure 32 indicates, L1 users of English were not willing to meet either of the melancholic samples; therefore, no bars appear in the graphical illustration. Similarly, German natives also reported low means, slightly favoring the American speaker (M=0.1) over the recording from England (M=0.06).

Briefly, participants with German as their L1 reported more positive attitudes regarding their desire to have coffee with the chosen speakers except for the SSE file. This was especially evident in the evaluation of the melancholic samples, where no English native participant wanted to meet the two speakers.

7.9 Analysis on the influence of participants' self-identified English

Participants reported using three different varieties: RP, GA and Irish English (see Figure 33). The majority identified with RP (n=94), closely followed by GA (n =85). Only one respondent claimed to speak Irish English. It needs to be emphasized that this section includes native and non-native speakers of English, which may lead to differences in the active use of the language as well as variance in pronunciation and proficiency level. The personally used variety of non-native English speakers is regarded as self-identified English, meaning that their pronunciation might not fully depict a native-like model but aims to imitate patterns of the target language.

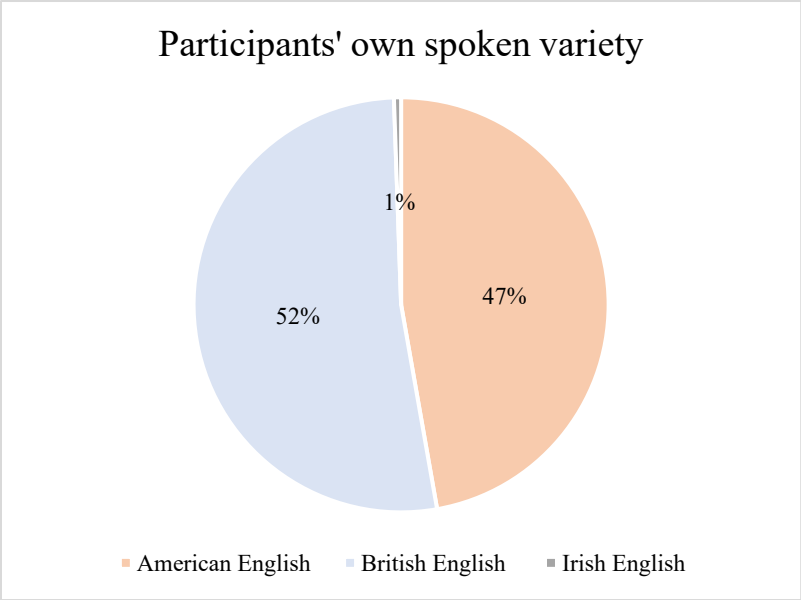


Figure 33. Participants‘ self-identified English

Table 13. Mean scores character traits by GA speakers

Variables	GA happy	RP happy	SSE happy	GA sad	RP sad
trustworthy	2.68	2.53	1.99	1.13	0.54
friendly	2.71	2.39	2.04	1.14	0.6
polite	2.76	2.43	1.86	1.19	0.6
intelligent	2.71	2.79	2.39	1.13	0.78
hardworking	2.56	2.52	2.32	1.14	0.6
outgoing	2.67	2.51	2.2	0.86	0.41
snobbish	0.49	1.87	1.82	0.71	1.21
reliable	2.65	2.51	2.19	1.12	0.71
wealthy	2.55	2.66	2.14	1.06	0.8
funny	2.34	1.91	1.65	0.59	0.31
confident	2.8	2.71	2.13	0.86	0.44

Table 14. Mean scores character traits by RP speakers

Variables	GA happy	RP happy	SSE happy	GA sad	RP sad
trustworthy	2.56	2.73	1.98	0.98	0.75
friendly	2.58	2.85	2.09	0.92	0.82
polite	2.64	2.75	2	0.98	0.67

intelligent	2.44	2.88	2.29	0.94	1.24
hardworking	2.43	2.71	2.1	0.89	1.01
outgoing	2.47	2.72	2.18	0.71	0.54
snobbish	0.80	1.33	1.83	0.6	0.86
reliable	2.6	2.81	2.14	0.94	0.83
wealthy	2.25	2.75	2.09	0.8	1.06
funny	2.39	2.48	1.77	0.55	0.43
confident	2.61	2.76	2.1	0.74	0.56

A comparison of Table 13 and 14 revealed a remarkable influence of participants' own spoken variety on their evaluations of the given character traits. Through a first glimpse at the column presenting mean scores of the happy GA voice, it becomes evident that respondents, who personally use American English, reported higher scores for all adjectives besides *funny*. The overall greatest discrepancy appeared in the analysis of *wealthy.* While respondents, favoring GA, assigned a mean score of M=2.55, those, preferring RP, distributed an average of M=2.25, resulting in a difference of 0.3 points; nevertheless, the perceived politeness of the concerned speaker reached the highest scores by both data sets. Participants with GA as their self-identified English distributed a mean of M=2.76, whereas those with RP as their preferred English reported an average of M=2.64. These results indicated that respondents show more positive attitudes towards their own spoken variety. This trend also continued for the conversational speaker from England, who obtained greater values for all character traits by the group personally using the variety from England. Considerable differences emerged, particularly in the field of *funny*. While GA users reported a mean of M=1.91, the RP group showed a higher score with M=2.48, resulting in a difference of 0.57 points. Greatest census between both data sets appeared in the voice's perceived confidence, outlining mean scores from M=2.71 to M=2.76. As visible in the tables above, the SSE sample revealed a smaller variance among all characteristics.

The sad speaker from the US demonstrated similar findings as the happy counterpart. All character traits were evaluated more positively by those using the American English variety compared to the English group. Differences varied between 0.04 and 0.26 points, mirroring the results of the conversational GA voice. Interestingly, both groups distributed their highest ratings to the politeness of the concerned speaker, but with a difference of 0.21. Another similarity was the low rating of the category of *funny*, covering average scores from M=0.59,

assigned by GA users, to M=0.55 by participants preferring the variety from England. Lastly, the influence of respondents' own spoken variety also applied to the sad voice from England. Here, respondents using American English distributed lower scores for the concerned sample than the RP group, outlining discrepancies from 0.07 to 0.46.

In summary, participants' self-identified English had a remarkable impact on their evaluations regarding the character traits of the concerned speakers.

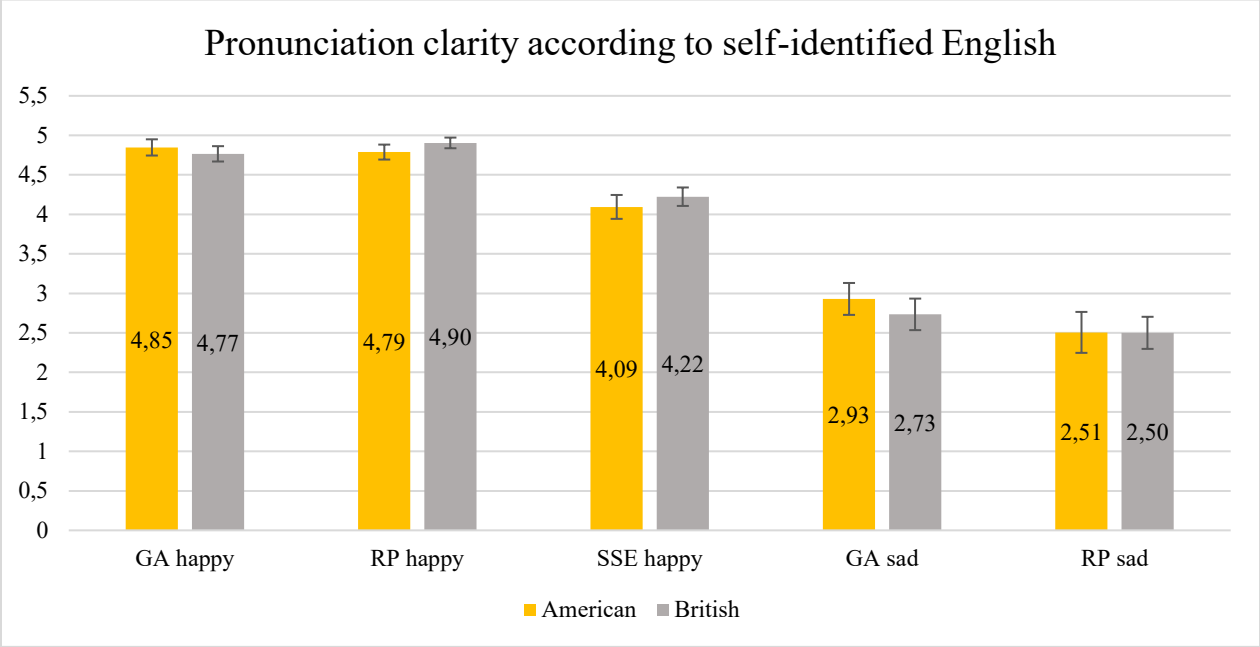


Figure 34. Comparison pronunciation clarity according to participants' self-identified English

By a closer look at Figure 34, interesting results can be observed. While the grey bars indicate the distribution of mean scores ascribed by participants, who claimed to use the RP variety, the ones in orange present average values from respondents with GA as their preferred choice. As immediately visible, both GA samples, regardless of emotional tone, received higher scores by American English speakers; nevertheless, a comparison of the American files showed a remarkable difference of 1.92 in means, indicating that mood was a consistently influential factor. The happy GA voice reached its peak score of M=4.85 within this group. In contrast, the lowest score of GA participants emerged in the analysis of the sad RP speaker.

Evaluations by the group with RP as their own spoken variety showed a different trend. While they also described most favorable attitudes in terms of pronunciation to the speaker from England (M=4.9), results for the sad voices diverged. Although users of American English still distributed higher values to the melancholic GA sample, respondents preferring RP claimed greater scores for the sad American (M=2.73) than the English counterpart (M=2.5). Both

groups recorded similar mean scores towards the SSE voice, which was put in the middle rank with values ranging from M=4.09 to M=4.22; thus, a slight preference of participants personally using the RP variety appeared. The overall greatest numerical discrepancy of 0.2 was obtained by the melancholic speaker from the US. In contrast, only minimal differences (0.01 points) occurred in the calculations of the RP counterpart. Both happy recordings were evaluated similarly by their corresponding groups. Regarding the conversational GA sample, participants with American English as their preferred variety assigned a mean of M=4.85, whereas respondents who prefer RP distributed lower values (M=4.77), differing by 0.08. A marginally higher variance became evident for the happy voice from England with values covering M=4.79 to M=4.9 points.

In short, participants' own spoken variety appeared to partly influence their evaluations of the individual speakers' pronunciation clarity. While respondents favoring the American variety, showed more positive outcomes for the GA samples, speakers of RP did not display a clear pattern. Users of the concerned variety reported higher values for the happy RP file compared to the American group, but lower scores in respect to the melancholic recording from England.

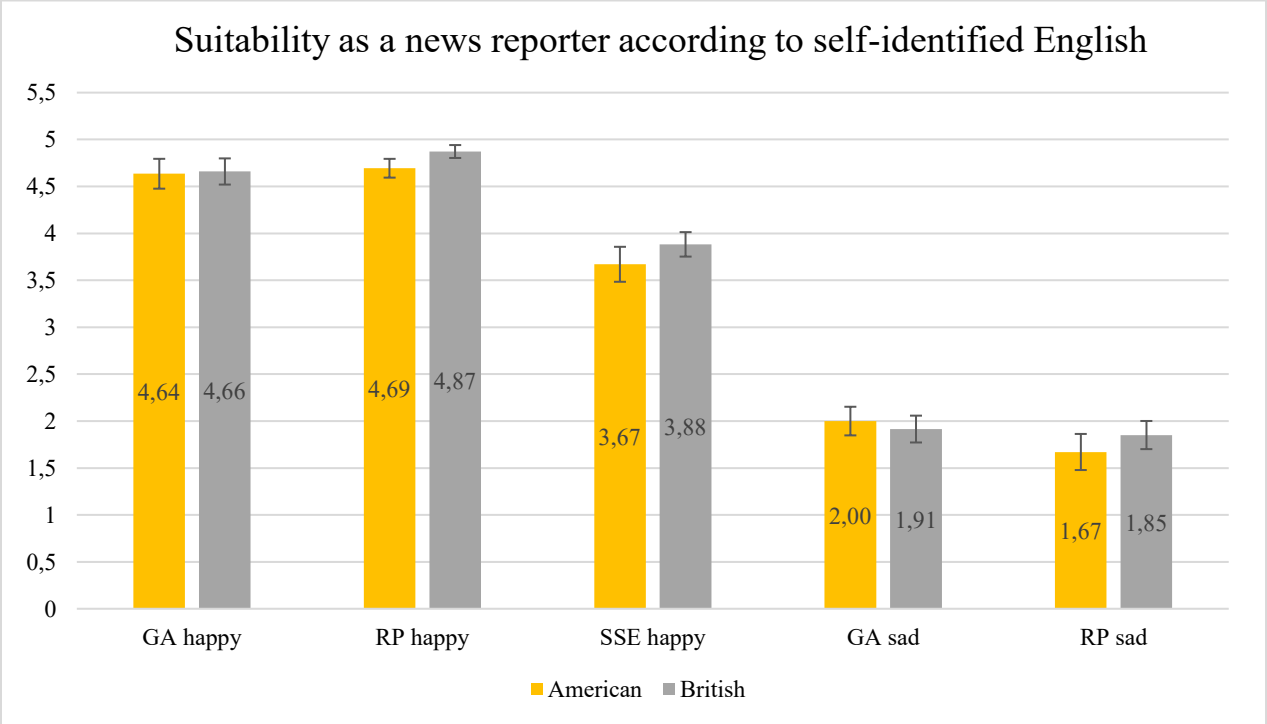


Figure 35. Comparison suitability as a news reporter according to participants' self-identified English

Figure 35 shows the distribution of mean scores based on participants' own spoken variety. The color scheme follows the same principles as in the previous analysis. A first glance already

reveals that respondents who personally use RP assigned higher values to all speakers except the sad GA voice. Marginal discrepancies appeared in the comparison of the evaluations regarding the conversational American voice. Those with American English as their most favorable variety ascribed a mean of M=4.64, while the RP group outlined an average of M=4.66. Both data sets rated the happy speaker from England as most suitable for reading the news with the overall maximum score (M=4.69), given by participants using GA. A mean score of M=4.87 from respondents who use the accent from England led to a difference of 0.18. Similar to the analysis of the pronunciation clarity, the SSE voice was put in third place with average points ranging from M= 3.67 to M=3.88. Participants, speaking American English, awarded higher scores for the sad GA speaker (M=2.0) compared to the English group (M=1.91). Means ranging from M=1.67 to M=1.85 for the melancholic RP sample only sufficed for the last place and was therefore considered as inappropriate for reading the news.

To conclude, it became transparent that users of RP assigned greater means for all speakers except the sad voice from the US in terms of the perceived suitability as a news reporter compared to GA speakers.

Confidence intervals showed marginal differences, indicating a high degree of statistical certainty and reliability.

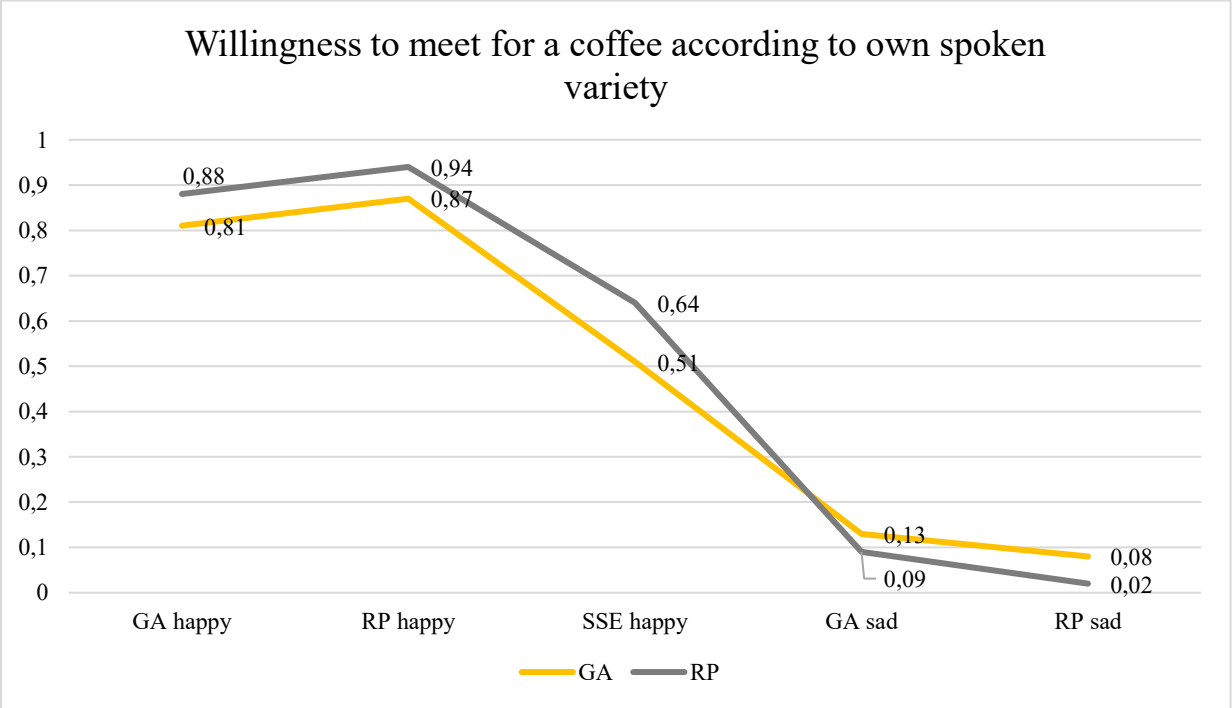


Figure 36. Comparison willingness to meet for a coffee according to participants‘ self-identified English

Figure 36 above outlines differences in participants' willingness to meet the individual speakers of the survey based on their own spoken variety. To obtain numerical data, the same principles were applied as from chapter 7.4 onwards regarding the analysis of having coffee with the concerning voices. Overall, respondents with RP as their preferred variety reported higher values for all happy samples. The conversational American speaker presented values ranging from M=0.81 to M=0.88, whereas the highest mean M=0.94 out of 1 possible points appeared in the calculations of the conversational sample from England by respondents preferring the RP variety. In comparison, American English users distributed an average of M=0.87; thus, as visible in Figure 36, both data sets were most willing to meet the happy RP speaker for a coffee. The SSE sample was ranked third. While respondents, favoring GA, assigned an average value of M=0.51, RP speakers showed M=0.64 points.

An opposite trend occurred in the analysis of the sad speakers. Participants who claimed to use GA ascribed higher average scores compared to RP users; A closer look at Figure 36 also revealed that smaller discrepancies emerged for the melancholic voices compared to the happy audio recordings. Participants favoring the GA variety distributed a mean value of M=0.13 for speaker 4, while a slightly lower score of M=0.09 was reported by RP speakers. Lastly, based on means ranging from M=0.02 to M=0.08, respondents were not willing to meet the melancholic voice from England, suggesting negative perceptions in terms of the perceived solidarity and warmth.

7.10 Judgements according to participants' age

Participants' age ranged from 15 to 89 years. The median was at 29 years; therefore, to avoid potential influences of the data set, immediately centering around the median area, this chapter focuses on the following groups: 15 to 25 years and 35 to 89 years. Figure 37 presents the distribution of the forementioned dimensions. While 55% (n=63) reported to be younger than 25 years, 45% (n=52) stated to be over the age of 35.

