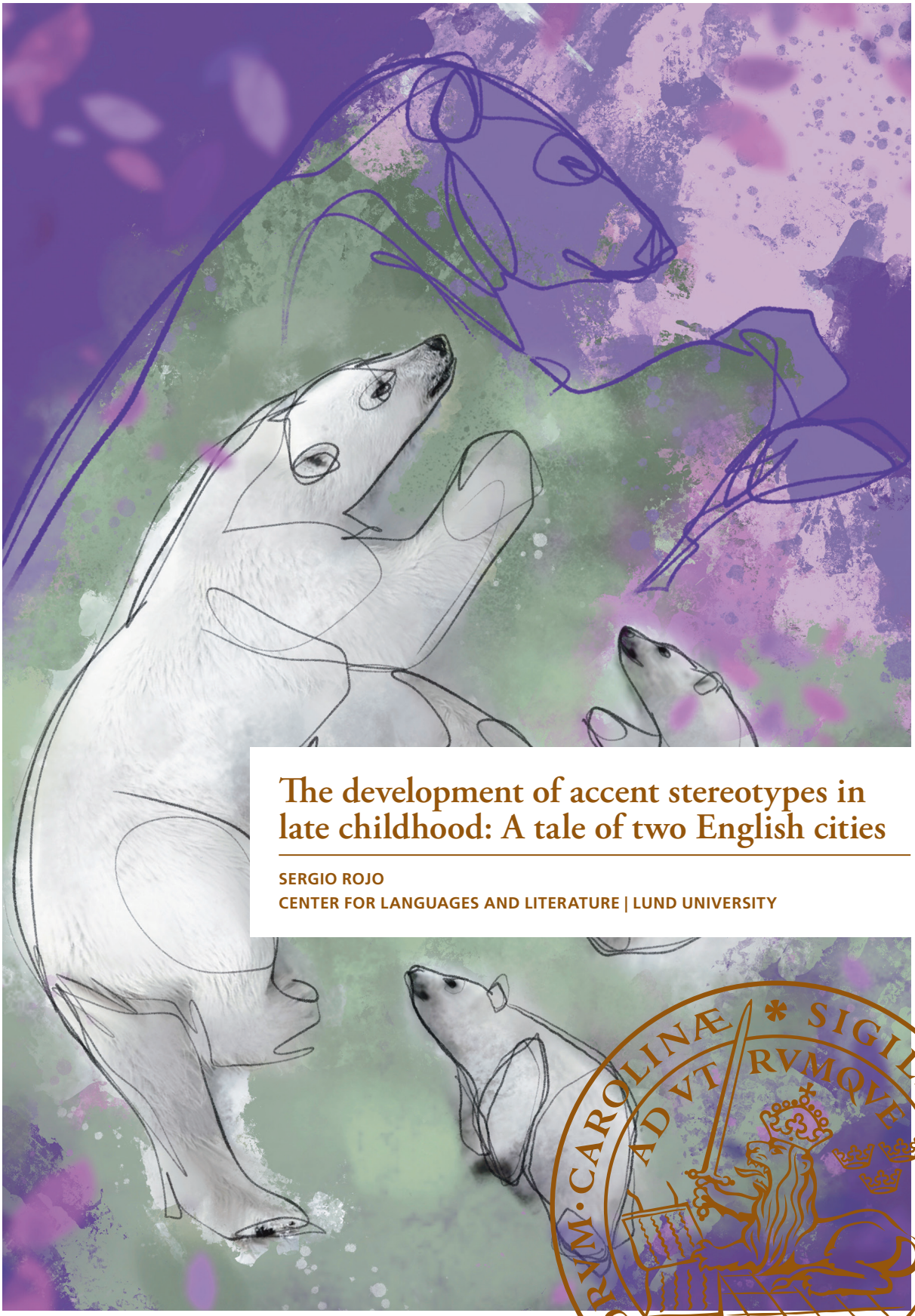
The development of accent stereotypes in late childhood: A tale of two English cities

Rojo, Sergio

2025

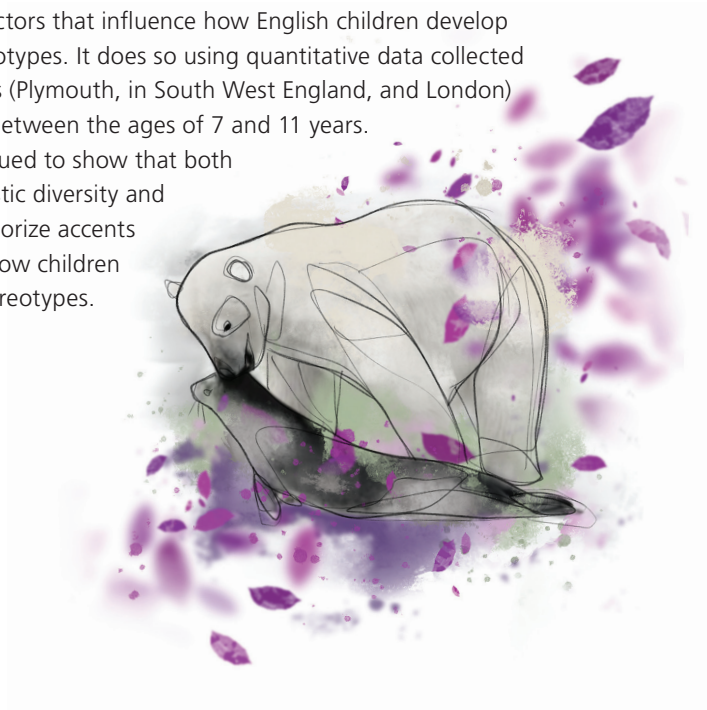


The development of accent stereotypes in late childhood: A tale of two English cities

SERGIO ROJO
CENTER FOR LANGUAGES AND LITERATURE | LUND UNIVERSITY

The development of accent stereotypes in late childhood: A tale of two English cities

English is spoken in a multitude of accents all over the world. But these accents are not neutral. They are a source of stereotypes. For example, people who speak with a Birmingham accent are considered less intelligent and less suitable as a university lecturer than people speaking with a standard British accent. This thesis investigates the factors that influence how English children develop these accent stereotypes. It does so using quantitative data collected in two British cities (Plymouth, in South West England, and London) with participants between the ages of 7 and 11 years. The results are argued to show that both exposure to linguistic diversity and the ability to categorize accents affect when and how children develop accent stereotypes.



Printed by Media-Tryck, Lund 2025 NORDIC SWAN ECOLABEL 3041 0903

LUND UNIVERSITY

Joint Faculties of Humanities and Theology
Center for Languages and Literature

ISBN 978-91-89874-84-8

The development of accent stereotypes in late childhood: A tale of two English cities

The development of accent stereotypes in late childhood: A tale of two English cities

Sergio Rojo

LUND
UNIVERSITY

DOCTORAL DISSERTATION

Doctoral dissertation for the degree of Doctor of Philosophy (PhD) at The Joint Faculties of Humanities and Theology at Lund University to be publicly defended the 27th of September 2025 at 10:00 in LUX C121, Helgonavägen 3, Lund.

Faculty opponent
Prof. Adriana Hanulíková (Heidelberg University)

Organization: LUND UNIVERSITY

Document name: Doctoral dissertation

Author: Sergio González de la Higuera Rojo

Date of issue: 27th of September 2025

Title and subtitle: The development of accent stereotypes in late childhood: A tale of two English cities

Abstract:

English is spoken with many different accents. As an example, someone born and raised in Manchester is likely to sound different from someone born and raised in Melbourne. In turn, someone born and raised in Tokyo who learned English as a foreign language is likely to speak English in a different way than the first two speakers. These accents are not neutral. Sociolinguistic research has shown that people use how others say something (rather than what they say) to judge them. For instance, people from the UK rate a person speaking with a standard British accent as more suitable to be a lawyer than someone with a London accent. In a parallel way, people consider someone speaking with a Birmingham accent to be less intelligent and less fitting for a job as a university lecturer than someone who speaks with a standard accent. These associations between accents and speakers' personal qualities are known as accent stereotypes.

This thesis investigates the factors that drive the development of accent stereotypes in late childhood. Specifically, it focuses on whether this development depends on (i) exposure to linguistic diversity, (ii) the type of accent used as stimuli (i.e. first-language vs. second-language accents), (iii) accent categorization and (iv) accent intelligibility.

A cross-sectional, quantitative investigation was carried out to answer these questions. Data from 249 children between the ages of 7 and 11 years was collected. This included children from Plymouth (Southwest England) and London. The experimental set-up included a survey for the children's caregivers and a battery of tests for children that included: an intelligibility task (speech-in-noise), a verbal-guise task (e.g. how smart they think a speaker of a given accent is), an accent classification task (grouping multiple speakers of each accent), and the British Picture Vocabulary Scale (a receptive vocabulary task).

The results suggest that exposure to linguistic diversity affects when and how children develop accent stereotypes. In contrast, whether an accent is a first-language or second-language accent does not affect this development. The ability to categorize accents appears to influence when children show accent stereotypes, while accent intelligibility seems to be unrelated to how children learn to match accents to speakers' personal traits.

Keywords: Sociolinguistics, Language acquisition, Social psychology, Language attitudes, Accents, Stereotypes, Linguistic stereotypes, Developmental sociolinguistics

Classification system and/or index terms (if any):

Supplementary bibliographical information:

Language: English

Number of pages: 235

ISBN (print): 978-91-89874-84-8
ISBN (digital): 978-91-89874-85-5

Recipient's notes:

Price:

Signature

Date: 2025-07-29

The development of accent stereotypes in late childhood: A tale of two English cities

Sergio Rojo

LUND
UNIVERSITY

Centre for Languages and Literature

The Joint Faculties of Humanities and Theology

Lund University

ISBN (print): 978-91-89874-84-8

ISBN (digital): 978-91-89874-85-5

Printed in Sweden by Media-Tryck, Lund University, Lund, 2025

Media-Tryck is a Nordic Swan Ecolabel certified provider of printed material. Read more about our environmental work at www.mediatryck.lu.se

MADE IN SWEDEN

"When the stars were right, They could plunge from world to world through the sky; but when the stars were wrong, They could not live."
H.P. Lovecraft, The Call of Cthulhu

Table of Contents

Acknowledgements

No one ever said doing a PhD would be easy. But I could have never imagined what a rollercoaster I said yes to back in 2019. In a world of constant chaos and change, one thing is certain: I would not have been able to withstand the ride if it were not for my amazing supervisors: Carita Paradis and Joost van de Weijer. Words cannot do justice to how grateful I feel for your unwavering support throughout my studies. Your faith in me, your guidance, patience, inspiration and encouragement on academic, administrative and personal matters have been invaluable. I would not have been where I am now if it wasn't for you. For that, and more things that I cannot begin to express, I am deeply and eternally grateful.

I also want to express a most sincere and heartfelt thank you to two people whose names may not be formally attached to this thesis, but whose help, beyond all expectations, made this thesis possible: Kathleen McCarthy and Caroline Floccia. I will never forget the generosity, knowledge, time and advice you have given me since the very beginning of my PhD journey.

I would also like to offer a big thank you to Victoria Johansson for accepting the role as my mock opponent. Her insightful and valuable comments were of great importance to the final phase of this thesis.

The recording of stimuli and data collection were only possible thanks to the invaluable assistance of Anna Caunt, Scott Kunkel, Yas Merrett de Souza, Rachel Riley and Roxana Rosculet. Thank you for all your time and all the zoom chats! I also don't know what I would have done without the help of Billy Jones, Simon Coxe, Mel Pinet, Laura Street, The Birkbeck BabyLab & ToddlerLab, and the Goldsmiths InfantLab for the data collection. I am truly thankful! But, of course, the most important people are the 249 children (and their parents!) who participated in this investigation. Your enthusiasm and willingness to help made the data collection a treat in its own right. To all of you I offer a warm and heartfelt thanks. I also want to thank Erez Levon, Devyani Sharma, Zoe Adams and Christian Ilbury. Your insightful and constructive discussions shaped this thesis at the beginning of my PhD studies.

During my years as a PhD student, I have been lucky to be able to present my (preliminary) results at the English Linguistics research seminar and the LAMiNATE research platform. Thank you all for your invaluable and constructive comments and for taking an interest in my research! Warm thanks go to my colleagues at the English unit and at the HT faculties broadly. I would like to especially thank Fabian and Lene for their kindness and assistance when I needed it most, Satu and Eva for making me a better teacher, as well as Annika L., Fanni F., Kristina T., Jane N., and *everyone* at the HT library's Forskningsstöd for always being welcoming and guiding me through the nightmare that LU's administration can sometimes be.

Of course, the hardships (and joys!) of the PhD life were made significantly better by being surrounded by wonderful, amazing people: Elin, Jorunn, John, Mette, Sophia,

Mari, Ben, Elena, Freja, Lena, Eleni, Kajsa, Tomas, Vlad. Thank you for being there despite my long (and constant) absences from Lund. You will always have a special place in my heart.

I would also like to thank *Stiftelsen Olof Sagers stipendiefond för studier och forskning i engelska* and *Lundborgska Idofonden* for their generous economic support during my studies.

Og selvfølgelig kan jeg ikke skrive de her linjer uden at nævne de vidunderlige, omsorgsfulde, sjove, kloge og uvurderlige mennesker, som jeg er så heldig at kende herhjemme. De sidste seks år har været nogle af de bedste og værste. Men I har altid (også langt inden jeg begyndte på den her ph.d.) været ved min side, uanset om jeg har haft brug for at grine, fejre, græde eller bare for at være mig selv. Maria, Steffen, Anne, Karoline, Line, Nico, Johanne, Vibe, Danny, Benno, Vibe, Emilie, Hjalte og Thora: *Tak*. Og Jens, du har set og oplevet *det hele*. Tak for din kærlighed og for din ubetingede og urokkelige støtte.

Dice el refrán que lo mejor se deja para el final. Así que ahora toca mencionar a la gente maravillosa de Madrid. Primero la familia elegida. Estimadas LNT (Sara, Fani, Tamara, Yas): se hace difícil poder expresar cuánto os quiero agradecer el haber estado siempre a mi lado pese a la distancia. Gracias por las tortitas y bollos, el chino y todo lo demás. Por último, la familia dada: Ángeles, Jose, Loren y Lucky. Desde pequeño, me habéis apoyado en todo lo que quería hacer y esto no ha cambiado en los últimos seis años. Vuestro cariño internacional me ha dado fuerzas cuando más las necesitaba. Mil gracias de corazón.

Any shortcomings, if present, are my own.

Abbreviations

AAVE	African American Vernacular English
ANOVA	Analysis of variance
BPVS	British Picture Vocabulary Scale
CI	Confidence interval
ERP	Event Related Potential
GA	General American
GB	General British
IAT	Implicit Association Test
L1	Accent result of first language acquisition
L2	Accent result of foreign language acquisition
M	Mean
Mdn	Median
MLE	Multicultural London English
PCA	Principal Component Analysis
SD	Standard deviation
SE	Standard error
SSBE	Standard Southern British English
X_{chi}	Result involving Chinese-accented English
X_{fn}	Result involving the FUN question
X_{fr}	Result involving the FRIENDLY question
X_{fre}	Result involving French-accented English
X_{hw}	Result involving the HARDWORKING question
X_{mle}	Result involving MLE
X_{mn}	Result involving the MONEY question
X_{ply}	Result involving the Plymouth accent
X_{sm}	Result involving the SMART question
X_{ssb}	Result involving SSBE
X_{tr}	Result involving the TRUST question

List of figures

List of tables

1 Introduction

December 13th, 2009, Copenhagen. The Danish *Green Left* (*Socialistisk Folkeparti* in Danish) chairperson Villy Søvndal gives a speech at the *Copenhagen Climate Change Conference*. In it, he aims at raising awareness of climate change and utters the words *the ice is melting at the poles*. This sentence became an instant social media meme and was subject to mockery (e.g. Klarskov & Henriksen, 2011; Søndergaard, 2011). To this day, he is still ridiculed for it and the quote has become an everyday expression (Amdisen, 2019; Nyvold, 2024; Theilgaard, 2024). His offence? He produced this sentence and the entire speech with a (some would say) strong Danish accent[1].

People being judged in this way for the way they speak is not uncommon. In 2016 in the UK, Shadow Education Secretary Angela Rayner received abusive emails after an appearance on Channel 4 in which she spoke with her Northern British accent (Gill, 2016). Furthermore, she was also subject to snobbery from her Labour peers because of her accent. More recently, presenter Katie Owen has spoken up about how her Welsh accent has led to mockery and discrimination (Bevan, 2023). Sometimes people anticipate these reactions and act preventively to mitigate them. For instance, Danish media has reported how Danish people with "foreign" names change their names to increase the chances of getting a job (e.g. Kjempff, 2022) or perform better at their telemarketing job (e.g. Kryger, 2017)[2].

Language is considered the human communication tool *par excellence*. However, as the examples above show, what one communicates is not always what one intends. People use how others say something (rather than what they say) to create an opinion about them. This knowledge is not new to sociolinguists. The field of language attitudes, which typically falls under sociolinguistics, studies how people evaluate others based on their use of language as well as how people evaluate languages (or dialects/accents) as a whole.

There is a plethora of research, starting from the mid-twentieth century, investigating the dynamics of these attitudes. There is just one caveat. The vast majority of this research has focused on the language attitudes of adults. Language attitudes are

[1] This portion of the speech can be seen here: https://www.youtube.com/watch?v = D6zu6zKibEI&t = 1m1s

[2] See The Accentism Project (https://accentism.org) for more stories in media dealing with how people are judged negatively because of their accent.

a learned behavior. Nonetheless, sociolinguists have typically not focused on how these attitudes are acquired during childhood. One could perhaps expect that researchers from the field of language acquisition have focused on how language attitudes are acquired. This has not been the case. Researchers within the field of language acquisition have typically focused on the traditional core areas of language: vocabulary, grammar and phonology. This state of affairs has resulted in very little research on how children learn to evaluate people based on their use of language (see similar sentiments in Barbu et al., 2013; DeJesus et al., 2019; Dossey et al., 2020; Fantini, 2012; Jones et al., 2017; Kinzler, 2013, 2021; Kinzler & DeJesus, 2013a, 2013b; Nardy et al., 2013; O'Shannessy, 2013; Rosenthal, 1974). There are individual studies on the subject dating back to the 1960s (e.g. Anisfeld & Lambert, 1964; Labov, 1964). But it is only relatively recently that the field of developmental sociolinguistics has emerged and that it has been possible to start unveiling how children acquire language attitudes. Nevertheless, much is still unknown.

The present thesis aims at advancing our knowledge of this field. Its focus is on how children develop stereotypes toward speakers of different accents. Simply put, accent stereotypes are understood as associations between the accent someone uses and their cognitive skills, personality traits, social status, etc. For instance, people are known to judge speakers with a Birmingham accent to be less intelligent and less appropriate for a job as a university lecturer than speakers of standard English (Giles et al., 1975). Similarly, people form the UK consider a person speaking with a standard British accent as more suitable to be a lawyer than someone with a London accent (Levon et al., 2021). As the following Chapter shows, these kinds of associations have important consequences for speakers. Therefore, understanding how accent stereotypes are transmitted will not only help us understand how children learn to use language as a social tool, but will also bring us a step closer to understanding the nature of these kinds of attitudes.

1.1 Aim and research questions

The superordinate aim of this thesis is to investigate *how English children develop accent stereotypes in late childhood and what factors influence this process.* To explore it, a cross-sectional data collection was carried out in two English cities: Plymouth (Southwest England) and London. This broad question was narrowed down and operationalized as the four research questions (RQs) below. This was done in light of the complex nature of language attitudes revealed in the adult research. This thesis cannot explain all different facets of how children learn to associate accents with speakers' traits. Therefore, these RQs aim at bringing us a step closer to understanding the subject. They were framed as exploratory RQs due to the under-researched nature of the field, as mentioned in the previous section. Moreover, it was believed that either a positive

or negative answer to each of them would uncover a further piece of the puzzle that is the development of accent stereotypes.

The four RQs below were chosen to paint a holistic picture of the factors that influence the acquisition of accent stereotypes in childhood. This is the result of the Cognitive Linguistics approach followed in this investigation (see Section 2.4). Briefly put, this thesis relies on the theoretical assumption that, to fully understand any linguistic behavior, one must take an approach that foregrounds its relations to its cognitive bases and sociocultural context. As a consequence, this thesis focuses on four RQs that view the development of accent stereotypes from different angles. The first two RQs focus on aspects external to the individual. RQ1 focuses on whether the sociocultural context of a child influences the way they develop accent stereotypes and, if so, how, whereas RQ2 centers on whether the properties of the stimuli influence this development. In contrast, the last two RQs ask whether the development of accent stereotypes relies on other sociolinguistic skills: accent categorization (RQ3) and accent intelligibility (RQ4). These RQs are discussed in turn.

RQ1 Does linguistic diversity in a child's environment while growing up shape how or when they develop accent stereotypes?

RQ2 Do children develop stereotypes toward L2 accents at an earlier age than toward L1 accents?

RQ3 Is there a relationship between the development of accent stereotypes and the ability to categorize accents?

RQ4 Is there a relationship between the development of accent stereotypes and the ability to understand accents?

RQ1 pays attention to whether a child's environment during development affects how they learn accent stereotypes. The sociocultural environment in which a child grows up is of theoretical relevance. Specifically, exposure to linguistic diversity has been reported to be a key variable in their linguistic and social development (Durrant et al., 2015; Howard et al., 2014). Therefore, it cannot be expected that children develop accent stereotypes at the same time or the same way when facing different levels of language variation.

RQ2 focuses on whether children acquire stereotypes toward different kinds of accents at different ages. This is based on the fact that children do not appear to develop attitudes toward all forms of linguistic input at the same time (see Chapter 3). Moreover, children are able to differentiate their own accent from L2 accents (sometimes also called foreign accents; see Section 2.1) at an earlier age than from L1 accents (sometimes also called regional accents; see Section 2.1) (Floccia, Butler,

Girard, et al., 2009; Girard et al., 2008; Goslin et al., 2012; Wagner, Clopper, et al., 2014). Therefore, it may be possible that children develop accent stereotypes toward L2 accents at an earlier age than toward L1 accents. The ability to distinguish accents could be considered a precondition for the ability to show stereotypes toward them (Dossey et al., 2020; Wagner, Clopper, et al., 2014). As Wagner, Clopper, et al. (2014, p. 1068) pointed out, "[c]hildren may well be prepared to use language to make social judgements, but until they can reliably perceive differences among different language variants, there would seem to be little basis for those judgements to be made." Furthermore, children develop a sense of national identity and a preference for people with their same nationality relatively early during development, around the age of 5 years. This takes place before they develop a sense of a more local, region-based identity (Barrett, 2007; Barrett & Short, 1992; Barrett et al., 2003; Brown, 2011). Hence, this may further cause an earlier development of accent stereotypes in childhood toward L2 accents.

RQ3 and RQ4 focus on whether the development of accent stereotypes is related to other cognitive and (socio)linguistic skills and, if so, how. On the one hand, RQ3 leans on RQ2 and asks whether, more generally, the ability to categorize accents is necessary for developing stereotypes toward them. It thus takes a broader approach beyond a simple L1–L2 divide. On the other hand, RQ4 is based on contradicting results in the literature. Intelligibility has provided conflicting results with regard to its relationship to language attitudes. Research on adult participants has shown that there might be a relationship between intelligibility and attitudes (Dragojevic & Giles, 2016; Dragojevic & Goatley-Soan, 2022a). Specifically, higher levels of intelligibility are correlated with more positive attitudes. In turn, lower levels of intelligibility lead to more negative attitudes. In developmental research, however, intelligibility has been shown to be irrelevant regarding children's attitudes toward different accents and languages (see Chapter 3). Nevertheless, there are significant methodological differences between these kinds of studies. For adults, speech-in-noise is typically used. For children, speech is typically altered significantly. The stimuli typically include low-pass filtered texts, nonsensical sentences (such as the Jabberwocky poem by Lewis Carroll) or backwards speech. Thus, these differences in results may simply reflect methodological disparities.

Finally, an auxiliary aim of the study, though not an overt RQ, is how children growing up in England develop accent stereotypes. The UK has a long tradition of sociolinguistic research in general and research on language attitudes in particular (e.g. Bishop et al., 2005; Coupland & Bishop, 2007; Giles, 1970, 1971; Pear, 1931; Sharma et al., 2022; see also Mugglestone, 2007; Garrett, 2010 for overviews). However, there is a lack of research investigating the development of accent stereotypes in children in the UK, particularly in contrast to the US. Where a child grows up is important due to the social nature of accent stereotypes. Therefore, this study focuses on children growing up in England.

1.2 Structure of the thesis

This thesis consists of eight chapters, the first of which is the current introductory chapter. The next two chapters provide the theoretical core of the investigation. Chapter 2 anchors the thesis within the fields of sociolinguistics, on the one hand, and social psychology/social cognition, on the other. It bridges the gap between these subjects and problematizes current terminology. Specifically, it describes how *accent* and *stereotype* are conceived of in this investigation and how the study of accent stereotypes has been partially neglected in social cognition. Finally, Chapter 2 closes by arguing that anchoring the study of accent stereotypes in Cognitive Linguistics provides important theoretical tools to understand their nature. Chapter 3 provides an overview of the field of the development of language attitudes. This area of study is characterized by fragmented research from different academic traditions that have not typically informed each other. Therefore, Chapter 3 brings together the key findings from the field in one place. It also highlights recurring themes in the literature and the difficulties in creating generalizations.

The next four chapters constitute the empirical side of this thesis. Chapter 4 describes the methodological choices and considerations behind the data collection. For this purpose, it focuses on methodological drawbacks from the literature in Chapter 3 as well as the importance of choosing proper techniques in developmental research. Specifically, it shows how the design of the present study was created to be used with younger participants and to enable comparisons with current, relevant literature. The results of this data collection are described in the following three chapters. Chapter 5 covers the results from the participants from Plymouth, while Chapter 6 does so for London. Chapter 7 compares the results from both cities to uncover how differences in children's sociocultural contexts affect how they develop accent stereotypes.

Finally, Chapter 8 concludes the thesis. It briefly summarizes this investigation, highlights its main contributions and proposes future lines of research that can be carried out to further understand how children develop accent stereotypes.

2 Background and terminology

Chapter 1 established that the overarching aim of this thesis is to investigate when and how English children develop accent stereotypes. There are two main keywords in this statement: *accent* and *stereotype*. Despite their common use in everyday speech, these terms are frequently used in the literature with conflicting meanings. One author's definition of the term *accent* or *accented* frequently differs from how another author uses these terms. Therefore, this chapter outlines the terminological choices of this thesis. These choices place this investigation within its general context and ensure transparent use of terms. For this purpose, Section 2.1 presents the term *accent* within a sociolinguistic framework. Section 2.2 provides a summary of how social psychology frames stereotypes as part of the broader concept of attitudes. Afterwards, Section 2.3 discusses how both these concepts are applied to sociolinguistics by illustrating how language attitudes and accent stereotypes are conceptualized in this investigation. Finally, Section 2.4 addresses how the study of accent stereotypes can benefit from being anchored in the framework of Cognitive Linguistics.

2.1 What are accents?

The term *accent* is used with different meanings both in everyday speech and linguistics. This has sometimes led to confusion in the literature (see a discussion in, for instance, van der Hulst, 2011). As a technical term, it can be used to describe the syllable or syllables that have prominence in a word due to pitch, loudness, and/or the quality and quantity of the segments in it (Cruttenden, 2014, p. 242). Based on this definition, the first syllable in *tiger* can be said to be accented[3]. Second, it is also used in the phrase *pitch accent* — sometimes also referred to as nuclear accent or nuclear tone. In this instance, it refers to the peak of prominence in a tone unit (Cruttenden, 2014, p. 242; van der Hulst, 2011). For example, in the sentence *she was running*, the first syllable in the word *running* has the main pitch accent. Third, the term is also used to describe pitch accent languages (Cruttenden, 1997, pp. 10-12). This type of language is characterized by the presence of one syllable in a word that has prominence. Such

[3] Linguists also label this concept *stress*, leaving the term *accent* to the other meanings described below (e.g. Cutler, 1984).

prominence is marked by specific changes of pitch; that is, they are associated with a specific tone (Ambrazaitis & Bruce, 2006; Bruce, 2005; Cruttenden, 1997, p. 11). An example of this kind of language is Swedish, in which the word *anden* has different meanings depending on which accent the first syllable has (Cruttenden, 1997, p. 11). The word *´anden* has what is typically referred to as Accent 1 and means *the duck*, while *`anden* has Accent 2 and means *the spirit*.

In contrast to these meanings, this thesis employes the term *accent* with its sociolinguistic meaning. From this point of view, *accent* is defined as the differences in segmental and suprasegmental features across spoken varieties of a language (Chambers & Trudgill, 1980). For instance, a person from the UK may pronounce the word *water* as [ˈwɔːʔə], someone from the US may realize it as [ˈwɔːɾɨ] and a French speaker who speaks English as a foreign language may say [wɔˈtɛʁ]. This does not mean that all differences in pronunciation are related to accents (Honey, 2017). In some cases, different pronunciations of a word may be the result of lexical idiosyncrasies and, therefore, fall in the realm of dialectal differences rather than accent differences (see the end of this section for the difference between *dialect* and *accent*). For example, most speakers of British English pronounce the word *schedule* as /ˈʃedjuːl/, while the majority of speakers of American English pronounce the word as /ˈskedʒuːl/. A further example of pronunciation differences that are not regarded as accents involves respiratory illnesses. A person sounding different due to having the influenza would not be categorized as having a different accent.

A key variable for the sociolinguistic definition of accent is that changes in pronunciation need to be generalized and regularized. Bringing back the case of British and American English, there are numerous systematic differences across these accents (Trudgill & Hannah, 2017). For example, American English does not have the vocalic phoneme /ɒ/ present in British English. As a consequence, words such as *hot, top* and *pond* (all of which are pronounced with /ɒ/ in the UK) are pronounced with /ɑ/ in American English. Similarly, the phoneme /t/ is realized as a voiced alveolar tap [ɾ] in between a stressed and unstressed vowel in North American English in words like *water*, *better* and *pottery*. In contrast, the Standard British accent realizes this phoneme as [tʰ] in these contexts. These phonemic (/ɒ/ versus /ɑ/) or phonetic ([ɾ] versus [tʰ]) differences do not lead to differences in so-called propositional meaning (but see Section 2.4).

Despite the examples used above, accents do not only refer to differences as a function of a person's geographical background. Accents also vary as a function of the context of an interaction or the social group the speaker belongs to (based on socioeconomic status (SES), ethnicity, etc.). With regard to the former, speakers may alter their accent based on whether they are talking to a friend at home or a stranger at a formal job interview. For SES, the pronunciation differences between Received Pronunciation (RP) and Cockney are also examples of different accents, the former associated with higher SES while the latter is equated with speakers from a lower SES.

In a British context, a person's SES and geographical background have a closely intertwined relationship when it comes to their accent (e.g. Wells, 1982). This relationship has often been depicted in the shape of a pyramid, as in Figure 2.1. The area inside the pyramid illustrates how much accent variation there is at a given intersection between region and SES. The greatest variation in regional accents in England[4] is found among speakers of lower SES. In contrast, the higher a person's SES is, the fewer regional traits their accent has. This leads to fewer distinguishing accent features for middle-class speakers than for speakers of lower-class backgrounds. Finally, speakers from the highest classes speak the same language variety. In the context of England, this refers to RP, which is often said to be characterized by its "non-localizability" (Wells, 1982, p. 14). In other words, it is categorized as an accent that does not reflect a speaker's geographical background.