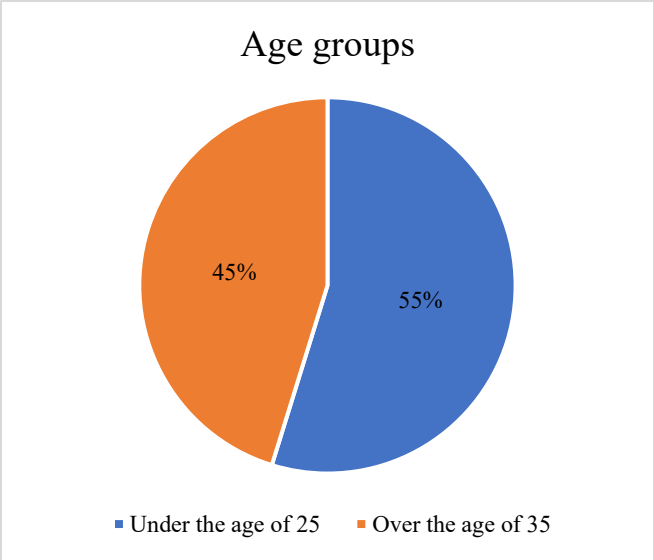


Figure 37. Age groups

Table 15. Mean scores character traits by participants under 25

Variables	GA happy	RP happy	SSE happy	GA sad	RP sad
trustworthy	2.7	2.56	2.03	1.17	0.68
friendly	2.6	2.49	1.95	1.1	0.63
polite	2.81	2.48	1.86	1.16	0.60
intelligent	2.65	2.81	2.27	1.15	1.02
hardworking	2.56	2.49	2.19	1.08	0.87
outgoing	2.52	2.46	2.1	0.83	0.35
snobbish	0.67	1.73	1.79	0.68	1.13
reliable	2.68	2.62	2.17	1.13	0.79
wealthy	2.49	2.71	2.03	1.16	0.95
funny	2.22	2	1.54	0.54	0.25
confident	2.73	2.65	2.06	0.87	0.49

Table 16. Mean scores character traits by participants over 35

Variables	GA happy	RP happy	SSE happy	GA sad	RP sad
trustworthy	2.43	2.77	2	0.98	0.69
friendly	2.59	2.82	2.12	1.04	0.86
polite	2.52	2.78	2.02	1.02	0.7

intelligent	2.46	2.9	2.35	1.02	1.12
hardworking	2.33	2.76	2.2	0.94	1
outgoing	2.51	2.78	2.24	0.78	0.61
snobbish	0.76	1.37	1.8	0.69	0.9
reliable	2.53	2.8	2.18	0.92	0.86
wealthy	2.22	2.84	2.18	0.65	0.96
funny	2.47	2.53	1.94	0.63	0.55
confident	2.57	2.86	2.22	0.75	0.57

The tables presented above (see Tables 15 and 16) demonstrate mean scores regarding a set of character traits, depending on participants' age. An examination already reveals that respondents under 25 attributed slightly higher values for the happy GA speaker across all fields, except *snobbish* and *funny*, compared to the older group. Respondents under 25 showed their highest scores in the evaluations of *polite* (M=2.81), whereas people over 35 reached their peak for *friendly* (M=2.59). Notably, both adjectives belong to the field of warmth. In contrast, an opposite trend was observable for the speaker from England. Here, participants older than 35 reported consistently greater values than the younger data set. Unlike in the evaluations of the conversational US voice, respondents did not ascribe their highest scores for the same category. While younger people expressed the most favorable attitudes towards the perceived intelligence of the speaker (M=2.81), associated with the category of social status, those over 35 assigned their maximum score for *confident* (M=2.86), falling under the dimension of dynamism. Overall, greater differences could be observed in the comparison to the happy equivalent from the US, ranging from 0.09 to 0.53. Similarly to the English sample, the SSE file achieved higher ratings from people over 35, except for *trustworthy*. As seen in the analysis of the happy GA file, respondents younger than 25 revealed a greater number of points for the sad variant, again with the exception of *funny*. In this case, participants under 25 assigned their maximum value for the speaker's trustworthiness (M=1.17) while the older group ranked their friendliness highest (M=1.04). Both were included in the dimension of warmth. Lastly, the older group, which assigned greater scores for all character traits, showed a preference for the melancholic voice from England compared to the younger data set.

In summary, participants' age had an influence on the evaluations of the specific character traits, as clear preferences among the groups were detectable.

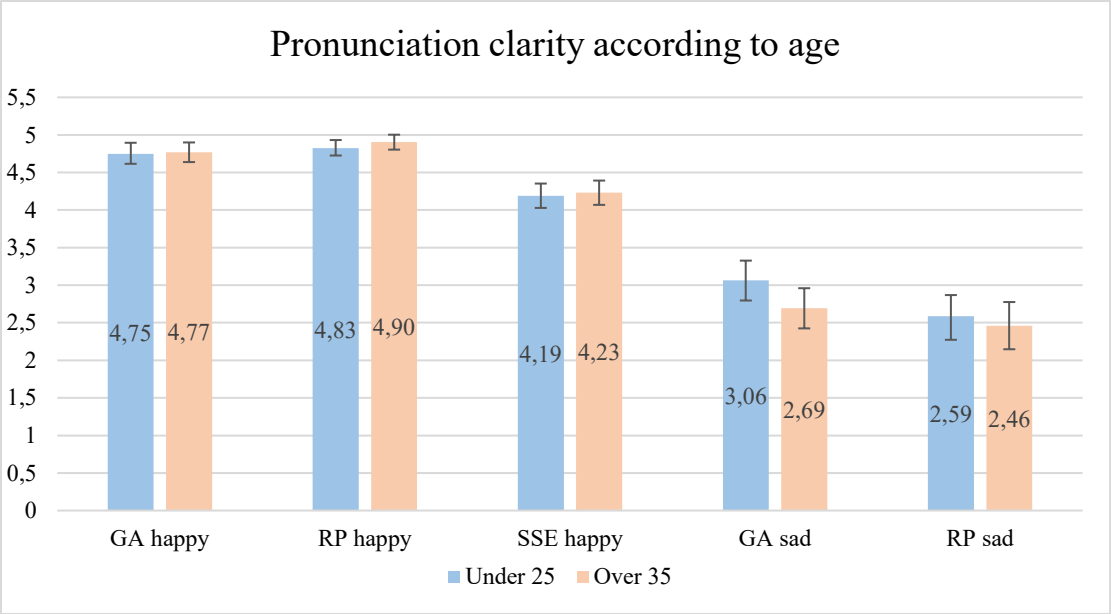


Figure 38. Comparison pronunciation clarity according to participants' age

Figure 38 above illustrates differences in the evaluation of the speakers' pronunciation clarity based on respondents' age. While the blue bars represent the group under 25 years, the orange ones indicate responses from participants older than 35. As the diagram shows, the first three recordings, namely GA, RP and SEE in a happy mood, did not exhibit remarkable differences. Both groups assigned their overall highest scores to the conversational speaker from England with means of M=4.83 among the younger data set and M=4.9 by participants over 35.

For the American speaker, ratings from M=4.75 (under 25) to M=4.77 (over 35) appeared. A comparable pattern emerged for the means of the SSE recording, where only marginal discrepancies appeared. With a mean of M=4.23 participants over 35 years rated the Scottish speakers' pronunciation as slightly clearer than the age group under 25 (M=4.19); however, this trend changed for the sad voices. The sad GA sample traced the greatest discrepancies based on means (0.63 points). Whereas respondents under 25 assigned values of M=3.06, the group older than 35 only assigned M=2.69 points and thus, considered the pronunciation as less intelligible. Similar results were revealed in the analysis of the sad speaker from England. Again, the younger age group ascribed higher scores (M=2.59) than the older data set (M=2.46), but differences narrowed down to 0.13.

Overall, as the bars together with the confidence intervals in Figure 38 show, no substantial differences appeared based on the influence of participants' age; nevertheless, a slight preference of the group younger than 25 occurred towards the sad voices, while an opposite pattern was detected for the happy voices.

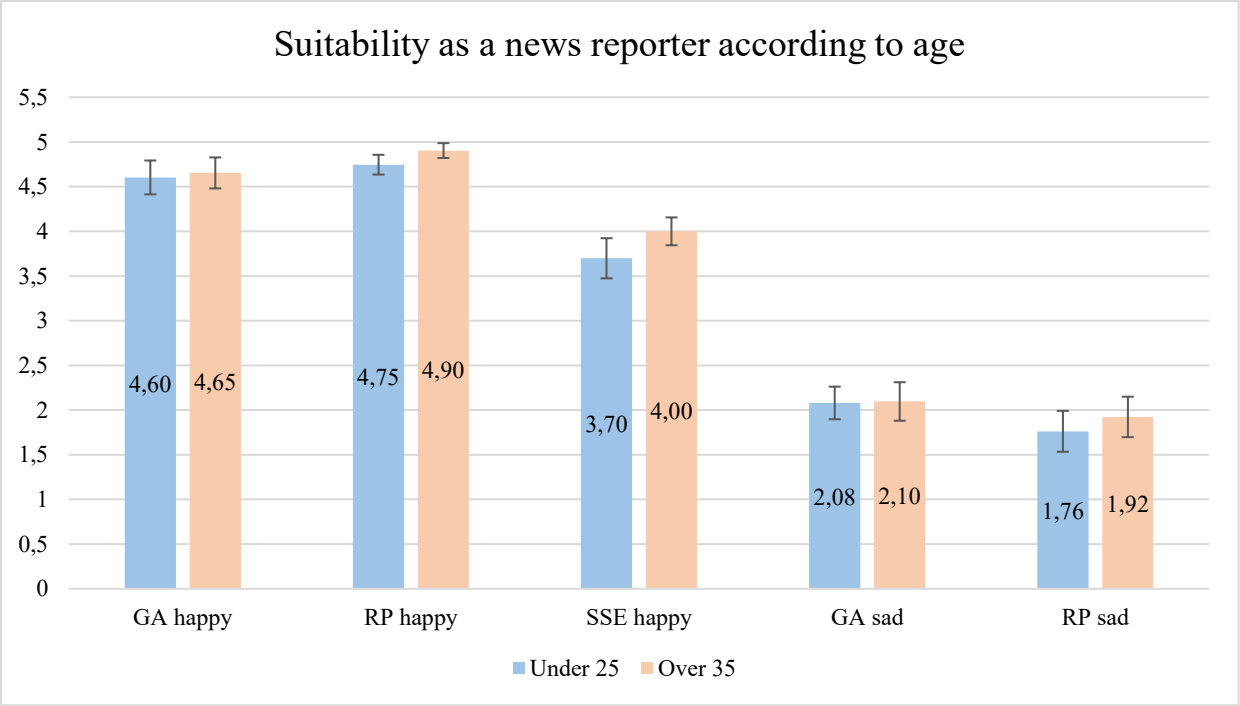


Figure 39. Comparison suitability as a news reporter according to participants‘ age

As the bar chart above illustrates (see Figure 39), an only minor discrepancy of 0.05 was found for the happy GA speaker among both groups in terms of the perceived suitability as a news reporter. A closer analysis reveals that participants over 35 perceived the respective voice as marginally more appropriate for reporting the news than the those under 25. The highest overall rating was given to the conversational RP sample with a mean of M=4.9 out of 5 possible points. In comparison to the American recording, a bigger difference could be observed for the happy RP speaker with slightly more favorable attitudes from respondents older than 35. This pattern continued in the case of the SSE file, outlining a variation of 0.3 points. While participants under the age of 25 assigned a value of M=3.7, those over 35 dedicated an average score of M=4. 0. Similar to the conversational American speaker, the sad counterpart did not display a considerable difference between the groups with a difference of only 0.02; however, means remained significantly lower for both melancholic samples, not exceeding M=2.1 points. The sad RP speaker was considered least suitable for reading the news, achieving values from M=1.92, distributed by the group over 35, to M=1.76, allocated by respondents under 25.

The greatest divergences were observed in the evaluations of both RP samples. In contrast, the American recordings, regardless of the mood change, did not show any considerable influence of participants’ age.

Overall, confidence intervals remained steady, as mainly similar values were obtained. Narrow intervals appeared for both dimensions, suggesting consistency across both groups and low variance in participants' answers.

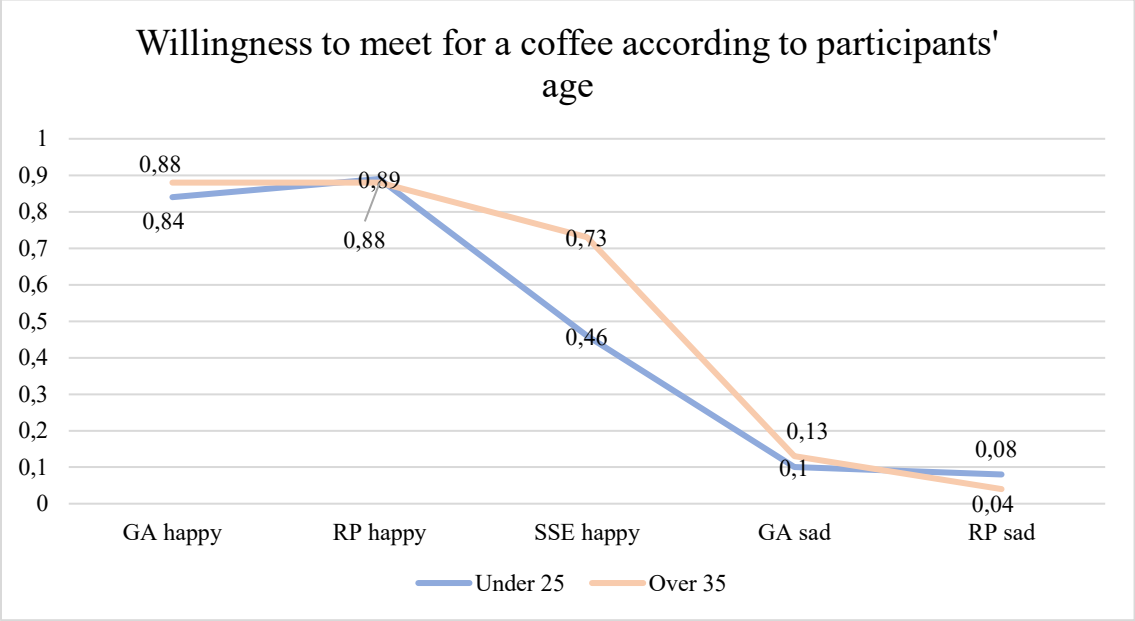


Figure 40. Comparison willingness to meet for a coffee according to participants' age

Figure 40 shows a considerable decline in evaluations among the age group under 25. While the happy speakers from the US and England received high values, the remaining three recordings were perceived less favorably in terms of willingness to meet for a coffee. In addition, it became evident that participants older than 35 assigned higher scores for four out of five samples. An almost identical score with a difference of 0.01 was obtained by the happy RP voice, indicating that age did not affect the attitudes towards this particular variety.

An opposite trend emerged for the SSE speaker. While people under 25 assigned a mean score of M=0.46, the older group over 35 allocated M=0.73; hence, the largest discrepancy of 0.27 appeared, suggesting that younger respondents were less inclined to have coffee with the speaker from Scottland. The sad counterpart also reported a minor difference of 0.03 points with mean scores covering M=0.1 to M=0.13 points. Similar to the other varieties, older respondents appeared more willing to meet speaker 4. Finally, the lowest values were garnered by the melancholic RP file; nevertheless, both groups only outlined a discrepancy of 0.04 points, suggesting that age did not considerable impact participants' perceptions.

In short, age seemed to particularly influence attitudes towards the SSE sample. In contrast, the happy GA and RP file as well as the sad sample from England and the US recorded only marginal variance.

7.11 Evaluations of *snobbish*

The descriptor *snobbish* was not assigned to any categories introduced in chapter 7.1 due to its negative connotation compared to the other adjectives; thus, it was considered more appropriate to discuss its findings in a separate section. Calculations followed the same principles as the previous analyses.

As visible in Figure 41, the SSE sample was perceived as the most snobbish with a mean score of M=1.82 out of 3 possible points. The happy speaker from England was ranked second, receiving an average value of M=1.58. In contrast, the sad counterpart was regarded as significantly less snobbish (M=1.02). A different trend emerged for the American English speakers. No differences were observable in the evaluations of the concerned samples, as both revealed a rating of M=0.66 points. These findings suggest that a change in emotional tone had an impact on the assessments of the speakers from England, as evidenced by a discrepancy of 0.56 points between the two files. Notably, snobbishness remained the only category for which better results were achieved by the melancholic than their conversational samples; however, identical values for the American English voices imply that mood and emotions did not influence participants' attitudes towards in this specific context.

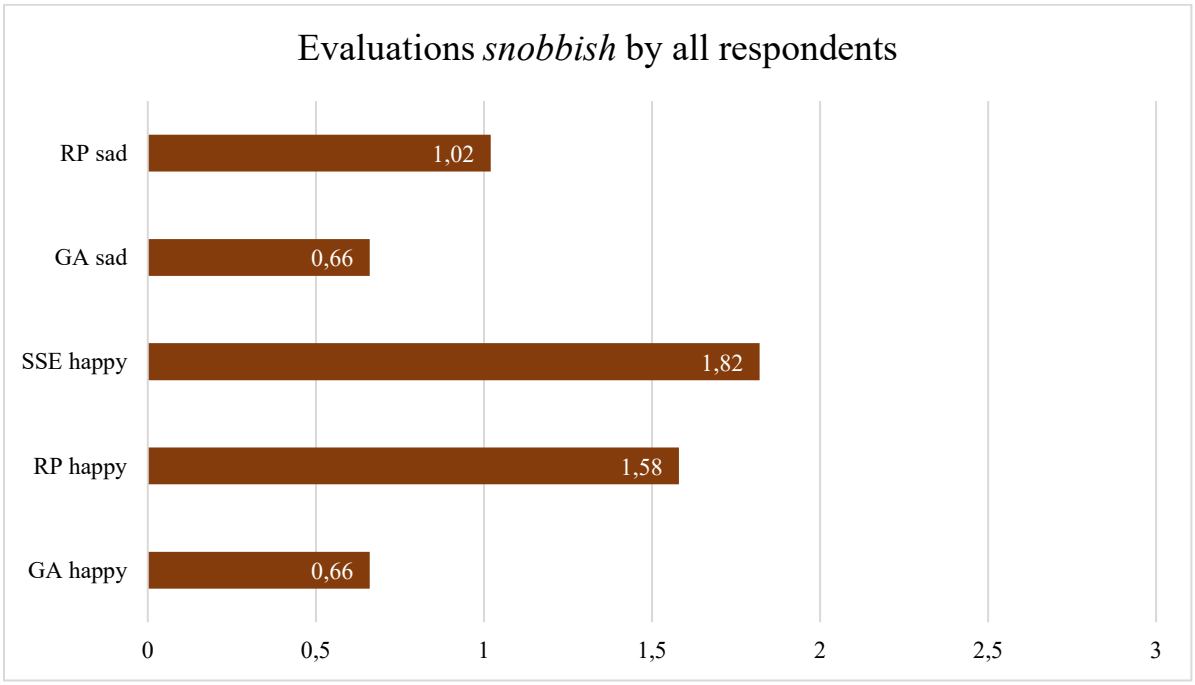


Figure 41. Evaluations snobbishness by all respondents

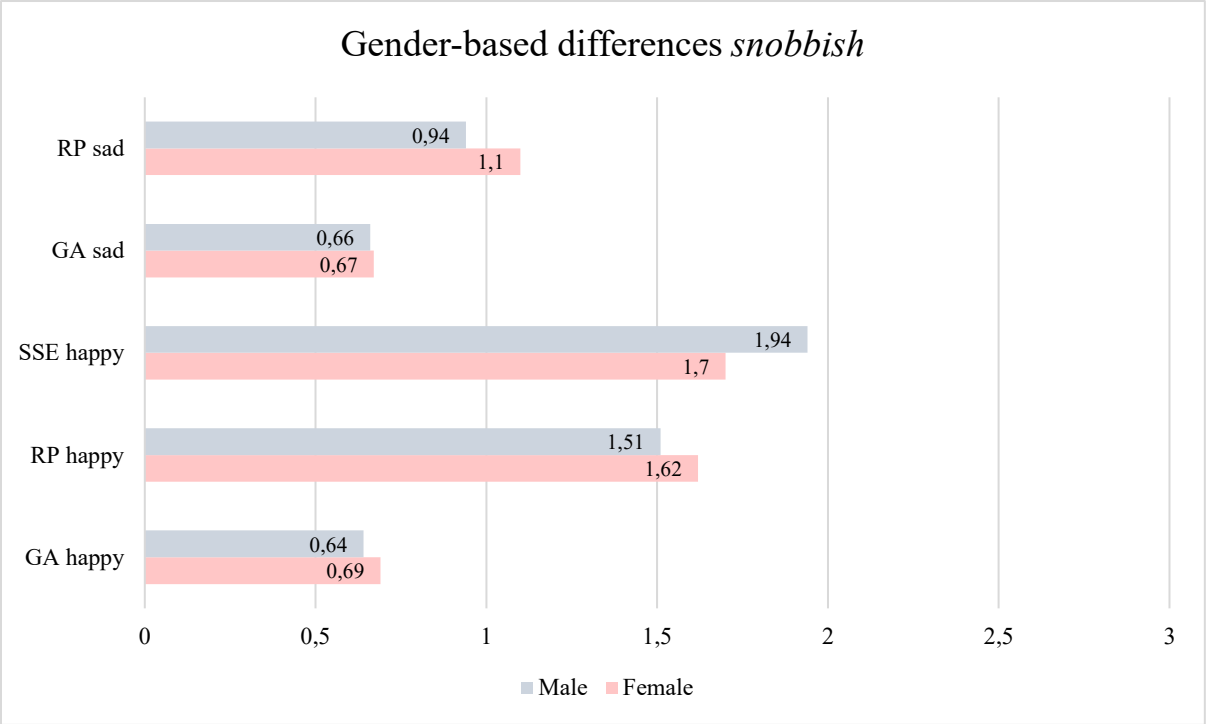


Figure 42. Gender-based differences snobbishness

The bar chart above (see Figure 42) illustrates gender-based differences in the evaluations of the speakers' perceived snobbishness. In general, men tended to report slightly lower values across all voices except for the SSE file. While females gave an average score of M=1.7, males ascribed the highest overall points of M=1.94.

An inverse pattern emerged for the mean values of the happy speaker from England, who was perceived as more snobbish by women (M=1.62) than men (M=1.51). The conversational GA voice received significantly lower points compared to the accent from England with only a slight difference of 0.05 between female (M=0.69) and male participants (M=0.64). A comparison of both American recordings, regardless of emotional tone, did not reveal any significant discrepancies. The sad version recorded values from M=0.66, by male respondents, to M=0.67, by the female group; hence, women reported marginally higher average points. Overall, respondents described the melancholic speaker from England as more snobbish than the American counterpart. While males assigned a score of M=0.94, women recorded a total of M=1.1. Generally, the sad RP sample was perceived as less snobbish compared to their happy equivalent, while mean scores of the US recordings were almost identical.

To conclude, the greatest gender-based variation in mean scores appeared for the SSE voice, with a difference of 0.24 points. In contrast, the sad GA sample yielded almost equal means. Due to the mainly minimal discrepancies, gender did not have a substantial impact on the perceived snobbishness of the speakers.

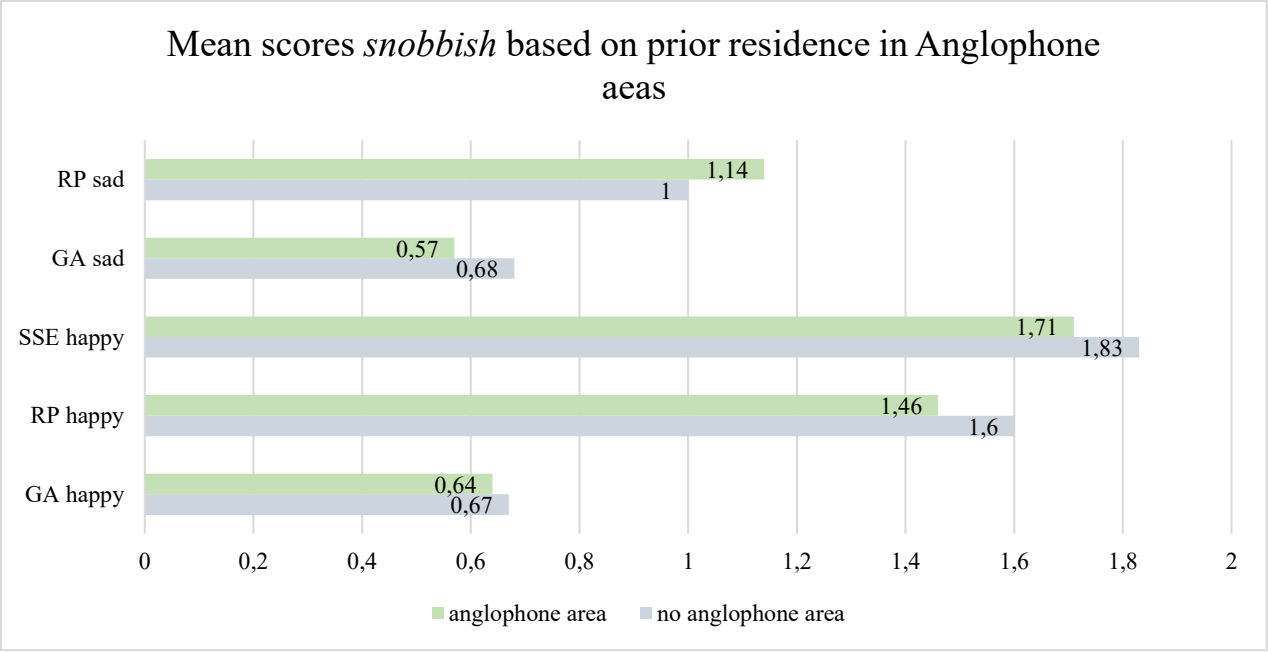


Figure 43. Differences snobbishness based on prior residences in Anglophone areas

Figure 43 presents average scores regarding participants' perceived snobbishness across the five chosen speakers. The analysis aims to further examine potential effects of prior residences in Anglophone areas. The bars in green represent respondents' answers who reported having lived in an English-speaking area, while the grey ones symbolize responses from those who have not resided in countries with English as an official language.

A more detailed analysis reveals that participants without a prior residence in Anglophone regions generally tended to rate the individual voices as more snobbish except for the sad RP sample. Once again, the SSE voice was considered the most snobbish by both groups with mean scores ranging from M=1.71 to M=1.83. The conversational speaker from England achieved an average of M=1.46 by people with Anglophone experience, compared to a mean of M=1.6 from those without. Substantially lower points of M=0.64 and M=0.67 appeared for the American counterpart, outlining a difference of only 0.03 points. Regarding the sad samples, it is notable that the US accent was described as less snobbish than its English equivalent. While internationally experienced respondents assigned a mean of M=1.14, the remaining group reported lower values (M=1.0), resulting in a difference of 0.14. Finally, emotional tone did not have an impact on the American speakers, as similar findings were observable for both speakers.

To conclude, the greatest variance in means between the two groups emerged for both RP voices, indicating that evaluations of these varieties was most affected by prior experiences living in Anglophone regions.

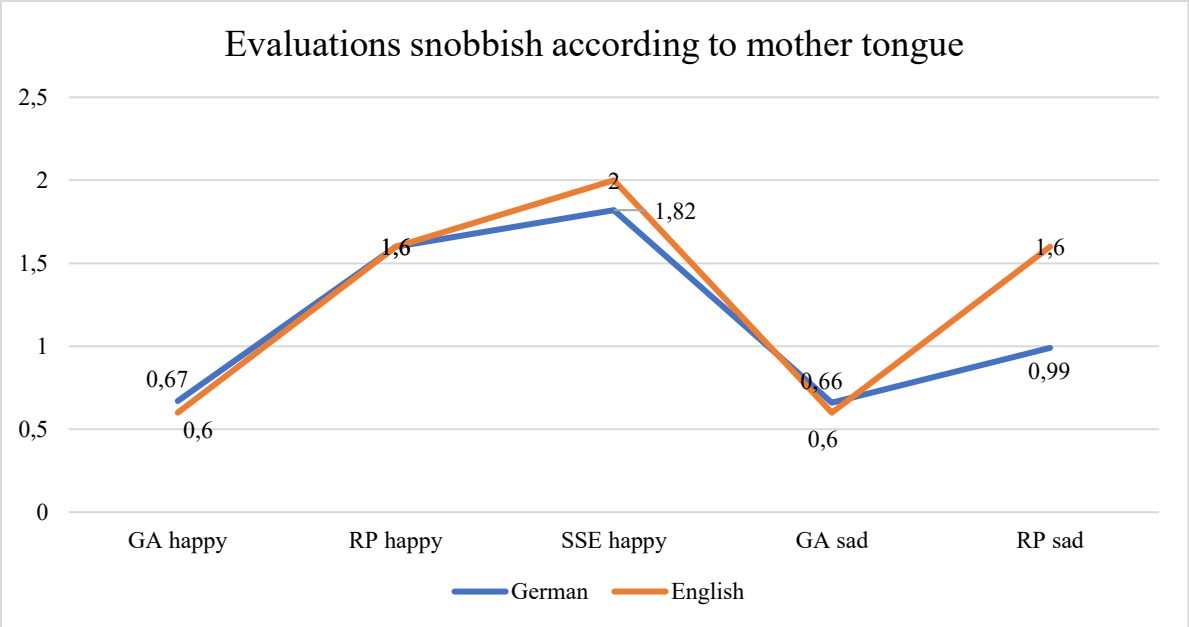


Figure 44. Differences snobbishness according to participants‘ mother tongue

Interesting results can be obtained by a closer look at Figure 44, illustrating the effects of participants' mother tongue on the judgements of the speakers' snobbishness. It is crucial to consider that the sample size of native English speakers was significantly smaller than that of L1 users of German.

Overall, the happy RP voice showed the greatest agreement between the two groups (M=1.6). Slightly larger discrepancies were observable for the conversational GA sample. While participants with German as their L1 reported an average score of M=0.67, L1 users of English assigned a slightly lower value of M=0.6. This trend continued in the evaluations of the SSE file, where differences were more remarkable. The concerned sample yielded the maximum score across all audio recordings with a mean of M=2.0, assigned by English native speakers, whereas respondents with German as their L1 only ascribed a total of M=1.82, resulting in a difference of 0.18 points. For the sad GA voice, discrepancies narrowed. The German group reported a mean of M=0.66, while L1 users of English assigned M=0.6 points. For the sad RP sample the German natives allocated M=0.99 points, but the English natives gave a considerably higher score of M=1.6, differing by 0.61.

In summary, respondents' mother tongue appeared to partially influence the perception of the speakers' snobbishness. Notable variances were recognizable for the sad voice from England; however, no significant effect was detected for both samples from the US as well as the happy RP speaker.

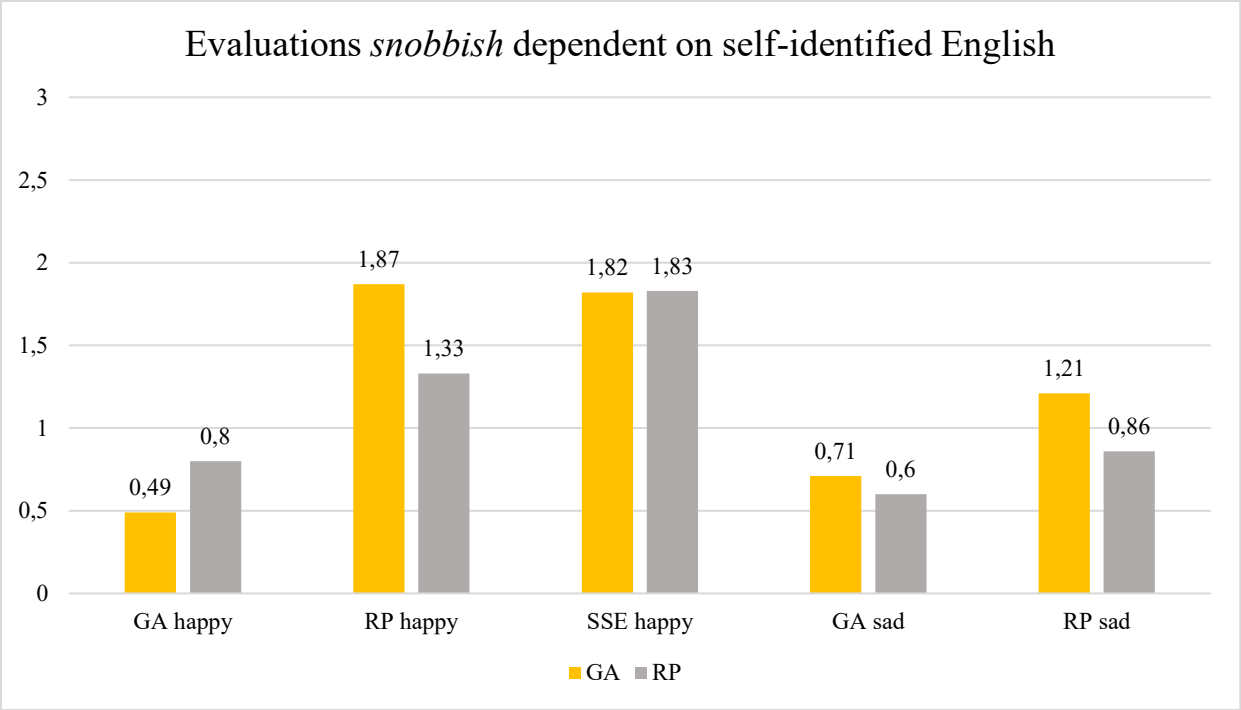


Figure 45. Differences snobbishness based on participants‘ self-identified English

Figure 45 above presents differences in mean scores of the speakers' perceived snobbishness based on participants' self-identified English. As immediately visible, several differences emerged. While the x-axis represents the individual speakers, the y-axis indicates the average number of points.

Overall, users of American English reported higher scores for three out five voices, specifically for the happy RP file and both sad samples. The conversational RP audio file achieved the overall highest scores of M=1.87 and was therefore regarded as most snobbish among all speakers. In contrast, participants with RP as their self-identified English, assigned a mean of M=1.33, resulting in a discrepancy of 0.54. A smaller difference of 0.35 appeared for its sad counterpart. For the melancholic voice from the US means ranged from 0.6 to 0.71, yielding a variation of 0.11. Regarding the remaining two files, respondents identifying with RP, reported greater values than speakers of GA. The conversational American English file received the lowest points with M=0.49 from users of GA, while participants RP speakers ascribed a greater score (M=0.8). The greatest agreement was observed in the analysis of the SSE speaker.

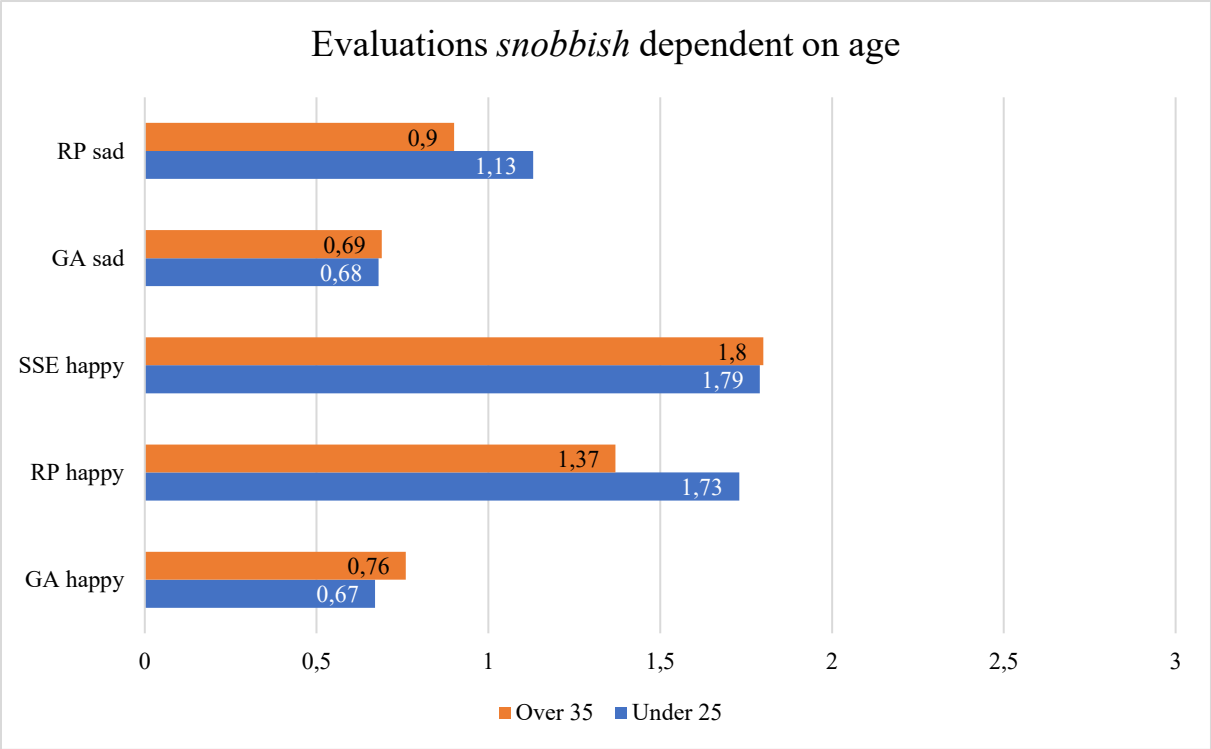


Figure 46. Differences snobbishness based on participants' age

As illustrated in Figure 46, participants' age influenced their assessments of the perceived snobbishness of certain speakers. The largest differences between respondents under 25 und those over 35 years were observable for the happy voice from England. While the younger group reported a mean of M=1.73, the older group perceived the recording as significantly less snobbish with M=1.37, resulting in a discrepancy of 0.36. The sad counterpart exhibited the second highest variance with a difference of 0.23 points. Again, older respondents regarded the melancholic recording as less snobbish. While people under 25 assigned an average value of M=1.13, those over 35 again ascribed a lower score (M=0.9). In contrast, both American English speakers reached remarkably lower scores. The happy US voice received an average of M=0.67 by the younger and a mean of M=0.76 by the older data set. Similar results were obtained for the melancholic counterpart with an average of M=0.68 from respondents under 25, and a mean of M=0.69 from those over 35, indicating that both American samples were evaluated more critically by the younger age group. The overall maximum scores emerged in the analysis of the SSE speaker.

To conclude, respondents' age primarily had an influence on the evaluations of the speakers from England. Apart from that, no considerable differences could be detected.

7.12 Rank ordering results

Figure 47 presents rank ordering results for all five varieties. In the online survey, participants were asked to rate each speaker on a scale from one to five stars, which were subsequently converted into numerical scores ranging from one to five points. These are represented on the y-axis in the illustration below.

In line with the dimensions evaluated above, the two sad speakers were perceived significantly less favorably than the conversational samples. The happy voice from England received the overall highest rating of M=4.49, securing first place. With a difference of 0.11, the conversational GA sample received the second highest score (M=4.38). The SSE audio file was ranked in third place (M=3.64). A sharp decline of 1.99 points emerged in the evaluations of the SSE and sad GA recording (M=1.65). The sad speaker from England received the overall lowest score (M=1.51).

These findings revealed clear preferences for the conversational voices, particularly the American and English voice. This also became evident through the distribution of maximum and minimum means, which were both obtained by RP samples with a considerable discrepancy of 2.98.

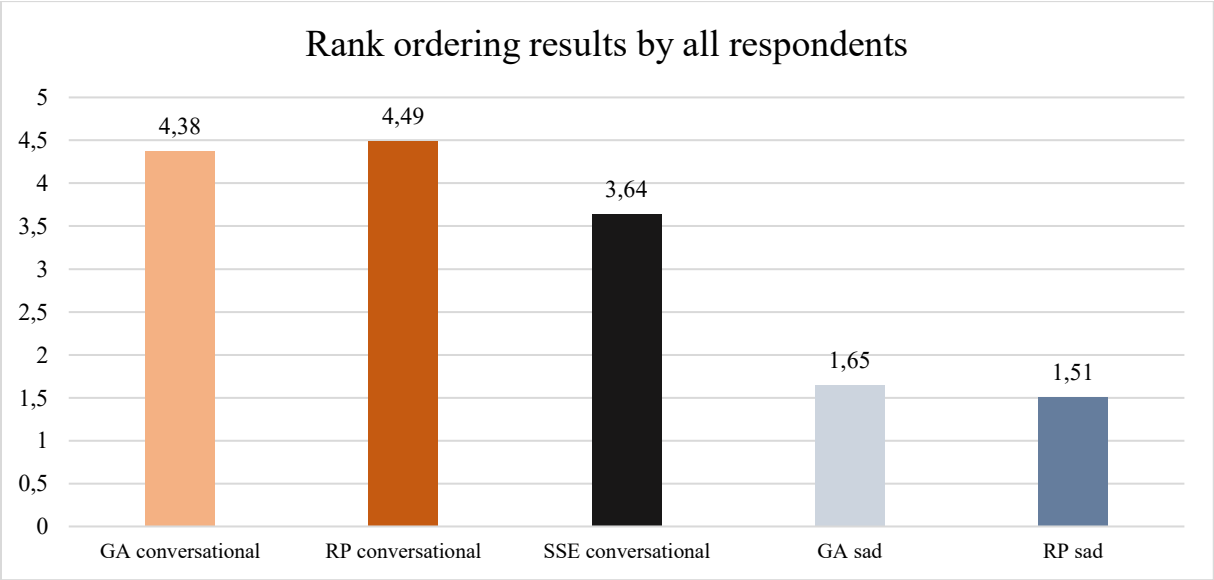


Figure 47. Rank ordering results by all respondents

The remaining part of this section examines potential influences of participants' gender, potential former residences in Anglophone countries, mother tongue and age, on their rank ordering results.