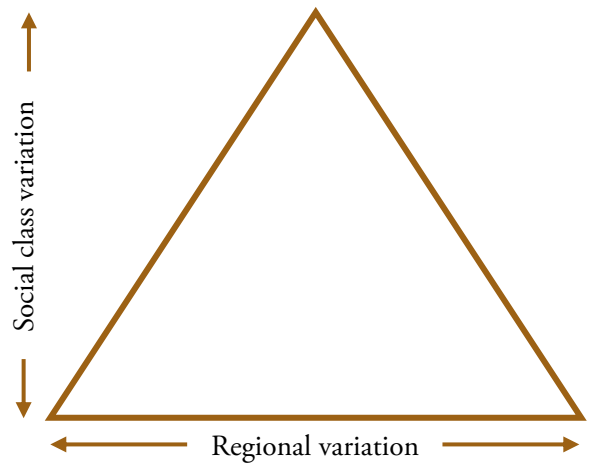


Figure 2.1. Relationship between class and geography for accents
Figure based on Wells (1982, p. 14). © 1982 CUP. Reproduced with permission of Cambridge University Press through PLSclear.

A key aspect of the definition of accent above is the linguistic consensus that everyone has an accent (Esling, 1998; Lippi-Green, 2012; Matsuda, 1991). That is, there are no such things as speakers without accents nor unaccented speech. All speakers of a language pronounce their utterances in a certain way. Therefore, everyone has an accent. An equivalent non-linguistic situation is that of gender. One can say that someone is, for example, a man or a woman. However, one would not say that one of

[4] This pyramid model appears to be exclusive to England and Wales, as pointed out by Wells (1982, p. 15). It does not capture the sociolinguistic diversity in other English-speaking countries such as Ireland, Scotland, the US, Australia or New Zealand.

them is gendered or has a gender while the other one is un-gendered or does not have a gender. Classifications such as (i) accented and unaccented speech or (ii) who has or doesn't have an accent are ideological without linguistic bases (Lippi-Green, 2012; Matsuda, 1991). As Matsuda puts it:

> When […] parties are in a relationship of domination and subordination we tend to say that the dominant is normal, and the subordinate is different from normal. And so it is with accent. […] People in power are perceived as speaking normal, unaccented English. Any speech that is different from that constructed norm is called an accent (Matsuda, 1991, p. 1361)

The literature also uses adjectives such as *native* (accent), *foreign* (accent) and *regional* (accent) with different (and even conflicting) meanings. *Foreign accent* is sometimes used to refer to the accent a native speaker has from the point of view of a native speaker from a different country. In this context, an American speaker may be said to have a foreign accent by an Australian speaker. In contrast, the phrase is sometimes understood as referring to the accent a speaker has because of speaking a foreign language. Thus, someone born and raised in Italy, who learns English as a foreign language, may be said to speak English with a foreign accent. There are also similar contradictory uses of *regional accent*. In some cases, the phrase refers to different accents within the same country, excluding accents from other countries. For example, from the point of view of a London speaker, a person with a Birmingham accent has a "regional accent", but someone from the US does not. In other cases, the phrase is understood as referring to any accent someone has because of learning a language as their native language.

To avoid confusion, this investigation categorizes accents in a two-step fashion: first, from the point of view of the speaker and, second, from the point of view of the hearer. First, the adjectives *L1* and *L2* are used to categorize accents from the point of view of the speaker. The term *L1* refers to accents that speakers have due to acquiring a language as their native language (but see Davies, 2003 for a critique of the term "native speaker"). Hence, someone born and raised in Sydney would have a specific L1 accent of English, while someone born and raised in Edinburgh would have a different L1 accent of English. The term *L2* refers to accents that speakers have due to learning a foreign language. For instance, a native speaker of Finnish who learned English at school has a specific L2 accent when speaking English.

Second, accents can be further categorized from the point of view of the hearer using the adjectives *native* and *nonnative*. A speaker is said to have a native accent when the hearer has the exact same accent. Following the same logic, a speaker is said to have a nonnative accent when the hearer has a different accent. For instance, one can imagine a context in which person A speaks English with a General American accent and listens to two people speaking English: person B (with a General American accent) and person C (with a New Zealand accent). From the point of view of person A, person B has the same native accent, since they have the same accent. In contrast, person C has a

nonnative accent, since the accents person A and person C have are different[5]. These accents could then be categorized from the point of view of the speaker. In this case, both accents would be referred to as L1 accents of English. This use of the adjective *native* matches that of *accented* in Cristia et al. (2012).

A further crucial dimension for the understanding of accents (as well as dialects and languages more generally) is that they do not have clear-cut boundaries. This relates to the sociolinguistic concept of dialectal continua (Heeringa & Nerbonne, 2001). Labels such as *General American accent*, *Northern English accent* or *French-accented English* act as categories. Each of these labels includes different varieties considered to belong together based on some linguistic features. These accents are similar to each other in some respects and different in others, creating a network of accents that look like a family-resemblance structure (Wittgenstein, 1986). A non-linguistic example involves color terms. The term *red* is a category including colors that differ physically in terms of hue, brightness and saturation. Therefore, if one labels two colors as *red*, it does not entail that they are the exact, same, physical color. Similarly, terms such as *Northern English* include accents that differ in their exact realizations. As a consequence, two speakers of Northern English do not necessarily speak the exact same way.

The lack of unambiguous boundaries between accents does not only apply to geography; that is, the impossibility of marking the exact location where, for instance, a Northern British accent ends and where a Southern British accent starts. An important implication of the fuzziness of accent boundaries for attitudinal research is the concept of accent strength. Accent strength can be defined as the degree to which an accent differs from a given baseline accent. One can thus speak of a strong/broad or weak accent. Accent strength has important implications for sociolinguistics. Adults' views of accents are modulated by how strong the accent is. Typically, adults have more negative attitudes toward stronger accents than toward weaker versions of the same accent (Brennan & Brennan, 1981a; Giles, 1972; Hendriks et al., 2017; Nesdale & Rooney, 1996; Ryan et al., 1977). Furthermore, accent strength is also relevant to how early children are able to distinguish different accents (Floccia et al., 2006; Goslin et al., 2012). How strong (or weak) an accent is considered to be is nevertheless not objective, as already noted by Giles (1972). It depends on the perception of the hearer. What may be categorized as a strong accent by a North American English speaker need not be categorized as such by a Scottish English speaker and vice versa.

Two final notes on terminology are in order. First, in both everyday speech and in the relevant literature, the term *accent* is sometimes used interchangeably with *dialect*. Nevertheless, this thesis uses the term *dialect* in its technical, sociolinguistic meaning. *Dialect* is understood as referring to different varieties of a language that diverge in terms of pronunciation as well as grammatical, lexical and semantic differences

[5] See, for instance, Cutler (2012) and (Guo et al., 2023) for how differences across L1 accents affect speech processing the same way people process a second language differently.

(Wolfram, 2017). Thus, the term *dialect* can be understood as a hypernym including the term *accent*. Second, the term *code* (Swann et al., 2004) is used in this investigation as an umbrella term when the specific kind of variety referred to is not relevant. *Code* is used as a neutral label to refer to languages, dialects and accents in a general way. That is, when a discussion does not involve a specific language, dialect or accent, it functions as a hypernym, like in *different codes evoke different attitudes*.

2.2 What are stereotypes?

Social psychology includes the concept of stereotype under the umbrella term of attitudes. Therefore, this section first provides a general account of attitudes (Section 2.2.1) followed by a description of how stereotypes are conceptualized specifically (Section 2.2.2).

2.2.1 The tripartite structure of attitudes

Research on attitudes has been important in social psychology since the early 20th century. For example, Allport (1935, p. 798) stated that "the concept of attitude is probably the most distinctive and indispensable concept in contemporary American Psychology". The roots of the field can be tracked back to Lippmann (1922), Bogardus (1927), Thurstone (1928) and Katz and Braly (1933) (see Fiske et al., 2016; Stangor, 2016 for an overview). Back then, attitudes were defined as a "readiness" for feelings, beliefs or behaviors toward an object (Allport, 1935, pp. 800, 805, 815; 1954).

Already back then (see Lippmann, 1922), it was established that attitudes cannot be observed directly. They are "hypothetical mediating variable[s]" (Fiske & Taylor, 2017, p. 423). They are thus a cognitive construct. Nevertheless, they have real consequences for their targets. When attitudes (particularly negative ones) guide behavior, people at the receiving end see a detriment to their physical (Williams, 1999; Williams & Rucker, 2000) as well as mental well-being (Corning, 2002; Williams & Williams-Morris, 2000). Negative attitudes can affect different aspects of someone's life such as job applications and performance evaluations (Riach & Rich, 2004).

Despite being understood as cognitive constructs, attitudes are social and cultural artifacts (Morling & Masuda, 2012; Stangor, 2016). This is because the attitudes an individual has are the result of their living in a specific sociocultural context. For example, Human Rights Watch (Singh, 2002) reported that the amount of violence aimed at the Muslim population in the US increased tenfold after 9/11. In a similar fashion, anti-Asian hate crimes increased in 2020 as a result of the COVID-19 pandemic (Gover et al., 2020; US Department of Justice, 2023).

Attitudes are learned both directly and indirectly from multiple sources. These sources include parents and peers (Degner & Dalege, 2013; Kiesner et al., 2003; Poteat, 2007; Rodríguez-García & Wagner, 2009), media (Burgess et al., 2011; Newman, 2006; Reid, 1979; Weisbuch et al., 2009; Zuckerman et al., 1980) and language (Hill, 2008; Maass et al., 2014; Ng, 2007). In other words, the attitudes a person has toward an individual or a group are not innate but learned. Notwithstanding, research has found correlations between a number of personality traits (such as social dominance orientation, authoritarian personality traits, need for closure and structure and egalitarianism) and the strength of different attitudes (Adorno et al., 1950; Bäckström & Björklund, 2007; Jost et al., 2003; Moskowitz et al., 1999; Plant & Devine, 1998; Roets & Van Hiel, 2011; Rowatt et al., 2006; Sidanius & Pratto, 2001). Moreover, some studies have shown that attitudes may have a genetic component (Lewis & Bates, 2010; Lewis et al., 2014; Martin et al., 1986; Orey & Park, 2012).

All this may be taken to imply that attitudes are something people are born with, but there are two main flaws in this reasoning. First, establishing causal relationships between genes and behaviors is not a straightforward issue (Lewis & Bates, 2010; Lewis et al., 2014). For example, a person's environment can activate or inhibit the expression of a given gene (Kean, 2012). This can result in a person genetically predisposed to a certain trait never exhibiting it. Second, the effects of genetic or personality traits on attitudes are weaker than those of environmental factors and personal experiences (Kite et al., 2023). Therefore, this study sees them as the result of (mainly) socio-cultural forces.

From the very beginning, attitudes have been understood as performing social and cognitive functions (Lippmann, 1922). Our (social) world has vast amounts of information. Humans therefore categorize objects and people to efficiently process the constant stream of input they receive (Macrae et al., 1994; Tajfel & Wilkes, 1963). Once these categories are created and people associated with them, humans attach beliefs and emotions to them. These associations help people guide future interactions in an effective fashion. This was summarized by Macrae and Bodenhausen (2000, p. 94) in the following way: "knowing what to expect — and exactly where, when and from whom to expect it — is information that renders the world a meaningful, orderly, and predictable place." When building these social categories, the most important distinction is that of ingroup (people who are like me) and outgroup (people who are not like me) (Sumner, 1906, p. 12; Tajfel, 1969, 1978). Nevertheless, many kinds of social categories (such as race, gender and age) are important in social cognition (Macrae & Bodenhausen, 2000). The second nature of this social categorization process can be seen by how fast it takes place. ERP studies demonstrate that, within the first 100-200 milliseconds of observing an individual, people categorize them based on gender and race (Ito & Urland, 2003; Kubota & Ito, 2007).

These categories, and the attitudes they evoke, also help us fill in the blanks when information is ambiguous or missing. For instance, Plant et al. (2004) showed

participants facial expressions that could be interpreted either as angry or sad. When the expression was portrayed by a man, participants labeled the emotion as anger. If the expression was portrayed by a woman, it was categorized as sadness. These results parallel those of gender stereotypes about emotions.

As the discussion above shows, attitudes are conceptualized as including three facets: behavior, affect and cognition (Zanna & Rempel, 1998). These three facets are closely related. However, they are independent theoretical constructs (Dovidio, 2001; Dovidio et al., 1996; Fiske & Taylor, 2017). They are therefore treated as such in this investigation.

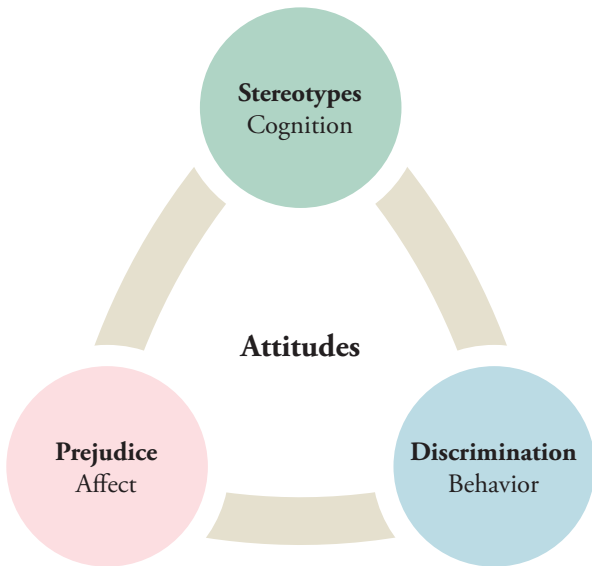


Figure 2.2. Tripartite structure of attitudes

This tripartite structure of attitudes (behavior, affect and cognition) corresponds to the terms *discrimination*, *prejudice* and *stereotypes*, respectively (see Figure 2.2 for the relationship between them). Discrimination (the behavioral component) is understood as treating people differently from others based on their social group (Sue, 2003). This can include very explicit behaviors (like hate crimes) to more implicit ones (like sitting next to people of specific races on public transport) (Jacobs, 1999). Prejudice (the affective component) is the emotion an individual experiences when interacting with a member of a social group because they are part of that group (Brewer & Brown, 1998). This can include the feelings of disgust people experience toward drug addicts, homeless people or obese people (Vartanian, 2010). Finally, stereotypes (the cognitive component) encompass "the beliefs about the characteristics, attributes, and behaviors of members of certain groups" (Hilton & von Hippel, 1996, p. 240). An example of a stereotype is the belief that women are more caring than men (Deaux & Lewis, 1984).

These last two components (prejudice and stereotypes) are elicited very fast, which arguably shows their important role in social cognition. Both are associated with a P300 component in ERP studies (Bartholow et al., 2001; Ito et al., 2004).

Two final terminological remarks close this section. First, a fundamental element of the understanding of the terms *discrimination*, *prejudice* and *stereotypes* is that they are not intrinsically negative. One can speak of (i) positive and negative discrimination (when someone is hired for a job due to being a minority or being subject to a hate crime, respectively), (ii) positive or negative prejudice (liking or disliking someone because of their nationality) and (iii) positive or negative stereotypes (thinking someone of a specific race is intelligent or assuming someone of a specific nationality is dangerous). Second, in everyday speech, the terms *attitudes*, *prejudice* and *stereotype* are frequently used interchangeably. Nevertheless, in this investigation they are used with their technical definitions outlined above.

2.2.2 Stereotypes

This section narrows down the discussion above and focuses on stereotypes. The history of the term *stereotype* can be traced back to Lippmann (1922). He took the term from the printing industry, as stereotypes (in the word's original meaning) created fixed images every time they were applied (Ashmore & Del Boca, 1981). Back then, he defined them as mental pictures of given groups. Broadly speaking, this definition matches the one provided in the previous section. Nonetheless, research during the past hundred years has provided insights into how stereotypes can be viewed from both structural and processing points of views[6]. This is the subject of this section. First, different views about how stereotypes are structured in cognition are discussed. Then, the difference between implicit and explicit stereotypes is introduced followed by a brief depiction of different theories on how they are processed. Finally, a general conceptualization of the content of stereotypes is described.

First, the definition in the previous section described stereotypes as cognitive structures. They have sometimes been conceptualized as schemas (Cox et al., 2012), prototypes (Brewer et al., 1981) or exemplars (Bodenhausen et al., 1995). Nevertheless, their specific mental representations have been subject to theoretical discrepancies. On the one hand, some researchers posit that stereotypes are a stable, symbolic system (Fazio, 1995; Petty et al., 2007). In his discussion promoting a closer collaboration between different theorical approaches, Smith (1996, p. 895) described the symbolic view with the metaphor of stereotypes as "static filing cabinet[s]." This view helps explain the stability of different stereotypes through history as well as the difficulty of changing stereotypes (Weber & Crocker, 1983).

[6] The following discussion can be applied to attitudes in general. However, due to the aims of this thesis, it focuses on how it applies to stereotypes specifically.

On the other hand, other scholars have defended connectionist models (McClelland & Rumelhart, 1986; Schwarz & Bohner, 2001). Representations in memory are discarded in favor of "time-dependent states of the [cognitive] system" (Conrey & Smith, 2007, p. 718). Stereotypes are thus not stored but created. This perspective accounts for the fact that stereotypes are fluid. For instance, Wittenbrink et al. (2001) reported that participants showed stronger biases against black people when they were shown in a ghetto than at church (see also Barsalou, 1982). However, as Nayakankuppam et al. (2018) highlighted, there must be some sort of representation stored in memory for the system to create a stereotype under different contextual forces. Hence, most current theories assume that stereotypes can be seen as stored representations, parts of which will be activated under different circumstances.

This view leads to a second vital aspect of stereotypes, the differentiation between implicit and explicit stereotypes. Social cognition has established that results from self-report methods and implicit tests do not always correlate (e.g. Cameron et al., 2012). One of the main reasons for this effect is social desirability bias, namely the fact that individuals do not want to appear to have attitudes toward people. As a consequence, implicit procedures such as the Implicit Association Test (Greenwald et al., 1998), the Sequential Priming Task (Fazio et al., 1986), the Extrinsic Affective Simon Task (De Houwer, 2003), the Implicit Relational Assessment Procedure (Barnes-Holmes et al., 2006) or the Go/No Go Association Task (Nosek & Banaji, 2001) are said to show the implicit stereotypes individuals have. Nevertheless, despite the frequent use of the adjectives *implicit* and *explicit* to describe stereotypes, their specific definitions have been subject to disagreements (see Payne & Gawronski, 2010 for a review).

A typical instance of an implicit stereotype is said to be automatic and take place outside conscious awareness, while a typical explicit stereotype is said to be one that individuals can control and are aware of (Devine, 1989). Automaticity and awareness are thus seen as key parameters. With regard to automaticity, Bargh (1994) posited that it can be decomposed into four different features: awareness, intention, controllability and efficiency. One can thus talk of an implicit stereotype if the individual is not aware of it (awareness), cannot control its activation (intention) or its deactivation (controllability) and if the stereotype requires few-to-no attentional resources (efficiency). In contrast, a typical explicit stereotype can be said to take place when these four conditions are not met.

With regard to awareness, Gawronski et al. (2006) explained that individuals may lack awareness of a stereotype in three different ways. They may be unaware of (i) the origin of the stereotype (source awareness), (ii) the content of the stereotype (content awareness) or (iii) the effect the stereotype has on other cognitive processes (impact awareness). Of these three, impact awareness has been the aspect most strongly correlated with implicit stereotypes (Gawronski et al., 2006).

Despite the use of these parameters for characterizing implicit and explicit stereotypes, a significant portion of the literature agrees that a strict divide between

them may not be necessary (Devine & Sharp, 2009; Gawronski & Bodenhausen, 2011; Kubota et al., 2013; Stangor, 2016). Therefore, rather than two different kinds of stereotypes, the difference between implicit and explicit stereotypes has been turned into a discussion of how stereotypes are elicited and processed.

In this regard, modern theories of stereotypes rely on general theories of social cognition (see Fiske & Taylor, 2017). Specifically, dual-process theories (Chaiken & Trope, 1999) have been key in shaping our current understanding of how different kinds of stereotypes are elicited under different circumstances. Broadly speaking, dual-process theories posit that the cognitive processes behind social judgements and their links to behavior can be categorized into two kinds. One kind is automatic, while the other is performed in a controlled fashion by the individual.

There is a plethora of theories and models under the umbrella of dual-process theories (see Gawronski & Creighton, 2013 for a review). Of these, the Associative-Propositional Evaluation model (APE) (Gawronski & Bodenhausen, 2006, 2011) has been of significant relevance in both social cognition and sociolinguistics (see Section 2.3). Gawronski and Bodenhausen (2011, p. 61) described associative processes as "the *activation* of mental associations in memory, which is assumed to be driven by spatiotemporal contiguity between stimuli and the similarity between the features of input stimuli and available memory representations" (emphasis in original). This thus describes how encountering an attitudinal object activates a given association. For example, upon seeing a person with facial tattoos, the association with CRIMINAL may be automatically activated (Funk & Todorov, 2013). Importantly, this model expects that the specific associations activated are subject to contextual parameters. This is because its definition explicitly includes the "features" of the stimuli, which encompass its context.

Next, propositional processes encapsule "the *validation* of the information that is implied by activated associations, which is assumed to be guided by the principles of logical consistency" (Gawronski & Bodenhausen, 2011, p. 62, emphasis in original). This process takes the associations above and translates them into statements, for example "people with face tattoos are dangerous." These kinds of statements are then contrasted with other propositional beliefs relevant to the situation, such as an individual's wish not to be prejudiced (Monteith et al., 2016). In the end, a propositional belief is evoked and can be verbally reported by the individual.

A key aspect of this model is that these processes are not independent of each other. As Gawronski and Bodenhausen (2011, p. 63) pointed out, associative processes can influence which specific propositional processes operate, since they rely on contextually activated associations. Similarly, propositional processes can influence associative processes by activating new associations in their course. The model helps us understand why not all attitudes measured through indirect methods need to be unconscious as well as why not all attitudes measured through direct methods need to be explicit. People can be aware of their automatic associations (Nier, 2005). Likewise, people may

not necessarily be conscious of propositional processes, especially when they simply validate an activated evaluative association.

Finally, the content of stereotypes has also received a significant amount of attention in the literature. Researchers from social psychology have shown that different social categories evoke different stereotypes. For instance, men are judged to be better at math than women (Sue, 2010) and Latina women are thought to be feisty and good cooks (Ghavami & Peplau, 2013). In contrast, there has also been an interest in trying to categorize the fundamental components of the content of stereotypes. In this regard, the Stereotype Content Model (Cuddy et al., 2007; Durante et al., 2013; Fiske et al., 2007; Fiske et al., 2002) has had significant influence. It posits that the specific traits associated with different social groups can be categorized on the basis of two dimensions: competence and warmth, which date back to Osgood et al. (1957). The competence component includes perceived social status, success, intelligence, etc. The warmth component involves perceived sociability, trustworthiness, friendliness and the like. These two dimensions create four different categories: [+competence,+warmth], [+competence,-warmth], [-competence,-warmth], [-competence,+warmth]. These are argued to influence how people perceive and behave toward different individuals. In the US, white, Christian, heterosexual people are seen as having [+competence,+warmth]; homeless people are stereotyped as having [-competence,-warmth]; older and disabled people are judged as having [-competence,+warmth] and Asian people as having [+competence,-warmth] (Cuddy et al., 2007). Importantly, the two dimensions of competence and warmth are not categorical. They are better conceptualized as being two continuous dimensions along which social groups can be placed (see Cuddy et al., 2007, p. 638).

2.3 What are *language* attitudes?

This section describes how the concepts of accent and stereotype described in the two previous sections combine and illustrates how accent stereotypes are viewed in sociolinguistic research. As can be inferred from Section 2.2.1, accent stereotypes fall under the umbrella of language attitudes. Accent stereotypes can thus be defined as "beliefs about the characteristics, attributes, and behaviors" of speakers due to the accent they use (Hilton & von Hippel, 1996, p. 240). For example, upon hearing someone speaking with a Standard Southern British English, one may judge that person to be intelligent and of high SES. In contrast, an individual may judge a speaker with a broad Irish accent to be less competent but more friendly than the standard British speaker. In this section, a brief overview of what language attitudes are is first provided, with a special focus on accent stereotypes. Then, the focus is placed on the special role accents have enjoyed in the field. Finally, how sociolinguistics and social psychology differ in their treatment of attitudes is highlighted.

Social cognition has typically focused on issues regarding race, age and gender (Gluszek & Dovidio, 2010b; Macrae & Bodenhausen, 2000). Nevertheless, sociolinguists have long been aware that different codes prompt attitudinal responses among individuals. In fact, some researchers have argued that linguistic differences may have played a more important role in human evolution than racial differences (Kinzler et al., 2010; Kinzler & Dautel, 2012; Pietraszewski & Schwartz, 2014a, 2014b; Rakić et al., 2011a). The first sociolinguistic study investigating how a speaker's accent influences how listeners perceive them is Pear (1931). He asked BBC radio listeners to judge different speakers reading the same short text. Listeners were asked to assess speakers' place of birth, gender, age, occupation and whether they were good leaders. This was one of the first studies positing that listeners judge speakers based on their accent. Nevertheless, it was not until the 1960s that the field took off. It was specifically Lambert et al. (1960), Anisfeld and Lambert (1964) and Labov (1966b) who laid the ground for the importance of linguistic variation in social interaction. This paved the way for a proliferation of research from the 1960s until the 1980s. The field was dormant during the 1990s and early 2000s (Dragojevic et al., 2013, p. 14), but it has seen a renaissance since then (e.g. Dovidio & Gluszek, 2012; Sharma et al., 2022).

Language attitudes exist because linguistic variation exists. Variation is in fact at the core of language. There are approximately 7,000 living languages (Eberhard et al., 2024), each of which includes different dialects and accents. Even more, the way speakers use language changes by situational context (see Section 2.1). All these sources of variation have sometimes been described as "different ways of saying the same thing" (Coupland, 1985, 2007; Labov, 1972; Lavandera, 1978; Macaulay, 2007). For instance, the sentences *jeg vil spise kagen* and *I want to eat the cake* are thought to be different ways of conveying the same meaning. The same applies to the differences between *I don't want nothing* and *I don't want anything* or the different pronunciations of *water* provided in Section 2.1 These pairs of examples have the same propositional meaning. But they do not "say the same thing" (Coupland, 2007). This is because meaning encompasses more than just propositional meaning (see Section 2.4). People use all sources of linguistic variation to infer qualities about the speaker or decide how to behave toward them. In other words, language variation provides the grounds for language attitudes.

There are many examples of language attitudes in the literature. Harari and McDavid (1973) showed that students' names affected the grades their essays received. They generated some essays and randomly distributed names to them. These texts were handed to experienced teachers who were asked to rate them. Teachers gave higher grades to the essays with common names than to the essays that were given an uncommon name. Grammatical constructions may also evoke attitudes among language users. Double negatives (*I don't want nothing* versus *I don't want anything*) have received a great deal of attention. The use of the construction with *nothing* has been shown to elicit strong negative prejudice among listeners (Chesire, 1998; Milroy

& Milroy, 1999). The use of different languages as a whole can also elicit reactions. Jørgensen and Quist (2001) showed how speakers of immigrant languages in Denmark are discriminated against.

People's names, what grammatical structures they use or what language they speak are not the only features on which language users are judged. The following (non-exhaustive list of) sources of linguistic variation are used to judge language users, whether intentionally or not: spelling (Hinrichs & White-Sustaíta, 2011), code-switching (Edwards, 2012), lexical diversity (the ratio of types-tokens, Bradac, 1990; Bradac et al., 1988), lexical etymology (e.g. the use of Germanic or Latinate word, Levin et al., 1994), speech rate (Brown, 1980; Brown et al., 1985), context of the interaction (Cargile et al., 1994; Carranza & Ryan, 1975), speech pathologies (Bettens et al., 2020; Lass et al., 1988; Lass et al., 1991), dialects (Garrett & Coupland, 1999) and accent (see below in this section). Furthermore, language attitudes in general and accent stereotypes in particular are widespread phenomena. Sociolinguistic studies have shown their prevalence in numerous countries, for instance and in no particular order: Guatemala (Tsalikis et al., 1992), The Netherlands (Nortier & Dorleijn, 2008), Belgium (Speelman et al., 2013), the UK (Giles, 1970), the US (Brennan & Brennan, 1981b; Heaton & Nygaard, 2011), Canada (Munro, 2003), Brazil (Gluszek & Hansen, 2013), Australia (Nesdale & Rooney, 1990, 1996), Sweden (Bijvoet, 2020; Boyd, 2003), Denmark (T. Kristiansen, 2009; Ladegaard, 1992, 2000; Maegaard et al., 2013), Germany (Klink & Wagner, 1999; Rakić et al., 2011b), Spain (Cabrera, 2011; Loureiro-Rodriguez et al., 2013), Cyprus (Papapavlou & Sophocleous, 2009), Finland (Kristiansen, 2013), Israel (Remennick, 1999) and more (see Garrett, 2010; Giles & Watson, 2013 for overviews).

Language attitudes can be understood as being the result of the following three interrelated ideologies. First, normative monolingualism (Fuller, 2018) defends the notion that a language user should only speak one language and that one nation should speak one language — a view that goes back to the Romantic period with Johann Gottfried Herder (Bauman & Briggs, 2003). Second, the ideology of nativeness (Shuck, 2006) illustrates the view that a language must be spoken the way its native speakers speak it, creating a strong divide between "us" (L1 speakers) and "them" (L2 speakers) (Giles, 2012; Schmidt, 2002). Finally, the standard language ideology encompasses the belief that, among native varieties of a language, there is one variety that is correct, aesthetic and pure. This variety is typically the standard variety of a language. In contrast, all other varieties are considered inferior, impure or even corruptions of it (Milroy & Milroy, 1999).

A fundamental aspect of language attitudes is that they are subject to both cultural norms and contextual factors, like attitudes in general (see Section 2.2.1). In fact, sociolinguists have sometimes been praised for acknowledging and incorporating this lack of universality into their theories before social psychologists (Cargile et al., 1994; Ryan & Giles, 1982; Soukup, 2012). This dependence on context can be understood

in two different ways. First, the same linguistic variable can be judge differently by different listeners. A classic example involves the pronunciation (or lack thereof) of postvocalic r in words such as girl, car or birch. In the US, the r-less pronunciation is frowned upon and associated with lower class speakers from the North-East of the US (Labov, 1972; Nagy & Irwin, 2010). In contrast, in England, it is the use of r in this context that is associated with lower-class speakers (Costa & Serra, 2022). Second, the same linguistic input can be judged differently depending on the context in which it is processed (Gluszek & Dovidio, 2010b; Maegaard & Pharao, 2021). Thus, the same language can be rated differently based on whether it is spoken at home or in a classroom (Carranza & Ryan, 1975).