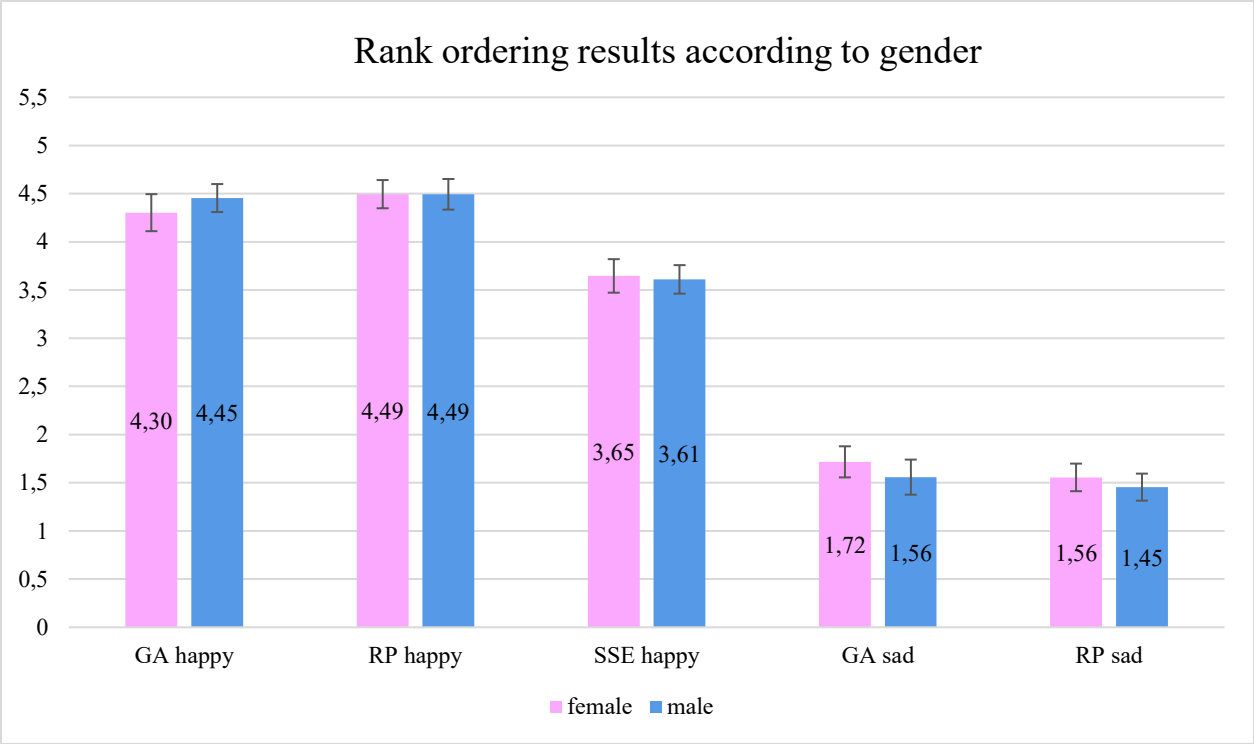


Figure 48. Rank ordering results based on participants's gender

The bar chart (see Figure 48) displays the distribution of mean scores from the rank ordering results based on participants' gender. As shown in chapter 7.6, 55% of the respondents (n=98) were women while 43% reported to be male (n=77). The remaining 2% (n=4) identified as none of these; however, this part focuses exclusively on the evaluations from women and men, since they represent the largest groups and thereby increase validity. The calculation of average points followed the same pattern as previously described.

A first glimpse at the illustration above reveals that overall, men and women did not report remarkable differences in their ranking of the individual varieties but some interesting results can be derived from the analysis of the happy RP voice. Both groups expressed comparable attitudes towards the concerned sample, as indicated by the identical mean of M=4.49 and similar confidence intervals, placing the sample in first place. In contrast, a slight discrepancy was detectable for the conversational American file. While male participants assigned a mean of M=4.45 out of 5 possible points, women tended to ascribe a lower score (M=4.3). The SSE voice was ranked third by both groups, with women giving higher scores (M=3.65) compared to men (M=3.61). The analysis of the sad GA speaker revealed the largest gender-based influence, amounting to a difference in means of 0.16 points. Similar to the Scottish voice, female respondents expressed more positive perceptions towards this variety (M=1.72) than the male group (M=1.56). Finally, the lowest values across both data sets were obtained for the

melancholic speaker from England. While women rated it with a mean of M= 1.56, men only reported M=1.45 points, resulting in a discrepancy of 0.11 points.

To conclude, gender did not significantly influence the rank ordering results of the five speakers; however, female participants expressed slightly more favorable evaluations, except for the happy GA sample. An equal score was garnered once by the happy RP file (M=4.49).

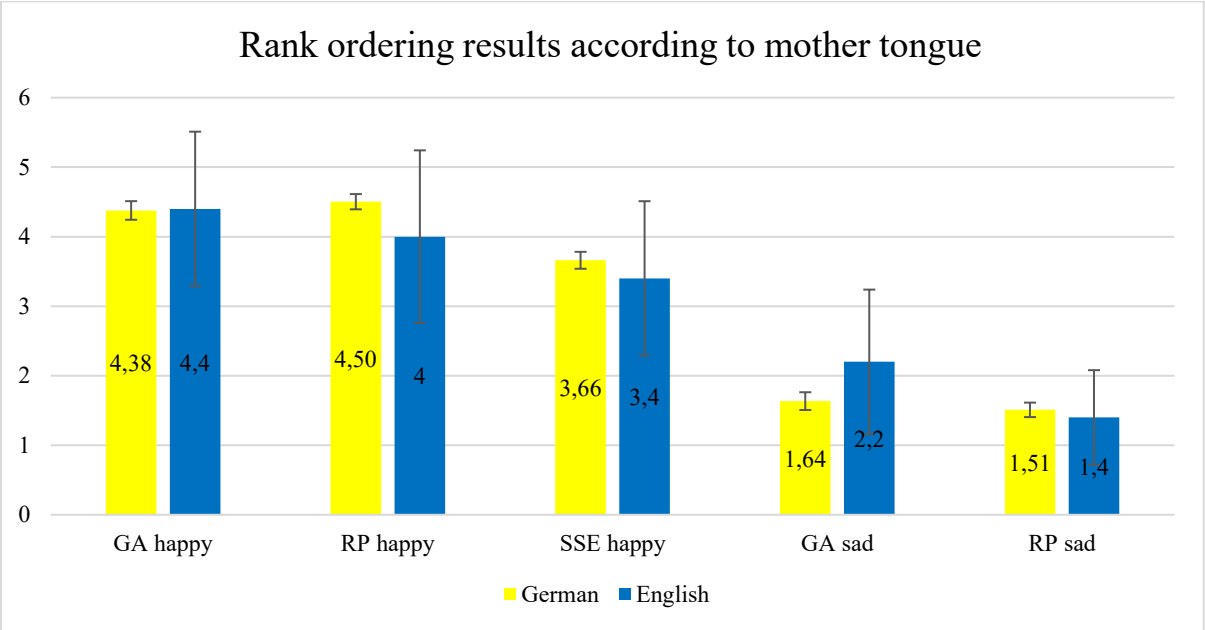


Figure 49. Rank ordering results according to participants‘ mother tongue

Interesting results can be obtained by a closer look at Figure 49 above. Overall, German native speakers reported higher mean scores for all varieties, except for both American samples regardless of the mood change. Notably, four out of five participants with English as their L1 reported American English as their standard variety, indicating a potential influence on their ratings. The two groups distributed their highest overall ratings to different speakers. L1 users of German awarded 4.5 out of 5 possible points to the RP sample, while the English native group preferred the American voice (M=4.4). In terms of the differences of the individual evaluations, it became evident that larger discrepancies emerged in the mean calculations of the RP (0.5 points) compared to the GA file (0.02). Complying with the previous observations, the SSE recording was put in third place. Once again, German natives gave higher means (M=3.66) than the English data set (M=3.4). The greatest variation was detected for the sad GA speaker. Similar to the happy counterpart, L1 users of English showed more favorable attitudes towards the concerned audio file. English native speakers ascribed a mean of M=2.2 compared to M=1.64 points by L1 users of German, differing by 0.56. The melancholic voice from England

reported lowest scores from both groups with means ranging from M=1.4 to M=1.51. As with the happy RP equivalent, the German natives assigned higher values.

In short, participants with English as their mother tongue preferred the American voice over the English and Scottish speakers. While discrepancies remained slight in the analysis of the happy GA sample, the sad counterpart exhibited the largest discrepancies overall. The conversational RP speaker followed, expressing more favorable attitudes of the German natives with a difference of 0.5 points; however, as indicated by the wide confidence intervals, statistical reliability remains low.

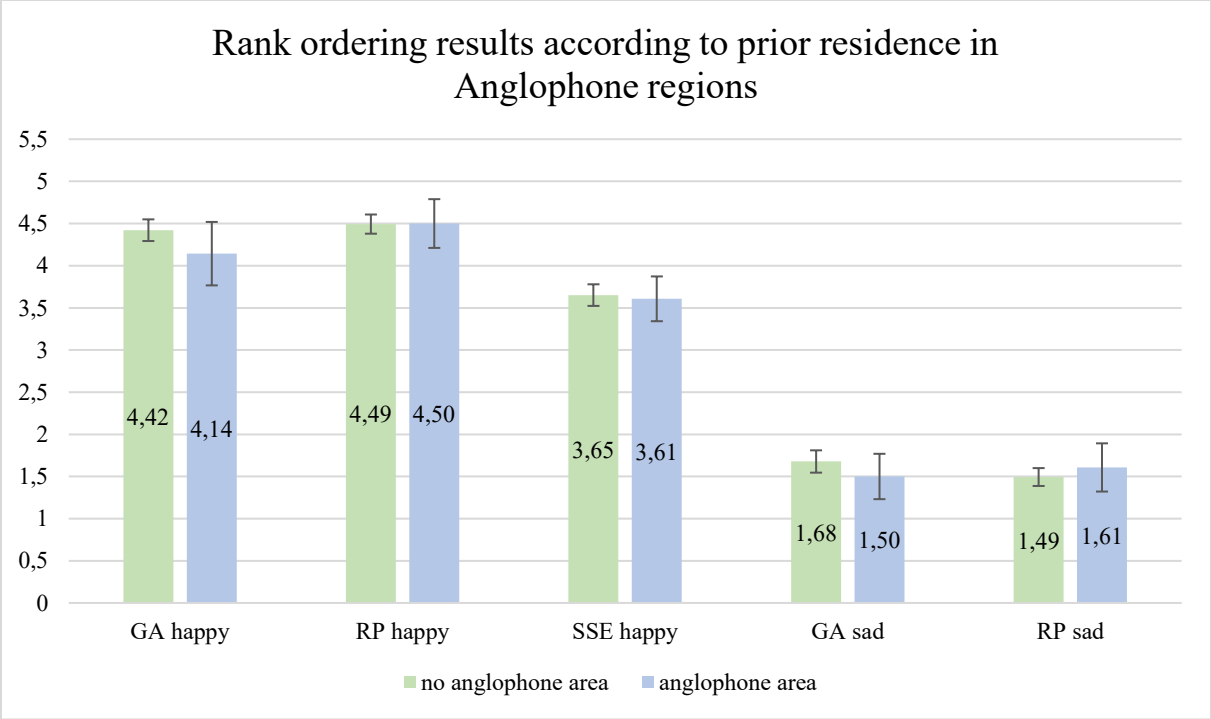


Figure 50. Rank ordering results according to former residences in Anglophone areas

Figure 50 above presents the distribution of mean scores based on participants' potential former residences in English-speaking countries. Notably, the online survey asked participants whether they have lived in Anglophone regions, but did not request detailed information on the duration of the residence. A first glimpse already reveals that respondents, who have not resided in regions with English an official language, tended to report more positive or equal perceptions towards the concerned speakers than those with such international experience. Both groups assigned their highest scores to the happy RP speaker. Larger discrepancies in means emerged in the evaluation of the conversational American file. Participants, who have not lived in an Anglophone country, expressed more favorable perceptions (M=4.42) than the remaining group

(M=4.14), leading to a difference of 0.28. The SSE voice was ranked in the middle by both groups with average scores ranging from M=3.61 to M=3.65.

Interestingly, participants with and without travel experience did not assign their lowest values to the same speaker. Respondents, who reported having lived abroad, distributed a mean of M=1.5 points to the American voice, while a slightly higher value of M=1.61 appeared for the English equivalent. In contrast, participants without the experience of residing in an Anglophone region ascribed M=1.68 points to the American voice and a mean of M=1.49 to the counterpart from England. Similar to the evaluations regarding the influence of participants' mother tongue, the smaller sample size of participants with prior residences in Anglophone areas affected confidence intervals, leading to higher statistical uncertainty.

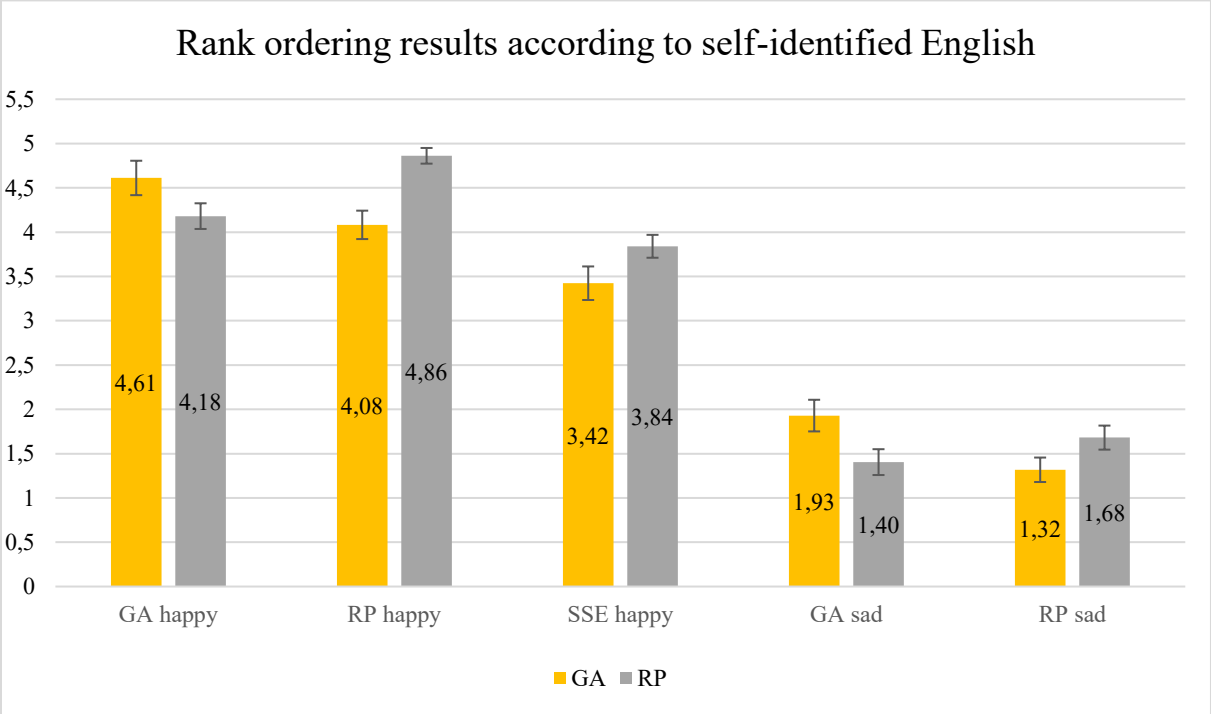


Figure 51. Rank ordering results according to participants' self-identified English

Interesting results emerge from a closer look at Figure 51, illustrating the differences in mean scores according to participants' own spoken variety. The orange bars symbolize answers of American English users, while the ones colored in grey represent mean scores assigned by RP speakers. Generally, respondents' preferred language varieties proved to be a highly influential factor in the ranking of the samples. The GA group showed better outcomes for both American English recordings, regardless of the emotional tone. In contrast, the RP data set distributed higher mean scores for the speakers from England. It is noteworthy that, compared to the

analysis of other potential influences, such as former experiences in Anglophone areas or gender, the personally favored variety revealed the most pronounced variance in average scores.

The conversational American sample received a total of 0.43 points more from respondents who prefer the GA (M=4.61) accent over the RP (M=4.18). The overall largest difference of 0.78 emerged in the evaluations of the speaker from England. While participants, identifying with RP, distributed the highest overall score of M=4.86, American English speakers only ascribed M=4.08 points. As confidence intervals do not overlap, this difference appeared to be statistically significant. A similarity in the rank ordering results between the two data sets was revealed through the analysis of the SSE recording, as it was ranked third by both groups with average scores from M=3.42 to M=3.84. The results of the melancholic audio files mirrored the calculations regarding the happy counterparts. Participants, who personally speak American English, rated the sad GA sample with a mean of M=1.93 while respondents, preferring RP, reported an average number of M=1.4. Lastly, the melancholic sample from England received mean scores from M=1.32, by GA users, to M=1.68, from people favoring RP.

Overall, a clear trend appeared reflecting language preferences based on participants' own spoken variety. Significant differences suggest that the personal use of either GA or RP had an impact on the evaluations of American or English voices with relatively narrow confidence intervals. It can be concluded that people who prefer a specific accent or dialect report more favorable attitudes towards speakers using the same variety.

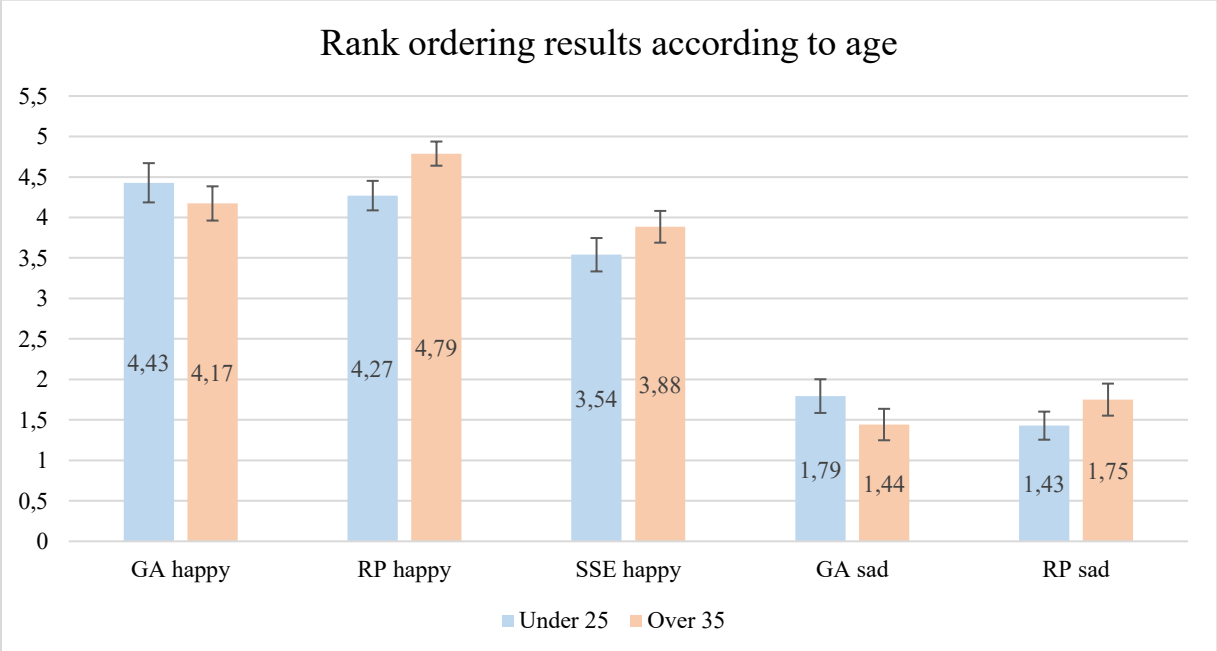


Figure 52. Rank ordering results according to participants' age

As can be taken from Figure 52, participants' age proved to be an influential factor on rank ordering results. The specific age groups were based on the same principles as presented in chapter 7.10. A closer look already reveals that both GA recordings scored higher average points by respondents under 25 compared to those over 35. While the younger group recorded their most favorable attitudes towards the conversational American sample (M=4.43), the older data set only reported a mean of M=4.17 points. An opposite trend emerged for the RP and SSE speakers. The greatest discrepancies became evident for the happy RP voice, with participants older than 35 awarded their highest score (M=4.79), whereas people younger than 25 only distributed an average score of M=4.27, differing by 0.52.

This comparison also showed that both age groups assigned their highest values to different speakers. The SSE speaker received more points (M=3.88) by respondents over 35 than by those under 25 (M=3.54), leading to a numerical difference of 0.34. Similar to the outcomes of the happy GA speaker, its sad counterpart outlined more positive perceptions of the younger participants (M=1.79) compared to older respondents, who ascribed a mean score of M=1.44. The melancholic voice from England achieved their lowest results from the age group under 25 (M=1.43), whereas people older than 35 assigned a mean of M=1.75 points. Comparing speaker 4 and 5 showed that respondents gave their lowest ranking to different speakers. While younger people described the sad RP sample as their least favorite, the older put the American counterpart in last place.

To conclude, these findings showed that participants younger than 25 preferred both American samples regardless of the change in emotional tone, while the over 35 group described more favorable perceptions of the RP as well as SSE samples.

7.13 Variety recognition rates

To assess participants' recognition rates, answers were classified as either correct or incorrect and subsequently grouped; however, this turned out highly complex as the specificity of the responses varied significantly. Broad descriptors appeared especially for the evaluations of the speakers from England and Scotland, such as "GB" or "UK". These could not be considered as entirely correct. Therefore, a third dimension, named peripheral, was required to incorporate all labels, which were not suitable for any of the predefined dimensions. In total, four categories were introduced: *correct*, *peripheral*, *incorrect* or *no answer*. As visible in Figure 53, results were converted into percentages. The illustration below presents the collected answers, using

the following color scheme: bars in green symbolize correct answers while the red ones indicate wrong responses. The yellow ones denote the peripheral descriptors. No received guesses for the respective speakers are colored in grey.

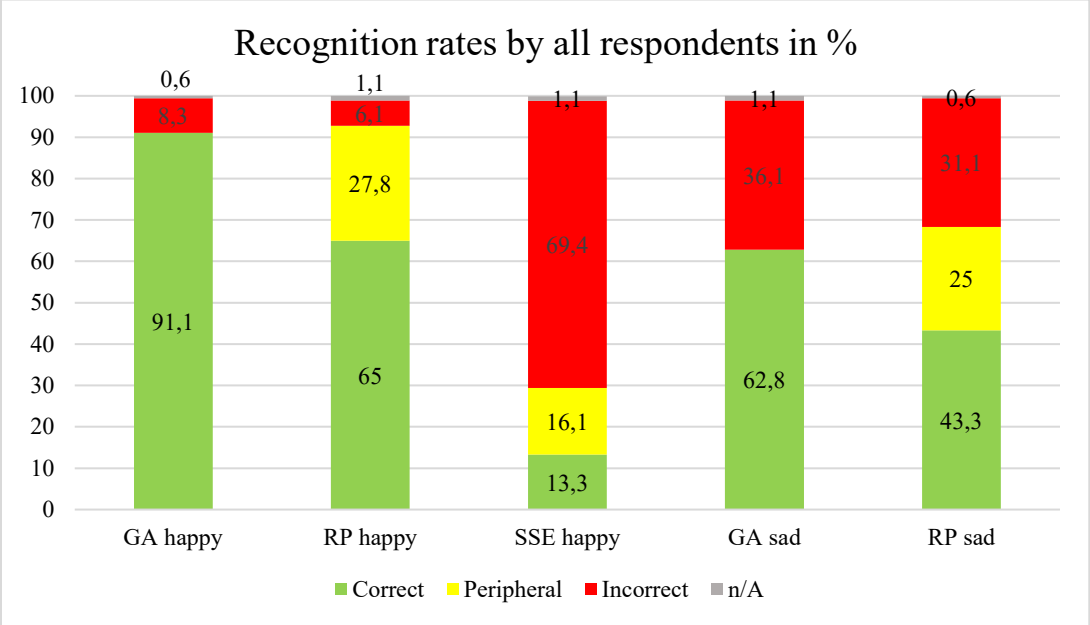


Figure 53. Variety recognition rates by all respondents

Interesting results can be obtained from a first glimpse at Figure 53 above. As immediately visible, recognition rates varied considerably across the different varieties. The highest number of correct responses was recorded for the happy GA voice with a total of n=164 out of 180, equaling 91.1%. Fifteen respondents (8.3%) misidentified the speaker's origin, while one participant (0.6%) did not submit any guesses. In comparison, the conversational RP sample showed a significantly lower correct recognition rate with 65% (n=117). Eleven people (6.1%) mentioned incorrect home countries, whereas 27.8% (n=50) provided partially correct answers, including "United Kingdom" and "Great Britain", which were deemed too broad. The illustration indicates that the SSE speaker reached the lowest score of accurate labels with only 13.3% (n=24). A total of 16.1% responses were considered as partly correct, including similar descriptors as in the analysis of the speaker from England. The majority of respondents (n=125), amounting to 69.4%, submitted wrong guesses, like "Ireland", "Australia" or "Canada".

In comparison to the happy American file, the sad counterpart was described significantly less accurately with a recognition rate of 62.8%; however, Figure 53 reveals that no peripheral answers were detected in either speaker analysis. For the sad voice, 36.1% (n=65) incorrectly labeled the place of origin and, similar to all pervious samples, 1.1% (n=2) did not respond to

the question. Lastly the melancholic speaker from England was put in fourth place in terms of accurate identification. A total of 43.3% (n=78) of participants reported the correct regional provenance, while 31.1% (n= 56) submitted wrong descriptions. Moreover, 25% (n=45) were partially correct and one person out of 180 did not contribute to the respective question.

To conclude, the happy GA voice achieved the highest recognition rate with 91.1% (n=164). The sad counterpart received 113 correct labels; therefore 62.8% of the participants could identify the speaker's country of origin based on its accent. A similar trend occurred for the English sample, as the conversationally read file outlined a higher number of accurate answers than the sad recording. These findings indicate that mood proved to be an influential factor.

7.13.1 Variety recognition rates based on self-identified English

In total, 85 out of 180 people reported General American as their preferred variety. A closer look at Figure 54 already reveals that these participants performed better in correctly determining the American voice compared to the other speakers. As shown in the first bar, users of Standard American English outlined a recognition rate of 90.6%; thus, the majority of respondents (n=77) was able to identify the speaker's country of origin, whereas 8.2% (n= 7) submitted wrong guesses, such as "Australia" or "UK". One participant (1.2%) responded with "I don't know", which was grouped into the n/a category.

The second highest recognition rates were obtained for the sad American voice with 71.8% (n=61); however, the percentage of inaccurate labels increased to 25.9% (n=22). The conversational RP file was correctly determined by 55.3% of GA users (n=47). Twenty-nine participants (34.1%) provided partly correct guesses, including "GB" and "UK", whereas 9.4% (n=8) mentioned the wrong regional provenance. The sad speaker from England reached a considerably lower score of 36.5% (n=31). Twenty-five people, using Standard American English, submitted peripheral responses, such as "UK" and "GB", while a total of 32.9% (n=28) misidentified the speaker's accent. The lowest recognition rate became evident in the analysis of the SSE speaker with only 11.8% (n=10). Fifty-five people (64.7%) misinterpreted the concerning accent and submitted incorrect answers, including "Ireland", "England" or "Australia".

In conclusion, a more detailed analysis of the recognition rates among participants with GA as their preferred variety outlined a similar trend to that presented in the evaluations by all respondents; thus, it can be assumed that respondents' own spoken variety slightly affects the recognition of an accent, particularly in the case of the sad American file.

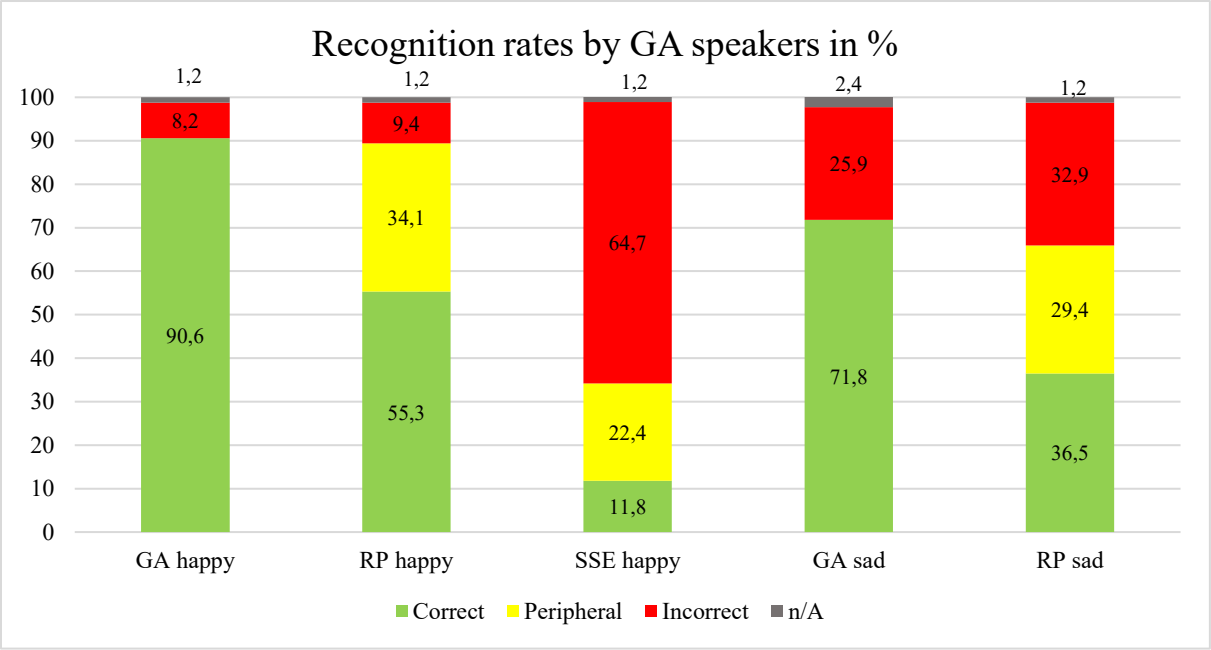


Figure 54. Recognition rates by GA speakers

The graphical illustration below (see Figure 55) presents recognition rates of participants who reported favoring and using RP (n=94). Interestingly, the conversational GA voice obtained the highest overall score with 86 correct guesses, resulting in an identification score of 91.5%. Only 8.5% (n=8) people submitted wrong answers. This pattern is also reflected in the analysis of the respondents preferring Standard American English (see Figure 54); however, more substantial differences emerged in the evaluations of the happy voice from England. Seventy participants (74.5%) correctly labeled the regional provenance, compared to only 55.3% (n=47) of the respondents with GA as their actively used variety. Three RP speakers (3.2%) gave incorrect guesses, while 21.3% (n=20) mentioned places like “GB” and “UK”, which were categorized under the peripheral dimensions.

The SSE sample received the lowest recognition rates from participants favoring Standard American English. A similar trend remained for the English users with a percentage of only 14.9% properly identifying the speaker’s regional background. In contrast, most of the participants (73.4%) were not able to determine the voice’s origin.

A comparison between the GA and RP data set revealed that users of Standard American English achieved significantly higher recognition rates for the sad GA sample. Specifically, 61 GA users submitted correct responses (71.8%), next to 51 (54.3%) respondents preferring RP. The fourth bar presented below (see Figure 55) indicates that 45.7% of those respondents misidentified the accent of the sad GA speaker. Forty-seven people, equaling 50%, provided accurate guesses

about the voice's place of origin, while 28.7% (n=27) provided wrong answers regarding the melancholic RP voice.

In summary, the comparison of participants' responses, either preferring GA or RP, demonstrated that participants' own spoken variety had an impact on their recognition rates. Respondents tended to determine the accent of a speaker more accurately, which resembles their own language habits; thus, speakers of Standard American English showed lower rates for all voices from England than the English group. A similar trend occurred among RP users except for the happy GA sample, where RP users outperformed speakers of General American in properly labeling the speaker's regional provenance.

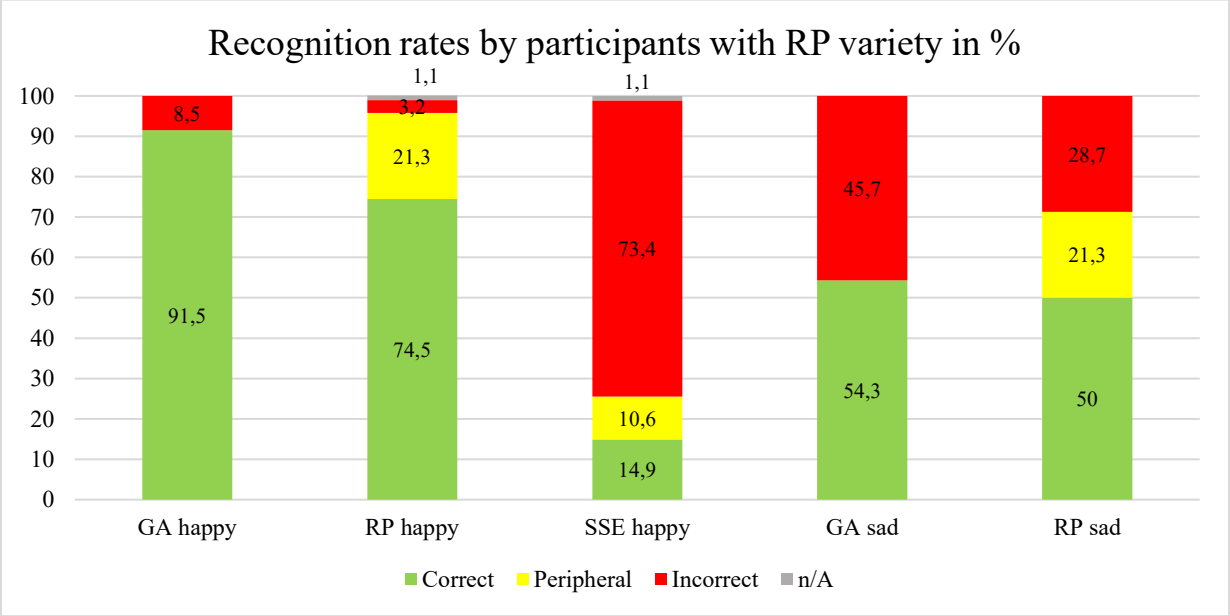


Figure 55. Recognition rates by RP speakers

7.13.2 Variety recognition rates based on participants' mother tongue.

The following section explores differences in recognition rates dependent on participants' mother tongue. As stated in chapter 7.8, 159 people reported German as their L1, while only five were English native speakers. Specially, most of them personally used GA and one participant prefers RP. It is important to highlight that the small number of participants with English as their mother tongue (n=5) reduces the reliability of the findings; nevertheless, the results still offer valuable insights into their rating behavior, although generalizability is remarkably limited and outcomes require careful consideration.

Figure 56 below presents noteworthy results regarding recognition rates of English native speakers for the five voices. Firstly, all five respondents correctly determined the happy GA

sample. The same pattern emerged for the sad counterpart, leading to an accurate identification rate of 100% for both files from the US; however, a different trend was observable for the recordings of the speaker from England. The correct and peripheral answers were balanced with 40% each (n=2), while one respondent was unable to properly define the speaker's place of origin. The same regional descriptors were reported for both voices from England (speaker 2 and 5). The lowest rates became evident for the SSE speaker, as none of the participants could exactly define the speaker's accent. Answers included "Ireland", "Australia", "GB" and "Europe", which were either incorrect or too general; As a result, four people provided an incorrect response and one participant guessed the regional provenance partly correct.

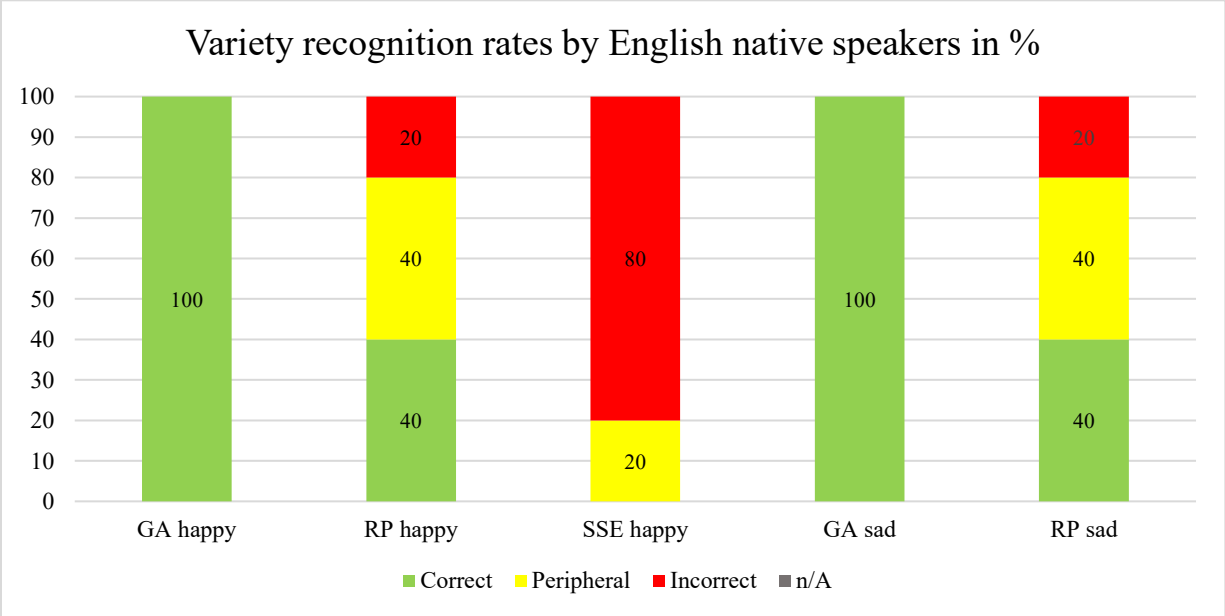


Figure 56. Variety recognition rates by English native speakers

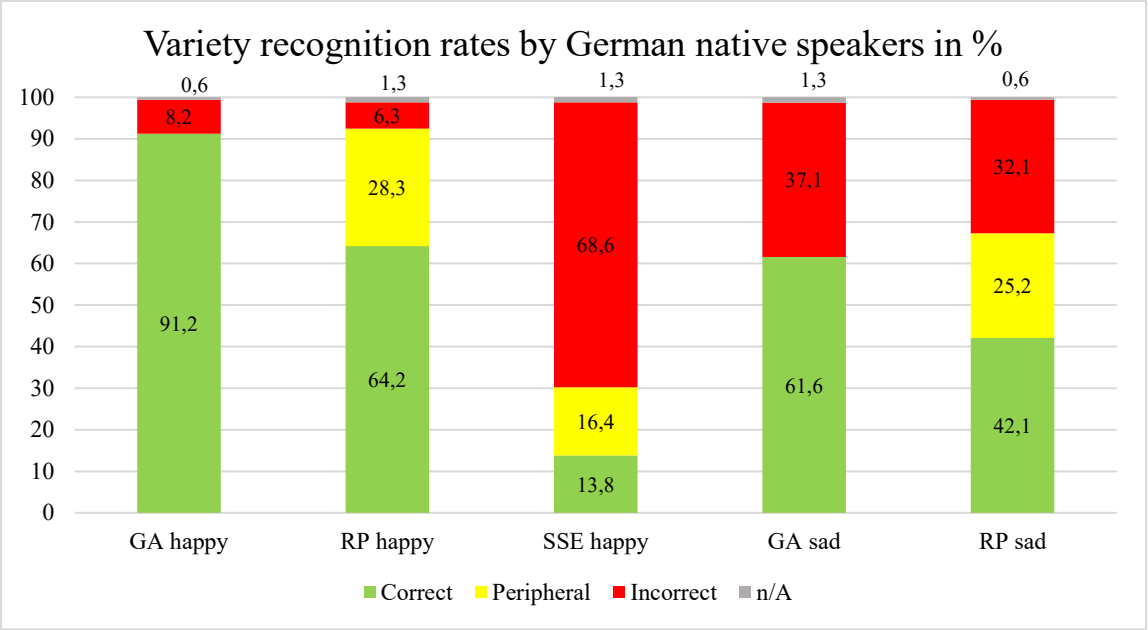


Figure 57. Variety recognition rates by German native speakers

Interesting insights can be gained from a closer look at Figure 57, which presents variety recognition rates by German native speakers (n=159). As immediately visible, the American voices, regardless of emotional tone, tended to be identified more correctly compared to the equivalents from England, mirroring the evaluations of participants with English as their L1. The happy GA voice reached an accurate recognition rate of 91.2% (n=145), while 13 participants, amounting up to 8.2%, submitted incorrect answers. In contrast, 64.2% (n=102) were able to determine the regional provenance of the sample from England. Forty-five of the German L1 users (28.3%) provided partially correct responses, including "GB" and "UK", while 6.3% (n=10) could not identify the concerning accent. Another resemblance between both language groups consisted of the low values for the SSE speaker. Only a minority of 13.8% (n=22) submitted the accurate response, while a large majority (n=109) assumed that the speaker came from "Ireland", "England" or "Australia" for instance. Moreover, the sample outlined two respondents, not providing any guesses about the speaker's variety. The identification rates of the melancholic voices were lower compared to the happy counterparts. Ninety-eight out of 159 German native speakers managed to properly identify the GA speaker's accent, while 59 people submitted wrong guesses. The percentage of the properly described responses regarding the sad voice from England remained lower with 42.1% (n=67). Forty participants stated partly correct descriptors; however, they were regarded as too general, such as "GB" or "UK". A total of 32.1% (n=51) misinterpreted the speaker in terms of their country of origin.