Language attitudes are sometimes rationalized by language users. However, these rationalizations lack any linguistic grounds. For example, Algeo (1998) described how British speakers held very negative views of the word *controversy* when pronounced with the primary stress on the first syllable (/ˈkɒntrəvɜːsi/) rather than on the "correct" syllable, the second one (/kənˈtrɒvɜːsi/). The "incorrect" use was said to be avoided because it was a "vulgar Americanism." Nevertheless, such pronunciation is not present in the US and is a British innovation (Algeo, 1998, p. 177). A word's etymology has sometimes been used for similar purposes. Trudgill (1998) illustrated how speakers of British English rejected use of the word *alternatives* in phrases like *plenty of alternatives*. This is due to the fact that the etymology of *alternative* is from the Latin word *alter*, meaning two. Therefore, the word should not be used when there are more than two choices. Nevertheless, Trudgill (1998, p. 2) noted that the etymology of the word *nice* is "not cutting" and states that "no-one in their right mind though would argue that the 'real' meaning of 'nice' is 'not cutting'." Language attitudes are sometimes seen as simple "truths" (Johnson, 2012, p. 75). For instance, the belief that only English should be spoken in the US is often held without any strong reasoning other than "This is the US" (Johnson, 2012).

The negative views listeners have toward accents specifically have sometimes been defended on intelligibility grounds. That is, non-standard accents are judged negatively because they are said to be difficult to understand. However, this argument has not received empirical support. A classic example is Rubin (1992) (but see also Fiedler et al., 2019; Flege & Fletcher, 1992; Hanulíková, 2018, 2020; Kang & Rubin, 2009; Kutlu, 2020; Levi et al., 2007; Lippi-Green, 2012; Munro & Derwing, 1995; Rubin et al., 2016; Rubin, 2012; Rubin & Smith, 1990; Yi et al., 2013; Zheng & Samuel, 2017). Their study was based on previous literature demonstrating that undergraduate students in the US give lower ratings to teaching assistants with L2 accents. To investigate the different factors behind this, they recorded a lecture by a native speaker of American English. This recording was the only linguistic stimulus in the study. Participants were also shown two different faces together with the survey. They were either told that (i) the speaker is from China and shown the picture of a Chinese woman or that (ii) the speaker is from the US and shown a picture of a Caucasian woman.

Participants who were shown the picture of a Chinese woman reported that the recording had a strong accent and was difficult to understand. The linguistic stimuli, however, did not differ in the two conditions. Only the picture and the nationality of the speaker varied. This shows that how people judge accent strength or comprehensibility is not only a linguistic matter.

Regardless of whether or not they are rationalized, language attitudes have real consequences for language users. In short, users of non-standard varieties or less prestigious languages are negatively discriminated in different settings: in courts (Reinard, 2013; Seggie, 1983), assessment of primary school teachers as well as of students in primary school (Boyd, 2003; Seligman et al., 1973 respectively), ratings of university teachers (Rubin, 1992; Rubin, 2012; Rubin & Smith, 1990), ratings of university students (McKenzie & Gilmore, 2017), in diagnostic interviews in health contexts (Fielding & Evered, 1980; Rubin et al., 1997), employment selection (Hosoda & Stone-Romero, 2010; Lippi-Green, 2012; Rakić et al., 2011b) and housing (Zhao et al., 2006), to name a few. This linguistic discrimination, in turn, has a negative impact on the physical, psychological and economic lives of speakers (Formanowicz & Suitner, 2019; Gluszek & Dovidio, 2010a; Lamberton, 2002; Wated & Sanchez, 2006). One aspect that differentiates discrimination based on language from other kinds of discrimination (e.g. based on race or gender) is that it is, in some contexts, an accepted form of bias encouraged by different sociocultural and legal systems (Esling, 1998; Lippi-Green, 2012; Matsuda, 1991; Nguyen, 1993).

Within the field of language attitudes, accents have enjoyed special status. So much so that attitudes toward accents have been argued to play a stronger role in human evolution than race, as mentioned at the beginning of this section (e.g. Kinzler et al., 2010). An example of the high status of accents is Levin et al. (1994). They examined how accent (RP or South-East Wales) and lexical provenance (Germanic or Latinate origin) affected listeners' opinions of speakers. They found that accent accounted for most of the results. In fact, even in the conditions in which there were no differences across the lexical origin of words in the stimuli (that is, the readers read the exact same text, just with different accents), participants commented that the RP speaker used longer and more eloquent words. Accent can thus cause us to perceive things not present in the stimuli, like attitudes were described to do in Section 2.2.1[7] (see Buson & Billiez, 2013; Stewart et al., 1985 for similar results). The special role of accents can also be seen in results showing that listeners only need short stimuli in order to infer properties from the speaker (Baugh, 2000; Flege, 1984; Scharinger et al., 2011). For example, Boberg (1999) reported on how American listeners judge speakers differently based solely on their realization of the vowel in the word *last*: [læst] or [la:st]. Specifically, they rated the second option, typically associated with British English, as

[7] This relationship is not unidirectional; social information can also affect how accents are perceived (e.g. Hanulíková, 2018; Hay & Drager, 2010; Hay et al., 2006; Niedzielski, 1999)

more sophisticated and educated. This shows how even just a difference of two phones can elicit different attitudinal responses.

When it comes to how they are viewed socially, accents can be categorized in a two-step fashion. First, accents can be labeled as standard or non-standard. Then, non-standard accents can be further split into L1 or L2 accents[8]. Overall, speakers of standard varieties are subject to positive prejudice, stereotyping and discrimination. So much so that L1 and L2 speakers have been reported to take "accent-erasing" lessons to get rid of their specific accents (Hernandez, 1993; Morrish, 1999; Mugglestone, 2007). This is known to have occurred as early as in 18th century England, when a Scottish lawyer only achieved success in his field after attending lessons to "lose [his] Scots accents" (Mugglestone, 2007, p. 39 addition in original). Importantly, listeners need not properly categorize the accent in order to have attitudinal responses to it. The mere labeling of an accent as non-standard (either L1 or L2) will elicit negative views of individuals (Lindemann, 2003; Milroy & McClenaghan, 1977; van Bezooijen, 1994).

With regard to L1 and L2 accents, L2 accents are generally subject to more negative attitudes than L1 accents (Bayard & Green, 2005; Birney et al., 2019; Edwards, 1982; Gluszek & Dovidio, 2010b; Kavas & Kavas, 2008; Marvasti, 2005). For example, speakers with an L2 accent are viewed as less intelligent, loyal and competent (Bradac, 1990). This picture is subject to nuances, however. As described in Section 2.2.2, stereotypes can be viewed as involving different factors. When it comes to language attitudes, different factors have been proposed (Fuertes et al., 2012; Mulac, 1975; Mulac et al., 1974; Zahn & Hopper, 1985). Most of them propose a tripartite factor structure (with minor differences across studies) that includes prestige, warmth and dynamism. First, prestige (often called superiority, competence or status) includes characteristics such as intelligence, socioeconomic status, kind of job, confidence, or ambition. The second factor, warmth (also referred to as social attractiveness or solidarity), includes traits such as friendliness, trustworthiness, kindness and beauty. Finally, dynamism has not been included in studies as often as the other two (Zahn & Hopper, 1985, p. 119). This factor refers to traits such as talkativeness, aggressiveness, energeticness and laziness. Both L2 and non-standard L1 accents are sometimes rated higher on non-competence factors, illustrating what some sociolinguistics have called covert prestige (e.g. Edwards, 1982; Giles & Marlow, 2011). This hierarchy is, nevertheless, not completely rigid. For instance, French-accented English is frequently rated above other L1 accents in the UK across all factors (Bishop et al., 2005; Coupland & Bishop, 2007; Giles, 1970; Sharma et al., 2022).

Sociolinguists have frequently categorized language attitudes as either direct/explicit or indirect/implicit (e.g. T. Kristiansen, 2003, 2009; Kristiansen & Grondelaers, 2013;

[8] Due to the use of different terms in the literature, the labels introduced in Section 2.1. are also used here.

Pharao & Kristiansen, 2019). The former refers to attitudes obtained by overtly asking participants about their views of a given code. For example, Coupland et al. (1994) (see also Garrett et al., 2003; Williams et al., 1999) performed a large-scale study of language attitudes in Wales. They directly asked participants to rate different Welsh accents, together with RP, on how prestigious, pleasant, dynamic and truly Welsh they were. In contrast, what sociolinguists call indirect methods aim at concealing the fact that the topic of the study is the participants' language attitudes. The most popular method is the Matched-Guise Test (Garrett, 2010; Garrett et al., 2003; Giles & Watson, 2013; Lambert et al., 1960; Loureiro-Rodriguez & Acar, 2022), a variant of which is used in this study (see Section 4.1.2). In a matched-guise task, a person reads the same text either in different languages or different accents. This thus creates two "guises" of the same speaker, giving the method its name. Afterwards, participants are asked to rate each guise on a number of variables, typically using differential semantic scales or Likert Scales. With this design, participants are not aware they are rating the same person with two different guises. Therefore, the differences in ratings are associated with the different attitudes participants have toward the different codes. This method is argued by sociolinguists to elicit the "true" attitudes people have (Soukup, 2012).

The sociolinguistic distinction between direct and indirect methods appears to be parallel to the one introduced in Section 2.2.2 from social cognition. In both cases, direct attitudes are those that reflect the conscious opinions of individuals, such as *I don't like people with face tattoos* or *speakers with an Irish accent are funny*. Nevertheless, their conceptualization of implicit/indirect attitudes differs significantly. Specifically, a number of sociolinguists have taken a more cognitive approach to language attitudes (Álvarez-Mosquera & Marín-Gutiérrez, 2018; Campbell-Kibler, 2012; McKenzie & Carrie, 2018; Pantos, 2014; Pantos & Perkins, 2013; Redinger, 2010). These studies utilize methods from social cognition such as the Implicit Association Task mentioned above to investigate the automatic, unconscious and fast responses people have toward different codes. From this point of view, the matched-guise technique is not viewed as a tool to evoke implicit attitudes. Instead, it is thought of as an indirect method to elicit explicit attitudes (in the sense used in social psychology) (see also Adams, 2018, pp. 60-61, 151-152; Pantos, 2019). This does not mean that the Matched-Guise Test cannot capture these automatic responses. In fact, results from both traditions correlate with each other (compare Pantos, 2014; Reinard, 2013; Seggie, 1983). Labelling the matched-guise as a method that elicits explicit attitudes simply means that it does not necessarily capture these automatic associations in the same way the Implicit Association Task can be argued to do.

2.4 Accent stereotypes as meaning

This section deals with how accent stereotypes can be regarded as meaning. First, the traditional distinction between linguistic and encyclopedic knowledge is outlined. Afterwards, theoretical and empirical grounds for specifically viewing stereotypes as meaning are summarized. Finally, Cognitive Linguistics is introduced and argued to provide fruitful theoretical tools for sociolinguistic research.

Section 2.2.2 above describes how stereotypes have been seen as cognitive constructs. This investigation takes this view a step further. Specifically, accent stereotypes are seen as meaning. This interpretation is not completely new. Sociolinguistic research has already argued for the inclusion of sociolinguistic perception within the realm of communicative competence in the sense of Hymes (1972) (e.g. Chambers, 2003; Day, 1982; de Vogelaer & Toye, 2017; Rosenthal, 1974). However, this view is not equally widespread within linguistics as a whole or related disciplines, which is the result of the history of how language and meaning have been understood. Traditionally, knowing a language was understood as being able to produce and understand sentences based on its lexicon as well as its phonological and grammatical rules. This view can be observed in Saussure's (2015) distinction between *langue* and *parole* as well as Chomsky's (1965) differentiation between competence and performance. For example, Chomsky (1965, p. 3) claimed that:

> linguistic theory is concerned primarily with an ideal speaker-listener, in a completely homogenous speech community, who knows its language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of the language in actual performance.

When it comes to meaning, this theoretical standpoint led to the demarcation between semantic/linguistic meaning and encyclopedic meaning. The former refers to features or brief descriptive statements of solely linguistic character (Langacker, 2008, p. 38). Hence, the word *girl* can be theorized to have the meaning of [+human, -adult, +female] (Murphy, 2003, p. 12). This purely linguistic meaning is posited to be different from the general knowledge that, stereotypically, girls wear dresses or like the color pink. This perspective has sometimes been labeled the dictionary view of meaning (e.g. Haiman, 1980), since it ascribes the meaning one would find in a dictionary to words. In contrast, the general knowledge behind linguistic units has been labeled encyclopedic meaning. Thus, knowing that girls might wear dresses or knowing the role of bulls in bullfights is not part of the meaning of the words *girl* and *bull*, respectively (Langacker, 2008, p. 38). These aspects of meaning are thought of as being independent aspects of speakers' linguistic structures.

Nevertheless, the division between semantic and encyclopedic meaning has been subject to strong criticisms since the mid-20th century (Haiman, 1980; Keesing, 1979; Lakoff, 1971; Russell, 1969). On the one hand, it is not possible to draw precise boundaries between these kinds of meaning in a principled fashion (Bolinger, 1965; Croft & Cruse, 2004; Haiman, 1980; Quine, 1951). For example, a phrase such as *boys will be boys* or the lyrics *he was a boy, she was a girl, can I make it any more obvious?* (Lavigne et al., 2002) would be senseless if the words *boy* and *girl* were to be understood with only their so-called semantic meaning: [+human, -adult, +male] and [+human, -adult, +female], respectively.

On the other hand, there is empirical evidence supporting the notion that there is not a clear-cut difference between these so-called different kinds of meaning. Neurolinguistic research provides relevant results in this regard, specifically research on the N400 component[9]. N400 refers to a component of negative activity in an event-related potential framework (ERP). This component takes place approximately 400 milliseconds after the presentation of stimuli and is observed in the centro-parietal regions of the scalp. Briefly put, the N400 shows semantic violations (Kutas & Federmeier, 2000; Kutas & Hillyard, 1980a, 1980b). For example, the word *dog* in the sentence *I take coffee with cream and dog* elicits an N400 when compared to the same sentence ending with *sugar*. It also appears with words much more semantically related. Thus, the word *hour* in *the bill was due at the end of the hour* elicits an N400 in contrast to the word *month*, as in *the bill was due at the end of the month* (Kutas & Federmeier, 2011, p. 624).

The N400 is also present when the violations include so-called encyclopedic knowledge (Baetens et al., 2011; Proverbio et al., 2018; Proverbio et al., 2017; Rodríguez-Gómez et al., 2020; White et al., 2009). For instance, White et al. (2009) presented participants with word pairs following the set-up in Bentin et al. (1985). The first item in each pair included either the word *men* or *women*. The second item constituted their target words. The target words included traits (nurturing, aggressive) or concepts (makeup, mechanic) stereotypically associated with the gender categories in the word pairs. These target words were chosen based on previous gender stereotype research (Bem, 1974; Oakhill et al., 2005). The combination of both kinds of items gave rise to both stereotypically congruent (women-nurturing, men-aggressive) and incongruent (men-nurturing, women-aggressive) pairs of words. Their ERP results showed a statistically significant N400 effect for incongruent pairs compared to the corresponding congruent ones.

The theoretical and empirical results described in the previous paragraphs have led scholars to abandon the semantic/encyclopedic divide. Meaning is thus argued to be

[9] This is not the only methodology that supports this view. Similar results can be found, for instance, using eye-tracking methodologies. See Borovsky and Creel (2014); Creel (2018); Creel and Jimenez (2012); Creel and Seubert (2015).

conceptual structures (Croft & Cruse, 2004; Langacker, 1987; Paradis, 2003, 2015; Talmy, 2000, 2015) rather than a cognitively independent component, as mentalist approaches to language propose (Chomsky, 1986). This does not mean that all knowledge associated with a given linguistic unit is equally important, as highlighted in the literature (Langacker, 1987). The main point is that a clear-cut divide between these two poles is impossible to create. It is the actual use of a linguistic unit in a specific situation, with a specific purpose, by a specific person that helps create meaning and, thus, highlights different aspects of the meaning of a particular linguistic unit.

Returning to the case of accent stereotypes, there are different arguments that advocate for their categorization as meaning. From a purely theoretical point of view, accent stereotypes refer to conceptual structures activated due to the way a person speaks. Therefore, they satisfy the definition of meaning described above. From a more empirical perspective, two kinds of results can be described. On the one hand, accent stereotypes have been shown to be a productive tool speakers exploit in actual language use (e.g. Coupland, 2007). Hence, they are posited to form part of our communicative competence, as mentioned above (Hymes, 1972). For instance, Bell (2001) showed how the same newsreader from New Zealand would realize the phoneme /t/ differently depending on what station they worked at. At a national station, the /t/ phoneme was realized as [t^h], while at a local station it was realized as [ɾ]. Bell (2001) argued that the newsreader made use of the attitudes associated with the different realizations of the phoneme /t/ to portray a different image. A different productive use of accent stereotypes has been seen in the film and TV industry (Ahrens, 2004; Dobrow & Gidney, 1998; Dragojevic et al., 2016; Hodson, 2014; Lippi-Green, 2012). Dobrow and Gidney (1998) demonstrated how a character's dialect or accent was used to index their role as villain or hero. Specifically, protagonists were typically given standard American accents while villains tended to speak with non-standard accents, specifically L2 accents.

On the other hand, a different line of evidence for the view of stereotypes as meaning comes from neurolinguistics. The N400 component has also been found to result from a mismatch between accents and their associated stereotypes (e.g. Hansen et al., 2017; Pélissier & Ferragne, 2022; van Berkum et al., 2008). For instance, Pélissier and Ferragne (2022) studied accent stereotypes associated with the Parisian upper-class accent and the Banlieue accent (the accent associated with suburban, underprivileged areas around large cities in France). Half the sentences in their stimuli were consistent with the upper-class accent — *I love playing golf with my associates* — and the other half were consistent with the lower-class accent — *I always listen to rap music in my car*. Each sentence was recorded by speakers of both accents. Their results showed a statistically significant N400 effect in sentences with a mismatch between accent and sentence content (e.g. a person with a Banlieue accent saying *at Christmas I eat caviar from Russia*) versus sentences with a match (e.g. a person with an upper-class accent saying the same sentence). This implies that accents activate conceptual structures

about the speaker in an extremely short time, in the same way that so-called linguistic meaning does. All these results together make it possible to argue that accent stereotypes are not extralinguistic or part of a language-independent system.

A theoretical approach that takes this broad view of meaning as its core is Cognitive Linguistics. Its origins can be traced back to the 1970s and 1980s (e.g. Fillmore, 1975; Lakoff, 1977, 1987; Langacker, 1979, 1987; Rudzka-Ostyn, 1988; Talmy, 1978), when a number of linguists objected to the then standard approach of formal semantics and generative grammar (Nerlich & Clarke, 2007). Cognitive Linguistics views language as intrinsically interwoven with the general conceptual knowledge of the speaker, their social, cultural and historical background as well as their general cognitive processes (Dabrowska & Divjak, 2015; Geeraerts & Cuyckens, 2007). It hence takes a maximalist approach to language (Paradis, 2003), viewing it as embodied in a physical, psychological and cultural sense (Barsalou, 2008).

Cognitive Linguistics offers a useful theoretical framework for sociolinguistic research in general and accent stereotypes in particular[10]. This becomes apparent with the emergence of the field of Cognitive Sociolinguistics in the 2000s and early 2010s (see Geeraerts, 2018; Geeraerts et al., 2010a; Kristiansen & Dirven, 2008; Kristiansen et al., 2022; Moreno Fernández, 2016; Pütz et al., 2014). Cognitive Linguistics brings two main assets to sociolinguistics: their understanding of meaning and their work on categorization, both of which are now discussed.

With regard to the former, sociolinguistics can benefit from the long tradition within Cognitive Linguistics of performing detailed analyses of meaning, how it is structured, evoked and created in actual interaction (Berthele, 2010; Geeraerts et al., 2010b; G. Kristiansen, 2003, 2009; Paradis, 2008; Põldvere et al., 2022; Põldvere et al., 2021). This is particularly the case because sociolinguists have traditionally not focused on meaning. For instance, the concept of the linguistic variable has played a key role within sociolinguistics since the very beginning. The term was introduced by Labov (1966a, 1966b) and is meant to describe a linguistic unit that can be realized in different ways. For example, the word *water* can serve as a linguistic variable given the fact that it can be pronounced [ˈwɔːʔə] or [ˈwɔːɾɻ], among many other ways. The linguistic variable is typically defined as "different ways of saying the same thing" (Coupland, 2007, p. 88). However, this claim has been subject to criticisms (see, besides the discussion in Section 2.3, e.g. Coupland, 2007; G. Kristiansen, 2009; Lavandera, 1978). Within the so-called third wave of sociolinguistics (Eckert, 2012), the different forms that a linguistic variable can take have been shown to be used by speakers actively to portray specific identities, as in Bell (2001) above. Even though using different accents does not alter the referential meaning of an utterance, there are in fact different meanings associated with them, as argued in Section 2.3.

[10] Sociolinguistic research has also been argued to benefit research within Cognitive Linguistics; see Geeraerts et al. (2010b).

When it comes to categorization, sociolinguistics necessarily deals with social groups. Originally, these groups were based on broad demographic categories such as sex, geographical region, SES, etc. (Eckert, 2012). Then, more ethnographically relevant categories, such as the now classic example of jocks and burnouts (Eckert, 1989), were introduced in variationist work. More recent work has focused on how speakers portray different identities by linguistic means (such as Bell, 2001 above; see also Podesva, 2006) based on how they categorize situations and interlocutors. All these lines of research have in common the importance of social categories and how humans categorize situations. Cognitive Linguistics, heavily influenced by cognitive psychology, has a long tradition of studying how humans categorize the world and how categories are organized from a cognitive standpoint. This theoretical knowledge can be relevant for sociolinguistics. Humans categorize their social world to the same extent as they categorize the natural world (G. Kristiansen, 2009; Macrae et al., 1994; Taylor, 2003). Hence, research from Cognitive Linguistics about how humans perform this process can only help provide a more cognitively realistic view of how speakers deal with the social world around them and how they employ this knowledge in language use.

Narrowing it down to accent stereotypes specifically, Cognitive Linguistics can help us understand how accent stereotypes are interwoven with our general conceptual knowledge, how they are processed and evoked by linguistic structures and how they are stored in long-term memory. As Gries (2013, p. 7) put it:

> […] something can by definition only be sociolinguistically relevant if, at some point of time, it has passed through the filter of the human mind, where it was either readied for production or comprehended and interpreted. The reverse may also be true to some extent — what is cognitively/psycholinguistically relevant often arises once it is produced within some socially-defined situation.

Hence, this thesis emphasizes how, to understand accent stereotypes, we not only need to understand their social bases and functions, but also what they are and how they work form the point of view of the cognitive system of the language user. This hence brings an integration of research from sociolinguistics, Cognitive Linguistics and social cognition.

2.5 Summary

This chapter provides the theoretical foundations of this thesis and bridges the gap between the fields of sociolinguistics, social cognition and Cognitive Linguistics. First, it establishes how the term *accent* is understood in this investigation as the regularized differences in pronunciation across varieties of a language. Furthermore, it highlights how categorizing accents is not a straightforward matter. On the one hand, there is not

a clear-cut linguistic divide across different accents of one language; on the other hand, the concept of accent strength entails that two people speaking the "same" accent may speak in significantly different ways. Second, the term *stereotype* is framed under the umbrella term *attitude* and defined as the properties and behaviors associated with certain social groups. The traditional divide between symbolic and connectionist models of stereotypes as well as that between explicit and implicit stereotypes is described. This is followed by emphasizing the importance of how stereotypes are evoked and processed in specific contexts rather than by drawing upon categorical distinctions of types of stereotypes. The next section introduced the field of language attitudes. It highlights how the field has been neglected in discussions within social cognition and describes the pervasiveness of language attitudes and how they impact the lives of speakers. Finally, this chapter closes by arguing for the advantages of anchoring the study of accent stereotypes in Cognitive Linguistics due to its maximalist view of language.

3 What we (do not) know about the development of language attitudes

This chapter provides an overview of the literature on the development of linguistic attitudes. As mentioned in Chapter 1, individual studies within the field can be found as early as Anisfeld and Lambert (1964). Nevertheless, the existing literature is rather fragmented. Researchers typically approach the field from (social) psychology or sociolinguistics without much of a dialogue between them. Two issues can be argued to arise from this situation.

First, there is a lack of standard terminology in the field as a whole. Similar to the situation described in Chapter 2, authors sometimes use the same terms with what can be interpreted as different meanings. An example of this can be observed by comparing G. Kristiansen (2010) and Kinzler and DeJesus (2013b). The study reported in G. Kristiansen (2010) investigated the formation of accent stereotypes in Spanish children. However, they define *accent stereotype* as the association between a social group and the way in which they speak (G. Kristiansen, 2010, p. 226). This is in line with how sociolinguists understand the term *stereotype* (Labov, 1972). In contrast, Kinzler and DeJesus (2013b) understood the term *accent stereotype* to mean the association between an accent and the speaker's personal characteristics, such as intelligence, SES or trustworthiness (Kinzler & DeJesus, 2013b, pp. 1146-1147). In this case, the definition of stereotype is derived from social psychology (see Section 2.2.1). Both G. Kristiansen (2010) and Kinzler and DeJesus (2013b) can be classified as belonging under the development of language attitudes. Nevertheless, despite using the same term, one could argue they are investigating different topics.

Second, the field of the development of language attitudes lacks a clear delimitation of its sub-branches. This can pose a problem, particularly when a study cites the results of a different investigation as evidence for children's language attitudes, but their topics are technically different. For instance, Cremona and Bates (1977) is sometimes referred to as one of the earlier studies on language attitudes. In their study, the social judgement task was a binary choice (Standard Italian or the Valmontonese dialect) and the question, asked in Italian, was "which one spoke better?" (Cremona & Bates, 1977, p. 225). This question does fall under language attitudes, namely the Standard Language Ideology (see Section 2.3). However, just because children are able to rate different dialects as better or worse, it does not follow that they rate the speakers in the same

fashion, as done in Giles et al. (1983, p. 141). These two topics (what children think of dialects and what children think of speakers of said dialects) arguably belong to different subfields of developmental sociolinguistics. Therefore, one cannot infer that they will be subject to the same developmental trajectories.

Currently, it is not known whether all the different subfields under this broad field develop in parallel, in a chain reaction fashion or independently of each other. This makes positing generalizations across studies challenging. Therefore, this section describes the main results in each subfield individually. At the end, similarities and parallelisms (or a lack thereof) are summarized.

Individual studies within the development of language attitudes are sorted into three different sub-categories in order to facilitate comparison of results. These categories are based on thematic similarities. The first sub-category (Section 3.1) includes research on the development of the Standard Language Ideology (see Section 2.3). The second sub-category (Section 3.2) contains studies about when and how children start associating different groups of people with different ways of speaking. In this context, different groups of people can involve different ethnicities or people from different geographical locations. The third sub-category (Section 3.3) encompasses literature investigating when children begin to use different codes to infer qualities about the speaker. The main difference between the Standard Language Ideology and this last sub-category is that the former deals with attitudes about codes. For example, someone may believe that Standard English is the only right way to speak or that a Scottish accent is uglier than the Standard British accent. In contrast, the third sub-category deals with attitudes about speakers: a hearer may dislike a speaker because of the dialect they use, or they may think a speaker is less intelligent because of the language they speak. This sub-category follows the tripartite structure of attitudes described in Section 2.2: prejudice (Section 3.3.1), discrimination (Section 3.3.2) and stereotypes (Section 3.3.3).

3.1 Standard Language Ideology

From a young age, children believe that some codes are better than others. Rosenthal (1974) showed that children in the age range 3 through 5 years show signs of Standard Language Ideology. Children said that an anthropomorphized box speaking General American (GA) spoke better than a similar box speaking African American Vernacular English (AAVE). These results held regardless of the participants' race. In contrast, Day (1980) presented different results. They compared reactions to GA and Hawaiian Creole English following the same set up as Rosenthal (1974). Participants in Day (1980), who were attending kindergarten, rated the speaker of their native variety as speaking well. That is, native speakers of Hawaiian Creole English rated the Hawaiian Creole English speaker higher and native speakers of GA rated the GA speaker higher. It was only in the first grade that both groups rated the GA speaker higher. However,

the lack of inferential statistical analyses in these two studies makes comparisons difficult.

More recent research has confirmed the early signs of Standard Language Ideology mentioned in the previous paragraph. One example is Barbu et al. (2013). They showed that socioeconomic status (SES) affects the development of negative views toward French liaison (the realization of a consonant sound between two words). Specifically, children from families with a higher SES considered speakers producing mandatory liaison as better speakers at the age of 3 and a half years. In contrast, children from families with lower SES only showed the same pattern after the age of 4 and a half years. Similar disparities were found with regard to variable liaison. Participants with higher SES rated speakers using the non-standard variable liaison worse (following the adult norm) at the age of 5 and a half years.

Another example of children's early views on codes is present in Satraki (2015). Between the ages of 5 and 7 years, children's ratings matched the acrolect, mesolect and basilect continuum found in Cyprus. They rated standard Greek highest (the acrolect), followed by the local variety most similar to the standard (mesolect) and, lowest of all, the local variety most different from the standard one (the basilect). Interestingly, Fotiou and Ayiomamitou (2021) did not find signs of Standard Language Ideology in Cyprus pupils aged 9–12 years. In their study, children from both rural and urban cities in Cyprus were asked about their opinions regarding Cypriot Greek and Standard Modern Greek. Their responses showed an equally positive view of both dialects. This difference could be the result of the methodologies used in these studies. Satraki (2015) used a more indirect approach. Participants saw three cartoon characters on a computer. They had to rate them on different variables (e.g. who talks best). In contrast, Fotiou and Ayiomamitou (2021, p. 5) used a direct approach. Without the use of stimuli, participants were asked to what extent they thought speakers of different dialects matched different traits. The use of direct or indirect methods in attitudinal research has been shown to produce different results (see Chapter 2). Therefore, it is not unreasonable that the different results in Satraki (2015) and Fotiou and Ayiomamitou (2021) reflect this.

There is also evidence of the Standard Language Ideology in children between the ages of 6 and 7 years. Cremona and Bates (1977) showed how children at the age of 6 years consider Standard Italian to be better than the Valmontonese Italian dialect. Their data showed how the effect grows stronger with age. Likewise, Stamou et al. (2015) reported that children this age rated Standard Modern Greek better than rural Greek dialects used in children's shows. Odato (2010) showed similar results in this age bracket for discursive uses of *like*, with children from the age of 7 years negatively judging non-standard uses of *like* (such as *it's* like *a girl thing*). van Bezooijen (1994) reported that 6- to 8-year-olds rated standard Dutch as more beautiful than other regional accents. Finally, other studies (like Ramirez et al. (1976), Lewis (1975), Ubalde

et al. (2017) and Ioannidou (2004)) have also shown the presence of linguistic ideologies toward different codes in the age range 10–12 years.

In sum, children show signs of Standard Language Ideology from approximately the age of 3 years. This ideology is present when the stimuli involve dialectal differences or even a single feature. However, the precise timeline is subject to sociocultural factors, specifically SES. Children from higher SES families reproduce these views at an earlier age than children from lower SES backgrounds.

3.2 Different people speak differently

Children seem to view language as an intrinsic property of people, just like eye color. This is known as the essentialist view of language (Medin & Ortony, 1989). Early signs of this view are present at the age of 6 months, as shown by Uttley et al. (2013). In their study, infants expected people from other races than their own to speak differently (just as they associate differences in gendered faces to differences in speech, Patterson & Werker, 2002). Children looked longer in trials in which a speaker introduced with an East Asian face spoke English than in trials in which the speaker spoke Chinese. This pattern was not observed with Caucasian speakers. Furthermore, looking rates were random when backwards Chinese and English was used instead of actual speech. These results imply that infants matched faces of other races with other languages, not just "unfamiliar sounds" with other races.