Briefly, a comparison of answers, submitted by L1 users of English as well as German did not reveal remarkable differences. The greatest discrepancies appeared in the recognition rate of the sad GA voice. While 100% of the native English speakers (n=5) were able to correctly determine the regional prevenance of the concerned sample, only 61.6% of the German natives (n=98) labeled the variety accurately. Interestingly, 64.2% of those (n=102) provided accurate guesses about the happy voice from England, compared to only 40% (n=2) of respondents, using English as their L1; still, as stated in the beginning of the section, the sample size of participants with English as their mother tongue was significantly smaller; thus, it seemed almost impossible to draw reliable conclusions and examine, whether mother tongue was an influential factor regarding the recognition rates.

7.14 AI-detection by all respondents

As outlined in chapter 6.6, all audio files were created with an AI tool called Murf AI; hence, a further analysis of participants' answers was conducted to examine whether they recognized the speakers as computer-generated and describe their way of talking accordingly. To gather relevant data, the following questions were adhibited: *How would you describe speaker XY's presentation of the news in word?* and *Where do you think speaker XY comes from?*.

Regarding the first item (see Figure 58), interesting results could be revealed. The happy GA voice elicited one response, claiming that the reading style seemed very automated. Apart from that, no assumptions about a possible use of AI occurred. Likewise, none of the participants described the happy RP speaker as computer- or AI-generated.

However, a different trend was detectable for the sad GA voice. Seven participants, equaling 3.9%, provided descriptors such as "AI", "artificial", or "robotic". Three respondents (1.7%) perceived the melancholic voice from England as "artificial" and "robotic". The SSE sample recorded one assumption about artificial intelligence; hence, it became evident that only a small percentage of all participants noted a potential utilization of AI.

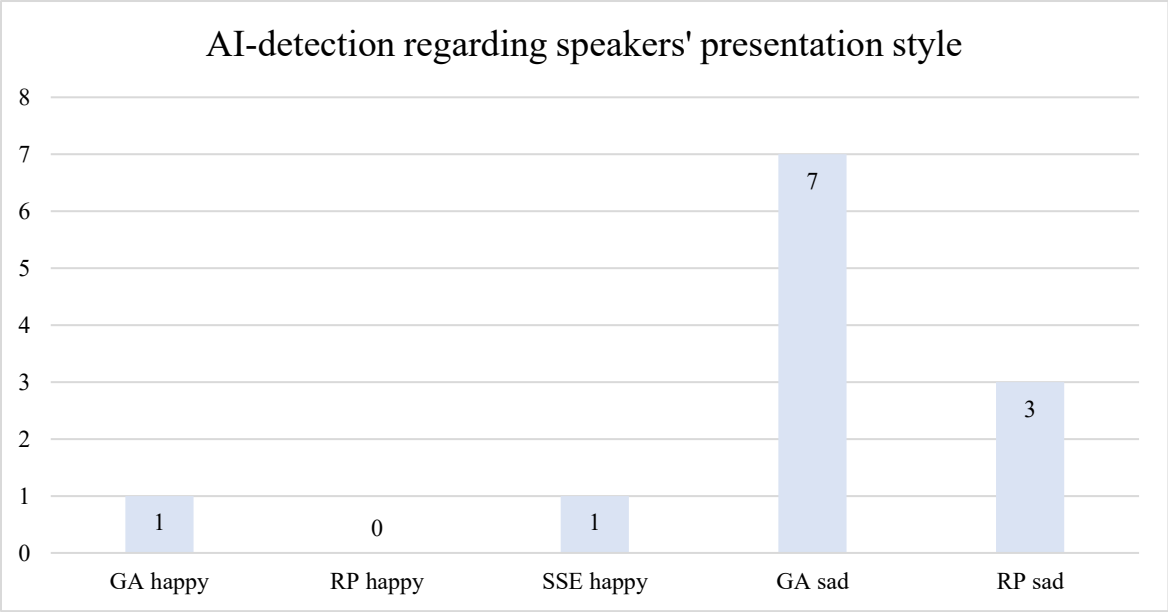


Figure 58. AI-detection regarding speakers‘ presentation style

Question 2 received even fewer responses, in which respondents assumed that AI generated the individual voices. The results are presented in Figure 59 below. The conversational American sample reported one comment, explicitly claiming that the speaker was computer-produced. Once again, its sad counterpart yielded the highest number of respondents (n=3), amounting up to 1.7%, suggesting an AI use. Aligning with the evaluations above, the voice from England did not outline any assumptions regarding AI; however, the melancholic equivalent showed one instance of AI detection. The SSE file was not associated with artificial intelligence; nevertheless, the recognition rate remained marginal.

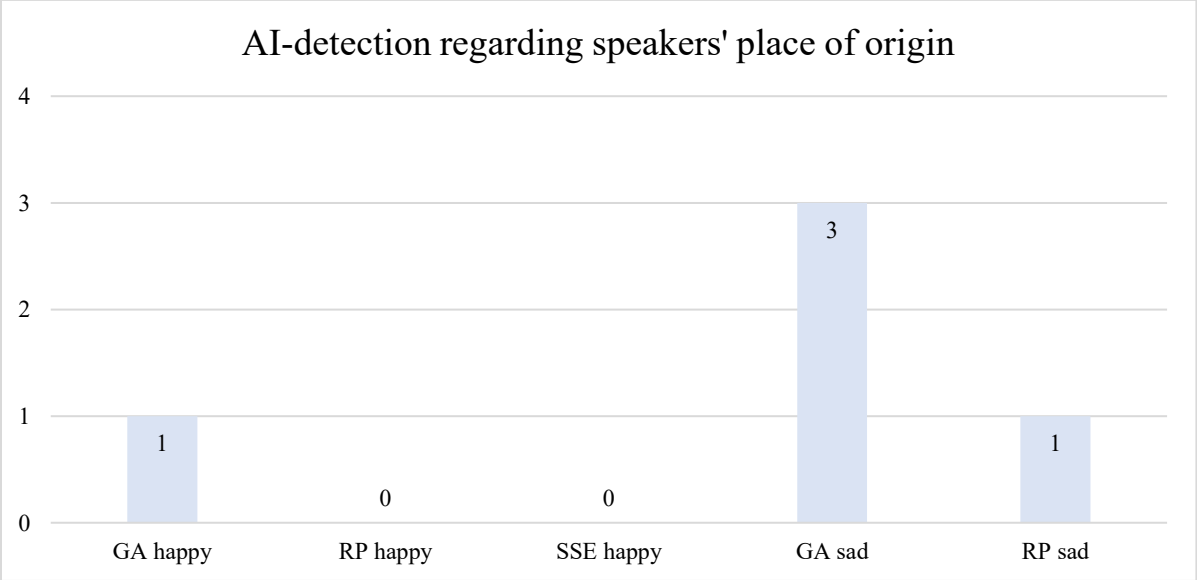


Figure 59. AI-detection regarding speakers‘ place of origin

In short, it can be assumed that the majority of participants did not identify the individual voices as computer-generated. The highest number of responses, claiming that AI was adhibited, was found in descriptors regarding the sad GA voice for both questions, ranging from three to seven submissions. In contrast, the conversational sample from England did not receive any labels indicating AI; nevertheless, it is noteworthy that this section is solely based on participants' spontaneous answers. As no specific question was provided with regards to AI, it cannot be ruled out that some might have assumed a potential use of AI, but did not report on it.

8 Discussion

The following chapter presents the revealed results based on former research conducted in the field of language attitudes. In view of the fact that all audio files were AI-generated, results require careful consideration, since their presentation style might differ from real-life speakers and thus, may affect participants' perceptions, particularly in the case of the sad voices.

8.1 General evaluations

Overall, high mean scores were observed regarding all character traits for both the happy GA and RP voice; notably, the speaker from England achieved greater scores in the field of social status, including *intelligent*, *hardworking* and *wealthy*, compared to its counterpart from the US. These findings align with previous research by Hans J. Ladegaard, observing language attitudes towards British, American and Australian English in Denmark (1998), revealing that samples from England were considered as most prestigious (Ladegaard 1998: 258). A similar trend was also observed by Dalton-Puffer et al. (1997: 122), explaining that RP speakers were awarded with great means for terms such as *intelligent*, *educated*, *successful*, *determined* or *ambitious*. Giles (1970: 223, 225) also explained the same concept, pointing out that RP tends to be associated with a high degree of prestige. While the other two dimensions, more precisely warmth and dynamism, only outlined marginal differences between the GA and RP voice, respondents reported greater discrepancies in terms of social class. Interestingly, slightly higher values were obtained by the happy US sample next to the English equivalent for some adjectives of the categories of warmth and dynamism, such as *polite* or *funny*, which was therefore also partly confirmed in Ladegaard's research (Ladegaard 1998: 259). To put it in the author's words, the RP variety here "is down-graded relative to the other speakers" (Ladegaard 1998: 259). Another similarity to the study by Ladegaard consisted of the open answer item, where respondents were asked to describe the speakers' presentation style in one word. As stated by Carrie and McKenzie (2018: 324), descriptors regarding RP frequently contained adjectives relating to a high social status. This phenomenon also emerged in participants' responses within this project, such as *posh*, *rich* and *upper class*. In contrast, terms for the American speaker belonged to the field of solidarity, as labeled by Carrie and McKenzie (2018: 324). Again, a similar pattern appeared for these concerned AI-generated voices with words, like *nice*, *warm*, *friendly* and *positive*. In summary, the authors conclude that RP indicates high class, whereas GA symbolizes a high degree of solidarity, correlating with the findings from this specific project (Carrie & McKenzie 2018: 325).

However, significantly lower means were recorded by the sad samples, indicating that a mood shift is highly influential on participants' evaluations across all domains. These outcomes contradict the previously cited conclusions of conversational voices, demonstrating that RP is associated with most prestige, as the melancholic speaker from the US surpassed the RP files in all traits; nevertheless, with regards to the US sample, the category of warmth remained in the highest rank, aligning with the literature, which suggests that American English is frequently related to solidarity (Carrie & McKenzie 2018: 325). Although the sad American English voice was linked to a higher social status, this dimension showed the overall smallest differences between the two varieties; thus, the RP file achieved higher means here, while warmth and dynamism yielded comparatively lower values (Ladegaard 1998: 259). Scherer (2003: 228) describes a great effect of a speaker's emotion on their interlocutors, as voices naturally change depending on mood and context. This was further supported by Briñol et al. (2007: 711), explaining that emotion affects perceptions of the confidence of a person. In the context of this thesis, the character trait *confident* supported the authors' argument, as remarkably lower scores were obtained by the sad compared to the happy speakers (Briñol et al. 2007: 711).

The Scottish sample was consistently put in the middle rank for all character traits except *snobbish*, contradicting the findings from Ladegaard (1998: 259). In his study, he concluded that the SSE speaker received high values in terms of helpfulness and friendliness (Ladegaard 1998: 259). Notably, this specific project did not ask participants for their perception of the speakers' perceived helpfulness, but the evaluations of *polite* placed it in third place. Overall, Standard Scottish English was evaluated less favorably compared to the other two happy samples, which might be a result of participants' unfamiliarity with the variety, as it is less present next to GA and RP.

As stated before, evaluations of the speakers' snobbishness were presented in a separate section due to its negative connotation compared to the other ten character traits; hence, lower scores in this category represent more positive perceptions. Interestingly, *snobbish* was the only instance, with regards to RP, where the sad voices achieved better results than the happy ones. All in all, the Scottish voice obtained the highest scores, but only a minority was able to correctly label the accent. The second greatest points were yielded by the voice from England. Even though the happy RP file was described as *intelligent*, *wealthy* and *hardworking*, respondents also perceived it as *snobbish*, indicating a certain ambivalence. While the English variety was connected to many positive character traits, it also revealed some criticism likely due to its perceived arrogance; however, aligning with former research, the GA sample achieved remarkably lower scores (Carrie & McKenzie 2018: 325). No differences in the ratings among

the American files emerged, suggesting that emotional tone did not influence participants' perceptions, but this could not be confirmed for the RP equivalent, as considerably lower points were dedicated to the melancholic (M=1.02) compared to the happy sample (M=1.58). Mood might have a stronger impact here, since RP is frequently regarded as more distant than GA. Consequently, a sad presentation style contrasts with its usual performance, affecting participants' judgements.

Overall, results of the perceived suitability as a news reporter aligned with findings from a previous study by Dalton-Puffer et al. (1997). The RP speaker was ranked in first place and therefore considered as most appropriate for reading the news, which might be attributed to its dominance in educational instructions in Austria, as the majority of participants were of Austrian origin (Dalton-Puffer et al. 1997: 126). In addition, this particular accent is commonly used by news reporters on BBC, which may further contribute to its high rating (Roach 2004: 239). An opposite pattern appeared for the sad samples, where the lowest scores were reported regarding the speaker from England. In comparison, the US equivalent achieved greater values; nevertheless, these were still significantly lower than the happy samples, suggesting that mood highly affected respondents' evaluations; however, in the domain of language attitudes with a focus on mood shifts, the frame of reporting the news may be highly influential and thus, different results were obtained if another context was chosen.

The following findings could be revealed in terms of pronunciation clarity: The happy RP voice received the highest score, while the GA sample was ranked second. This aligns with findings from Mugglestone (2012: 1900) who points out that RP has always been associated with a high degree of pronunciation clarity. Both sad recordings received remarkably lower values than the happy ones. Specifically, the variety from England achieved the overall minimum score. Positive attitudes towards the RP speaker were also confirmed by Ladegaard (1998: 266), claiming that many speakers aim to follow this model, as it is frequently regarded as 'correct English'. Interestingly, the pronunciation clarity of all included speakers yielded higher scores than the suitability as a news reporter. Marked discrepancies could particularly be detected for the sad voices in both fields, indicating that emotional tone had a greater impact on respondents' judgements regarding the appropriateness for reading the news.

A similar pattern became evident for participants' willingness to meet the respective speakers for a coffee as in the dimensions presented above. The highest value was obtained by the happy RP voice, closely followed by the GA equivalent. The previously described positive attitudes

towards the happy sample from England throughout all categories might contribute to the high number of partakers, wanting to meet the concerned voice.

Altogether, different results were observable for the happy compared to the sad speakers across character traits, pronunciation clarity as well as news reporter suitability. An inverse trend was specifically detectable in both RP recordings, as the happy file frequently received the maximum score, while the melancholic counterpart often reported the lowest score of all samples; thus, it can be assumed, respondents did not recognize that the GA and RP speaker occurred twice, differing only in emotional tone; hence, these findings confirm hypothesis H_1-II.

8.2 Analysis dependent on moderating variables

The effects of participants' gender, age, mother tongue, self-identified English and potential former residences in Anglophone regions on their evaluations will be discussed in the following section.

Generally, no clear influence of respondents' gender was observable, as only marginal discrepancies emerged. Results revealed that men rated both American speakers slightly higher than women. In contrast, the female group reported greater values for the sad RP sample, while a balance emerged in the evaluations of the happy counterpart. Similar accounted for the SSE sample. According to Lausen and Schacht (2018: 17), women tend to be more sensitive to different emotions "due to their ascribed nurturing, affiliative, and less dominant role"; nevertheless, this assumption can only be partially confirmed by the collected data. Comparable findings emerged for the speakers' pronunciation intelligibility and suitability as a news reporter. Results did not report explicit variety preferences; still, women ranked the melancholic voices higher compared to men, confirming the argument by Lausen and Schacht (2018: 17).

When comparing evaluations by participants with and without a former residence in an Anglophone area, it became evident that respondents who have lived in an English-speaking country assigned greater mean scores to both RP speakers. In contrast, respondents without Anglophone experiences abroad, tended to favor the GA variety, regardless of emotional tone. As the majority of internationally experienced respondents stayed in Great Britain, the higher scores for the RP accent may be a result of greater exposure to the variety; nevertheless, it needs to be stressed that no duration of the stay abroad was given; however, as the question asked respondents, if they have lived in an Anglophone area, it can be assumed that they have resided there for a longer period of time. Findings of the pronunciation clarity indicated more positive

attitudes towards all speakers except the sad GA voice by participants with a former residence in Anglophone regions, slightly contradicting the forementioned results. The perceived suitability as a news reporter supported the argument that respondents with international experience tend to favor RP, while participants without a longer stay in an English-speaking country prefer GA. This can be explained by the *Contact Hypothesis*, describing that personal relations to a certain country and the corresponding language affect language attitudes (Dörnyei & Csizer 2005: 328); nevertheless, this did not apply for the sad RP file.

A correlation between language attitudes and respondents' self-identified English appeared; hence, those, wo personally speak GA tended to report higher means for most adjectives of the voices from the US, regardless of emotional tone. The same pattern emerged for speakers of RP. To put it in the authors' words (Dalton-Puffer et al. 1997: 117), people usually react more favorable towards language varieties, which sound familiar to them, supporting the results of this specific project. In summary, respondents' self-identified English had an impact on their evaluations regardless of emotional tone; however, means were considerably lower for all melancholic voices. Interestingly, evaluations regarding pronunciation clarity and the perceived suitability as a news reporter revealed a slightly different pattern, as respondents did not automatically assign greater means for their own spoken variety. Regarding the first mentioned dimensions it became evident that participants with GA as their self-identified English distributed higher scores for all speakers, except the happy RP and SSE sample, supporting Roberts' (2020: 202) argument, since the personally used accent must not necessarily influence ratings of voices. In addition, this would also be confirmed through judgements about the appropriateness for reporting the news. Users of RP reported more positive attitudes for all speakers except the sad GA sample. These findings contradict former outcomes by Dalton-Puffer et al. (1997: 117).