Similar results at the age of 16 months are shown in Weatherhead (2017) and Weatherhead and White (2018) with accents rather than languages. In their studies, they created artificial accent differences by shifting the vowel of certain words in either a systematic or random way. Children accepted unfamiliar pronunciations of familiar words if the speaker was of a race other than their own, but not if the speaker was the same race as them. Importantly, this effect only took place when the differences in pronunciation were systematic, not random vowel changes. Moreover, children only accepted unfamiliar pronunciations of familiar words after the first block of trials. This indicates that participants learned to associate different races with different ways of speaking. The fact that they learned to associate a difference in race with a difference in accent is highlighted in Weatherhead's (2017) data from 24- to 26-month-old infants. Their learning of novel vowel realizations generalized from one speaker of a different race than their own to another speaker of the same race as the original other-race speaker. Infants did not generalize from a speaker of a different race to a speaker of a third race unless the two speakers of different races were depicted as friends (they wore matching t-shirts, waved at each other and hugged). Hence, infants view race and more abstract relationships as cues to linguistic group membership.

The nature of children's associations between different groups of people and different codes can be better understood by looking at the switched-at-birth paradigm

introduced in Gelman and Wellman (1991). In this kind of task, participants are shown two families. These families are shown to use different codes. Furthermore, they are sometimes shown to differ in another key variable, typically race. Thus, a (white) family may be portrayed speaking English, while a (black) family may be depicted speaking French. Then, participants are told that a child born in one family is adopted by the other family. Participants are then asked what language the adopted child will speak when they grow up and/or what race the adopted child will be when they grow up. If the child selects the language/race of the original family, it implies they view that trait as fixed at birth. If the participant selects the language/race of the adopting family, however, it can be inferred that they view it as the result of their environment. As a non-linguistic example, 4-year-olds know that if a kangaroo baby is adopted by a family of goats, the kangaroo will grow up to be a kangaroo and not a goat (Gelman & Wellman, 1991).

This switched-at-birth paradigm has been a fruitful method for studying children's understanding of language. When only using different languages (English and Portuguese), Hirschfeld and Gelman (1997) found developmental effects. Children aged 2–3 years answered at chance level, that is, they did not view language as forming natural kinds. Children aged 4–5 years, however, answered that the adopted child will grow up to speak the language of the original family, not the language spoken by the adopting family. That is, they viewed language as an intrinsic property of people that is fixed at birth.

These results match those in Dautel and Kinzler (2018). Nevertheless, they show that this development is affected by both the lived experiences of the participant and the context of the data collection. Dautel and Kinzler (2018) used both race and language in their switched-at-birth task. They showed participants a black or white child speaking in either English or French. Then, participants were shown a picture of two adults: one black and one white. The language the adult faces spoke varied across trials. Participants had to say who the child will grow up to be. Bilingual 5- to 6-year-old children from the greater Chicago area (none of whom spoke French) consistently chose the speaker who spoke the same language, even if there was a mismatch in race (see similar results in Kinzler and Dautel (2012) for white participants). In a follow-up study, Dautel and Kinzler (2018) recruited bilingual French speakers from a French immersion school. Dautel and Kinzler (2018, p. 292) reported that the answers of participants tested in English followed a random pattern. In contrast, those tested in French chose the adult who matched the race of the child, not the language.

Similar results are shown by Byers-Heinlein and Garcia (2015) and Kinzler and Dautel (2012). In the former, 5- to 6-year-old sequential bilingual children were more likely to say that an adopted child would speak the language of the adopting family than were simultaneous bilingual children. In the latter, black participants aged 5 to 6 years chose the adult speaker who matched the race of the child, not the language.

With age, children's knowledge about the relationship between social groups and language becomes more precise. Kinzler and Dautel (2012) showed that 9- to 10-year-olds consistently chose adult speakers who matched the race of the child speaker across white and black participants. Similarly, Odato (2010) showed that 9- to 10-year-old children could attribute a single linguistic feature (the use of *like*) to female speakers.

In conclusion, humans appear to have an essentialist view of language from an early age. In fact, they can learn correlations between social groups and differences in code quickly, such as during an experimental task. How this view of language interacts with children's views of other social categories, such as race, seems to vary with age. First, other-race-ness is correlated with other-language-ness. Second, language seems to be considered more stable than race. Finally, they learn that race is not mutable, unlike language. Importantly, this last step seems to vary based on children's sociocultural context. This includes variables such as the child's race, kind of bilingualism and even task setting. Specifically, black children, sequential bilinguals and conducting the task in a child's second language seem to highlight the mutable nature of language through development.

3.3 Attitudes toward speakers

This section deals with studies on children's attitudes toward speakers based on the code they use. It is split into three sub-headings based on the tripartite structure of attitudes from Section 2.2. First, Section 3.3.1 includes studies on linguistic prejudice, Section 3.3.2 on linguistic discrimination and, finally, Section 3.3.3 on linguistic stereotypes. As a reminder, prejudice refers to the affective component of attitudes, discrimination to their behavioral component and stereotypes to their cognitive component. For example, a person may like or dislike someone because of the language they use (prejudice). They may treat them differently because of a linguistic feature they employ (discrimination) or they may have a particular belief about them due to the dialect they speak (stereotype).

Grouping the studies into these three categories was not always a straightforward matter. Specifically, the younger the participants in a study were, the more challenging the classification was. This is because the methods employed with newborns and infants need to be more indirect in order to accommodate their cognitive skills. Therefore, the definitions of the terms prejudice, discrimination and stereotype were applied in a technical sense. For example, a study was categorized as involving discrimination if a participant treated a speaker differently, regardless of whether differently means better, worse or neutral. An example of this involves the head-turn preference procedure and the preferential looking paradigm. Studies employing these methods were categorized as discrimination studies. This is because participants treated one speaker differently:

by choosing to listen to them. More examples of tasks included in each category are mentioned in the relevant section.

3.3.1 Prejudice

The literature on the development of linguistic prejudice typically uses two kinds of tasks. On the one hand, there are friendship tasks. In these, children are asked "who would you rather be friends with" or variations thereof. On the other hand, there are "liking" tasks. In this case, participants are asked "who do you like best," "who do you think is nicest" or similar questions.

At the age of 3, there seem to be no general signs of prejudice in children. While Rosenthal (1974) showed a tendency for white children to prefer a speaker of GA, black children seemed to answer randomly. Due to the lack of inferential statistical analyses, it is difficult to draw strong conclusions. More conclusive results are shown in Creel and Seubert (2015) and Creel (2018). Both studies reported that children at this age did not prefer a GA speaker over a Dutch-accented speaker. A trend toward liking GA started to appear at the age of 4 years in Rosenthal (1974), Day (1980), Creel and Seubert (2015) and Creel (2018).

This trend seems to consolidate at the age of five. van Avermaet and McClintock (1988) showed that children preferred a classroom with children speaking their same native language (Dutch) over one with children speaking another language (French). In a Canadian context, children also preferred speakers of English (their ingroup) over French (Powlishta et al., 1994). Similarly, children in the US preferred speakers of English over either French or French-accented English (Kinzler et al., 2007).

Kinzler et al. (2010) controlled for intelligibility, race and other physical attributes to understand the nuances of these preferences. They showed that this preference held even though children understood the French-accented speakers. When including race, children preferred a black speaker of English over a white speaker of French-accented English, even if they would choose a white person in silent conditions. In a parallel fashion, children preferred a computer-generated distorted face speaking English over a non-altered face speaking French-accented English, despite choosing the non-altered face in silent conditions. Similar results have been found for this age group in Satraki (2015) for Greek dialects in Cyprus, Souza et al. (2013) for French and English versus Haitian Creole, and Paquette-Smith et al. (2019) for English with a Canadian accent, over RP and Korean-accented English.

Children's linguistic prejudice at this age is nevertheless subject to sociocultural factors. Kinzler, Shutts, et al. (2012) collected data in Langa, a multilingual community in Cape Town, South Africa. In a rural area of Langa, children aged 5–11 years preferred a Xhosa speaker over a speaker of French, a foreign language. When tested using Xhosa and English, the local lingua franca, there was a non-significant trend in favor of English. However, children who attended an English-speaking school chose

English more frequently, while children who attended a Xhosa-speaking school chose Xhosa. Kinzler, Shutts, et al. (2012) also tested children in a racially diverse school in Cape Town where English was the language of instruction. Here, children chose English speakers over Xhosa speakers and, in turn, Xhosa speakers over Sesotho speakers, a non-native, but familiar, local language. A certain parallelism to these results is also shown in Kinzler and DeJesus (2013b). In this study, 5-year-old speakers of a more prestigious northern variety of American English chose speakers of their same accent over speakers of a southern, less prestigious accent. In contrast, 5-year-old children from the south of the US chose randomly. Thus, growing up in an environment where a socially prestigious code is spoken leads to earlier development of a preference for it. In fact, similar results were also found in Rosenthal (1974), with white children preferring GA and black children answering at chance.

A similar, nuanced picture of language prejudice among 5-year-olds was shown in DeJesus et al. (2017). French–English bilingual children from a French immersion school in Chicago did not prefer an English speaker over a French speaker. However, they chose a native English speaker over a French-accented English speaker. In fact, they found parallel results with Korean–English bilinguals in Chicago: children's preferences did not vary between English speakers and Korean speakers. However, they chose L1 English speakers over Korean-accented English speakers. This favoritism toward L1 codes does not only apply to accents. Hwang and Markson (2018) showed that children preferred an L1 speaker of GA over both a Korean-accented English speaker and speakers with a GA accent but with non-standard syntactic or semantic features (such as *four seasons there are and is coldest one winter* for a non-standard syntax and *at night, people can see the stars in the chair* for non-standard semantics) (Hwang & Markson, 2018, p. 1034).

At the age of 6 years, children seem to consider the overall behavior of speakers when deciding who they like best. Durkin and Judge (2001) created a two-by-two design involving scenes with (i) English speakers or people speaking gibberish and (ii) being prosocial (e.g. politely asking for food) or antisocial (e.g. taking food from each other in a hostile or aggressive manner). In the prosocial condition, 6-year-olds did not prefer either the English or the gibberish speaker. In the antisocial condition, in contrast, they preferred the English speaker. These results changed with 8-year-old participants. In that case, they preferred English speakers regardless of behavior. In the prosocial condition, they chose the English speakers, while in the antisocial condition, they dis-preferred the English speakers less. A similar increase in dis-preference for non-L1 varieties at this age was shown in Hwang and Markson (2018): from the age of 8, children began to prefer Korean-accented speakers over semantically and syntactically anomalous sentences, an effect that got stronger with age. However, the preference of an American over a Korean accent did not change.

The results on this timeline for prejudice are not uniform. For example, Creel and Seubert (2015) did not find significant differences for American English and Dutch-

accented English at the age of 5 years. It is not until the age of 6 and 7 years that their participants began to prefer American English speakers. A key variable for understanding this result may be that of accent strength. Creel (2018) noted that English and Dutch are typologically similar in phonology. This may have imposed a "ceiling effect" on how strong a Dutch accent can be compared to, for instance, a French accent. More discrepancies are found in Cohen and Haun (2013). They collected data in four Brazilian Amazonian towns. Two of them had high accent diversity while the other two had one dominant accent. They tested children aged 5 to 10 years. Regardless of location, children began preferring the local accent at the ages of 9 and 10 years. This can possibly be explained by the accents used. The accent contrasting the local accent was a European Portuguese accent from Madeira. The comparatively late age for signs of linguistic prejudice may be due to the fact that Brazilian accents are typically lower in prestige than their European counterparts (Bagno, 2003; Bugel, 2009). The fact that speakers of less prestigious varieties do not show the same favoritism toward their own accent was also shown in Kinzler and DeJesus (2013b) above. In fact, in Kinzler and DeJesus (2013b), 9- and 10-year-olds from the southern US showed no prejudice for speakers of either the Northen (prestigious) or Southern (non-prestigious) accents.

To conclude, signs of linguistic prejudice seem to appear between the ages of 3 and 4 years and consolidate at the age of 5 years. After this age, language is not the only tool children use in forming their responses. However, this development is modeled by different factors. The first one is the role codes play in a society. Specifically, growing up in an area in which a prestigious code is spoken appears to speed up the development of linguistic prejudice. In contrast, children growing up in an area where a less prestigious code is used develop comparable attitudes at a later age. The second factor is what code children react to, which interacts with the first factor. It appears to be the case that, for children speaking with a low-prestige accent, linguistic prejudice toward languages seems to appear at an earlier age than toward L1 accents. Nevertheless, this has not been directly tested. It can only be inferred by contrasting results from Kinzler and DeJesus (2013b), Cohen and Haun (2013) and Kinzler, Shutts, et al. (2012).

3.3.2 Discrimination

The study of linguistic discrimination with younger participants has taken several forms. Typical examples of tasks in the literature include (i) who participants prefer to give gifts to or receive them from, (ii) head-turn tasks and (iii) whose instructions children imitate. As mentioned above, there is a sharp contrast between the methods used with newborns and infants, on the one hand, and toddlers and young children, on the other, due to their differences in cognitive abilities.

Children's linguistic discrimination seems to be shaped by their linguistic environment even before birth. Newborns 0 to 5 days old prefer listening to a language

the mother spoke during pregnancy or a rhythmically similar language (Byers-Heinlein et al., 2010; Moon et al., 1993). This also applies to bilingual mothers who spoke two different languages during pregnancy. Similar patterns are found at the age of 3 months. Australian newborns preferred listening to an Australian accent over an American accent (Kitamura et al., 2013). Moreover, 4- to 5-month-olds preferred listening to their native language over a different language of the same rhythm class (Bosch & Sebastián-Gallés, 1997). Analogous results are present in Butler et al. (2011). Five-month-old infants preferred their local English accent over another L1 variety of English. In contrast, children showed no preference between two nonnative L1 accents of English. At the age of 6 months, Kitamura et al. (2013) showed that children prefer listening to an Australian accent over a South African English accent. In contrast, their preference for an Australian accent over an American accent was lost, although they were still able to distinguish between the two.

The keyword in the results above is accent, specifically L1 accent. Namely, when the contrast is between a nonnative L1 accent and an L2 accent, infants show signs of linguistic discrimination. In Kinzler et al. (2007), 5- to 6-month-old infants preferred listening to an L1 speaker (either English in the US or French in France) over a speaker with an L2 accent (French-accented English in the US or English-accented French in France). Further instances of discrimination at this age involve different languages. For instance, Kinzler et al. (2007) showed that children preferred listening to English speakers over backwards speech or Spanish speakers. Similarly, in Soley and Sebastián-Gallés (2015), 6- and 7-month-old infants preferred listening to music introduced by a speaker of Spanish (their native language) over a speaker of German. The same pattern was also found when 10-month-old infants were asked which toy they wanted. They took toys from speakers of their own language rather than from speakers of a foreign language (Kinzler et al., 2007; Kinzler, Dupoux, et al., 2012). Moreover, 1-year-old infants also choose food introduced by speakers of their native language versus another language (Shutts et al., 2009).

These early signs of linguistic discrimination appear to be modulated by exposure to linguistic diversity. Howard et al. (2014) tested 19-month-olds' willingness to imitate either English speakers (their native language) or Spanish speakers. They found that infants growing up in linguistically diverse neighborhoods were more likely to imitate the Spanish speaker than were infants growing up in more linguistically homogenous neighborhoods. Importantly, it was not direct exposure to Spanish specifically that mattered, but "incidental exposure" (Howard et al., 2014, p. 477) to linguistic diversity during early development.

At the age of 3 years, there is further evidence of linguistic discrimination in favor of different dialects and accents. White children took gifts from speakers of GA, while black children took gifts from speakers of AAVE (Rosenthal, 1974). In a similar fashion, children at this age accepted the label of a novel item when offered by a speaker of GA over a speaker with a Spanish accent (Corriveau et al., 2013). Nevertheless, this

age group chose randomly if the item was a common object, but the label was wrong (e.g. they called a doll *a cup*). More nuances regarding linguistic discrimination in this age group were found in Kaiser and Kasberger (2021). They asked children between 3 and 11 years which doctor they would like to visit (for the older participants) or which doctor their puppet should go to (for the younger participants). The key variables in their results were (i) participants' SES and (ii) the doctor's gender. For female doctors, children between 3 and 6 years from medium and higher SES chose the doctor speaking Standard Austrian German. In contrast, children from lower SES families chose the doctor speaking their local accent. For male doctors, answers were at chance level.

At the age of 4 years, Day (1980) reported parallel results to those of Rosenthal (1974). In a Hawaiian context, English speakers accepted a gift from English speakers, while speakers of Hawai'i Creole English chose the gift from their own native language. Kinzler et al. (2011) also reported a preference among children for speakers of the same native accent. Children between the ages of 4 and 5 imitated how native GA speakers used a novel item over how Spanish-accented speakers used the same item. In fact, this difference was still present when the speakers were introduced by reading the *Jabberwocky* poem by Lewis Carroll, that is, a text that follows English syntax and phonotactics but is nonsensical (e.g. *all mimsy were the borogoves*, and *the mome raths outgrabe*).

Children from the age of 5 years display more nuanced reasoning when discriminating in favor of different speakers. In contrast to their 3-year-olds, Corriveau et al. (2013) reported that children at the ages of 4 and 5 years accepted new labels from speakers who previously gave correct labels, regardless of accent. That is, children preferred a Spanish-accented speaker who called a cup a *cup* rather than a native English speaker who called a cup a *doll*. Comparable results can be seen with regard to speaker certainty (Wagner, Dunfield, et al., 2014). When children were told to imitate a novel action, they preferred a native GA speaker over a Spanish-accented speaker. However, they also controlled for certainty. In some cases, the speaker acted as if they were not certain of how to perform the novel action. In other cases, they acted confidently. In this case, children preferred to imitate the actions of the speaker who acted in a confident manner, regardless of accent.

From the age of 7 years, linguistic discrimination seems to consolidate. Participants in Kaiser and Kasberger (2021) chose the doctor with a standard accent regardless of gender and SES. Similarly, children gave more coins to speakers of their native accent than to speakers of French-accented English, even when controlling for the race of the speaker (Spence & Imuta, 2020). Results parallel to these were found in Cohen and Haun (2013), where 7-year-olds chose to give a sweet to a local speaker. Moreover, 7-

to 10-year-olds preferred getting a sticker from a native speaker than from speakers with either an nonnative L1 or L2 accent (Hanulíková, 2024)[11].

Not all results investigating linguistic discrimination in children are in agreement. In Spence and Imuta (2020), the preference for L1 over L2 accent regardless of race only became significant at the age of 7 years. Prior to that age, participants gave more coins to white L1 speakers than to black French-accented speakers. If race and accent did not match, results were random. Furthermore, Cohen and Haun (2013) found no instances of linguistic discrimination in their 5-to 6-year-old participants. At this age, children simply chose the option that was more beneficial to them.

Discrimination, in its broadest sense, appears very early in life and becomes stable around the age of 7 years (see also Spence et al., 2021 for a meta-analysis). The specific timeline for its development is subject to different influences. Exposure to linguistic variation as well as belonging to lower SES families seems to delay signs of linguistic discrimination. Task difficulty has also been proposed as an effect delaying signs of linguistic prejudice (see Spence and Imuta (2020) and its comparison to Kinzler et al. (2010)). However, other societal factors may also be relevant. The data collection in Spence and Imuta (2020) took place in Australia. In contrast, Kinzler et al. (2010) tested children in the US. Therefore, the potential influence of cultural differences cannot be ignored.

3.3.3 Stereotypes

The literature on the development of language stereotypes is divided into three categories: first, research investigating expectations of friendship between different speakers is described; second, associations between code and speakers' general background (such as clothing, location, housing etc.) are introduced; third, the focus is placed on the development of connections between code and speakers' personal attributes, such as intelligence, friendliness, etc.

3.3.3.1 Who is friends with whom?

Liberman et al. (2017) showed that, at the age of 9 months, children expected language to dictate relationships. They tested infants' gaze while watching a video of two people interacting. Their two-by-two design included speakers of English and Spanish, on the one hand, and these speakers being friendly or unfriendly toward each other, on the other. Participants looked more at videos where English speakers were being unfriendly toward each other. Participants looked more at videos in which a Spanish and an English speaker were friendly to each other.

[11] Interestingly, Hanulíková's (2024) participants preferred getting a sticker from an L2 speaker than from a nonnative L1 speaker.

A more direct approach is found in Arredondo and Gelman (2019). They overtly asked children who they thought a speaker would be friends with among two options. The variables included were code, gender and race. When the linguistic variable was L1 Spanish accents, English–Spanish bilingual children from the US aged 4–12 years chose the speakers of the same gender or race, regardless of accent. In contrast, if the variable was language (English or Spanish), bilingual speakers aged 4–6 years believed the speakers of the same language would be friends, regardless of race or gender. However, this result differed when the participants were monolingual English speakers. In this case, they chose gender or race over language.

The results presented in the previous paragraph appear to be subject to children's knowledge about speakers. An example includes Weatherhead, White, et al. (2018). They showed 4- to 6-year-old children two speakers speaking with different accents (either (i) Spanish-accented English and South African English or (ii) Turkish-accented English and Mandarin-accented English). Then, participants were shown a speaker with an accent that matched one of the previous two. They were told that the three speakers lived in either the same place, in different places or were not given information about location. The task had participants say whom they thought the third speaker would want to be friends with: the one with a matching accent or the one with a different accent. Those who were told that the three speakers lived in different places answered at chance level. The other two groups were significantly more likely to answer that the third speaker would want to be friends with the speaker of the same accent. Importantly, the effect of accent without location information was significantly stronger than when participants were told they all came from the same place.

From an early age, humans seem to use language as a cue to infer relationships. Distinctions of language, rather than accents, are the relevant keys at this stage. At the age of 4, bilingual children seem to choose language over race and gender. When it comes to accents, 4- and 6-year-olds assume speakers of two different L2 accents will want to be friends with speakers of the same accent. Interestingly, if the variable involves L1 accents, children choose race or gender over linguistic differences in assumed friendships.

3.3.3.2 Code and speakers' general background

The previous section shows that very young children assume social relationships between speakers of the same code. In contrast, it is not until the age of 3 years that children exhibit more contentful stereotypes. Hirschfeld and Gelman (1997) tested toddlers aged 3–5 years with a task in which they needed to match people speaking either English (their native language) or Portuguese (unknown to the participants) with pictures. The pictures were split into four categories: race (black or white), clothing (western or non-western), housing (western or non-western) and orientation (a control condition with an adult facing forwards or backwards,). All participants matched the English speaker with the white face and the western items. This effect grew significantly

stronger with age. In the case of the 3-year-olds, they were the only ones to match a front-facing face with English. Similar results were found for this age group in Weatherhead et al. (2016), which included L2 accents: Spanish-accented English and Turkish-accented English. This implies that children did not generalize from one L2 accent to another L2 accent nor did they match "familiar item" to "familiar accent."

The results present in the previous paragraph also hold at the age of 4 years. Children mapped speakers with the same accent to liking the same food item unless they were overtly told the speakers live in different places (Weatherhead, White, et al., 2018). The attribution of accents to different locations at this age was also corroborated by Weatherhead et al. (2019). When asked to place speakers on a map based on where these speakers lived, they placed the speakers who share the participants' accent significantly closer than other speakers of L1 or L2 accents. These two categories did not differ from each other. In a modified task, their stimuli were low-pass filtered so that only prosodic information remained. In this case, L2 accents were placed significantly further away than L1 accents[12].

In contrast to these results, Creel (2018), Creel and Seubert (2015), McCullough et al. (2019a) and Dossey et al. (2020) all showed that 4- to 5-year-old children could not successfully categorize accents when asked "who is from here." This discrepancy in results can potentially be explained on two grounds. First, how direct the tasks in these studies were could have influenced the results. Weatherhead's team used a more indirect method compared to the overt geographical question of, for example, Creel (2018). As noted in Chapter 2, the difference between direct and indirect methods is of theoretical importance.

Second, the specific phrasing of the question in Creel (2018), Creel and Seubert (2015), McCullough et al. (2019a) and Dossey et al. (2020) may have caused the differences in results. Weatherhead, Friedman, et al. (2018) can shed some light in this respect. Their paper looked at whether 4- to 6-year-old children's responses about the relationship between accent and geography changed as a function of the way in which the question was phrased. Their results showed that children thought that speakers of L2 accents were *born* in a different place while speakers of L2 did not necessarily *live* in a different place. Thus, questions such as *is this person from here* may have been too ambiguous. It could be taken to mean both that the speaker was born here or that the speaker lived here. Hence, these differences in results may be attributed to methodological disparities.

At the age of 5 years, children can map differences in languages to nationality, not just broad differences in location. Santhanagopalan et al. (2021) showed that children in Chennai, India, considered speakers of Tamil the most Indian, followed by speakers

[12] This contrasts with the results from Alcorn et al. (2020), where participants performed best with (i) unaltered stimuli or (ii) when the segmental information was kept but differences in intonation were removed. Nonetheless, their participants were adults. It is a possibility that different phonetic cues have different levels of relevance at different ages.

of Hindi. DeJesus et al. (2018) had similar results. Monolingual English speakers in Chicago rated English speakers as American regardless of race, but not Korean speakers, French-accented speakers or Korean-accented Speakers. They also found parallel results with monolingual Korean speakers in Korea: they rated Korean speakers Korean regardless of race, but not French- or Korean-accented English speakers. In contrast, bilingual English–Korean speakers from Chicago rated the French- and Korean-accented speakers as American if they were white, but rated English speakers as American and Korean speakers Korean regardless of race. It can be concluded that monolingual children have a stronger association between language and nationality, whereas bilingual speakers also take race into account.

Further intricacies at this age were shown by Kinzler and DeJesus (2013b) and Kinzler and DeJesus (2013a). In the former, when asked *who is American* and *who lives around here*, 5-year-olds from the northern US chose the local accent for both questions. In contrast, responses from children from the southern US did not differ from chance. In Kinzler and DeJesus (2013b), the authors investigated how accent and social behavior interact in children's accent stereotyping. They included (i) French-accented English and GA as well as (ii) friendly and unfriendly speakers. In the case of friendly French-accented speakers and unfriendly GA speakers, responses were at chance level. Nevertheless, when the contrast was more subtle, i.e. (i) a friendly French-accented speaker against a neutral GA speaker or (ii) a neutral French-accented speaker against a mean GA speaker, children chose the GA speaker as being American and from here. These results suggest that accent is not the only source of information children use to judge nationality.

From the age of 6, associations between code and geography seem to become more nuanced. van Bezooijen (1994) showed that 6- to 8-year-olds could associate speakers of standard Dutch with pictures of affluent houses with vast gardens and high-end cars. In contrast, they could not match other Dutch accents to their stereotypically associated housing images. Similar results were reported by G. Kristiansen (2010) with Castilian Spanish accents. Six- to 7-year-olds from Madrid were able to correctly categorize the Madrid accent at the level of province, region and broad area. But they could only correctly categorize one other accent at the levels of region and broad area.

In the case of L2 accents, children could categorize GA speakers with western items and as being from here and Korean-accented speakers with non-western items and as not bring from here (Hwang & Markson, 2018). Cohen and Haun (2013) also reported that 7-year-olds from multi-accent sites could correctly categorize local or non-local speakers as from here or not from here. Furthermore, McCullough et al. (2019a) showed that, from the age of 8, their participants could successfully categorize L1 speakers as from here or not from here by accent.

Improvements in associations between code and stereotypes in 9- to 12-year-olds are found in the literature. Some examples include van Bezooijen (1994), Trowell (2007) and Bouchard Bouchard (1969). In this age rage, children could associate different

accents and dialects with differences in housings and jobs. Furthermore, children from Madrid (G. Kristiansen, 2010) could associate L1 Spanish accents from Spain, L1 Spanish accents from other Spanish-speaking countries and L2 accents of Spanish with the speakers' geographical background, a result matching findings from Anisfeld and Lambert (1964) with Canadian French and European French.

In short, 3-year-olds are already able to associate languages and accents with the speaker's general background. However, at this age children also overgeneralize. That is, they match a forward-facing face to their native language. From here, the association gets stronger with age. Interestingly, the L1-L2 divide does not seem relevant unless the stimuli have been low-pass filtered. From the age of 5 years, children also map differences in code to nationality. From the age of 6, children begin being able to map different L1 accents from their own country to specific regions. These associations appear to become stable between the ages of 9 and 12 years.

3.3.3.3 Code and speakers' personal traits

The literature shows that children can associate someone's code with their personal traits or capabilities from the age of 4 years. For example, Schneiderman (1982) showed that Canadian girls from Welland, Ontario, were pro-French while boys' answers were neutral. The article did not include any inferential statistics. Furthermore, their analysis grouped their ten different questions together. As a consequence, it is not possible to know whether some of the questions are driving the effect or whether it is an overarching result across all of them. Powlishta et al. (1994) employed a similar method for a study with English-speaking children in Quebec. They showed a significant decrease in stereotypes toward French speakers. Nevertheless, they also merged the different questions included and did not include inferential statistics. Dossey et al. (2020) provided a better insight into accent stereotypes at this age. In their study, 4-year-olds rated a southern US accent as less smart than a midland, Mid-Atlantic or northern accent. A parallel result was shown in McCullough et al. (2019a), in which another low prestige accent, the New England accent, was rated low in SMART. With regard to FRIENDLY, the Mid-Atlantic accent was rated low in Dossey et al. (2020), while in McCullough et al. (2019a) this age group did not show any rating differences.

Five- to 6-year-old children have stereotypes toward different codes similar to those detailed above. Satraki (2015) reported that Cypriot children at this age rated speakers of standard Greek as more intelligent than speakers of local dialects, but, once again, this study lacks inferential statistics. In Santhanagopalan et al. (2021), Tamil was rated higher than Hindi, Indian English and RP with regard to both INTELLIGENT and KIND, and Hindi and Indian English were rated higher than RP with regard to KIND. In Kinzler and DeJesus (2013b), northern speakers of American English rated speakers with a northern accent as smarter, nicer and more in charge than southern speakers. In contrast, southern speakers of American English rated northern and southern speakers as equally smart, nice and in charge (Kinzler & DeJesus, 2013b). At this age,

McCullough et al. (2019a) showed that children rated the northern accent higher than the New England accent with regard to INTELLIGENT. Furthermore, New England speakers began being rated as less honest than midland speakers.

Some studies have revealed that 5- to 6-year-olds' ratings of speakers are dependent on both accent and social behavior. First, Kinzler and DeJesus (2013a) included GA and French-accented speakers without information about social behavior. In this case, they found a significant difference in ratings of SMART and IN CHARGE. Nevertheless, if the contrast was between a friendly French-accented speaker and a mean GA speaker, the French-accented speaker was rated as smarter. This effect held even if the set-up included a more nuanced difference. That is, a friendly French-accented speaker was rated as smarter than a neutral GA speaker. Similarly, a neutral French-accented speaker was rated higher than a mean GA speaker. Moreover, Durkin and Judge (2001) reported children judging unfriendly English speakers higher than unfriendly gibberish speakers. Unfortunately, this study also merged the scores across different questions. Therefore, it is not possible to discern whether a specific variable is driving the effect.

At the ages of 7 and 8 years, more literature shows a strengthening of previously found stereotypes. Bristolian children rated RP speakers as funnier[13] than speakers with a Welsh accent (Giles et al., 1983). Satraki (2015) reported what can be interpreted as a strong correlation between Standard Greek and the Cypriot dialect (the native accent of the participants in the study.) The former was rated highest on INTELLIGENT and lowest on the variables FUNNY, NAUGHTY and LAZY. In the US, the southern accent was rated significantly lower than the midland and northern accent, while still above the New England one (McCullough et al., 2019a). A similar pattern was found for FRIENDLY in Dossey et al. (2020): the Mid-Atlantic and southern accents were rated lower than the midland and northern accents. Nevertheless, with regard to SMART, it was the southern and midland accents that were rated below the northern and Mid-Atlantic (Dossey et al., 2020). In contrast to these results, de Vogelaer and Toye (2017) did not show signs of associations between Dutch accents and stereotypes in this age range.

Finally, from around the age of 10 years, children's accent stereotypes begin matching those of adults. Kinzler and DeJesus's (2013b) data showed that children in both the southern and northern US rated a speaker with a northern accent as smarter and in charge, while a speaker with a southern accent as nicer. This pattern matches that found in adults (Labov, 2006; Preston, 1993, 1998, 1999b). Parallel results can be found in Dossey et al. (2020). McCullough et al. (2019a) also reported that ratings for SMART and HONEST stabilize at this age into adulthood. Nevertheless, the RICH responses only began differing significantly across accents and showing adult-like patterns from the age of 12 years. Similar adult-like results from the age of 12 years

[13] It is not specified whether the variable "funny" is understood as *amusing* or *odd*. They "assum[ed] […] that the children took the latter meaning" Giles et al (1983, p. 144).

could be found in Giles et al. (1983), Nesdale and Rooney (1996), Anisfeld and Lambert (1964).

Stereotypes toward different codes at this age are subject to sociocultural factors. Lambert et al. (1966) showed that SES and bilingualism are important at this age. French-English bilinguals from a higher SES background developed stereotypes against French speakers at an earlier age compared to monolingual speakers from a similar SES background. These two groups, in turn, showed signs of accent stereotyping against French speakers earlier than bilingual lower-class participants. The effects of children's background on their accent stereotypes at this age are also apparent if we contrast Bouchard (1969) and Koch and Gross (1997). Bouchard (1969) used upper-middle class white participants. Their stimuli included GA and AAVE. Their results showed that AAVE was rated negatively overall. In contrast, Koch and Gross (1997) included only black participants. In this case, they rated AAVE higher than GA.

To summarize, 4-year-old children can match different L1 accents to different degrees of SMART. This continues up to the age of 5. Here, variables such as KIND and IN CHARGE become relevant. However, when children develop these stereotypes is dependent on the child's sociocultural context. Specifically, children growing up in a context in which a high-prestige language is spoken develop these stereotypes at an earlier age. Later, at the age of 6 years, children rate accents differently with regard to the variables HONEST and FRIENDLY. The number of variables subject to accent stereotyping further increases at the ages of 7 and 8 years. Finally, from the age of 10 years, children's responses begin to match those of adults with a peak at the age of 12 years.

3.4 Common trends

The results from the previous three sections are briefly summarized below. The earliest signs of language attitudes appear in newborns. Already in their first week of life, newborns prefer to listen to different codes, as shown by head-turn preference tasks. Around the first half year of life, infants can also associate differences in code with third-party relationships as well as with differences in race. Nevertheless, it is not until the age of 3 years that we see a leap in children's development of language attitudes. At this age, children are aware that certain ways of speaking are considered socially better than others (Standard Language Ideology), trust new information of selected speakers (discrimination) and associate speaking different codes with differences in general background (stereotypes). Around the age of 4 years, children begin explicitly classifying people speaking the same code as friends, seeing language as a permanent, human feature (unlike race) and associating different personality traits with speakers of different codes. In contrast to these results, clear signs of linguistic prejudice do not seem to appear until the age of 5 years. It is also at this age that other areas subsumed

under the development of language attitudes start becoming more nuanced. Signs of adult-like attitudes can be seen around the age of 7 for discrimination and Standard Language Ideology. In contrast, it is at the age of 10 that children begin showing adult-like attitudes with regard to prejudice as well as speakers' general background and personal traits. Finally, at the age of 12, children have levels of language attitudes roughly similar to those of adults.

The findings in the three previous sections suggest that the development of language attitudes does not follow a one-size-fits-all trajectory. This is in line with the Cognitive Linguistics view that the way language (understood in its widest sense) is learned and used is shaped by the sociocultural context of speakers. Specifically, a number of variables appear to be key for when and how children develop language attitudes. These are briefly discussed below.

First, growing up speaking more than one language appears to influence this development. Briefly put, multilingualism causes children to display adult-like linguistic attitudes at an earlier age compared to their monolingual peers. Nevertheless, not just any kind of multilingualism causes these differences. More adult-like results are found when the participants in a study speak the languages included in the stimuli. An example of this trend was found in Dautel and Kinzler (2018) in Section 3.2. Similarly, this effect seems to be caused by sequential bilingualism and not by simultaneous bilingualism.

Second, exposure to linguistic diversity seems to reduce linguistic discrimination in young children (see Section 3.3.2). Interestingly, this need not be exposure to the codes included in the study. Exposure to linguistic diversity in general during development appears to cause this effect. Unfortunately, exposure has not been included in the literature very often. Therefore, the strength of its effects is more uncertain. Next, the societal structures in which children grow up have important implications as well. Specifically, two key variables appear to be SES and the prestige of the code children speak. Children from families with a higher SES develop language attitudes at an earlier age than children from lower SES families. Similarly, children who grow up in an environment in which a prestigious code is spoken develop language attitudes at an earlier age than children who grow up where a code with less prestige is spoken.

Perhaps surprisingly, participants' race has not been a common independent variable. When it is present, its effects vary depending on the task. With regard to the switched-at-birth paradigm, black children understand that language is mutable, unlike race, at an earlier age than do white participants. However, when it comes to linking code to people's personal traits, black children show signs of preference for their own code all the way into prepubescence and adolescence, a pattern that changes in adulthood (Doss & Gross, 1994; Garner & Rubin, 1986). This parallels the results of children growing up speaking a less prestigious code, who also develop adult-like attitudes at a later age. In the same way, gender has not often been included. The literature including the participants' or speakers' gender as an experimental variable has

shown that girls have attitudes toward codes at a younger age than boys, and that attitudes toward female speakers appear at an earlier age than attitudes toward male speakers. Finally, intelligibility appears not to be a prerequisite for positive or negative attitudes toward different codes. Specifically, preferences for a particular code are present even when it is not intelligible. This contrasts with research on adults positing that processing fluency is a potential cause of language attitudes (Dragojevic & Giles, 2016; Dragojevic et al., 2024).

3.5 Shortcomings

The previous section depicts the main takeaways from the relevant literature. Nevertheless, there are difficulties in trying to draw strong conclusions or painting a clear timeline for the development of language attitudes. This section outlines the main obstacles to achieving these goals.

First and foremost, the methodologies across studies are too diverse to enable robust generalizations (see McCullough et al., 2019a; Paquette-Smith et al., 2019 for similar sentiments). Every aspect of an experiment's design can affect its results, from the specific tasks included to the phrasing of the instructions. As an example, Weatherhead, Friedman, et al. (2018) in Section 3.3.3 above showed the importance of subtle differences in phrasing. Children were able to say that people with L2 accents were *born* somewhere else, while they answered at chance level when asked whether these same speakers *were from somewhere else*. Another important parameter is the difference between direct and indirect methods. Research with adults has established that people display different language attitudes when asked directly (e.g. *what do you think of someone speaking RP?*) than when asked in a more indirect way (such as using the matched-guise or verbal-guise test; Garrett, 2007; Garrett, 2010; Kristiansen, 2001; T. Kristiansen, 2003, 2009, 2010). Therefore, it is not possible to equate or directly compare developmental studies that use different approaches. For example, the differences between Satraki (2015) and Fotiou and Ayiomamitou (2021) in Section 3.2 were proposed to be the result of the differences in approaches between these studies.

A similar methodological issue is the specific personality traits included in studies involving accent stereotypes. Adult research has long established that language attitudes cluster around three main factors (e.g. Zahn and Hopper (1985); see Section 2.3). Nevertheless, developmental research cannot assume that children use these same factors throughout development, as also argued in de Vogelaer and Toye (2017). Thus, even though traits such as friendly, nice, and helpful all represent a warmth factor in adults, it may be the case that children organize these traits differently. Hence, when working on developmental research, results from one study using only one of these traits cannot necessarily be extended to a similarly rated trait in adults.

A further problem is studies in which researchers include questions referring to different personality traits but merge all the results together for the analysis. A couple of instances of this issue were mentioned in Section 3.3.3. It cannot be assumed that children develop all of the different attitudinal variables at the same time. In fact, McCullough et al. (2019a) and Dossey et al. (2020) showed that children develop certain stereotypical associations before others. Therefore, merging variables makes it impossible to know whether the presence or absence of statistically significant results is being driven by a specific variable or is the result of an overarching tendency in the data.

Another important factor impeding the direct comparison of results is the way different studies report age. Age is arguably one of the most important variables in developmental research. Nevertheless, there is a great disparity in how age has been reported in the existing literature. Some articles include participants' ages only as a range. While smaller ranges like 5 to 6 years are probably narrow enough, results from a population described as being from 5 to 9 years are difficult to interpret. Without further clarifications (such as the lack of age effects in a statistical model), it is difficult to know whether the presence or absence of language attitudes is due to the mudding of results across ages or a generalized effect across the relevant age span. Similar difficulties arise when studies only report an age average without providing the age range of the participant pool. This way of reporting age does not only suffer from the issues above; it is also not possible to know what that mean represents without knowing the specific age range.

Finally, the use of school years (or grades) instead of age is also a significant challenge. The age at which children begin different kinds of education (kindergarten, primary school, middle school, etc.) varies from country to country. For example, in Sweden and Denmark, children in their first year of primary school are between 6 and 7 years. In contrast, children at this age are typically in their year two in England and in the first grade in the US. Moreover, these ages may also vary within a single country. In Scottland, children aged 6 to 7 years are typically in Primary 3. This means that, from a developmental standpoint, it is not possible to compare results from "first graders" across different educational systems[14]. Even more, the age of children at a given stage, for instance first grade, may also change within the same country due to educational reforms. This means that even comparisons in the same country at different periods in time are not straightforward. Children may also start school sooner or later based on their individual circumstances. An additional difficulty is that different educational systems use different cutoff points for educational steps. As an example, in Denmark, children in a class are typically born within the same calendar year. In England, children in a classroom are born between the 1st of September and the 31st of August of the year after. Hence, even if two studies report their participant pool including first grade

[14] Unless the goal of a study is to investigate the effects of schooling on language attitudes.

students, the biological age span may differ. Finally, when age is reported as a range of educational steps (like grades one through three,) there are increased difficulties, similar to those described in the previous the paragraph.

A further issue when trying to generalize results is the use of different kinds of codes as stimuli. The fact that children show language attitudes toward one specific accent does not necessarily mean they display the same attitudes toward another accent, a language, a dialect or a specific linguistic feature. DeJesus et al. (2017) offered an example of how the specific kind of code included in the stimuli is of theoretical relevance, as described in Section 3.3.1. Children preferred a native speaker of English over a French-accented or a Korean-accented English speaker. Nevertheless, this preference was not present when the contrast included English-vs-Korean or English-vs-French. Therefore, results with one specific code may not be applicable to another code. A related challenge is that of accent strength. Accent strength influences adults' views on accents as well as how well children can categorize them (see Section 2.1). Hence, different studies using the "same" accent may lead to different conclusions solely due to how strong the accents of their specific stimuli are.

Next, the specific language being researched can hinder the generalizability of results. Generally speaking, it cannot be assumed that children acquiring different first languages develop language attitudes in the exact same way. A plethora of language-specific factors may affect the specific trajectory for this development. This includes things from disparities in the phonetic inventories of their native languages to differences in sociolinguistic variation.

An example of the importance of language-specific parameters involves how vowels and consonants are processed. Research has shown that vocalic and consonantal sounds are processed differently by both adults and children (Bouchon et al., 2015; Cutler et al., 2000; Floccia, Butler, Goslin, et al., 2009; Floccia et al., 2018, pp. 10-11; Nespor et al., 2003; van Ooijen, 1996). For instance, when adults are given a non-word such as *wemmen* and asked to change one sound to make it a real word, most adults propose changing the consonant (and create *lemon*) rather than changing the vowel (and create *women*) (Cutler et al., 2000; van Ooijen, 1996). Similarly, Nazzi (2005) demonstrated that children perform better at a word-learning task when the target words contain contrasting consonants rather than vowels.

These discrepancies in processing vowels and consonants have been posited as a potential explanation for why children can distinguish L2 accents at an earlier age than L1 accents (Floccia, Butler, Girard, et al., 2009). Floccia, Butler, Girard, et al. (2009) noted that the main differences between English L1 accents are vocalic ones. In contrast, L2 accents include both vocalic and consonant differences. They argue that this can contribute to reducing the level of redundancy in L2 accents, making them more noticeable at an earlier age. Nevertheless, it is not the case for all languages that L1 accents are distinguished by their vowels. In Spanish, L1 accents are primarily characterized by differences in consonants (Hualde, 2005; Lipski, 2012, 2018). If the

vowel/consonant divide is important for the development of language attitudes, which features each language uses to differentiate its L1 accents is of theoretical importance. Hence, generalizing results from one language to another is not a straightforward matter.

Finally, even with children learning the same language, the specific socio-cultural context can hinder generalizing results. This involves both comparing studies done with the same language in different countries as well as different areas or regions within a country. For example, Kinzler and DeJesus (2013b) demonstrated that even within the US, northern children develop attitudinal responses to accents at an earlier age than southern children (see also Kinzler, Shutts, et al., 2012).

3.6 Summary

This chapter summarizes the key findings within the field of the development of language attitudes and forms the bases for the RQs in Chapter 1 (see Section 1.1 for how the existing literature informs this thesis' RQs). One of the main takeaway messages is that children use language (broadly speaking) as a tool for making social judgements from a much earlier age than traditionally thought. Briefly put, the earliest signs of language attitudes can be found in newborns, who already show preferences for listening to some codes over others. Before the age of five years, children are also able to infer third-party relationships based on language, judge different ways of speaking as better than others, trust information from speakers of certain codes over others and show the first signs of accent stereotypes. From this age forwards, their language attitudes become more nuanced, with adult-like patterns beginning to consolidate at the age of 10 to 12 years. Nevertheless, the literature shows that not all children develop language attitudes the same way. Specifically, speaking more than one language, the child's SES background, exposure to linguistic diversity as well as what code the child speaks all appear to influence when and how children will develop language attitudes. Unfortunately, making strong generalizations across studies is not a straightforward matter due to the vast differences in methodologies and analyses. These issues are dealt with in the next chapter, where the methods employed in this investigation are described.

4 Methodological considerations

This chapter outlines and motivates the methodological choices made to investigate the RQs in Section 1.1. To understand how English children acquire accent stereotypes in late childhood, a data collection was carried out in two British cities: Plymouth (Southwest England) and London. Using a cross-sectional design, participants between the ages of 7 and 11 years completed four tasks that included five different accents, while their legal guardians completed a survey. The details of this data collection are the topic of this chapter.

The choice of methods used in developmental research across childhood is a non-trivial question. This is because significant social and cognitive changes take place between the ages of 5 and 12 years. As a consequence, research methods that would be appropriate for a prepubescent child are likely too complex for a 6-year-old. Conversely, a task designed for 5-year-olds might prove too easy for an 11-year-old, leading to ceiling effects. In these cases, if a study finds differences in results across ages, these differences may reflect a methodological issue rather than a variable of theoretical importance, such as cognitive development.

The importance of using appropriate methods is not new in developmental research. One example concerns research dealing with object permanence. Object permanence refers to knowing that objects still exist if they are out of sight. Back in the mid-twentieth century, Piaget (1954, p. 13) posited that young infants, below the age of 9 months, lack this ability. Nonetheless, this claim was subject to criticism. Researchers pointed out that the methods employed by Piaget were rather complex and did not allow infants to successfully complete the task. Later on, studies employing methodologies adapted to younger infants showed that the development of object permanence is more nuanced than previously thought (Baillargeon, 1987a, 1987b, 1993; Diamond, 1985; Perris & Clifton, 1988; Stack et al., 1989). For instance, Diamond (1985) indicated that varying the amount of time between the researcher hiding the object and the infant being allowed to reach for it affected the response. Six-month-olds reached for the object if allowed to do so immediately. Seven-month-olds could wait two seconds and 8-month-olds four seconds. Therefore, how infants perform in research is the result of, among other things, the cognitive demands of the task at hand.

Similar discussions about the importance of methodological choices are found in sociolinguistics. One of the earliest studies of development of language attitudes is

Labov (1964). He collected data from participants between the ages of 8 and 19 years. His set-up included sociolinguistic perception and production tasks with different lexical items, such as whether *length* is realized with [n] or [ŋ]. Based on his findings, he proposed one of the first timelines of sociolinguistic development (Labov, 1964, pp. 91-92). He posited that it is only after the age of 14 years that people begin being aware of the social significance of language. This claim has sometimes been described as the reason for the relative lack of research in the field of developmental sociolinguistics described in Chapter 1 (e.g. Bratož et al., 2021, p. 236; Day, 1982, pp. 125-126). Labov's developmental timeline has been subject to criticisms (e.g. Bouchard, 1969; Bratož et al., 2021; Day, 1982; de Vogelaer & Toye, 2017; Jones et al., 2017; Nardy et al., 2013; Rosenthal, 1974). While the specifics of the criticisms vary, the overarching theme is that the methods used in his study were not appropriate for younger participants. Therefore, more recent work on the development of sociolinguistic competence highlights the importance of adapting research methods to the age of the target participants (e.g. Beck, 2014; Bell, 2007; Creel, 2018; Creel & Jimenez, 2012; McCullough et al., 2019a; Paquette-Smith et al., 2019; Wagner, Clopper, et al., 2014).

Another issue specific to how children develop language attitudes is that of the multitude of different methodologies used in the existing literature. As highlighted by McCullough et al. (2019a, p. 1082) and Paquette-Smith et al. (2019, p. 817), the tools employed to study how children develop language attitudes are significantly varied in the literature. All these methodological differences render making generalizations difficult (see Section 3.5 for further discussion).

These two issues (the importance of age-appropriate methods as well as the use of experimental techniques that facilitate the comparison of results across studies) are the bases for the choice of this thesis' methods. First, the main goal was to create a set-up that matches the skills of the youngest participants while still being engaging for older ones. Second, inspiration was drawn from current research in the field (see references in the sub-sections below) to allow for maximal comparability across current and future studies.

Section 4.1 describes the design of the tasks. Section 4.2 deals with what accents and texts were employed as stimuli. Section 4.3 describes how the speakers for each accent were selected. Section 4.4 deals with what participants took part in this investigation. Finally, Section 4.5 covers how the data was analyzed.

4.1 Design

This section outlines the structure of the experimental session. The set-up comprised five tasks. Four of them were completed by the child[15]. These were the intelligibility task (Section 4.1.1), the verbal-guise task (Section 4.1.2), the categorization task (Section 4.1.3) and the British Picture Vocabulary Scale (Section 4.1.4). The verbal-guise is the main component of this investigation, while the other three are auxiliary tasks. The final task, a questionnaire (Section 4.1.5), was completed by the child's guardian. The data collection described below was approved by the *Swedish Ethical Review Authority*, The University of Plymouth's *Faculty Research Ethics and Integrity Committees* and the *Queen Mary Ethics of Research Committee*.

The entire session was carried out online. The experimenter (the author of the thesis[16]) met one child and, at least, one caregiver at a time on Zoom. Different security measures were taken to ensure the participants' privacy. First, LU Zoom was employed, with servers within the EU that comply with GDPR and a data protection agreement between NORDUnet and Zoom. Moreover, the meetings were carried out with end-to-end encryption. Finally, the meetings were secured with a "waiting room." This allowed the experimenter to control who accessed the meetings at any given time as well as to only allow one participant in at a time.

The course of a session was as follows. Upon establishing rapport with a participant, the experimenter asked for verbal assent from the child. The goal of the study was explained to them in an age-appropriate manner. They were told that they did not have to participate if they did not want to and that they could stop at any time. They were ensured that "no one would be angry" if they chose to discontinue the session. If the child agreed, they were led to the Gorilla Experiment Builder platform (www.gorilla.sc) (Anwyl-Irvine et al., 2020)[17]. There, the legal guardian provided written informed consent. Before starting the experimental session, the child and the guardian were reminded to be in a quiet environment and that "it is the child's opinions that matter"[18].

The following modules were used to build the tasks in Gorilla. The experimental tasks were built with Gorilla's Task Builder 2, while the parental survey used Gorilla's Questionnaire Builder 2. Moreover, Gorilla's *multiplayer* function was used. This

[15] See the entry for this thesis at https://portal.research.lu.se/en/ for the appendix including images of the entire flow of the session.

[16] The experimenter was a man and spoke with an L2 accent different from the ones included in the stimuli (see Dautel & Kinzler, 2018; Price et al., 1983 for the importance of experimenter traits in research).

[17] While the participant was on Gorilla, the zoom connection was maintained to allow for communication.

[18] This was heavily emphasized. Participants were told that "parents' opinions are boring," that the experimenter wanted to know "what children think" and "didn't care about what parents thought."

allowed the experimenter to control the flow of tasks. Nonetheless, the experimenter could not hear the stimuli or answer the questions.

The whole session was gamified to enhance engagement. The tasks children were asked to complete were framed as "missions" they had to do to help the experimenter. Each child was told that the computer the experimenter and the professor (a cartoon character introduced to guide participants through the session) used had gone "crazy." Therefore, they needed the child's help to "bring it back to normal." Participants were also informed that each task had a practice session and that they could ask any questions they had to the experimenter. The practice sessions, which were easier versions of the main tasks (see below), were included for two main purposes. First, they helped participants get used to the tasks. Second, they made it possible to see whether the child understood the task. If a child did not complete the practice session according to the instructions, it could be assumed that they would not be able to complete the main task. In this case, their data would not be included in the final analysis of the relevant task (see the description of the tasks below for the results of the training sessions).

Before describing each task, two further remarks about the study design are needed. First, the order of the tasks was kept constant across participants. Participants first completed the intelligibility task, then the verbal-guise task, the categorization task and, lastly, the British Picture Vocabulary Scale. Finally, parents completed the questionnaire. This strict order was motivated by the nature of the tasks themselves. Accent intelligibility has been shown to improve after exposure to just a few sentences of the target accent (Clarke & Garrett, 2004; Maye et al., 2008). Therefore, the intelligibility task was placed first to ensure that participants' responses were unaffected by exposure to the accents occurring in the other tasks. Second, the categorization task was placed after the verbal-guise task because the former actively asked participants to focus on accents (see more details in Section 4.1.3). For the verbal-guise, participants should ideally answer without being explicitly asked to focus on accent differences. Therefore, the verbal-guise was placed second. Finally, the questionnaire was completed last so that children could leave earlier if they wanted to, as it was the guardian who was asked to complete it.

Second, the experimental set-up was framed with a school setting. Participants knew they were going to meet someone from a university. The cartoon character was introduced as "the professor" and they were told that the people speaking were students. This was a conscious choice. Research on language attitudes has shown that context is a key factor in determining what attitudes are evoked (see Section 2.3). When it comes to language, Callan et al. (1983) and Carranza and Ryan (1975) are earlier studies that demonstrated how the context of an interaction affects what attitudes listeners have toward speakers. For instance, Carranza and Ryan (1975) told participants that an interaction had taken place either at home or at school. They reported that the effect of language is stronger in the school context than in the home context. That is, the difference in ratings between the two languages was greater when the participants were

told the speakers were at school than when they were told they were at home. This means that how an experiment is framed for participants can influence their responses.

In this investigation, the school frame was specifically chosen for two reasons. First, children knew they were going to meet someone from a university to do research-related tasks. Framing the tasks in a school setting matched the previous knowledge participants had about the goal of the session. Second, this framing device can better explain why the speakers in our tasks were saying the exact same things (see each task below for more details). Participants could consider it odd that several people having a conversation at, for example, a bus stop would say the *exact* same sentences (which was necessary to isolate the effect of accent on the verbal-guise task). In contrast, it is more natural for students to read the exact same text from a book in class (see Section 4.1.2).

4.1.1 Intelligibility task

This first task investigated what accents children understand best. Participants were asked to listen to a word with background noise. Then, they were presented with 4 images (the target and three distractors; see Section 4.2.2). Participants had to choose which picture matched the word they heard (see Dossey et al., 2020 for a similar design). An example of the set-up is shown in Figure 4.1.

Figure 4.1. Set-up of the intelligibility task
Target word *house*; distractors *mouth, hose* and *blouse.* Images marked with CC0 1.0.

All the words were presented to participants as speech-in-noise with a signal-to-noise ratio of 6 dB (see Dossey et al., 2020). The goal of this added noise was to avoid ceiling effects since the task would have been too simple otherwise. During each trial, participants could listen to the target word as many times as they wanted to. This was done as a scaffolding method (Tabak & Reiser, 2022). Its goal was to reduce the cognitive load of the task and, therefore, make it more accessible to the younger participants.

There were 20 words, each functioning as a trial. These 20 words were divided equally across accents. Four words were included in each of the five accents in the investigation and, within each accent, two words were produced by each of the two speakers of each accent (see Section 4.3). All key variables of the task were randomized. The order of the trials was randomized across participants. Similarly, the two words each speaker read out loud were also randomized across participants. Finally, the placement of the target word and the distractors was randomized for each trial.

In the practice session, the target word was *triangle*, while the distractors were a square, a circle and a star. None of the participants failed the practice session.

4.1.2 Verbal-guise task

The second task focused on the main object of interest of this thesis, that is, how children rate people speaking with different accents. For this purpose, the verbal-guise task was employed (Markel et al., 1967; Tucker & Lambert, 1969; see Dragojevic and Goatley-Soan 2022b for an overview of this technique). Its design is based on the matched-guise test (see Section 2.3 for details). The main difference between these two is that, while the matched-guise uses the same speaker using different accents (or dialects/languages), the verbal-guise technique uses different speakers for each accent. The main benefit of using different speakers is that it yields more authentic guises than the matched-guise counterpart (Preston, 1996). This is especially relevant when the stimuli include more than two accents that are quite different from each other (Dragojevic & Goatley-Soan, 2022b, p. 206).

The task ran in three steps. First, participants were told that they had to help the professor and the experimenter fill out the reports of some students. Then, participants simply listened to five speakers, one of each accent, reading a story about polar bears out loud (see Section 4.2.2). At that point, they were not required to answer any questions. There were two reasons why this section was introduced. On the one hand, it ensured that participants had sufficient exposure to the text to enable them to understand it in all five accents. On the other, listening to all five accents beforehand would make it easier for them to compare accents when answering the different questions. This was expected to lower the cognitive load of the task and to make it more accessible. Afterwards, participants listened to the speakers reading the same text again. Then they were asked to answer six questions about them, one speaker at a time. The

questions included the following three prestige variables and three warmth variables (see Section 2.3):

- SMART: How smart do you think this person is?
- MONEY: How much money does this person have?
- HARDWORKING: How hardworking do you think this person is?
- FRIENDLY: How friendly do you think this person is?
- TRUST: How much can you trust this person?
- FUN: How fun do you think this person is?

These specific variables were chosen based on previous research with adults to represent the typical dimensions found in the literature (Zahn & Hopper, 1985; see also Section 3.3.3 for the variables used with young participants, in particular Dossey et al. 2020; Kinzler and DeJesus 2013b). The MONEY question may require further commenting. In England, accents are closely associated with SES (see Section 2.1 and Section 4.2.1). Therefore, it was considered important to include a variable that enabled investigating when children begin associating accent with SES. Of course, how children understand SES is different from how adults do. Therefore, the specific wording of the question was key. It needed, on the one hand, to make children understand what the variable referred to and, on the other, capture the desired meaning. The literature on children's perceptions of SES has established that they view SES as closely related to how much money people have (Kustatscher, 2017; Ridge, 2002; Sutton et al., 2007; Weinger, 2000). As a consequence, the phrasing *How much money does this person have?* was chosen. It was considered simple enough and to capture, somewhat indirectly, perceived SES.

The structure of each trial is described below. An example of the set-up for the questions is shown in Figure 4.2. To answer the questions, participants were given a four-item image-and-text scale, as they have been shown to work best in experiments with children (Chambers & Johnston, 2002; Mellor & Moore, 2014; Royeen, 1985). Participants were able to listen to the speakers as often as they wanted to lower the demands of the task and allow participants to focus on the answer. Which speaker of the two speakers from each accent participants listened to (see Section 4.3) was randomized across participants, as was the order of the questions.

The conceptual structure of the questions and answers was kept constant across questions. For the questions, the relevant adjective was impartial regarding the degree of its scale. For example, *how smart is this person?* allows for answers that indicate that the person is smart or not. In contrast, *how dumb is this person?* implies that the person is dumb to some degree (see Croft & Cruse, 2004, pp. 175-176 for a discussion of impartiality and committedness with regards to antonyms). For the answers, the option including *very* always denoted the highest value of the relevant property (e.g. *very smart* refers to a high level of SMARTNESS, while *very dumb* refs to low levels of

SMARTNESS) (see Croft & Cruse, 2004, pp. 173-174 for a discussion of how degree modifiers interact with the meaning of antonyms). These measures were included to reduce the cognitive complexity of the task.

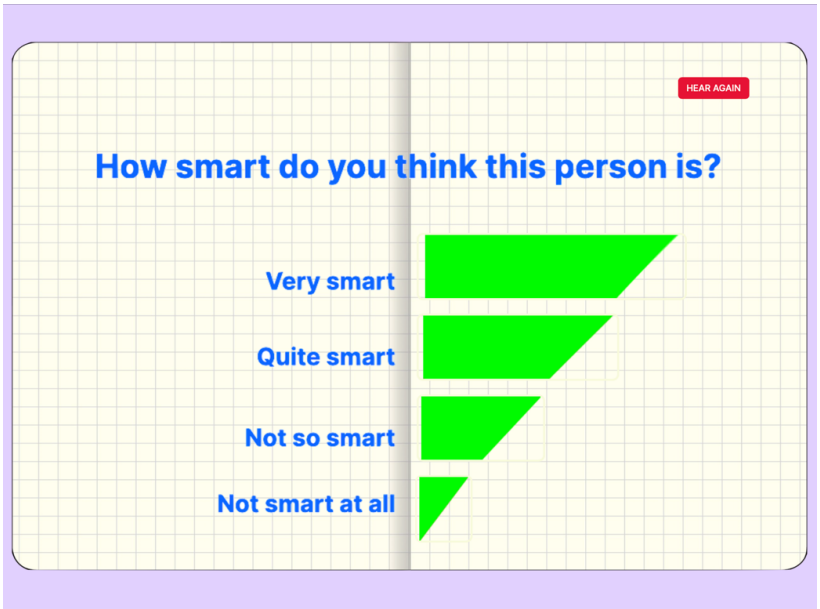


Figure 4.2. Set-up of the verbal-guise task
Background image marked with CC0 1.0.