The influence of participants' mother tongue did not show an explicit trend regarding the given character traits. The evaluations of the happy GA speaker revealed mostly lower values, assigned by German natives compared to L1 users of English, except for the domain of social status, including *intelligent*, *hardworking* and *wealthy*. Doubtlessly, mother tongue is tied to a high degree of emotionality and familiarity (Huguet & Janes 2008: 257). As most English native speakers in this study personally use GA, it is plausible that they assigned greater scores to the fields of warmth and dynamism, since these dimensions are related to feelings instead of social status; still, this argument did not apply for the sad counterpart, as respondents with English as their mother tongue assigned higher points to all characteristics, except funny. Both samples from England did not indicate clear preferences by a certain group but scores were balanced;

thus, mother tongue did not affect the judgements of the two files, regardless of emotional tone; still, values of the melancholic voices again remained remarkably lower than the happy equivalents. With regards to pronunciation clarity, outcomes showed that the native English group distributed higher points for all speakers, except the happy RP voice. As previously mentioned, Austrians are mostly familiar with RP; hence, the higher rating of the German native group might be due to its frequent usage in educational institutions (Dalton-Puffer et al. 1997: 126).

Respondents' age was found to exert an influence upon their evaluations of the concerned language varieties. Both American English speakers were perceived more favorably by the age group under 25 in all dimensions; yet, tremendous differences emerged within the GA ratings, effected by mood. The preference of young people to use American English might be due to the US (social) media presence. Especially on platforms, such as TikTok or Instagram, GA dominates; thus, it is possible that this exposure results in more positive judgements of the perceived warmth and dynamism of the speakers, as these fields are usually connected to relatable performances. An opposite trend appeared for respondents over 35, who ascribed higher means for the speaker from England. Since RP has traditionally been associated with the norm and 'correct English', it seems likely that older people favor this English accent (Ladegaard 1998: 266).

In short, participants' self-identified English influenced their judgements regarding the given character traits, affirming hypothesis H_1-I. Additionally, the assumption that former residences in Anglophone areas have an impact on the character trait level (hypothesis H_1-V) can be confirmed, as a preference for the RP speakers emerged, which could be due to the majority have resided in Great Britain. Overall, the sad voices received lower scores in all dimensions, except *snobbish*; consequently, hypothesis H_1-III, claiming that the melancholic samples are more negatively evaluated on every character trait level, is only partly supported through the collected data.

8.3 Rank ordering

The rank ordering results of the given varieties revealed that most participants favored the RP speaker complying with former research (Dalton-Puffer et al. 1997: 126). As stated before, schools in Austria follow the RP pronunciation model; thus, it can be concluded that the majority outlines positive attitudes towards the pronunciation norm in educational institutions (Dalton-Puffer et al. 1997: 126). Overall, Dalton-Puffer et al. describe that people tend to favor the varieties, closest to their own speaking habits (Dalton-Puffer et al. 1997: 126). This

argument also applies for this project, as the majority reported to personally use RP. The happy GA voice received the second highest score, whereas the SSE sample was ranked in third place, indicating that respondents showed least favorable attitudes among the happy files towards the concerned file. Interestingly, an opposite trend occurred for the melancholic voices. While the happy speaker from England received the maximum score, the sad counterpart achieved the overall lowest points. Consequently, the voice from the US obtained marginally greater values suggesting that the previously described argument by Dalton-Puffer et al. (1997: 126) does not apply for the samples with a shift in emotional tone.

A comparison of respondents under 25 and over 35 outlined that younger people generally preferred the US samples. Ladegaard (1998: 265) sees this as a result of widespread American youth culture online or on television. On the contrary, the older age group revealed higher means for both RP samples as well as for the Scottish voice. Although respondents, dependent on their age, assigned higher scores for one specific variety, points were still remarkably lower for all sad recordings. In summary, a clear pattern emerged that younger participants tended to prefer the US samples, while the older data set favored the accent from England, regardless of emotional tone; nevertheless, the ratings of the conversational and melancholic speakers within one variety differed greatly.

Regarding respondents' self-identified English, results showed that people favor varieties, which sound most familiar to them (Dalton-Puffer et al. 1997: 126); thus, participants, personally using GA, assigned higher means for both speakers from the US than people with RP as their preferred accent. Discrepancies were notable, but even more significant variances appeared for the samples from England; hence, an in-group preference effects were observable.

However, data did not show a consistent pattern of the influence of respondents' gender on rank ordering results. Trudgill (1974: 87) outlines the concept of status-consciousness, claiming that women usually tend to adapt varieties, which are associated with a high status; thus, it might be assumed, that the variety from England would be ranked higher by female participants, since it was frequently regarded as high class; nevertheless, this could not be confirmed, as men and women reported the same number of points for the happy voice from England. The sad counterpart revealed slightly higher values by the female group, but differences remained marginal. Similarly, the melancholic GA voice also received better scores by women than men, suggesting that emotional tone had a greater impact on evaluations by the female group.

Sample size of native English speakers was significantly smaller compared to L1 users of German; thus, results need to be treated with caution and cannot be generalized. Nevertheless,

the analysis of rank ordering results based on respondents' mother tongue yielded a clear trend. Participants with English as their mother tongue reported greater scores for both GA samples, which could be attributed to the fact that four out of five native English speakers use American English. These findings support the arguments by Huguet and Janes (2008: 257), claiming that mother tongue is a highly influential factor in the domain of language attitudes due to its presence in family context, for instance. On the contrary, German natives ascribed higher ratings for the RP files, regardless of emotional tone, which might be a result of their high familiarity with the variety through educational institutions (Dalton-Puffer et al. 1997: 126).

Another dependent variable were potential former residences in Anglophone areas. Notably, the majority of participants (n=17) have resided in Great Britian, whereas nine participants lived in the US for a certain period of time. Others moved to Australia or South Africa. Generally, it was observable that respondents, who spent some time in an English-speaking country, ranked both RP varieties higher compared to those without such experience. Given that most participants lived in GB, these had greater exposure to the specific variety, potentially leading to more favorable perceptions. Still, mood had a significant impact, as scores remained remarkably lower for both melancholic voices, regardless of the variety in question. Participants, who have not resided in Anglophone areas, ranked the speakers from the US more highly.

In summary, hypothesis H_1-VI, claiming that respondents' age, self-identified English and mother tongue affect rank ordering results, can be confirmed by the given data.

8.4 Recognition rates

Especially for non-native speakers of English, correctly identifying a person's accent, may be challenging, particularly without enough exposure to the language; thus, misinterpretations can easily occur, which subsequently influence attitudes (McKenzie 2008: 139).

As presented in chapter 7.13, respondents' descriptors of the assumed speakers' place of origin were categorized as *correct*, *peripheral*, *incorrect* or *n/a*. Too general terms, as in the case of United Kingdom and Great Britian instead of England, were regarded as peripheral. Since answers were open-ended and not limited, grouping turned out complex, corresponding with McKenzie (2008: 146).

Overall, the highest correct recognition rates emerged for the happy GA sample, followed by the RP voice. These outcomes align with Ladegaard (1998: 266), stating that GA and RP are commonly the most accurately identified English varieties due to their frequent occurrence in the media as well as entertainment industry. Besides that, Carrie (2017: 431) suggested that the

great exposure to both varieties might also be a result of the "social, cultural, political and economic power of the US and UK". McKenzie (2008: 146) reported the same outcomes for his project in Japan, arguing that high recognition rates for American English speakers can be attributed to the dominance of US media. Moreover, another study revealed that the majority of first- and second-year English students were able to correctly identify the GA speaker's accent, yet with a difference in preciseness (Stephan 1997: 102). The author linked this to the influence of educational institutions, as both, RP and GA, are primarily taught as standard English languages with a focus on their linguistic features (Stephan 1997: 102, 105). Additionally, Carrie and McKenzie (2018) examined recognition rates of L2 learners of English in Spain. In line with the previously described projects, results presented a high degree of accurately identified labels for RP and GA speakers; however, in contrast to the findings of this thesis, Carrie and McKenzie (2018: 313) showed that participants more frequently provided proper answers regarding the origin of the RP voice compared to samples from the US. They relate this to the geographical proximity of England to Spain alongside its generally high status across Europe (Carrie & McKenzie 2018: 325). Since most participants in the context of this specific research project were based in Austria, the argument of proximity does not apply in this case; nevertheless, as frequently emphasized, AI-generated samples may differ in their performance, potentially affecting respondents' identification behavior.

As presented by Ladegaard (1998: 261), significantly fewer participants correctly labeled the regional provenance of the Scottish speaker. The most common misidentification was (Northern) Ireland largely due to its linguistic resemblance to other British varieties, particularly for LX users of English (Ladegaard 1998: 261). The present study yielded similar findings. Only a minority of the respondents were able to provide accurate guesses, while most provided incorrect places of origin. Interestingly, the majority of respondents assumed that the voice was originally from Ireland, mirroring Ladegaard's outcomes. McKenzie (2008: 149) initially proposed that such misinterpretations had to be arbitrary. This assumption was challenged as most responses turned out consistent, centering around inner circle English, including the standard varieties.

A comparison of the recognition rates of both GA and RP recordings showed that the melancholic voices received remarkably lower scores than their conversational counterparts; thus, it can be assumed that participants did not recognize the speakers and evaluated them as separate examples, supporting hypothesis H_1-II. These outcomes might partly be due to the presentation style of the chosen AI-website, as the sad GA speaker was, compared to the other voices, most frequently identified as artificial. As Scherer (2003: 232) emphasizes, studies on

emotions are highly complex, since the research methodology easily influences the results. It is therefore possible that differences in recognition rates emerged, if real speakers were included instead of computer-generated voices. Moreover, the discrepancies in identification rates depending on emotional tone could also be a result of acoustic changes, arising through specific feelings, such as sadness. Concrete examples refer to influences on "respiration, phonation, and articulation" (Scherer 2003: 229). Consequently, sad voices might seem more unnatural and could therefore lead to misinterpretations in terms of the speakers' place of origin. Williams et al. (1999: 358) also confirm this argument by explaining that "cognitive mapping" is influenced by affections and feelings; nevertheless, there are no theoretical principles, precisely elaborating on the influence of emotions on variety recognition.

The analysis of moderating variables, such as mother tongue or respondents' own spoken variety, revealed interesting insights in recognition rates. Participants generally tended to label their personally used accent more accurately, supporting the argument by McKenzie (2008: 147) that a greater exposure to language enhances accent recognition. For instance, participants with GA as their self-identified English, reported greater scores for the American voices, regardless of emotional tone. Interestingly, the only partial exception appeared for users of RP. They still reported a higher degree of accurately labeled regional provenance for both speakers from England; however, compared to the GA group, a slightly higher recognition rate for the conversational US file appeared. On the one hand, it is possible that linguistic features of American English were pronounced more prominently compared to the other samples. On the other hand, higher rates could also be due to the influence of American culture across Europe (Carrie & McKenzie 2018: 325). Even though respondents tended to perform better for their own variety, these rates remained significantly lower for the sad compared to the happy recordings.

Doubtlessly, the sample size of English native speakers (n=5) was considerably smaller compared to L1 users of German (n=159); nevertheless, it can still be concluded that respondents' mother tongue affected recognition rates. Participants with English as their mother tongue assigned the overall highest scores of 100% to both GA speakers, indicating that mood was not an influential factor. Remarkably lower values appeared for the voices from England. Since four out of five native English speakers reported using GA, the influence of their own spoken variety became evident. In summary, emotional tone had a greater effect on LX users of English compared to natives. L1 German speakers assigned considerably lower values for the sad than the happy voices, while English natives did not show any differences in their recognition rates.

In summary, recognition rates tended to be lower for the melancholic speakers, indicating that participants did not realize that the voices occurred twice with a difference in emotional tone; hence, hypothesis H_1-II is confirmed.

8.5 AI-detection rates

AI was highly useful within the frame of this project, as it would have been difficult to find speakers, capable of all required accents; nevertheless, it needs to be stressed that some AI performances may not be authentic and could therefore falsify results; thus, the choice of an appropriate platform requires careful consideration. Although text-to-speech programs are easily accessible, their use still creates a dependance on technology, which, in the long run, can limit real-life interactions (Dugošija 2024: 278).

Doubtlessly, software has made great progress over the last years, enabling an imitation of real-life situations (Amin 2024: 888; Khanam et al. 2022: 398). Only a minority of participants indicated a potential usage of AI, particularly for items asking for the speakers' place of origin as well as a one-word description of their perceived presentation style. Regarding the latter, the melancholic samples received more responses, assuming an importation of AI, compared to the happy recordings.; nevertheless, number remained low with seven submissions for the sad GA voice and three descriptors for the counterpart from England. Even fewer answers were detected for the speakers' place of origin. While three participants believed that the sad GA file was computer-generated, only one person associated the melancholic RP sample with AI; thus, the most frequent submissions emerged for the sad GA speaker in both questions, claiming that certain linguistic features or the intonation sounded less natural and therefore, differed from the happy counterpart. Additionally, it is plausible that mood was influential here, as the happy equivalents on average obtained lower means.

Hypothesis H_1-IV, stating respondents did not detect that the recordings are computer-generated due to its progress over the last years, can be affirmed through the collected data.

9 Conclusion

This thesis explored language attitudes towards three varieties, namely RP, GA and SSE with particular emphasis on the influence of a speaker's emotional tone. The study employed the MGT and consisted of five recordings with two of the voices occurring twice, representing a different affective state.

Findings revealed a remarkable influence of the speakers' mood on respondents' evaluations, with significantly more negative attitudes detected for the melancholic compared to conversational samples. It is noteworthy that the AI-generated presentation style may have had an impact on these judgements, as some performances, depending on the language variety, could differ from real-life speech patterns. In addition, the contextual framing of the research project, specifically reporting the news, might have affected evaluations, as the sad delivery in this domain tends to be perceived less favorable.

An analysis of the selected character traits showed that participants' gender did not influence their ratings. In contrast, former residence in Anglophone regions played a more considerable role. Respondents with this international experience reported higher means for the RP speakers, which can be attributed to the *Contact Hypothesis,* as the majority of them lived in GB (Dörnyei & Csizer 2005: 328). Moreover, a substantial impact of the self-identified English was observable, since speakers are most likely to prefer varieties, that sound familiar to their personally used ones (Dalton-Puffer et al. 1997: 117). Whereas participants' mother tongue did not appear as an influential factor, their age indicated clear preferences. Specifically, the younger age group showed more favorable perceptions towards GA, while older people tended to rank RP higher.

Similar to the analysis of the given characteristics, gender did not affect judgements in terms of pronunciation clarity and suitability as a news reporter. The same applied to former residences in Anglophobe areas. Respondents' mother tongue as well as self-identified English proved to slightly influence evaluations of the speakers. Lastly, age partially impacted attitudes towards the selected varieties.

With regard to the rank ordering results it is noteworthy that RP was ranked first, which might be due to its presence in the Austrian school system (Dalton-Puffer et al. 1997: 126). The second rank was achieved by GA. Interestingly, the sad speaker revealed different outcomes compared to the happy counterparts, as the melancholic voice from the US elicited more positive attitudes than the one from England. Once again, the influence of emotional tone became evident. An

observation of potential influential factors demonstrated a similar pattern as the evaluations of the character traits. A considerable impact could be attributed to respondents' mother tongue, their self-identified English, age as well as prior residences in English-speaking countries.

The identification rates showed that GA received the most accurate labels, followed by the equivalent from England, which may be a result of their tremendous influence in the media as well as educational institutions (Ladegaard 1998: 266). Both sad files outlined lower identification scores than the conversational recordings.

Further research in the domain of language attitudes may build on this research project, as the use of AI provides promising potential. A possible follow-up study could adapt certain features and be conducted with real-life participants, incorporating a shift in emotional tone and subsequently, observe differences and similarities between authentic versus computer-generated voices; however, it might still be a challenge to present an authentic mood, as real people, similar to AI, would most likely act them out instead of genuinely expressing them. To avoid this issue, one option would be to evoke real emotions, for instance by including personal related matters.

10 Bibliography

Ahn, Hyejeong. (2017). *Attitudes towards Englishes: implications for teaching English in South Korea*. London: Routledge.

Ali, Zuraina. (2020). "Artificial Intelligence (AI): a review of its uses in language teaching and learning". *IOP Conference Series: Materials Science Engineering* 769(1), 1-6.

Amin, Eman Abdel-Reheem. (2024). "EFL students' perception of using AI text-to-speech apps in learning pronunciation". *Migration Letters* 21(3), 887-903.

Ammon Ulrich; Dittmar, Norbert; Mattheier, Klaus J.; Trudgill, Peter. (2004). "Dialect and accent". In: Ulrich, Ammon; Norbert, Dittmar; Klaus J., Mattheier; Peter, Trudgill (eds.). *Sociolinguistics* (2nd ed.) Berlin/Boston: De Gruyter, 267-273.

Angle, John.; Hesse-Biber, Sharlene. (1981). "Gender and preference in language". *Sex Roles* 7 (4), 449-461.

Baker, Colin. (1992). *Attitudes and language.* Clevedon: Multilingual Matters.

Balasubramaniam, Nagadivya; Kauppinen, Marjo; Rannisto, Antti; Hiekkanen, Kari; Kujala, Sari. (2023). "Transparency and explainability of AI systems: from ethical guidelines to requirements". *Information and Software Technology* 159, 1-15.

Barten, Martijn. (n.d.). "Tourism industry; what are the sectors within tourism?". *Revfine*. https://www.revfine.com/tourism-industry/#cultural-tourism (27 Feb 2025).

Bieswanger, Markus; Becker, Annette. (2017). *Introduction to English linguistics*. (4thed.). Tübingen: Narr Francke Attempto Verlag.

Briñol, Pablo; Petty, Richard J.; Barden, Jamie. (2007). „Happiness versus sadness as a determinant of thought confidence in persuasion: a self-validation analysis". *Journal of Personality and Social Psychology* 93(5), 711-727.

Brown, Adam. (1988). "Linking, intrusive and rhotic /r/ in pronunciation models". *Journal of the International Phonetic Association* 18(2), 144-151.

Carrie, Erin. (2017). "'British is professional, American is urban': attitudes towards English reference accents in Spain". *International Journal of Applied Linguistics* 27(2), 427-447.

Carrie, Erin; McKenzie Robert M. (2018). „American or British? L2 speakers' recognition and evaluations of accent features in English". *Journal of Multilingual and Multicultural Development* 39(4), 313-328.

Chan, Cecilia Ka Yuk; Hu, Wenjie. (2023). "Students' voices on generative AI: perceptions, benefits, and challenges in higher education". *International Journal of Educational Technology in Higher Education* 20(1), 1-18.

Dalton-Puffer, Christiane; Kaltenboeck, Gunther; Smit, Ute. (1997). "Learner attitudes and L2 pronunciation in Austria". *World Englishes* 16(1), 115-128.

Datska, Tetiana. (2019). "General American: Codified word phonemic structure variation specifics". *Linguistische Treffen in Wroclaw* II(6). 229-235.

Dörnyei, Zoltan; Csizer, Kata. (2005). „The effects of intercultural contact and tourism on language attitudes and language learning motivation". *Journal of Language and Social Psychology* 24(4), 327-357.

Douglas, Fiona M. (2009). *Scottish newspapers, language and identity*. Edinburgh: Edinburgh University Press.

Dugošija, Tatjana. (2024). „Benefits and challenges of artificial intelligence in English language teaching". *Knowledge-International Journal* 62(2), 275-280.

Fisher Hurt, John. (2001). "British and American, continuity and divergence". In: John, Algeo (ed.). *The Cambridge History of the English Language*. Cambridge: CUP, 59-85.

Fitria, Tira Nur. (2023). "Using natural reader: a free text-to-speech online with AI-powered voices in teaching listening TOEFL". *ELTALL* 4(2), 1-17.

Fodde, Luisanna. (2013). "American English: diversity, ethnicity and the politics of place". in: Anna, Belladelli; Roberto, Cagliero (eds.). *American English(es): linguistic and socio-cultural perspectives*. Newcastle upon Tyne: Cambridge Scholars Pub, 10-23.

Fuertes, Jairo N.; Potere, Jodi C.; Ramirez, Karen Y. (2002). "Effects of speech accents on interpersonal evaluations: implications for counseling practice and research". *Cultural Diversity & Ethnic Minority Psychology* 8(4), 346-356.

Garrett, Peter; Coupland, Nikolas; Williams, Angie. (2003). *Investigating language attitudes: social meanings of dialect, ethnicity and performance*. Cardiff: University of Wales Press.