With regard to the practice session, its set-up was identical to that of the actual task shown in Figure 4.2. However, instead of accent, an easier variable was used. The selected variable was age, since young children are able to detect age from voice quality (Creel & Jimenez, 2012; Levi, 2021; Spence et al., 2002). Participants listened to four speakers. These four speakers had significant age differences. Two of them were women above the age of 70 years and two of them girls below the age of 15 years. All four speakers spoke with a standard British accent. The recordings were obtained from The Speech Accent Archive (http://accent.gmu.edu). The first two sentences of each recording were used: *Please call Stella. Ask her to bring these things with her from the store*. The question asked to participants was *how old do you think this person is?* and the answers followed the same structure as in Figure 4.2 replacing *smart* with *old*.

Participants passed the practice session if the older speakers received higher scores than the younger speakers. Participants failed the practice session in two possible cases. First, they failed if they gave the same rating to at least one of the old speakers and one of the young speakers. For instance, if they rated both an old and a young speaker as "quite old," they failed the practice session. Second, participants failed the task if they

gave higher ratings to at least one of the younger speakers than to one of the older speakers. In these two cases, participants still moved on to the task. However, their answers were not included in the final analysis[19], since it was not possible to establish that they understood the task. Nine participants from London and eleven from Plymouth did not pass this training session.

4.1.3 Categorization task

The third task focused on how children categorize speakers of different accents. Participants were told that the professor and the experimenter did not "know where the students were born." Therefore, the participants' task was to group the students who "were born in the same city." This specific phrasing was chosen based on Weatherhead, White, et al. (2018). In their study, they showed that children did not necessarily think people speaking accents different from their own *were* from different places. Nevertheless, children consistently think speakers of accents different from theirs were *born* in different places.

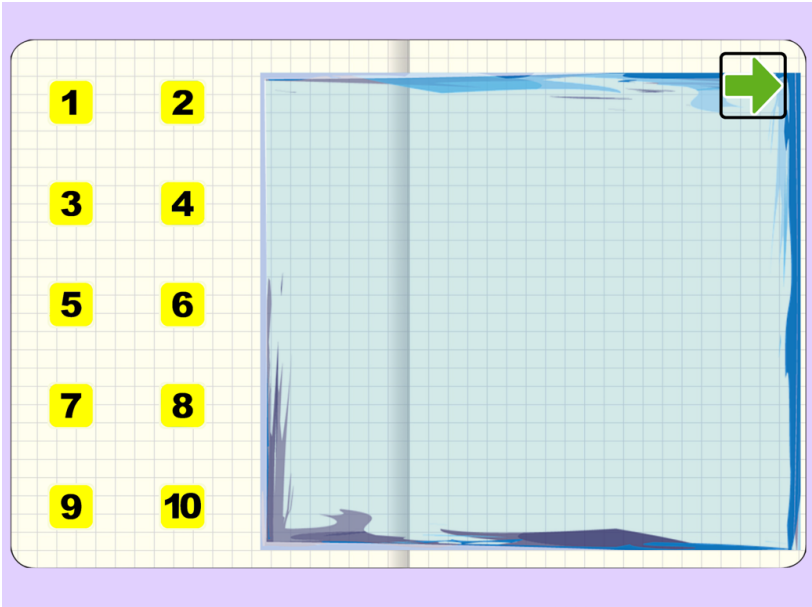


Figure 4.3. Set-up of the categorization task
Images marked with CC0 1.0.

[19] Their results from the other tasks were included if they passed their respective training sessions; see Sections 4.1.1 and 4.1.3.

The overall design of the task was based on Jones et al. (2017) and Dossey et al. (2020). The design is shown in Figure 4.3. Participants viewed ten numbers. When clicking on each of them, they heard one speaker reading the first two sentences of the polar bear story (see Section 4.2.2). There were two speakers for each of the five accents (see Section 4.3). The order of the numbers was constant across participants (in numeral order). Nevertheless, which speaker was associated with which number was randomized across participants. Participants could move the numbers with the mouse to the blue square on the right side of the screen and group them as they wanted. If the younger participants had difficulty using the mouse, parents were allowed to move the mouse for them. In such cases, it was repeated and emphasized that it was the child who should choose how to group the speakers. Finally, participants were explicitly instructed that they could make as many or as few groups as they wished and that the groups could be as big or small as they wished.

After grouping the speakers as they wished, participants were asked to state how many groups they made (see Figure 4.4). Then, a number of houses equal to the number of groups they said they made appeared on the screen. Participants were asked to move each group of speakers into one house. The purpose of this last step was to avoid confusion regarding how many groups participants made and which speakers were included in which groups.

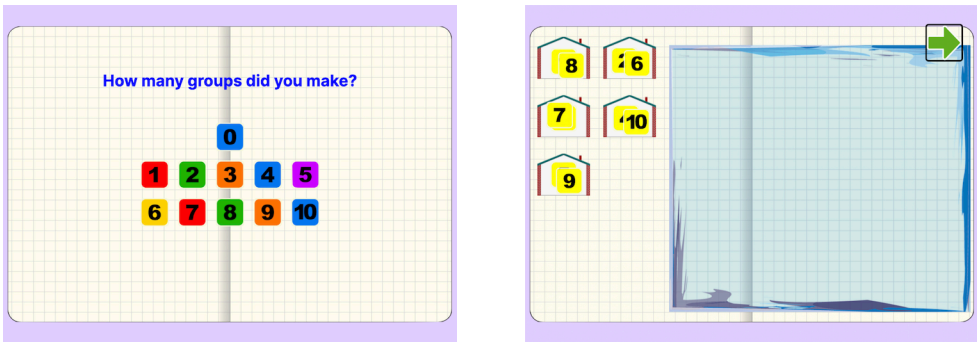


Figure 4.4. Confirmation of groups in the categorization task
Question asking participants how many groups they made (left) and fabricated example showing five houses after a participant answered they made five groups (right).

The practice session followed a structure similar to that of the actual task. Instead of ten numbers, they were shown four. Moreover, it did not include different accents. Instead, the stimuli included different languages. Two speakers spoke with a southern British accent and two speakers spoke Danish. The English speakers were selected from The Speech Accent Archive (http://accent.gmu.edu) and said *Please call Stella. Ask her to bring these things with her from the store*. The Danish speakers read the following text: *der lever mange krabber i havet. De små lever tæt på land. De store lever på dybt vand*

(Holm, 2018) (*There are many crabs in the ocean. The little ones live by the shore. The big ones live in deep-sea*, my translation). Since none of the participants were expected to speak Danish[20], it was not considered necessary to have both texts say the same thing. All speakers were women between the ages of 25 and 35 years. Participants were judged as passing the practice session as long as they did not put a Danish and an English speaker in the same group. Only one participant from Plymouth failed the practice session. Therefore, their data was not included in the final analysis of the categorization task[21].

4.1.4 British Picture Vocabulary Scale task

The final task participants completed was the third edition of the British Picture Vocabulary Scale (BPVS) (Dunn et al., 2009)[22]. This task was included as an indirect measure of linguistic development. Vocabulary development has been shown to correlate with linguistic development, particularly phonological development (Kehoe et al., 2018; Smith et al., 2006; Stoel-Gammon, 2011; Viterbori et al., 2018). Since the task is easy to administer (particularly online) and time-effective, it was chosen over more direct measures. The overall design of the task is similar to the intelligibility task in Section 4.1.1. Participants were shown four pictures. Then, they heard a word (spoken in a Standard Southern British English) and chose which picture matched the word they heard. Participants responded by naming the number of the corresponding picture (one through four). The task was administered online on Zoom via a high-definition webcam. Finally, the practice session comprised the practice slides from the BPVS itself. All participants passed the practice session.

4.1.5 Parent survey

The final component of the battery of tests was a survey to be completed by the participants' guardian. Its goal and structure can be divided into two parts. First, it enabled gathering of quantitative information on the child's exposure to different accents. Second, it collected general demographic information about the child and their legal guardians.

With regard to the first section, its overall design was inspired by the *BabyLab Language Exposure Questionnaire* (Cattani et al., 2014). First, parents were asked how

[20] Figure 4.10 in Section 4.4 shows that none of them did.

[21] Just as in the verbal-guise task, their data was included in the other two tasks if they passed their respective practice sessions. See Sections 4.1.1 and 4.1.2. The participant who failed the practice session in the categorization task did pass the practice session of the verbal-guise and the intelligibility tasks.

[22] The results of the BPVS are reported in Section 4.4 below and not in the result chapters.

much their child was exposed to the accents included in the stimuli (see Section 4.2). Parents had the option to listen to a speaker of each accent as an example of the relevant accent. After that, parents had the option to include other accents the child may be in regular contact with.

When inquiring about exposure, only bidirectional and regular contact with other speakers was included (as in Paquette-Smith et al., 2019). This included, for instance, relatives, friends, neighbors, etc. Sources of linguistic variation such as traditional media, videogames, social media, etc., were not included. This should not be understood as meaning that this type of input does not have an influence on the transmission of attitudes. Research has established that media affects our attitudinal responses (see e.g. Goorimoorthee et al., 2019 for language attitudes in videogames; Lippi-Green, 2012 for language attitudes in traditional media; and Thøgersen & Pharao, 2020 for the complex relationship between language attitudes and media in Denmark). This is perhaps even more relevant during the past decade due to the unavoidable presence of social and digital media in children's lives (e.g. Bozzola et al., 2022; Christakis & Hale, 2025). It would thus be naïve to suggest that children's accent stereotypes are not the result of, at least in part, exposure to linguistic ideologies through media.

The choice to limit the measurement of linguistic exposure to bidirectional and regular contact was made on practical grounds. How many accents children hear through media is more difficult to quantify. On the one hand, with short-form media (such as TikTok, Instagram Reels, YouTube Shorts, etc.), children may be exposed to a specific accent in 5- to 10-second spans. This makes calculating exposure time more complex for parents. On the other hand, a significant portion of children's media use is likely to be unsupervised. This makes it impossible for parents to know what accents children are listening to and for how long.

The second section of the questionnaire aimed at gathering demographic data. With regard to the child, this included linguistic variables such as languages spoken by the child, diagnoses of neuropsychiatric conditions and/or speech and language impairments[23] as well as whether the child had previously lived in other places. Some information on the guardians was also collected. Specifically, they were asked what languages they spoke to their child and their level of education, using the scale in Cattani et al. (2014)

[23] Only official diagnoses were considered.

4.2 Stimuli

In a study investigating how children develop accent stereotypes, it is of theoretical importance what specific accents are included and what texts are used as stimuli. For example, Weatherhead et al. (2019) demonstrated that children can categorize L2 accents from L1 accents if the stimuli have been low-pass filtered, but not if the stimuli are unaltered (see Section 3.2). Therefore, this section outlines what stimuli were chosen and why. First, it describes what accents were included in the investigation as well as why these specific ones were chosen (Section 4.2.1). Section 4.2.2 deals with what specific texts and words were included in each task.

4.2.1 Accents

The selection of accents for this investigation was based on two primary parameters. First, the overall number of accents was a significant concern. On the one hand, special focus was given to the length of the experiment. As mentioned at the beginning of this chapter, the design must be suitable for the youngest participants (7-year-olds; see Section 4.4). Therefore, too many different accents should not be included, as they necessarily make the experiment longer. On the other hand, not all accents evoke the same stereotypical reactions (Ball, 1983; Nesdale & Rooney, 1996; see also Section 2.3). Therefore, selecting, for instance, only two accents may result in too narrow a picture. It may show how children develop stereotypes toward these two accents, but it may not increase our understanding of the general patterns behind this process.

Second, the RQs in Section 1.1 also constrained what accents should be included. RQ1 (whether children exposed to different degrees of linguistic diversity develop accent stereotypes in a different way) requires that at least two different L1 accents be included, namely, those local to each city. Similarly, RQ2 (whether children develop stereotypes toward L2 accents at an earlier age than toward L1 accents) necessitates the inclusion of both L1 and L2 accents.

With these considerations in mind, five accents were selected: the Plymouth accent, Multicultural London English (MLE), Standard Southern British English (SSBE), Chinese-accented English and French-accented English. The Plymouth and MLE accents were included as the local accents for Plymouth and London, respectively, and SSBE was included due to its role as the standard accent in England (see below).

In addition, two L2 accents were included. Technically, only one L2 accent was necessary. Nonetheless, if only one had been chosen and its developmental trajectory differed from that of the L1 accents, it could be the case that the difference was not due to its L2 nature, but simply to the specific stereotypes this particular L2 accent evokes. By including two L2 accents, it is possible to differentiate between patterns specific to one L2 accent and more general L2-versus-L1 patterns (see Weatherhead et al., 2019,

p. 3 for similar arguments). Chinese- and French-accented English were chosen because they represent opposite ends of the spectrum in terms of prestige. French-accented English is typically rated positively by native English speakers[24]. In contrast, Chinese-accented English is rated negatively (Bishop et al., 2005; Coupland & Bishop, 2007; Magne, 2019; Sharma et al., 2022). This matches the contrast between SSBE, on the one hand, which is positively rated in England (Bishop et al., 2005; Cole, 2021; Coupland & Bishop, 2007; Levon et al., 2022; Levon et al., 2021; Sharma et al., 2022), and the Plymouth and MLE accents, on the other, which are negatively rated (Braber et al., 2024; Kircher & Fox, 2019; Levon et al., 2021).

In the remaining part of this section, the main distinguishing features of each accent, both phonetic[25] and/or sociocultural, are described. SSBE is discussed first, followed by MLE, the Plymouth accent, Chinese-accented English and, finally, French-accented English. As discussed in Section 2.1, it is crucial to keep in mind that accents are not clear-cut linguistic entities. Therefore, the following description focuses on some main features of each accent. This should not be understood as meaning that the speech of all speakers of each accent is characterized by these features or that these are the only features differentiating them.

An examination of language attitudes in England is perhaps not complete without including the *de facto* standard accent in England, Received Pronunciation (RP)[26]. Why SSBE is included then needs clarification. RP has a long history (see Mugglestone, 2007 for an overview). Nevertheless, its influence in linguistics can be argued to emerge in the early- and mid-twentieth century, when the first phonetic descriptions of it were made (Jones, 1909, 1917). Back then, RP was labelled Public School Pronunciation (PSP[27]) (Jones, 1917, p. viii). This is because of its origins as an accent used by upper-class speakers from public schools in and around London (Holmes-Elliott & Levon, 2024). This is why, despite it being labeled as a non-local accent (Agha, 2003; Mugglestone, 2007; Wells, 1982), it is strongly associated with the Southeast of England.

Despite being chosen to be the standard accent in England, linguists in the second half of the twentieth century already highlighted that different versions of RP exist. Gimson (1970) noted that three different varieties of RP could be distinguished. First, Conservative RP is the accent spoken by the oldest generations and associated with the highest end of the SES scale. Second, General RP describes the accent spoken by the

[24] Furthermore, French-accented English has been frequently included as the L2 accent in studies investigating the development of language attitudes; see Chapter 3.

[25] The description of the accents below uses the standard lexical sets of Wells (1982, pp. 127-167)

[26] This accent is sometimes also labeled the Queen's (now King's) English or BBC English. For non-linguists, it is also sometimes labeled the standard, neutral, normal, or correct accent in England.

[27] In a British context, the term *public school* refers, in fact, to private, fee-paying boarding schools for upper-class students.

BBC. Finally, Advanced RP is the version of RP spoken by younger people. Cruttenden (2014), in his revision of Gimson's work, uses Modern RP (or General British, GB; a term parallel to General American) to represent the changes RP underwent in the latter half of the twentieth century. Moreover, he also differentiates between Conspicuous General British, the commonly considered posh accent from upper-class families, and Regional General British, GB with "inclusion[s] of regional markers" (Cruttenden, 2014, p. 81). Wells (1982) offered a comparable typology of RP accents. He proposes U-RP for the posh, upper-class version of RP, Adoptive RP, the accent of those who did not grow up with RP as their native accent, Mainstream RP, the accent of upper-middle class and media, and, finally, Near RP, "accents which are not exactly RP though not very different from it" Wells (1982, p. 279). Finally, the volatile nature of RP's linguistic features is even reflected in the fact that the speech of Queen Elizabeth the Second changed with time (Harrington et al., 2000). Thus, not even the *Queen*'s English is an unmistakable term.

Use of the term RP is therefore not unproblematic. From a linguistic point of view, it is often used without a clear description of which linguistic variety it is meant to cover or used to only stand for the old, posh, upper-class accent (Cole & Strycharczuk, 2024; Cruttenden, 2014, p. 80). From the point of view of language users, the term RP is associated with snobbery, old speakers and even ridicule (Fabricius, 2018; Hughes et al., 2012; Lindsey, 2019). This has led a number of researchers to propose that the term RP (or modified versions of it) should be discarded (e.g. Cole & Strycharczuk, 2024). In its place, the term SSBE has gained popularity to label the "generational successor to RP" (Fabricius, 2018, p. 35). It highlights the fact that it is historically, linguistically and socially associated with the south of England (just as RP was) and that it is a different variety from the traditional, old-fashioned RP.

There are numerous linguistic features that differentiate SSBE (Fabricius, 2018; Lindsey, 2019). As the name indicates, it is phonetically a southern British accent: it is non-rhotic and uses /ʌ/ in STRUT and /ɑː/ in BATH. In terms of its vowels, its repertoire is not different from that of RP. However, how these vowels are realized phonetically has changed. One of these changes has been named the *anti-clockwise vowel shift* (Fabricius, 2019; Lindsey, 2019, pp. 17-22). It describes a process in which front vowels are lowered. Low vowels are realized more back. Back vowels are raised and, finally, back, high vowels are fronted. Two examples are given here. First, the high vowel in FLEECE is not realized as [iː], but as [ɪj], with a lower starting point of the diphthong. Second, the vowel in GOOSE is realized with a fronted and dipthongized [ʉw] rather than the traditional, back [uː]. A further change is the tensing of word-final [ɪ], which is now realized as [iː] in words like *happy*, a process labelled as *happY-tensing* (Fabricius, 2002; Lindsey, 2019; Wells, 1982). A final example of vowel changes includes the lowering of [ɪ] to [ə] in unstressed closed syllables like *item* or *civil* (Fabricius, 2002).

In comparison, there are fewer differences involving consonants. Nonetheless, they do exist. Two frequent contrasts involve the following. First, t-glottaling is present in coda position before a consonant (*fi*[ʔ]*ness* rather than *fi*[t]*ness*) or in word final position before a consonant [*par*[ʔ] *time* instead of *par*[t] *time*] (Badia Barrera 2015). Second, yod coalescence is a common process. In it, the phones [tj], [dj] are realized as [tʃ], [dʒ]. Therefore, words such as *tuna* and *duty* are pronounced as [tʃʉwnə] and [dʒʉwti], respectively (Hannisdal, 2006).

The next accent discussed is the L1 accent MLE. MLE is a multiethnolect (Cheshire et al., 2011)[28]. Multiethnolects are linguistic varieties that emerged at the end of the twentieth century in multilingual, working-class neighborhoods around urban centers, particularly in north-western Europe (Clyne, 2000). MLE is therefore comparable to the multiethnolects in Denmark (Zachariassen, 2022), Sweden (Kotsinas, 1988), Norway (Cutler & Røyneland, 2015), Germany (Wiese, 2009) or the Netherlands (Nortier, 2001). The origins of MLE can be traced back to the 1980s (Kerswill, 2014), while the term was first coined in 2006 (Cheshire et al., 2011).

Its linguistic influences span from varieties of English from Africa, the Caribbean and South Asia to Cockney, Jamaican Creole and L2 accents of English (Kerswill, 2014, p. 432). As a consequence, MLE has sometimes been described as the result of London children growing up in a linguistic environment with a dense feature pool (in the sense of Mufwene, 2004), that is, substantial linguistic diversity (Kircher & Fox, 2019). In terms of its speakers, MLE is historically the accent of the working class of east London (Cheshire et al., 2008). Importantly, despite its origins, MLE is an accent currently used by speakers of different ethnicities (Cheshire et al., 2011; Kircher & Fox, 2019). In fact, Fox (2015) posited that MLE will replace Cockney as the accent traditionally associated with London.

From a strictly linguistic point of view, MLE can be characterized as follows. Its diphthongs have been subject to substantial research. The vowel in GOAT is realized with a higher starting point as [oʊ] (Kerswill, 2014), so that it has a shorter trajectory. A similar pattern is found in PRICE, with a narrower [ɐɪ], and FACE, with [eɪ̯] (Kerswill et al., 2008). The diphthong in MOUTH has the back starting point [ɑʊ]. In terms of its monophthongs, its main characteristic is the "extreme fronting" of GOOSE, more so than in SSBE (Cheshire et al., 2011, p. 158). When it comes to consonants, four features are frequently said to characterize MLE (Cheshire et al., 2011; Kerswill, 2014). First, MLE is subject to L-vocalization, wherein the so-called dark L [ɫ] converts into a back vowel. Therefore, *milk* is pronounced [mɪɤk][29]. Second, t-glottaling is much more pervasive than in SSBE, also present intervocalically. Next, th-fronting also characterizes MLE, a process by which [θ] and [ð] are realized as [f] and

[28] The term *multiethnolect* is not without its critics; see e.g. Rampton (2011).

[29] The exact quality of the back vowel is subject to variation. It could also be [ʊ] or [o] (Wells, 1982, p. 258).

[v], respectively (Wells, 1982, p. 328). Thus, *thin* is pronounced [fɪn]. The last consonantal feature that characterizes MLE is k-backing. In this case, /k/ is realized as [q]. Finally, MLE has also been described as having a more syllable-timed rhythm, as opposed to the typical stress-timed one of standard English (Torgersen & Szakay, 2012).

The last L1 accent in this investigation is the Plymouth accent, popularly also known as Janner. This accent falls under the group of English varieties known as West Country English. Wells (1982) categorized the Plymouth accent under the Center subregion within the West Country, while Trudgill (1999) included it in a Lower Southwest category. In contrast to MLE and SSBE, the Plymouth accent has been less extensively described (but see Altendorf & Watt, 2008 for an overview). However, some typical characteristics can be identified. In terms of its vowels, short vowels are often lengthened, so that *did* is pronounced [dɪːd] or [dɪˑd] (Wells, 1982, p. 345). The diphthongs in FACE and GOAT can be realized as the monophthongs [eː] and [oː], respectively (Hughes et al., 2012). Furthermore, Plymouth speakers don't usually realize the distinction between TRAP-BATH typical of Southern English accents (Altendorf & Watt, 2008; Wells, 1982). Furthermore, PALM words often retain the [ɬ] sound and are realized with the [ɑ] vowel (Altendorf & Watt, 2008; Wells, 1982).

Its consonant-related features are discussed below. One of the most distinctive features of the Plymouth accent is its rhoticity. In contrast to the Southeastern accents, Plymouth speakers realize post-vocalic /r/. It is often realized as a retroflex [ɻ] (Wells, 1982), similar to the *r* in North American varieties of English. Next, a classic feature involves the change from voiceless [f], [θ], [s] and [ʃ] to their voiced counterparts [v], [ð], [z] and [ʒ]. This feature has been described as the Southwest's "feature *par excellence*" (Wakelin, 1986, p. 29). Nonetheless, it is recessive nowadays (Altendorf & Watt, 2008, p. 219). Finally, when other L1 accents of English would have a syllabic consonant in the environment of an [ə] followed by a nasal consonant, the Plymouth accent often pronounces the vowel fully, so that *happen* is pronounced *happ*[ən] and not *happ*[n̩] (Altendorf & Watt, 2008).

The last two varieties to be described are the L2 accents Chinese-accented English and French-accented English. A proper depiction of their characteristics is more troublesome when compared to the L1 accents in this study. As mentioned at the beginning of this section, accents are not homogenous linguistic entities. Of course, this also applies to L2 accents. Nonetheless, L2 accents are affected by different factors that do not necessarily apply to L1 accents. Specifically, L2 accents are subject to transfer effects from the speaker's native language. This applies to segmental and suprasegmental features as well to phonotactics. For example, Chinese-accented English can sometimes be characterized by a reduction of consonant clusters. This is the result of the differences in syllable structure between English and Mandarin (e.g. Neergaard

& Huang, 2022). While the maximum syllable size in English corresponds to CCCVCCCC[30] in *strengths*, the maximal syllable in Mandarin is CGVX[31].

The specific L1 accent a speaker has in their native language will also affect how this transfer will take place. As a result, it is not a straightforward matter what specific features of the native language will transfer to an L2 accent. For example, a French speaker from France may pronounce the English word *hair* differently than a French speaker from the Acadia region in Canada. The former may not pronounce the initial /h/, realizing it as /ɛʁ/, due to the lack of this phoneme in most varieties of French in France. In contrast, this phoneme is present in Acadian French (Cichocki, 2012, p. 221). A speaker from this region may thus pronounce the word as /hɛʁ/. Finally, an L2 speaker's proficiency will also affect how a given L2 accent is produced. The higher a speaker's proficiency in a foreign language, the closer their L2 will be to an L1.

Nonetheless, a few frequent features of Chinese- and French-accented English are mentioned below. First, Chinese-accented English often shows a reduction of the number of monophthongs compared to L1 accents of English. This is because Chinese has a smaller inventory of vocalic phonemes (Chen et al., 2001). For example, this may lead to the lack of a clear distinction between /ɒ/, /ɑ/ and /ɔː/ (Cruttenden, 2014, p. 127). Furthermore, the vowels in Chinese-accented English may be realized with different tones as a result of its tonal system (Zhang et al., 2008). In terms of consonant features, speakers may over-aspirate /p/, /t/ and /k/ in onset position (Cruttenden, 2014, p. 166), realize /r/ as [l] or [w] (Chan & Li, 2000) and produce [s] for both /z/ and /ʃ/ (Cruttenden, 2014, pp. 203-206). Finally, this accent is often characterized by its syllable-timed rhythm (Cruttenden, 2014, p. 276).

Second, some potential features of French-accented English include the following. The vowel in KIT is often realized as [i] (Cruttenden, 2014, p. 115) and the vowel in NURSE as [ø]. From a consonantal point of view, French-accented English frequently drops initial /h/ (Cruttenden, 2014, p. 27), only produces clear /l/ (Cruttenden, 2014, p. 222), /k/ and /g/ can be over-palatized before and after front vowels (Cruttenden, 2014, p. 182), /t/ and /d/ are dental, rather than alveolar (Cruttenden, 2014, p. 179) and /r/ is realized with its uvular counterpart [ʁ] (Cruttenden, 2014, p. 227). Finally, two of its potential suprasegmental features include its syllable-timed rhythm (Cruttenden, 2014, p. 276) as well as having stress on the final full syllable of each word (Fagyal et al., 2006).

[30] C stands for consonant and V for vowel.

[31] C refers to the initial consonant, G to a pre-nuclear glide, V a vowel and X to a post-nuclear glide or consonant.

4.2.2 Selection of texts and words

Two parameters determined which specific words and texts were to be used as stimuli. First, the accents used as stimuli served as the basis for selecting the specific words and texts in the tasks for this thesis. This was done to ensure that the stimuli in each task highlighted the different phonetic and phonological properties of each accent. Second, the words selected for the stimuli were required to be age appropriate. Thus, participants' answers should not be the result of their not knowing certain words. Therefore, all words had an average age of acquisition at or below the age of 7 years in the Age of Acquisition database (Kuperman et al., 2012). Finally, all stimuli in the study were levelled to an average intensity of 65 dB to ensure a comfortable experience.

The stimuli used in this investigation can be divided into two categories. First, a list of words was selected for the intelligibility task (Section 4.1.1). This list can be seen in Table 4.1. The term *target word* refers to the word participants heard that was produced by speakers and constituted the correct answer. The distractors were selected based on phonetic similarities to the relevant target words. They were meant to be similar to the target word and employed to avoid ceiling effects. The speakers of each accent were recorded reading each target word out loud (see Section 4.3 for more details).

Table 4.1. Intelligibility task, target words and distractors

Target word	Distractor 1	Distractor 2	Distractor 3
alligator	escalator	skater	calculator
bird	bread	girl	shirt
bowl	ball	doll	soap
branch	broom	lunch	plant
church	bench	chair	search
face	fence	skate	space
farm	arm	fire	lamb
foot	flute	food	hook
goat	boat	ghost	toast
goose	boot	fruit	juice
house	blouse	mouse	mouth
key	knee	sea	tree
leaf	bean	teeth	thief
mouse	blouse	house	mouth
palm	bomb	mom	park
porridge	orange	sausage	cabbage
salt	aunt	salad	sock
thorn	horn	thumb	corn
toy	boy	coin	toe
wolf	elf	roof	well

Second, the story in Text 4.1 was created to be used in both the verbal-guise (Section 4.1.2) and the categorization tasks (Section 4.1.3). The speakers recorded the whole

text. However, participants did not listen to the greyed-out sentences, only to the text in black. The purpose of recording more text than necessary was to make the target story sound more natural and as though it were part of a longer passage (cf. Dossey et al., 2020). As was the case with the word list above, all speakers were recorded reading the story out loud.

Text 4.1. Polar bear story
Story used as stimuli for the verbal-guise task and the categorization task

> There are many animals that live in the artic. One of them is the polar bear. Polar bears need milk to feed their beautiful cubs. To produce their milk, they look for food underwater at night-time. They are able to catch intelligent animals like whales thanks to the force of their paws and mouth. Then, polar bears take the whales home and eat them. Seals can also be found in the artic.

Text 4.1 was used for both the verbal-guise and categorization tasks in order to ensure that their results were comparable. Employing different stimuli would necessarily entail the presence of different accent features. This could potentially lead to differences in results that would not inform the study's RQs. For instance, it could be the case that children did not show accent stereotypes using text A, but they were able to categorize the same speakers using text B. This may simply reflect that one of the texts made the differences across accents more noticeable. Therefore, by using the same text, it can be expected that children will respond to the same linguistic features in both tasks, thus enabling a more direct comparison of the results.

For the verbal-guise task, participants listened to the whole story in black text. In contrast, only the two first sentences (from *polar bears* until *night-time*) were included in the categorization task. This decision was based on time concerns. As described at the beginning of this chapter, a vital aspect of the design was that the youngest participants should be able to complete the session without too much cognitive pressure. Furthermore, participants were already familiar with the speakers as well as the story, since the categorization task always followed the verbal-guise task (see Section 4.1). Therefore, using only half the story would result in a more comfortable experience for participants.

4.3 Choice of speakers: accent strength

This section describes the preparatory study conducted to select speakers of each accent. This study was based on Floccia, Butler, Girard, et al. (2009) and Creel (2018). Its rationale was as follows. As indicated in Section 2.1, accents are not homogenous entities. No two speakers of the same accent speak the exact same way. One of the

sources of this variability is accent strength, that is, the degree to which an accent differs from a given baseline accent (see Section 2.1). Accent strength has important consequences for how adults rate speakers in social judgement tasks as well as for how well children categorize them (see Section 3.5). As a consequence, it was crucial to ensure that speakers of each accent had similar levels of accent strength.

4.3.1 Methods

It has long been noted that ratings of accent strength are highly subjective (Giles, 1972; Grondelaers et al., 2018). For example, a speaker with a southern American accent may consider a speaker of northern Irish to have a strong accent. In contrast, a speaker from the south of Ireland may consider the northern Irish speaker to have a milder accent than the American speaker. Therefore, the accent strength ratings in this study were collected from adult populations from the same cities from which the participants in the developmental study originate: Plymouth (in Southwest England) and London (see Section 4.4).