Garrett, Peter. (2010). *Attitudes to language*. Cambridge: CUP.

Giles, Howard. (1970). "Evaluative reactions to accents". *Educational Review* 22(3), 211-227.

Giles, Howard; Watson, Bernadette. (2013). *The social meaning of language, dialect and accent. International perspectives on speech styles*. New York: Peter Lang.

Hammine, Madoka. (2021). "Educated not to speak our language: language attitudes and newspeakerness in the Yaeyaman language". *Journal of Language, Identity & Education* 20(6), 379-393.

Hannisdal, Bente. (2020). "A corpus-based study of /t/ flapping in American English broadcast speech". In: Anne, Przewozny; Cecile, Viollain; Sylvain, Navarro (eds.). *The corpus phonology of English*, 256-276.

Hill, Joseph Christopher. (2011). *Language attitudes in the American deaf community*. Washington DC: Gallaudet University Press.

Hillenbrand, James M. (2003). "American English: Southern Michigan". *Journal of the International Phonetic Association* 33(1), 121-126.

Hughes, Arthur; Trudgill, Peter; Watt, Dominic. (2005). *English accents and dialects: an introduction to social and regional varieties of English in the British Isles.* (4th edition). London: Hodder Arnold.

Huguet, Angel; Janes, Judit. (2008). "Mother tongue as a determining variable in language attitudes. The case of immigrant Latin American students in Spain". *Language and Intercultural Communication* 8(4), 246-260.

Kaminska, Tatiana Ewa. (1995). *Problems in Scottish English phonology*. Tübingen: Max Niemeyer Verlag.

Karatsareas, Petros. (2022). "Semi-structured interviews". in: Ruth, Kircher; Lena, Zipp. *Research methods in language attitudes.* Cambridge: CUP, 99-113.

Khanam, Fahima; Akhter Munmun, Farha; Afrin Ritu, Nadia; Kumar Saha, Aloke; Firoz Mridha, Muhammad. (2022). "Text to speech synthesis: a systematic review, deep learning based architecture and future research direction." *Journal of Advances in Information Technology* 13(5), 398-412.

Kingstone, Sydney. (2015). "'Scottish', 'English' or 'foreign': mapping Scottish dialect perceptions". *English Word-Wide* 36(3), 315-347.

Kircher, Ruth; Zipp, Lena. (2022). *Research methods in language attitudes*. Cambridge: CUP.

Kircher, Ruth. (2022). "Questionnaires to elicit quantitative data". in: Ruth, Kircher; Lena, Zipp. *Research methods in language attitudes.* Cambridge: CUP, 129-144.

Köchling, Alina; Wehner, Marius Claus. (2023). "Better explaining the benefits why AI? Analyzing the impact of explaining the benefits of AI-supported selection of applicant responses". *International Journal of Selection and Assessment* 31(1), 45-62.

Kovecses, Zoltan. (2000). *American English: An introduction*. Peterborough: Broadview Press.

Kraut, Rachel; Wulff, Stefanie. (2013). "Foreign-accented speech perception ratings: a multifactorial case study". *Journal of Multilingual and Multicultural Development* 34(3), 249-263.

Kuecker, Karrie; Lockenvitz, Sarah; Müller, Nicole. (2015). „Amount of rhoticity in schwar and in vowel plus /r/ in American English". *Clinical Linguistics & Phonetics* 29 (8-10), 623-629.

Kuznietsova, Tetiana. (2024). "Language attitudes in the mass media: what attitude towards the Ukrainian language was formed by the Odesa media (2024-2023)". *Zeitschrift für* Slawistik 69(3), 453-479.

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(4), 251-274.

Laroche, Michel; Li, Rong; Richard, Marie-Odile; Xu, Lu. (2022). "What is this accent? Effects of accent and language in international advertising contexts. *International journal of consumer studies* 46(4). 1209-1222.

Lausen, Adi; Schacht, Annekathrin. (2018). "Gender differences in the recognition of vocal emotions". *Frontiers in Psychology* (9), 1-22.

Loureiro-Rodriguez, Veronica; Acar, Elif Fidan. (2022). "The Matched-Guise Technique". In: Ruth, Kircher; Lena, Zipp (eds.). *Research methods in language attitudes.* Cambridge: CUP, 185-202.

McKenzie, Robert. (2008). "The role of variety recognition in Japanese university students' attitudes towards English speech varieties". *Journal of Multilingual and Multicultural Development* 29(2), 139-153.

McGuire, W. J. (1985). „Attitudes and attitude change". In G., Lindzey; E., Aronson (eds.), *Handbook of social psychology* (Vol. 2). New York: Random House, 223-246.

McKenzie, Robert M.; McNeil, Andrew. (2022). *Implicit and explicit language attitudes: mapping linguistic prejudice and attitude change in England*. New York, London: Routledge.

Milewski, Zadeusz. (2019). *Introduction to the study of language*. Berlin, Boston: De Gruyter.

Mugglestone, Lynda. (2012). "Varieties of English Received Pronunciation". In: Laurel, Brinton; Alexander, Bergs (eds.) *Historical English linguistics* (Vol. 2). Berlin, Boston: De Gruyter, 1899-1913.

Murf AI. (n.d.). Free AI voice generator: Versatile text to speech software. https://murf.ai (May 31 2025).

Peterson, Elizabeth. (2019). *Making sense of "Bad English": An introduction to language attitudes and ideologies*. London, Routledge.

Rangan, Pooja. (2023). *Thinking with an accent: toward a new object, method, and practice*. Berkeley: University of California Press.

Rastall, Paul. (2018). "What is language for? Interdisciplinary perspectives on the existence of language. *La Linguistique* 54(2), 3-20.

Rathcke, Tamara V.; Stuart-Smith Jane H. (2015). "On the tail of the Scottish vowel length rule in Glasgow". *Language and Speech* 59(3), 404-430.

Roach, Peter. (2004). "British English: Received Pronunciation". *Journal of the International Phonetic Association* 34(2), 239-245.

Roberts, Gillian. (2020). "Language attitudes and L2 pronunciation. An experimental study with Flemish adolescent learners of English". *English Text Construction* 13(2), 178-211.

Rochel, Johan. (2022). "Learning from the ethics of AI – a research proposal on soft law and ethics of AI". *Tilburg Law Review* 27(1), 37-59.

Rogers, Henry. (2014). *The sounds of a language: an introduction to phonetics*. Hoboken: Taylor and Francis.

Ryan, E. B., Giles, H.; Hewstone, M. (1988). "The measurement of language attitudes". In Ulrich, Ammon; Norbert, Dittmar; Klaus, Mattheier (eds.). *Sociolinguistics: An international handbook of language and society*. (Vol. 2). Berlin: De Gruyter, 1068–1081.

Scherer, Klaus R. (2003). "Vocal communication of emotion: a review of research paradigms". *Speech Communication* 40(1-2), 227-256.

Stephan, Christoph. (1997). "The unknown Englishes? Testing German students' ability to identify varieties of English". In: Edgar W., Schneider (ed.). *Englishes around the World*, 93-108.

Soukup, Barbara. (2009). *Dialect use as interaction strategy: a sociolinguistic study of contextualization, speech perception, and language attitudes in Austria*. Wien: Braumüller.

Trudgill, Peter. (1974). *Sociolinguistics: An introduction to language and society*. Baltimore: Penguin Books.

Van Hout, Roeland; Knops, Uus. (1988). *Language attitudes in the Dutch language area: An introduction*. Dordrecht: Foris.

Walsh, Olivia. (2021). "Introduction: in the shadow of the standard. Standard language ideology and attitudes towards 'non-standard' varieties and usages". *Journal of Multilingual and Multicultural Development* 42(9), 773-782.

Ward, Coleen A.; Bochner, Stephen; Furnham, Adrian. (2001). *The psychology of culture shock*. Hove: Routledge.

Wells, John C. (1982). *Accents of English II: the British Isles*. (Vol. 3). Cambridge: CUP.

Williams, Angie; Garrett Peter; Coupland, Nikolas; Preston, Dennis R. (1999). "Dialect recognition". In: Dennis, Preston. (ed.). *Handbook of perceptual dialectology*. Amsterdam: John Benjamins, 345-358.

Wilson, Rosemary; Dewaele, Jean-Marc. (2010) "The use of web questionnaires in second language acquisition and bilingualism research". *Second Language Research* 26(1), 103-123.

Yook, Cheongmin; Lindemann, Stephanie. (2013). “The role of speaker identification in Korean university students’ attitudes towards five varieties of English”. *Journal of Multilingual and Multicultural Development* 34(3), 279-296.

Yule, George. (2016). *The study of language*. (6th ed.). Cambridge: CUP.

Use of AI

Within this research project artificial intelligence, more precisely ChatGPT, was used for the purpose of proofreading and therefore served as a language refinement tool. In case of any uncertainties regarding vocabulary choice, grammar or coherence and cohesion, following prompts were applied:

- *Is the choice of vocabulary and grammar in this sentence appropriate for a master thesis?*
- *Is the sentence generally understandable and following a logical structure?*

Afterwards, ChatGPT offered some suggestions regarding the given dimensions; however, they sometimes did not seem suitable for this thesis and hence needed further improvement.

Besides that, it also proved to be a helpful platform in terms of finding synonyms; thus, I entered this prompt:

- *I need synonyms for…*

As a result, ChatGPT stated different alternatives together with several categories, such as general or academic and formal synonyms.

11 Appendix

11.1 Questionnaire

Hiring a news reporter

Dear participants,
I would like to invite you to take part in the following survey.
You are going to listen to five English speaking people reading a short text passage. All of these apply for a job as a professional news reporter. In order to find the most suitable one, questions are asked about the personal perception concerning the individual examples. Please listen to it once, while completing the questions. Besides that, you will find personal questions about your age, gender or highest completed level of education.
Your participation in the survey is voluntary. All parts are answered anonymously and will be treated confidentially. Your identity and privacy will therefore be protected and your responses will only be used for the purpose of this thesis.
You can refuse to participate without having to give a reason, or also break off once you have already started. In neither case will there be negative consequences for you. Should you feel uncomfortable during participation, do feel free to quit the survey. If there are any concerns, you want to share, feel free to get in touch with me (a11931421@unet.univie.ac.at).
The survey takes around 15 min.
If you have read the text above and agree to participate in the survey, please indicate your consent by clicking on the [NEXT] button below.
Thank you for your time and participation in advance.
Nadine Pfandl, BEd
If you have any questions do not hesitate to contact me: a11931421@unet.univie.ac.at

1. Age *

2. Gender *

Please Select

3. Country of birth *

4. Mother tongue *

5. Field of studies/work *

6. Highest completed level of education *

Please select

7. Have you ever lived in an English speaking country? *

◯ yes
◯ no

If you clicked yes, where did you live?

8. How would you describe your knowledge of English? *

	1	2	3	4	5	
insufficient	◯	◯	◯	◯	◯	very good

9. Which variety do you personally speak? *

◯ British
◯ American
◯ Scottish
◯ Irish
◯ Australian
◯ Other

10. If you have travelled to any of these countries, which did you enjoy the most?

◯ England
◯ USA
◯ Scotland
◯ Ireland
◯ Australia

Speaker A

Here you can find the first example (speaker A). Please click on it and listen to the short text passage, then answer the questions below.

1. Speaker A is.. (please tick the most suitable box):

	very	mostly	hardly	not
trustworthy	○	○	○	○
friendly	○	○	○	○
polite	○	○	○	○
intelligent	○	○	○	○
hardworking	○	○	○	○
outgoing	○	○	○	○
snobbish	○	○	○	○
reliable	○	○	○	○
wealthy	○	○	○	○
funny	○	○	○	○
confident	○	○	○	○

2. How clear was speaker A's pronunciation? *

	1	2	3	4	5	
unclear	○	○	○	○	○	very clear

3. How suitable do you think would speaker A be for presenting the news? *

	1	2	3	4	5	
not suitable	○	○	○	○	○	very suitable

4. How would you describe speaker A's presentation of the news in one word? *

5. Would you like to meet speaker A for a coffee? *

○ yes
○ no

Speaker B

Here you can find the second example (speaker B). Please click on it and listen to the short text passage, then answer the questions below.

1. Speaker B is.. (please tick the most suitable box):

	very	mostly	hardly	not
trustworthy	○	○	○	○
friendly	○	○	○	○
polite	○	○	○	○
intelligent	○	○	○	○
hardworking	○	○	○	○
outgoing	○	○	○	○
snobbish	○	○	○	○
reliable	○	○	○	○
wealthy	○	○	○	○
funny	○	○	○	○
confident	○	○	○	○

2. How clear was speaker B's pronunciation? *

	1	2	3	4	5	
unclear	○	○	○	○	○	very clear

3. How suitable do you think would speaker B be for presenting the news? *

	1	2	3	4	5	
not suitable	○	○	○	○	○	very suitable

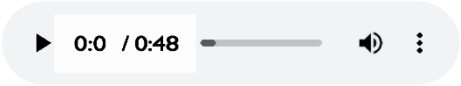
4. How would you describe speaker B's presentation of the news in one word? *

5. Would you like to meet speaker B for a coffee? *

◯ yes
◯ no

Speaker C

Here you can find the third example (speaker C). Please click on it and listen to the short text passage, then answer the questions below.



1. Speaker C is.. (please tick the most suitable box):

	very	mostly	hardly	not
trustworthy	◯	◯	◯	◯
friendly	◯	◯	◯	◯
polite	◯	◯	◯	◯
intelligent	◯	◯	◯	◯
hardworking	◯	◯	◯	◯
outgoing	◯	◯	◯	◯
snobbish	◯	◯	◯	◯
reliable	◯	◯	◯	◯
wealthy	◯	◯	◯	◯
funny	◯	◯	◯	◯
confident	◯	◯	◯	◯

2. How clear was speaker C's pronunciation? *

	1	2	3	4	5	
unclear	◯	◯	◯	◯	◯	very clear

3. How suitable do you think would speaker C be for presenting the news? *

	1	2	3	4	5	
not suitable	◯	◯	◯	◯	◯	very suitable

4. How would you describe speaker C's presentation of the news in one word? *

5. Would you like to meet speaker C for a coffee? *

- ◯ yes
- ◯ no

Speaker D

Here you can find the fourth example (speaker D). Please click on it and listen to the short text passage, then answer the questions below.

1. Speaker D is.. (please tick the most suitable box):

	very	mostly	hardly	not
trustworthy	◯	◯	◯	◯
friendly	◯	◯	◯	◯
polite	◯	◯	◯	◯
intelligent	◯	◯	◯	◯
hardworking	◯	◯	◯	◯
outgoing	◯	◯	◯	◯
snobbish	◯	◯	◯	◯
reliable	◯	◯	◯	◯
wealthy	◯	◯	◯	◯
funny	◯	◯	◯	◯
confident	◯	◯	◯	◯

2. How clear was speaker D's pronunciation? *

	1	2	3	4	5	
unclear	◯	◯	◯	◯	◯	very clear

3. How suitable do you think would speaker D be for presenting the news? *

	1	2	3	4	5	
not suitable	○	○	○	○	○	very suitable

4. How would you describe speaker D's presentation of the news in one word? *

5. Would you like to meet speaker D for a coffee? *

○ yes
○ no

Speaker E

Here you can find the fifth example (speaker E). Please click on it and listen to the short text passage, then answer the questions below.

1. Speaker E is.. (please tick the most suitable box):

	very	mostly	hardly	not
trustworthy	○	○	○	○
friendly	○	○	○	○
polite	○	○	○	○
intelligent	○	○	○	○
hardworking	○	○	○	○
outgoing	○	○	○	○
snobbish	○	○	○	○
reliable	○	○	○	○
wealthy	○	○	○	○
funny	○	○	○	○
confident	○	○	○	○

2. How clear was speaker E's pronunciation? *

	1	2	3	4	5	
unclear	○	○	○	○	○	very clear

3. How suitable do you think would speaker E be for presenting the news? *

	1	2	3	4	5	
not suitable	○	○	○	○	○	very suitable

4. How would you describe speaker E's presentation of the news in one word? *

5. Would you like to meet speaker E for a coffee? *

○ yes
○ no

Please rank the five speakers from 1 to five according to your preferences. Give the highest rank (5 stars) to the speaker you liked best and the lowest rank (1 star) to the person you liked least.

Speaker 1 *

1	2	3	4	5
○	○	○	○	○

Speaker 2 *

1	2	3	4	5
○	○	○	○	○

Speaker 3 *

1	2	3	4	5
○	○	○	○	○

Speaker 4 *

1	2	3	4	5
○	○	○	○	○

Speaker 5 *

1	2	3	4	5
○	○	○	○	○

Where do you think speaker A comes from? *

Where do you think speaker B comes from? *

Where do you think speaker C comes from? *

Where do you think speaker D comes from? *

Where do you think speaker E comes from? *

11.2 Abstract

This research project investigates language attitudes towards three varieties, specifically General American (GA), Received Pronunciation (RP) and Standard Scottish English (SSE), with particular emphasis on the effect of emotional tone on speaker evaluations. To examine this influence, the GA and RP recording occurred twice in a designed online survey, differing in mood (happy versus sad).

The study employed the matched-guise technique to minimize potential impactful variables, such as gender or pitch by using the same voice more than once; therefore, three female speakers, generated with an AI platform called Murf AI, were included.

Eleven character traits, assigned to the category of social status, warmth and dynamism were evaluated solely based on the concerned audio samples. In addition, dimensions including pronunciation clarity, suitability as a news reporter and willingness to have coffee with the speaker had to be assessed during the survey. Outcomes indicate that emotional tone appeared as a highly influential factor with considerably lower scores for the melancholic samples across all fields. Overall, the happy RP voice achieved the highest points in terms of intelligibility, appropriateness for reading the news and openness to have coffee. Moreover, most favorable rank ordering was also attested to the conversational RP voice, while the GA sample revealed the highest recognition rates. In contrast, the sad speaker from England frequently received the lowest ranking; however, these results vary depending on specific moderating variables, such as age, mother tongue, self-identified English or former residences in Anglophone areas.

11.3 Zusammenfassung

Diese Masterarbeit untersucht Haltungen und Einstellungen gegenüber drei englischen Sprachvarietäten, besonders General American (GA), Received Pronunciation (RP) und Standard Scottish English (SSE). Ein Schwerpunkt liegt auf den Auswirkungen von Emotionen der Sprecherinnen auf die Bewertungen der Teilnehmer*innen. Aus diesem Grund erschienen die Beispiele aus den USA und England zweimal, jeweils in einer glücklichen und traurigen Stimmung.

Die Studie nutzte die Matched-Guise-Technik, um mögliche einflussreiche Faktoren, wie Geschlecht oder Tonlage, durch mehrmaliges Verwenden von zwei Stimmen zu minimieren. Insgesamt waren drei weibliche KI-Sprecherinnen enthalten, die mit Hilfe der Website Murf AI erstellt wurden.

Teilnehmer*innen sollten elf Charaktereigenschaften, zugehörig zu sozialer Stellung, Wärme und Dynamik, nur anhand der gehörten Aufnahmen beurteilen. Weitere zu bewertende Bereiche waren Ausspracheklarheit, Tauglichkeit als Reporter oder Bereitwilligkeit, die jeweiligen Sprecherinnen persönlich zu treffen. Die Ergebnisse zeigten, dass die emotionale Lage der Stimmen sich drastisch auf die Haltungen der Zuhörer*innen auswirkte, wobei die traurigen Aufnahmen im Vergleich zu den glücklichen stets niedrigere Punkte erzielten. Allgemein wurde das fröhliche RP-Beispiel am besten in den Dimensionen Verständlichkeit, Angemessenheit sowie der Bereitschaft zu einem Treffen bewertet. Während die eben genannte Sprachvarietät auch in der allgemeinen Rangordnung den ersten Platz erlangte, wies das glückliche Gegenstück aus den USA die höchste korrekte Erkennungsrate auf. Im Gegensatz dazu konnten besonders negative Haltungen gegenüber der traurigen Stimme aus England erkannt werden. Nichtsdestoweniger gilt zu beachten, dass diese Ergebnisse, je nach zu untersuchender Variable, wie Alter, Muttersprache oder früheren Wohnsitzen in englischsprachigen Ländern, variieren können.

11.4 Data and audio files

Link to raw data and audio files: https://ucloud.univie.ac.at/index.php/s/jmAbWiiwxSmTpER