Participants for this preparatory study were recruited via Prolific (www.prolific.com). The requirements for participating included the following. First, participants had to be born and raised in the relevant region (Greater London for London and the Southwest of England for Plymouth). Second, participants had to live in the respective area at the time of the data collection and to not have lived abroad for longer than 6 months. Third, English had to be their native language, and they could not have language-related disorders. As compensation for their time, those who completed the study were given £5. Data from 30 participants in each area was collected. The London participants had a mean age of 37.90 years (*SD* = 11.9;0 range = 18–67; 17 women). Participants from Plymouth had an average age of 36.37 years (*SD* = 10.89; range = 21–58; 19 women).

The stimulus was the text in black from Text 4.1[32]. For each accent, five speakers were recruited. Speakers of the Plymouth accent were recruited and recorded in Plymouth, while the rest of the speakers (SSBE, MLE, Chinese- and French-accented English) were recruited and recorded in London. In both places, the recordings were made in Audacity® with a sampling frequency of 44100 Hertz and monaural sound in a soundproof booth.

At the recording session, the speakers were asked to read the story as if speaking to a friend to ensure the presence of the desired accent features. The context of the recording (in a soundproof booth, at a university, for a research study) is relatively formal. It could be the case that speakers would produce more standard forms (see Labov, 2002 for the

[32] For this preparatory study, only the black font text in Text 4.1 was used as stimulus. However, every speaker recorded the whole text (grey and black sections) and the entire list of words in case they were selected as the speakers for the developmental study.

relationship between style and accent). Therefore, the speakers were instructed to read the text in a way that ensured that the accent features were present. The speakers were all women between the ages of 20 and 35 years. Only speakers of one gender were chosen to accommodate the study to young participants. Gender is known to have an impact on language attitudes (Dennhag et al., 2019). Since the length of the final study was a key factor, it was not possible to include both male and female speakers to control for potential gender effects. Speakers were given £10 gift-cards as compensation for their time.

Recruitment of speakers was carried out based on demographic factors. First, they could not have any speech disorders. The L1 speakers had to have English as their native language. For the Plymouth accent, speakers had to be born and raised in Plymouth, and they had to self-report having a prominent Plymouth accent. For MLE, speakers had to be born and raised in East London and have a "London youth accent"[33]. The SSBE speakers had to be born and raised in the south of England and have a standard southern accent. Recruitment was slightly different for the L2 accents. These speakers had to be born and raised in the respective country (France for French-accented English and China for Chinese-accented English). They had to report their native language to be French and Chinese, respectively. Finally, they had to have learned English as a foreign language.

This preparatory study was created with PsychoPy Builder (Peirce et al., 2019) and run online via Pavlovia (https://pavlovia.org). Each participant listened to 15 speakers: 3 randomly selected speakers per accent. This reduction of speakers was done to keep the study short while still including the same number of speakers per accent. First, participants listened to all 15 speakers, one at a time, without providing any answers. This was done to make participants aware of the range of accents they had to rate. Afterwards, participants listened to each accent once more and were asked two questions. First, they were asked to rate how strong the accent of the speaker was on a 7-point scale. Option 1 was "no accent at all" and 7 "very strong accent." Second, they were asked to name where they thought the speaker came from with an open textbox. This question was included to ensure that the speakers were recognizable as speakers of the respective L1 and L2 accents. When answering these two questions, participants were allowed to listen to the recording of the relevant speaker as often as they wanted to.

4.3.2 Results

This section reports the results of the preparatory study. First, general analyses are described. These were carried out to ensure the quality and comparability of the data

[33] The term MLE was not used for recruitment since it is an academic term not widespread among non-linguists.

across cities. Then, the results of the accent ratings are introduced. Finally, the text answers from the second question are dealt with.

Which specific speakers participants listened to was randomly selected, as described in Section 4.3.1. Since each participant listened to three speakers of each accent, each individual speaker had a 60% chance of being chosen. This means that, on average, each individual speaker should have been assessed approximately 36 times. Each speaker was selected as stimulus on average 35.40 times (SD = 3.62; range= 29–42). This means that each speaker was rated a sufficient number of times.

Next, the accent strength ratings were compared across locations, genders and age to ensure comparability. With regard to city, an ANOVA on a linear model (see Section 4.5) with accent ratings as the dependent variable and city and accent as predictors reveals no effect of city ($F(1,835) = 2.56$, $p = .109$) or interaction between city and accent ($F(24,835) = 0.52$, $p = .974$). Similar results were found with regard to gender. There is no relationship between participants' ratings and their gender ($F(1,835) = 2.08$, $p = .155$) nor an interaction between gender and accent ($F(24, 835) = 1.22$, $p = .213$). Finally, no age ($F(1,835) = 3.04$, $p = .081$)[34] or age–accent interactions ($F(24,835) = 1.11$, $p = .323$) were found. This means that participants gave comparable ratings of accent strength regardless of their age.

The mean ratings for each speaker and their standard deviations can be seen in Table 4.2. Pairwise comparisons (see Section 4.5) reveal three main clusters. First, the SSBE accent (M = 2.66; SD = 1.43) was rated significantly lower than all other accents ($t_{ssb\text{-}chi}(880) = 23.22$, $p < .001$; $t_{ssb\text{-}fre}(880) = 20.58$, $p < .001$; $t_{ssb\text{-}mle}(880) = 9.85$, $p < .001$; $t_{ssb\text{-}ply}(880) = 10.06$, $p < .001$). This accords with the view that SSBE is the standard accent and, therefore, "accent-less" (see Section 2.1 and Section 4.2.1). Second, the two L1 accents (Plymouth, M = 4.00; SD = 1.45, and MLE, M = 3.97; SD = 1.45) do not differ from each other ($t(880) = -0.21$, $p = .999$). They differ from the two L2 accents, however, ($t_{ply\text{-}chi}(880) = 13.16$, $p < .001$; $t_{ply\text{-}fre}(880) = 10.53$, $p < .001$; $t_{mle\text{-}chi}(880) = 13.37$, $p < .001$; $t_{mle\text{-}fre}(880) = 10.74$, $p < .001$). Finally, the difference between French-accented English (M = 5.40; SD = 0.89) and Chinese-accented English (M = 5.75; SD = 0.90) is not statistically significant ($t(880) = 2.63$, $p = .065$).

The responses from the second question, namely, where participants thought the speakers came from, are discussed below. Based on the literature (Boughton, 2006; Preston, 1999a), perfect scores were not expected. Adult native speakers of a language are not always able to identify where an accent comes from, even when they are able to distinguish it from other accents.

Table 4.2 shows the percentages of correct answers for this question. Participants were able to write freely where they thought the speaker came from. Their answers were

[34] The effect of age almost reaches significance. Age was included as a continuous variable with age in years. The low p–value may reflect that, the younger the participant, the stronger they rated accents overall.

categorized as correct or incorrect in the following way. First, as mentioned in Section 4.2.1, SSBE is not bound to a specific location in England. In fact, it is highly coupled with SES (see Figure 2.1 for the relationship between SES and geography when it comes to accents). Therefore, it was not possible to establish an unequivocally correct answer for SSBE. For Plymouth, an answer that explicitly mentioned the region of Southwest England or its counties/cities was considered correct[35]. For MLE, explicit references to London were rated as correct[36]. In a parallel fashion, answers that included France/French or China/Chinese were considered correct for French-accented English and Chinese-accented English, respectively.

Table 4.2. Accent strength ratings and location guesses
Rows in bold indicate speakers chosen as stimuli for the developmental study

Speaker	Accent Strength		Location guesses
	Mean	SD	Percentage
SSBE – 2	**2.41**	**1.27**	**—**
SSBE – 5	**2.64**	**1.40**	**—**
SSBE – 4	2.71	1.47	—
SSBE – 3	2.75	1.61	—
MLE – 2	2.83	1.36	44.83
SSBE – 1	2.83	1.44	—
Plymouth – 5	3.42	1.42	5.41
MLE – 1	3.47	1.23	61.11
Plymouth – 2	3.47	1.39	31.58
Plymouth – 3	3.89	1.45	43.24
MLE – 3	4.03	1.44	37.84
Plymouth – 4	**4.46**	**1.28**	**55.26**
MLE – 5	**4.53**	**1.34**	**84.38**
MLE – 4	**4.71**	**1.11**	**69.05**
French – 3	4.89	0.98	25.00
Plymouth – 1	**4.97**	**1.15**	**68.97**
Chinese – 3	5.29	0.57	31.58
French – 4	5.36	0.76	40.48
French – 5	**5.44**	**0.80**	**68.75**
Chinese – 5	5.44	1.11	50.00
French – 2	5.65	0.82	18.92
French – 1	**5.73**	**0.91**	**56.67**
Chinese – 1	**5.91**	**0.91**	**54.55**
Chinese – 2	**5.97**	**0.87**	**67.74**
Chinese – 4	6.18	0.64	69.23

[35] This is not to say that all speakers from this region speak the exact same way. Both Wells (1982) and Trudgill (1999) provided different subclassifications of the accents in this region.

[36] As was the case with Plymouth, this should not be understood as meaning that all London speakers have the same accent.

4.3.3 Selection of speakers

The results of the previous section formed the basis for selecting which specific speakers of each accent were to be included in the developmental study. Two speakers per accent were to be chosen. This number was motivated on two different grounds. First, the results of the verbal-guise task (Section 4.1.2) could easily be affected by the idiosyncratic phonetic features of one speaker to the detriment of the effects of the accent itself. Therefore, two speakers of each accent allow us to level out such individual differences. Second, the categorization task (Section 4.1.3) requires more than one speaker per accent. However, to keep the experiment relatively short for the youngest participants, having many speakers was also not ideal. Therefore, the minimum necessary for the task to take place was chosen, that is, two speakers.

Which specific speakers were chosen was based on two parameters. First, the selected speakers had to have, at least, a rate of 50% correct responses concerning their geographical background. This was done to ensure that participants in the developmental study reacted to accents that matched the desired accent type. Second, the two speakers of each accent had to have accent strength ratings that did not statistically differ from each other.

Based on these constraints, the speakers in bold in Table 4.2 were chosen. With regard to SSBE, the accents with the lowest ratings were chosen. This was meant to reflect its status as a standard accent in England (see Section 4.2.1). The ratings between these SSBE speakers did not differ from each other ($F(1,73) = 0.55$, $p = .460$). The accent strength ratings of the two Plymouth speakers did not differ either ($F(1,64) = 2.77$, $p = .100$), nor did the ratings for the MLE speakers ($F(1,72) = 0.41$, $p = .523$). Furthermore, the speakers of the MLE and Plymouth accents had ratings that did not differ from each other ($F(1,138) = 0.05$, $p = .822$). Finally, the ratings of these two groups differed from the ratings of the SSBE speakers ($F(1,213) = 140.46$, $p < .001$). Next, the Chinese-accented speakers did not differ from each other in terms of accent strength ($F(1,62) = 0.07$, $p = .794$) and neither did the French-accented speakers ($F(1,60) = 1.86$, $p = .177$). These two accents did, however, differ from each other ($F(1,124) = 5.24$, $p = .024$), with the French-accented speakers having a lower mean rating (5.58) than Chinese-accented speakers (5.94). Furthermore, the two L2 accents differed in strength from the Plymouth and MLE speakers ($F(1,264) = 69.74$, $p < .001$). In this case, the L2 accents had a higher mean rating of accent strength (5.76) than the two L1 accents (4.66).

4.4 Participants

One of the RQs underlying this investigation is whether linguistic diversity during a child's upbringing affects when and/or how children develop accent stereotypes.

Therefore, special care was taken to find participants who differ on how much exposure they have had to different accents and languages while still being comparable. Two cities from the south of the UK were selected: Plymouth (Devon, Southwest England) and London.

The two populations differ on three demographic variables that were expected to lead to differences in linguistic diversity. The first variable is the number of residents in each city born outside of England, the second is ethnic diversity, and the third is number of languages spoken. These are discussed below. With regard to Plymouth, the Office for National Statistics reports that, as of 2021, 87.7% of the Plymouth population was born in England and 94% of its population consider their ethnic group as white (CENSUS 2021, 2023a). In contrast, 57.7% of London's population reports England as their country of birth. Similarly, 53.8% of London residents identify themselves as white (CENSUS 2021, 2023b). Moreover, Baker and Eversley (2000) estimate that there are approximately 300 different home languages in London. These differences in demographics were expected to reflect differences in exposure to linguistic diversity during development (see below in this section).

Participants were recruited via different university labs (primarily the *University of Plymouth Babylab*[37] and the *QMUL Language Acquisition Lab*[38]), social media and snowball sampling. In total, data from 249 participants was collected in these two locations. Of these, 136 are from Plymouth and 113 from London. All participating families were given a £15 gift-card as compensation for their time. Two participants from London were excluded because they were born and raised outside the UK. The rest of the participants were born and raised in their respective city and had not lived more than a year outside the UK. A further cause for disregarding data was if participants had neuropsychiatric diagnoses and/or speech and language impairments (7 participants from London and 17 from Plymouth). To be specific, only data from neurotypical children without a history of language or speech disorders was included in the final analyses. Figure 4.5 shows how many participants reported different kinds of diagnoses[39]. Developmental conditions included, for example, Autism Spectrum Disorder or Attention Deficit Hyperactivity Disorder. Examples of disorders in the category referred to as *other* include delayed language development or partial hearing loss.

[37] https://www.plymouth.ac.uk/facilities/babylab

[38] https://languageacquisitionlab.qmul.ac.uk

[39] The sum of the numbers of Figure 4.5 does not match the number of excluded participants (24), as a single participant could have more than one diagnosis.

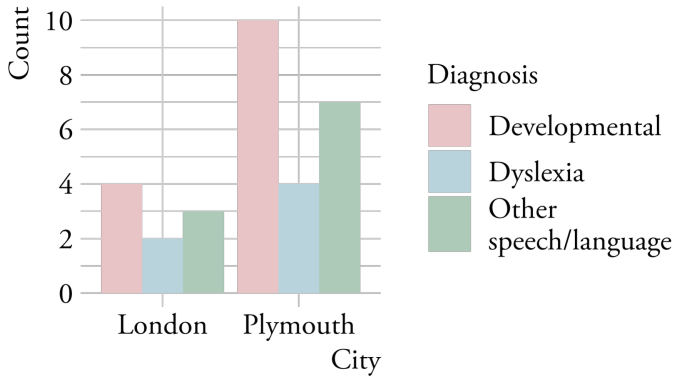


Figure 4.5. Number of participants with different types of diagnoses

The final number of participants was therefore 223. Of these, 119 came from Plymouth (66 girls) and 104 from London (48 girls). As mentioned above, the London population is much more diverse than the Plymouth population. Therefore, special focus was placed on recruiting children who lived in the east and south of London, where Multicultural London English is more prominent (see Section 4.2.1). The spread of participants across the different London boroughs is seen in Figure 4.6, where darker colors indicate more participants from the relevant borough.

Figure 4.6. Boroughs where London participants lived.
The darker the color of a borough, the more participants lived there. Source map: "London Boroughs" by Cartography Vectors, licensed under CC BY 4.0.

The participants from both cities were selected to represent comparable populations. This applies to three metrics: age, SES and vocabulary size. These are discussed in turn. First, the age span for both participant groups was from 7 to 11 years, thus creating a cross-sectional design. This specific range was chosen based on the literature available

at the beginning of this project. At the time, research showed that stereotypes toward L1 accents could be found at the age of 10 years. Since RQ2 in Section 1.1 asked whether the development of accent stereotypes occurs at an earlier age for L2 accents, participants younger than 10 years were selected. Table 4.3 shows how many participants were included by age in years. Moreover, Figure 4.7 depicts the distribution of participants based on months. Both Table 4.3 and Figure 4.7 show that the distribution of participants is comparable across ages and cities. Finally, all analyses were carried out with age as a continuous variable in months (see Section 4.5).

Table 4.3. Distribution of participants in final analyses based on age (years).

Age	London	Plymouth
7	19	25
8	22	26
9	22	23
10	21	20
11	19	25

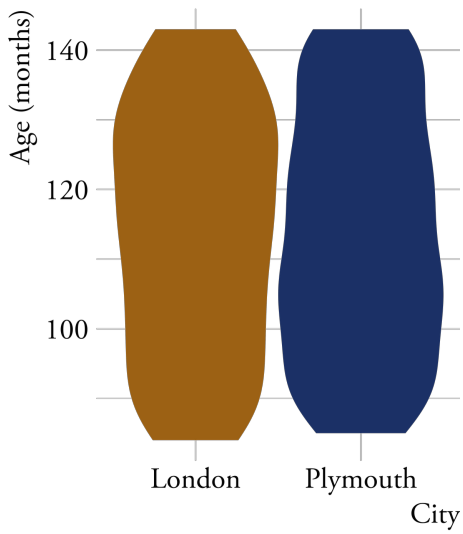


Figure 4.7. Distribution of participants in final analyses based on age (months).

Second, the two groups of participants did not differ on SES. SES has been shown to impact language development (Hackman & Farah, 2009; Hackman et al., 2010; Hoff, 2006; Pace et al., 2017). Children from families with a higher SES have more advanced language skills than children of the same age from families with a lower SES. It was therefore imperative to obtain data from children of comparable SES. SES was measured via parents' education level (see Section 4.1.5), as in Hoff (2006, p. 60) and Kaiser and Kasberger (2021).

Figure 4.8 shows the average education level of participants' parents across cities. The lowest value, zero, indicates no qualifications while the highest, six, a postgraduate degree. In both cities, participants mostly came from families with middle and upper-middle SES. The mean education level in London was 4.86 and, in Plymouth, 4.69. The visual similarities across Plymouth and London were corroborated by an analysis of variance that showed that the mean SESs did not differ across cities ($F(1,221) = 1.106$, $p = .294$). While these similarities are important for the comparability of the results across Plymouth and London, they entail that the results (see Chapter 5, 6, and 7) only show a specific view of the development of accent stereotypes. As highlighted in Section 3.4, SES has been shown to be relevant for how children learn to use language for social evaluations. Therefore, this thesis will not be able to tap into how SES affects the development of accent stereotypes.

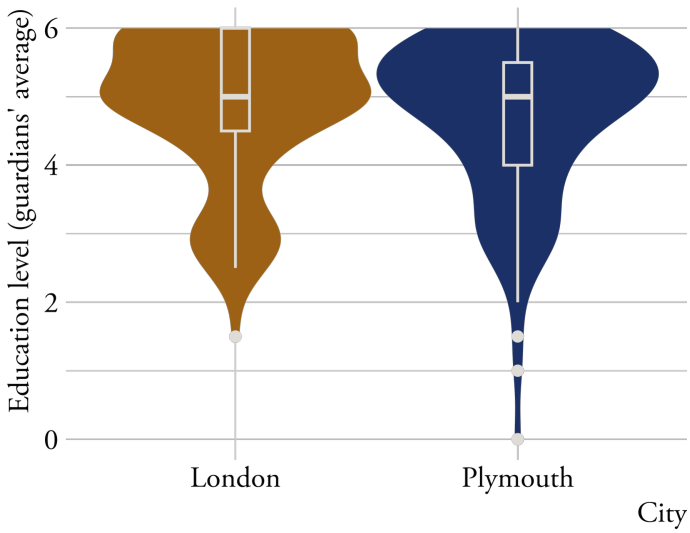


Figure 4.8. Socioeconomic status across cities

The third variable of comparability between the participants across cities is receptive vocabulary size relative to peers, measured with the BPVS (Section 4.1.4). Figure 4.9 shows the participants' standardized scores. A score of 100 indicates that a child had a normal receptive vocabulary size for their age. A score above 100 means that a child had a greater vocabulary than same-age peers. Conversely, a score under 100 indicates that a child had a smaller vocabulary size than same-age peers. As can be seen in Figure 4.9, most participants had a higher-than-average vocabulary score. This result was expected based on the SES pattern described in the previous paragraph. The children's higher-than-average vocabulary scores can be explained by the fact that they came from families with higher SES. Furthermore, all participants had a score within normal

values. Of special relevance is the fact that the scores from Plymouth (M = 107.17) and London (M = 104.95) did not differ from each other ($F(1,220) = 2.12$, $p = .146$).

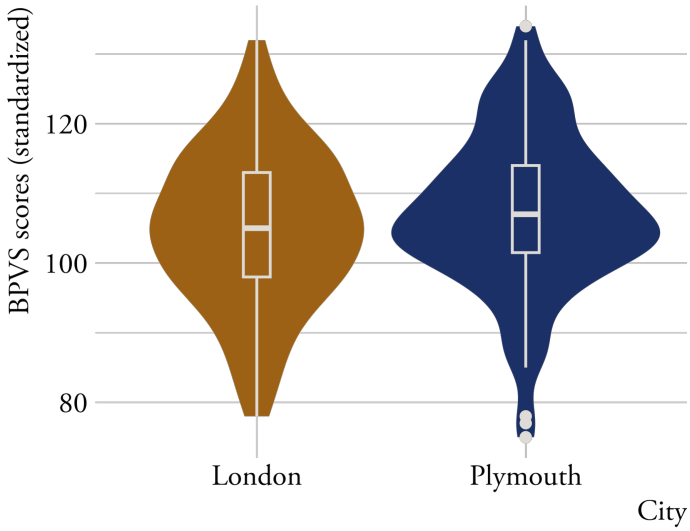


Figure 4.9. British Picture Vocabulary Scale, standardized scores, across cities

Up until now, the two populations were comparable based on age, SES and receptive vocabulary size. Nevertheless, they differed on key parameters that reflect the participants' exposure to linguistic diversity. These differences in exposure can be illustrated with three variables[40]: (i) how many languages the participants spoke, (ii) number of parents who spoke with non-local accents and (iii) how many hours a week participants spent listening to non-local accents. These three measures are discussed in turn.

Table 4.4. How many languages participants spoke
Values in parenthesis are percentages by city.

Number of languages participants spoke	London	Plymouth
1	63 (60.58%)	115 (96.64%)
2	32 (30.77%)	4 (3.36%)
3	9 (8.65%)	0

First, all participants were native speakers of English. Moreover, if they spoke other languages, English was the strongest one: the language they listened to and used the most. Nevertheless, the degree to which participants spoke other languages varied across cities. How many languages participants spoke is shown in Table 4.4 by city. For this

[40] These three measures were obtained via the parent questionnaire. See Section 4.1.5 for more details.

metric, languages learned at school (such as French as a foreign language as part of their primary education) were not included. Only languages spoken at home were considered. In Plymouth, 96.64% of participants were monolingual English speakers. In contrast, monolingual children comprised 60.58% of the participants from London. The differences between the two cities were statistically significant ($F(1,221) = 51.28$, $p < .001$). Which specific languages participants spoke is shown in Figure 4.10.

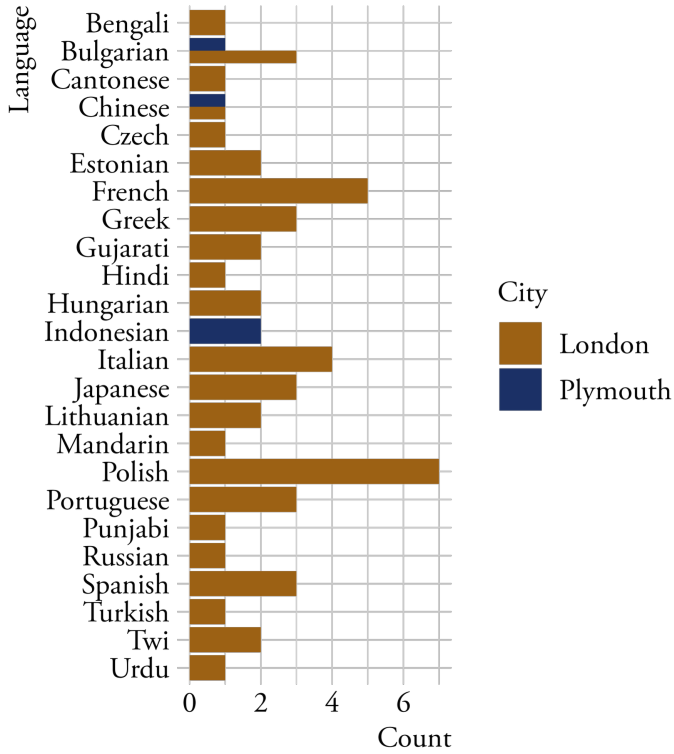


Figure 4.10. Languages spoken by participants

The second difference across populations with regard to linguistic exposure was how many participants had parents who spoke with non-local accents. This included both L2 accents and non-local L1 accents. For this analysis, SSBE was categorized as local in both Plymouth and London. This was due to its close relationship with SES and it often being characterized by its non-localizability in England (see more in Sections 4.2.1 and 4.3.2). Table 4.5 shows how many children had parents who only spoke with local accents (first row), how many had one parent who spoke with a non-local accent (second row) and how many had two parents who spoke with a non-local accent (third row). In Plymouth, 84.9% of participants (101) have only parents who spoke with a local accent. This diverged from London, where 57.7% of participants (60) had parents

who spoke with a non-local accent. An analysis of variance revealed that the differences across cities were statistically significant ($F(1,221) = 22.22$, $p < .001$).

Table 4.5. Number of participants with parents who speak with non-local accents

Number of parents with non-local accents	London	Plymouth
0	60 (57.69%)	101 (84.87%)
1	26 (25.00%)	13 (10.92%)
2	18 (15.13%)	5 (4.20%)

The third linguistic parameter that differentiated the two populations was how many hours a week participants spent listening to non-local accents. In a parallel fashion to the previous paragraph, SSBE was considered a local accent in both Plymouth and London. The differences between cities can be seen in Figure 4.11, which shows how many hours a week, on average, participants spent listening to non-local accents in both cities. A visual inspection reveals that the London median ($Mdn = 25$) was higher than the median in Plymouth ($Mdn = 4$). This difference was statistically significant based on an ANOVA on a linear model ($F(1,220) = 17.04$, $p < .001$).

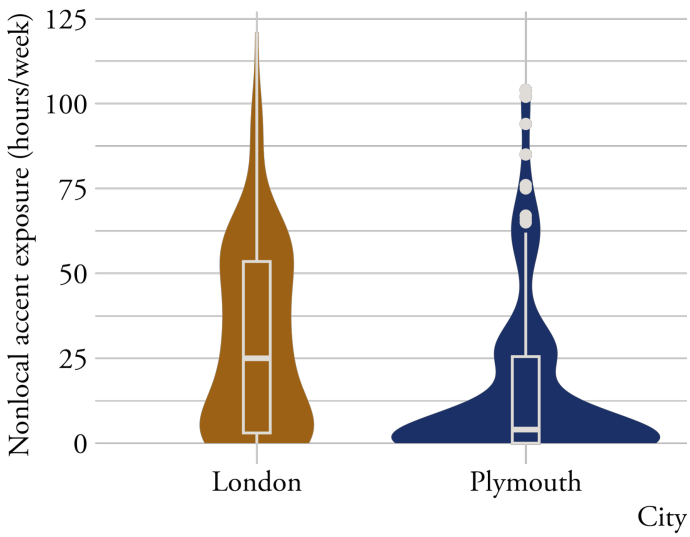


Figure 4.11. Number of hours/week participants were exposed to non-local accents

These three measures (number of languages spoken, number of parents with non-local accents and number of hours per week listening to non-local accents) made it possible to conclude that the participants from Plymouth and London differed with respect to their levels of exposure to linguistic diversity. Specifically, participants from London were significantly exposed to more linguistic diversity than their peers from Plymouth.

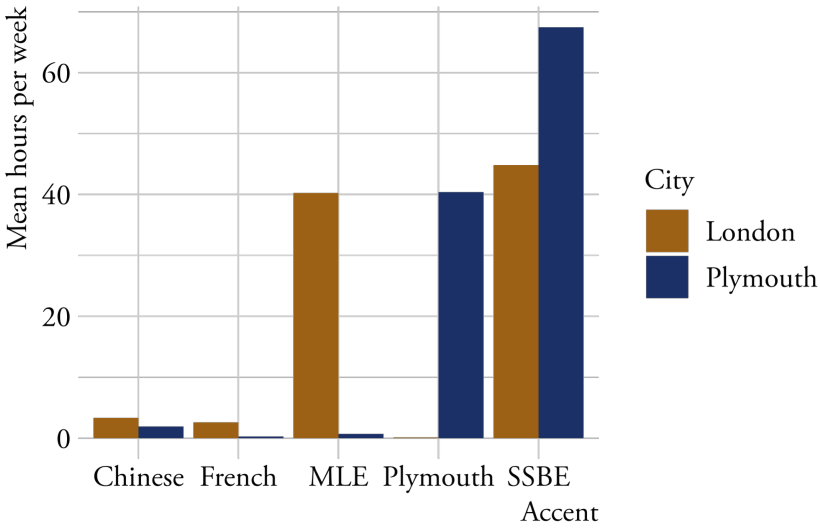


Figure 4.12. Average hours/week participants were exposed to the accents in the investigation

Finally, the participants' exposure to the accents in the stimuli (see Section 4.2.1) was also compared across cities. The average number of hours a week that children in each city listened to the accents included as stimuli is shown in Figure 4.12. The differences across cities were analyzed using pairwise comparisons (see Section 4.5). The differences involving Chinese- and French-accented English across cities were not statistically significant ($t_{chi}(1105) = 0.42$, $p = .999$; $t_{fre}(1105) = 0.69$, $p = .998$). The differences between MLE and the Plymouth accent across cities were clearly significant. However, there was not a significant difference between the average number of hours children from London listened to MLE and the average number of hours children from Plymouth listened to the Plymouth accent ($t(1105) = 0.04$, $p > .99$). In other words, participants from the two cities were exposed to their local accent to the same degree. Finally, participants from Plymouth were exposed to SSBE more than their London peers were ($t(1105) = -6.76$, $p < .001$). This was because more parents from Plymouth spoke SSBE than did parents from London, who were more likely to speak with non-local accents, as described above.

4.5 Analysis

This section deals with how the data collected with the methods described above was processed to obtain the results in the following chapters. First, the statistical software employed as well as its specific packages and functions are described. Then, Section 4.5.1 deals with how the data was processed for the analysis of specific tasks.

All descriptive and inferential statistics were carried out in R (version 4.4.1; R Core Team, 2024) with RStudio (version 2024.04.2+764; Posit team, 2024) unless otherwise stated. Analyses involving (i) linear models, (ii) analyses of variance (ANOVA) and (iii) binomial generalized linear models used R's native *stats* package (version 4.4.1; R Core Team, 2024). In contrast, linear mixed-effects models were created with the package *lme4* (version 1.1.35.4; Bates et al., 2015). Moreover, pairwise comparisons were made using the *emmeans* package (version 1.10.7; Lenth, 2025). For this purpose, Kenward-Roger approximations for the degrees of freedom and the Tukey's method for *p*-value adjustments were employed.

The data from the three experimental tasks (the intelligibility, verbal-guise and categorization tasks) was converted into z-scores for all statistical analyses. This was done to follow the procedure presented in Dossey et al. (2020). The three experimental tasks described above provided data in different metrics: the intelligibility task provided accuracy measured in percentages; the verbal-guise used ratings spanning from one to four; the categorization task measured accuracy on a scale from -100 to +100 (see Section 4.5.1 below). This complicated the direct comparison of results across tasks, which was a key aspect of this study. Converting the different metrics to z-scores allowed for a clearer comparison of how children performed on the different tasks. These z-scores were calculated by task. In the case of the verbal-guise study, the z-scores were calculated by question (see Section 4.1.2 above).

Finally, when these z-scores were analyzed together with age, age was included as a continuous variable in months. This was done to avoid creating categories such as 7-year-olds and 8-year-olds or second-year students and third-year students. These ordinal categories cluster participants in a way that does not allow differentiation between the younger and older participants in each one of them.

4.5.1 Task specific analyses

Two of the tasks for this study required further data processing prior to the final analyses. This section describes how the data was managed first for the intelligibility task, then for the categorization task (the participants' responses on the verbal-guise task were directly converted into z-scores since they did not require any preprocessing).

Besides the practice session for the intelligibility task (see Section 4.1.1), a further measure was conducted to ensure the quality of the data. The goal here was to ensure that participants did not complete the task by choosing an answer at random. The number of right answers per participant was matched against a cumulative binomial distribution, that is, how likely it is that participants would get a certain number of answers right if they answered at random. The task had twenty trials and four potential answers. This means there was a very small chance ($p < 0.04$) that participants would get nine or more answers correct if they answered randomly. All participants from Plymouth had nine correct answers or more. In the London data, only one participant

had fewer than nine. Therefore, it can be concluded that participants completed the task without answering randomly. The next step was to convert each participant's binary accuracy measures into a percentage score by accent. Since each participant heard four words in each accent, their scores could take the values 0, 25, 50, 75 or 100 per accent.

With regard to the categorization task, the groups the participants constructed were quantified following the procedure specified in Dossey et al. (2020). First, for each participant, the number of correct pairings within each accent was counted. That is, if a participant, for instance, placed the two speakers of MLE in the same group, it was coded as one "hit" for MLE. A participant would, therefore, score zero for MLE if the two MLE speakers were placed in different groups. Second, for each participant, the number of incorrect pairings for each speaker was also tracked. Hence, if a participant included (i) one speaker of French-accented English with a Plymouth speaker and an MLE speaker and (ii) the other speaker of French-accented English with an SSBE speaker, they would get three "misses" for French-accented English. Furthermore, if a participant included both speakers of French-accented English in the same group together with, for instance, one speaker of Chinese-accented English, the number of misses for French-accented English was calculated by counting the misses for each speaker. Therefore, this participant would get two misses: one miss for pairing the first speaker of French-accented English and the Chinese-accented English speaker and another miss for pairing the second speaker of French-accented English with the same Chinese-accented English speaker.

Third, the proportion of hits and misses out of all the possible pairwise hits and misses for each accent per participant was calculated. Since there were only two speakers per accent, a participant's proportion of hits was either 100 (they placed the two speakers of an accent together) or 0 (they placed each speaker in two different groups) for each accent. The maximum number of pairwise misses for a given accent was 16 (eight per speaker of an accent). Finally, the accuracy of a participant for a given accent was calculated by subtracting their proportion of misses from their proportion of hits for said accent. As Dossey et al. (2020: 10) note, this score helps in taking into account the number of groups a participant made. If a participant grouped speakers into few, but larger groups, they would receive both many hits and many misses. In contrast, if a participant did many, but smaller, groups, they would obtain very few hits as well as very few misses. However, this subtraction normalized these differences.

This subtraction provides a score between -100 and +100. A final negative score for an accent indicates that a participant placed the two speakers of an accent in two different groups that included speakers of other accents. For example, if a participant made the groups [SSBE, MLE] and [SSBE, Plymouth], they would get a negative score of -12.5 for the SSBE accent. A final score of zero for an accent means that the participant either (i) placed the two speakers of an accent in two individual groups on their own or (ii) placed all ten speakers in one group together. A final positive score

below 100 for an accent means that a participant grouped the two speakers of an accent together, in addition to speakers of other accents: the more speakers from other accents included in that group, the closer the score will be to zero. If the final score is exactly 100 for an accent, the participant placed the two speakers of that accent together in a group without any other speakers in it.

4.6 Summary

This chapter depicts and motivates the methodological choices made to study how children from two different English cities develop accent stereotypes in late childhood using a cross-sectional design. The methods used to study child development are a critical issue. Therefore, special care was taken to utilize age-appropriate methods for the participants in this thesis (7- to 11-year-old children) and to ensure that the results are comparable with those from the relevant literature. For this purpose, four tasks were chosen to be completed by the child (an intelligibility task, a verbal-guise task, a categorization task and the British Picture Vocabulary Scale), and a survey was created to be completed by their legal guardians. The stimuli included Chinese-accented English, French-accented English, SSBE, MLE and the Plymouth accent. Participants came from comparable backgrounds when it came to SES, language development (indirectly measured by the BPVS) and age. Importantly, the Plymouth and London children differed when it came to exposure to linguistic diversity, with children from London being significantly more exposed to different accents and languages than their Plymouth peers. Finally, to ensure transparency and comparability with future research, this chapter also describes how the data obtained using these methods was pre-processed and analyzed.

5 Accent stereotypes in Plymouth, a linguistically homogenous city

The previous chapter described the methods used in this thesis to investigate when and how children develop accent stereotypes. The results of their implementation are reported in three chapters. This chapter focuses on how children from Plymouth develop stereotypes to accents and what factors may influence such process. Chapter 6 reports on the corresponding findings from the data collection in London. Finally, Chapter 7 compares the data from the two cities. The results of the tasks described in Section 4.1 are reported below in the following way. First, the results from the two auxiliary tasks are described: the intelligibility task in Section 5.1 and the categorization task in Section 5.2. These sections are followed by the main task, the verbal-guise study, in Section 5.3. This chapter finishes with a discussion of the findings from the Plymouth data (Section 5.4). First, the results from each individual task are commented on. Afterwards, Sections 5.4.1 and 5.4.2 review how the findings in this chapter inform the RQs in Section 1.1.

5.1 Intelligibility task

In this task, participants heard 20 words with background noise and were asked to choose a picture that matched the word they heard (see Section 4.1.1 for details). The data from this task is organized in two ways. First, the accuracy scores by target word are shown in Table 5.1. The percentages were calculated based on the total number of times each word was shown to participants. Since each word was only presented once to each participant, the maximum number of correct answers per word is 119, one per participant. The percentage of correct responses spans from 97.48% with *church* down to 27.73% with *thorn*. These results show that some words were more difficult to understand than others when presented in noise.

Second, the data was analyzed by accent. The overall accuracy scores per accent can be seen in Table 5.2. The percentages were calculated based on the total number of words produced per accent. Since the stimuli always included four words in each of the

five accents for each participant, the maximum number of correct answers per accent is 476 (119 participants times four words per accent).

Table 5.1. Intelligibility task, Plymouth, correct answers per word
Correct answers (raw and percentage) per word in the Plymouth data. The maximum number of correct answers per word is 119.

Target word	Correct answers	Percentage
church	116	97.48
branch	110	92.44
house	110	92.44
alligator	109	91.60
mouse	109	91.60
wolf	107	89.92
foot	104	87.39
porridge	104	87.39
bird	102	85.71
farm	98	82.35
leaf	98	82.35
key	96	80.67
toy	93	78.15
salt	88	73.95
goat	82	68.91
goose	81	68.07
bowl	78	65.55
face	60	50.42
palm	42	35.29
thorn	33	27.73

Table 5.2. Intelligibility task, Plymouth, correct answers per accent
The maximum number of correct answers per accent is 476.

Accent	Correct answers	Percentage
Plymouth	389	81.72
SSBE	383	80.46
French	369	77.52
MLE	346	72.69
Chinese	333	69.96

Table 5.2 shows a clear gradience in results. Participants scored highest on the local, Plymouth accent. This was followed by, in order, SSBE, French-accented English, MLE and, in last place, Chinese-accented English. Yet Table 5.2 does not show whether the differences across accents are significant or whether there were developmental effects or interactions between age and accent. Therefore, the results were analyzed using a linear mixed-effects model (see Section 5.4). The dependent variable was the participants' accuracy scores, whereas the independent variables were (i) age and (ii) accent. These two terms were included with an interaction between them.

Furthermore, two random effects were included: participant ID and individual speaker per accent. An ANOVA reveals that there was a significant effect of age ($F(1,119) = 4.41$, $p = .038$) as well as of accent ($F(4,2261) = 6.98$, $p < .001$). However, there was no significant age–accent interaction ($F(4,2261) = 0.69$, $p = .601$).

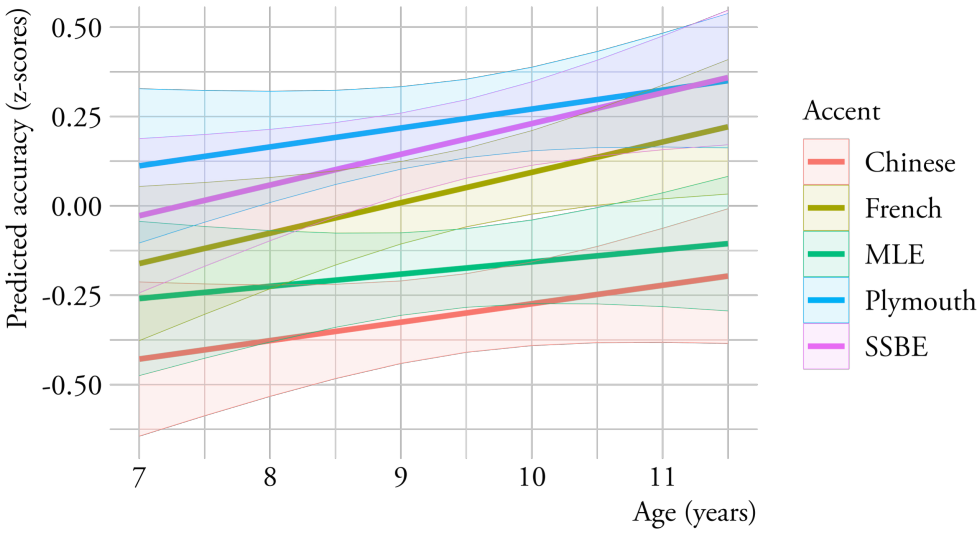


Figure 5.1. Intelligibility results, with age, Plymouth

The effects of age and accent from the model in the previous paragraph can be seen in Figure 5.1. The X-axis depicts age, while the Y-axis represents accuracy. Bold lines show the predicted mean accuracy for a given age, while the ribbons indicate the predicted 95% confidence intervals. It can be observed that younger children had a lower overall accuracy (M_{7yo} = -0.15, *95% CI*$_{7yo}$ = [-0.36, 0.06][41]) than older children (M_{11yo} = 0.09, *95% CI*$_{11yo}$ = [0.08, -0.27]). As the lack of a significant interaction between age and accent illustrates, the accuracy for the different accents remained relatively stable throughout development. Therefore, a new model was created that included no interaction between the predictors of age and accent. The other terms included were the same as in the model described in the previous paragraph. The results of this model are shown in Figure 5.2. In this case, the predicted average accuracy across all ages (Y-axis) is shown against the accents included in the stimuli (X-axis).

[41] Confidence Intervals calculated as: $[\bar{x} - 1.96*SE, \bar{x} + 1.96*SE]$

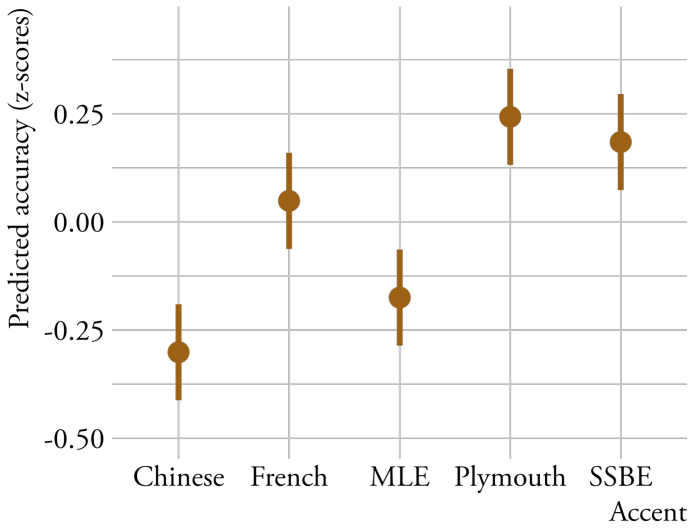


Figure 5.2. Intelligibility results, without age, Plymouth

Pairwise comparisons (see Section 4.5) of the five accents reveal that two main clusters appear. First, Chinese-accented English (M_{chi} = -0.30; *95% CI*$_{chi}$ = [-0.19, -0.41]) and MLE (M_{mle} = -0.17; *95% CI*$_{mle}$ = [-0.06, -0.29]) did not significantly differ from each other (t(5.02) = -2.22, p = .302). Second, Plymouth (M_{ply} = 0.24; *95% CI*$_{ply}$ = [0.13, 0.35]) and SSBE (M_{ssb} = 0.18; *95% CI*$_{ssb}$ = [0.07, 0.30]) did not differ from each other either (t(5.02) = 1.03, p = .834). These two groups were significantly different from each other ($t_{chi\text{-}ply}$(5.02) = -9.58, p = .002; $t_{chi\text{-}ssb}$(5.02) = -8.56, p = .002; $t_{mle\text{-}ply}$(5.02) = 7.36, p = .004; $t_{mle\text{-}ssb}$(5.02) = -6.33, p = .008). As can be seen in both Figure 5.1 and Figure 5.2, French-accented English (M_{fre} = 0.05; *95% CI*$_{fre}$ = [-0.06, 0.16]) ranked between those two poles. While the accuracy of French-accented English was significantly higher than that of Chinese-accented English (t(5.02) = -6.16, p = .009), it was not significantly different from SSBE (t(5.02) = -2.40, p = .253) and on the edge of significance from both MLE (t(5.02) = 3.94, p = .053) and (to a lesser degree) the Plymouth accent (t(5.02) = -3,42, p = .088).

In conclusion, younger children from Plymouth scored lower than their older peers. Moreover, children from Plymouth were not equally capable of understanding all accents when the stimuli were presented to them in noise. The Plymouth and SSBE accents were easier for them to understand than Chinese-accented English or MLE. The ranking of accents across development remained constant as reflected by the lack of interaction between age and accent.

5.2 Categorization task

In this task, children were asked to listen to two speakers of each accent and group them based on where they thought they were born (see Section 4.1.3 for details). The results from this task are reported in two steps. The accuracy results are first presented, followed by the co-occurrence matrices of the actual groupings participants made.

5.2.1 Accuracy scores

Accuracy is understood as how well participants were able to group the ten speakers into their five respective groups (one per accent). Accuracy was quantified following the procedure suggested in Dossey et al. (2020) (see Section 4.5.1). Their analysis provides scores ranging from -100 to 100. The higher the value, the more accurate the groups were.

Table 5.3. Categorization task, Plymouth, average scores
The scores span from -100 to 100.

Accent	Average score
SSBE	41.21
French	40.20
Plymouth	36.28
Chinese	24.79
MLE	19.39

The average accuracy by accent is shown in Table 5.3. A hierarchy is present, with SSBE and French-accented English at the top, the Plymouth accent in the middle and Chinese-accented English and MLE at the bottom. Like the case of the intelligibility task in Section 5.1, these scores do not allow for inferences about developmental effects or interactions. Therefore, the scores were analyzed using a linear mixed-effects model. The dependent variable was the accent scores. The independent variables were the interaction of age and accent with participant ID included as a random effect. In this case, it was not possible to include individual speaker for each accent as a random effect. This is because the final accuracy score merges both speakers (see Section 4.5.1).

An ANOVA reveals a significant effect of age ($F(1,118) = 31.41$, $p < .001$) as well as of accent ($F(4,1062) = 12.91$, $p < .001$). Just like in Section 5.1, there was no interaction between accent and age ($F(4,1062) = 0.36$, $p = .838$). The predicted results from this model can be seen in Figure 5.3. The X-axis depicts age, whereas the Y-axis represents accuracy. Bold lines represent the predicted mean score for a given age, while the ribbons show the 95% confidence intervals. As the age effect highlights, younger children ($M_{7yo} = -0.45$, $95\%\ CI_{7yo} = [-0.66, -0.24]$) scored overall lower than older children ($M_{11yo} = 0.28$, $95\%\ CI_{11yo} = [0.12, 0.45]$).

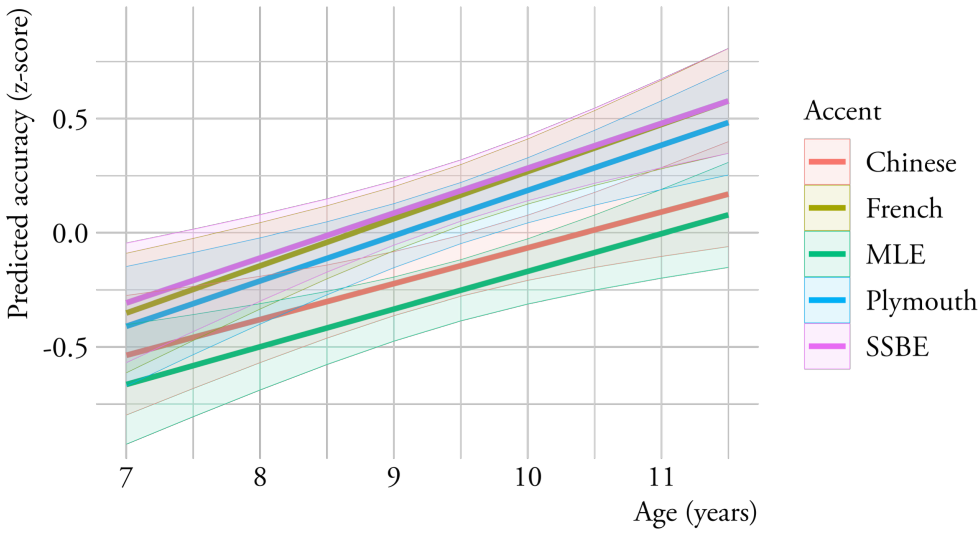


Figure 5.3 Categorization results, with age, Plymouth

The ranking of accents remained stable across development. Therefore, a model without an interaction between age and accent was run. The results of such a model are seen in Figure 5.4. The predicted average scores across all ages are shown on the Y-axis and accent on the X-axis. It parallels the results depicted in Table 5.3, with SSBE and French-accented English scoring highest and Chinese-accented English and MLE lowest.

Pairwise comparisons (see Section 4.5) of all five accents reveal two clusters of two accents, with one accent between them. First, Chinese-accented English (M_{chi} = -0.15; *95% CI*$_{chi}$ = [-0.29, -0.01]) and MLE (M_{mle} = -0.26; *95% CI*$_{mle}$ = [-0.40, -0.12]) had the lowest overall scores and did not differ from each other (t(5.04) = 1.40, p = .650). At the opposite end, French-accented English (M_{fre} = 0.16; *95% CI*$_{fre}$ = [0.01, 0.30]) and SSBE (M_{ssb} = 0.18; *95% CI*$_{ssb}$ = [0.03, 0.32]) had the highest overall scores and did not differ from each other either (t(5.04) = -0.26, p = .998). These two poles differed significantly from each other ($t_{mle\text{-}fre}$(5.04) = 5.41, p = .015; $t_{mle\text{-}ssb}$(5.04) = -5.67, p = .013; $t_{chi\text{-}fre}$(5.04) = -4.01, p = .049; $t_{chi\text{-}ssb}$(5.04) = -4.27, p = .039). In between these poles, there were the scores for the Plymouth accent (M_{ply} = 0.08; *95% CI*$_{ply}$ = [-0.06, 0.22]). It scored significantly higher than MLE (t(5.04) = -4.39, p = .035). Nevertheless, it did not differ from French-accented English (t(5.04) = 1.02, p = .837) or SSBE (t(5.04) = -1.28, p = .713). With regard to the Plymouth accent and Chinese-accented English, the difference in scores was not statistically significant either (t(5.04) = -2.99, p = .136), but greater than that between the Plymouth accent and both SSBE and French-accented English.

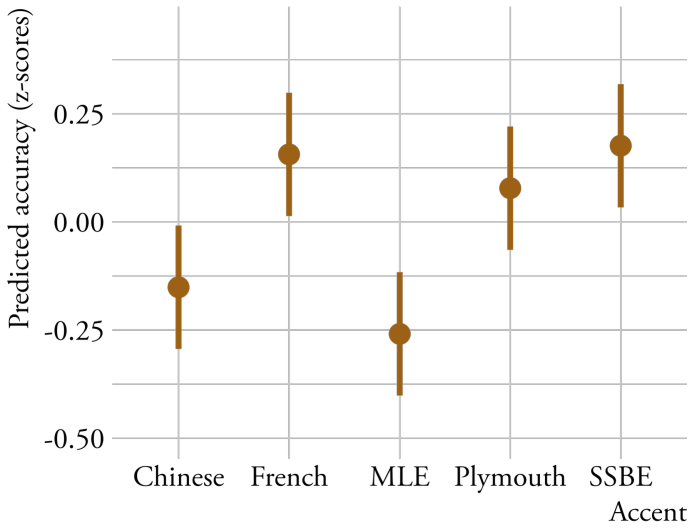


Figure 5.4. Categorization results, without age, Plymouth

5.2.2 Co-occurrence matrices

The results presented in the previous section show that children were better at grouping SSBE and French-accented English than MLE and Chinese-accented English. However, this does not allow us to see what speakers participants grouped together. To answer this question, the individual speakers participants included in each group were transformed into a co-occurrence matrix.

The overall percentage of times an accent was grouped with other accents or itself is shown in Figure 5.5. These percentages are to be understood row-by-row, that is, out of all the times a SSBE speaker was grouped with other speakers, 8.17% of the times was with a speaker of Chinese-accented English, 7.36% of the times with a French-accented English speaker, 25.34% of the times with an MLE speaker, etc. A chi-square test of independence with the raw cell counts reveals that the distribution was significantly different from chance ($X^2(16) = 610.84$, $p < .001$). In order to obtain a clearer view of the internal structure of the groups, a complete-linkage clustering analysis with the R package *pheatmap* (Kolde, 2019) was performed. Its results can be seen in Figure 5.6. It shows two main clusters: L2 accents, on the one hand, and L1 accents, on the other. Furthermore, within the L1 branch, the Plymouth and MLE accents were grouped together more frequently. In turn, these two accents clustered more frequently with SSBE than with the L2 accents. Children from Plymouth thus distinguished L2 from L1 accents and, in the latter category, standard (SSBE) from non-standard (MLE and Plymouth) accents.

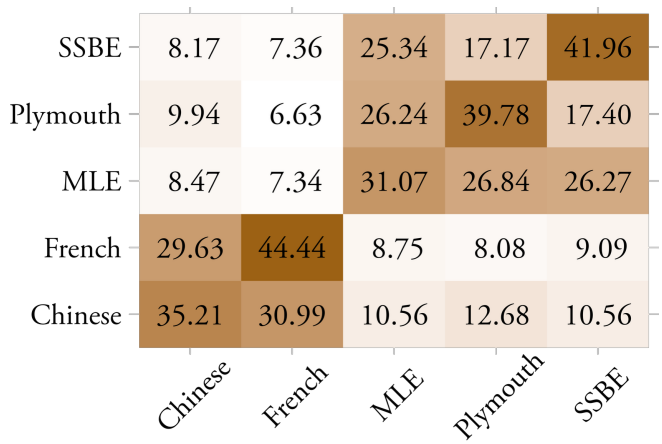


Figure 5.5. Co-occurrence matrix, without age, Plymouth
Values in percentage row-wise. The darker the color of a cell, the higher the percentage of groups.

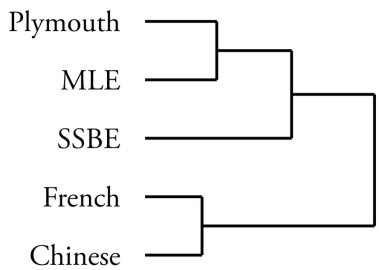


Figure 5.6. Co-occurrence dendrogram, without age, Plymouth

The developmental effect in the previous section reveals that older participants were better than younger ones at grouping speakers. Thus, co-occurrence matrices were made for each year to test whether the content of the clusters participants made changed as they got older.

Figure 5.7 shows the co-occurrence matrices by year. Chi-square tests of independence with the raw cell counts (rather than the percentages shown in the figures) indicate that all five subplots were distributed in a statistically significant way ($X^2_{7yo}(16) = 49.05$, $p < .001$; $X^2_{8yo}(16) = 55.74$, $p < .001$; $X^2_{9yo}(16) = 203.25$, $p < .001$; $X^2_{10yo}(16) = 147.13$, $p < .001$; $X^2_{11yo}(16) = 332.53$, $p < .001$). Furthermore, year-by-year dendrograms were also created with the help of complete-linkage clustering analyses. These can be seen in Figure 5.8.

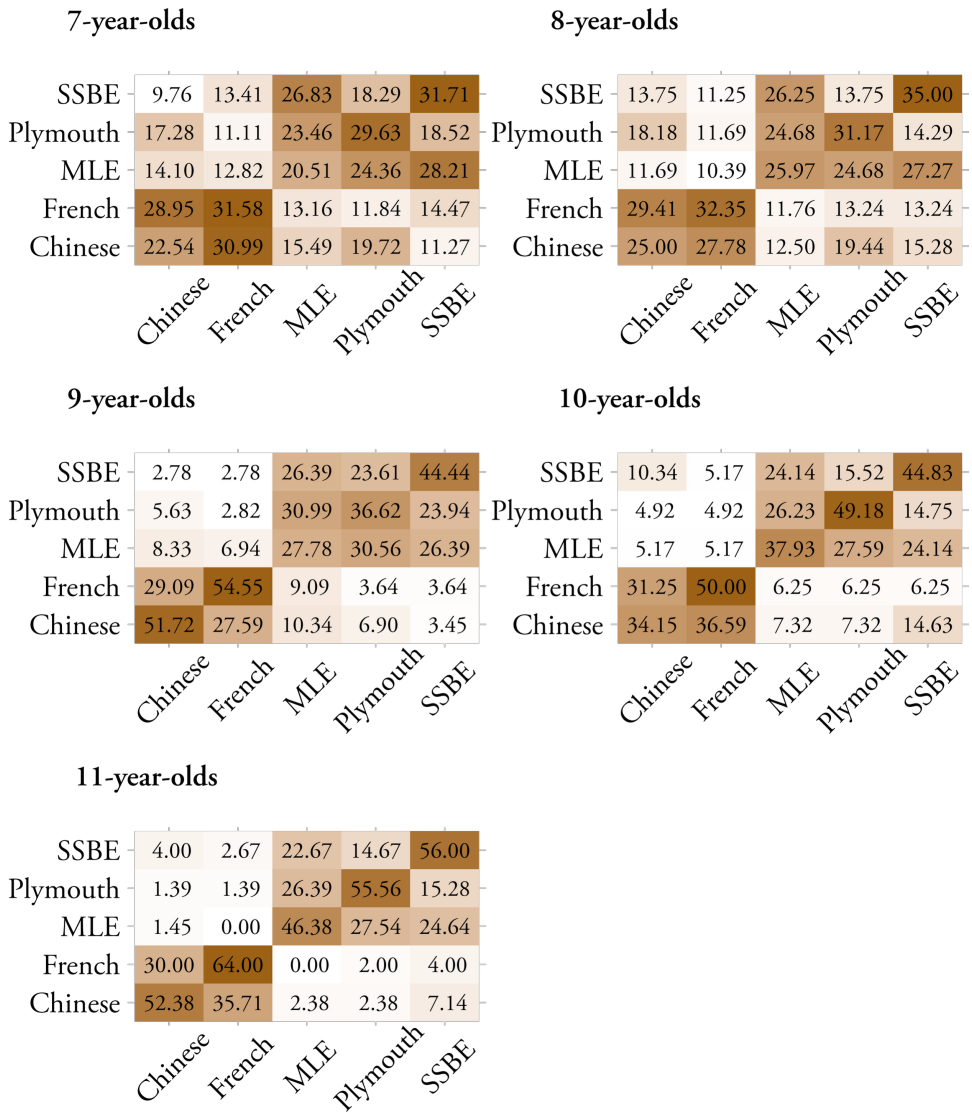


Figure 5.7. Co-occurrence matrices, with age, Plymouth
Values in percentage row-wise. The darker the color of a cell, the higher the percentage of groups.

Across all five years, both Figure 5.7 and Figure 5.8 show that the L2 accents were more frequently grouped with each other and, in turn, L1 accents were clustered together more frequently. Importantly, there appeared to be age-related differences in the internal structure of the branch including L1 accents, as can be seen in Figure 5.8. For the 7- and 8-year-olds, SSBE and MLE were more likely to be grouped together. The

Plymouth accent was clustered with these accents more frequently than with L2 accents, while not to the same extent that SSBE and MLE clustered with each other. Hence, the L1 accents can be categorized into the native Plymouth accent, and nonnative, L1 accents (SSBE and MLE). This changed from the age of 9 years onwards. From this age, the L1 branch was structured the same way as we saw in Figure 5.6. Namely, the L1 branch split into the standard SSBE accent and the non-standard MLE and Plymouth accents.

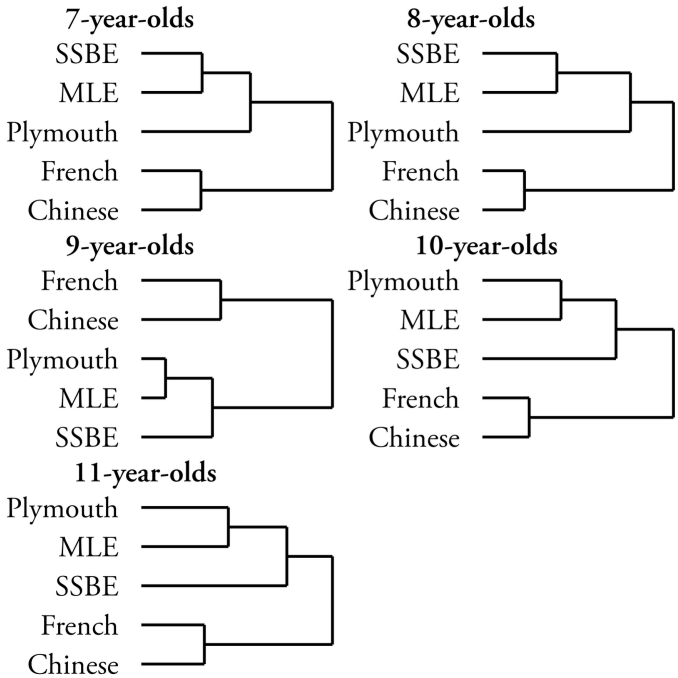


Figure 5.8. Co-occurrence dendrograms, with age, Plymouth

5.3 Verbal-guise task

The data for the verbal-guise task, measuring accent stereotypes, is reported in this section. Participants listened to one speaker of each accent reading the same text out loud. They were then asked different questions about them, such as *How smart do you think this person is?* (see more details in Section 4.1.2). The findings from this task are presented in two ways. First, whether participants rated speakers differently across questions was tested. The goal was to see whether children perceived unknown speakers differently on different variables regardless of accent. Second, a model was created with

the goal of testing whether, within each question, participants rated accents differently. This matches the analysis conducted in Dossey et al. (2020).

A linear mixed-effects model (see Section 4.5) was compiled to test whether children's ratings differed across questions. The ratings were the dependent variable. The independent variables were question type (such as SMART or FRIENDLY; see Section 4.1.2) and age. Moreover, participant ID and individual speaker of each accent were included as random effects. An ANOVA provides a significant effect of question type ($F(5,3185.9) = 4.19$, $p < .001$), but no effect of age ($F(1,3169) = 0.12$, $p = .728$) or question–age interaction ($F(5,3185.9) = 1.38$, $p = .230$). It can therefore be inferred that children rated speakers differently on each question in a similar way across all ages.

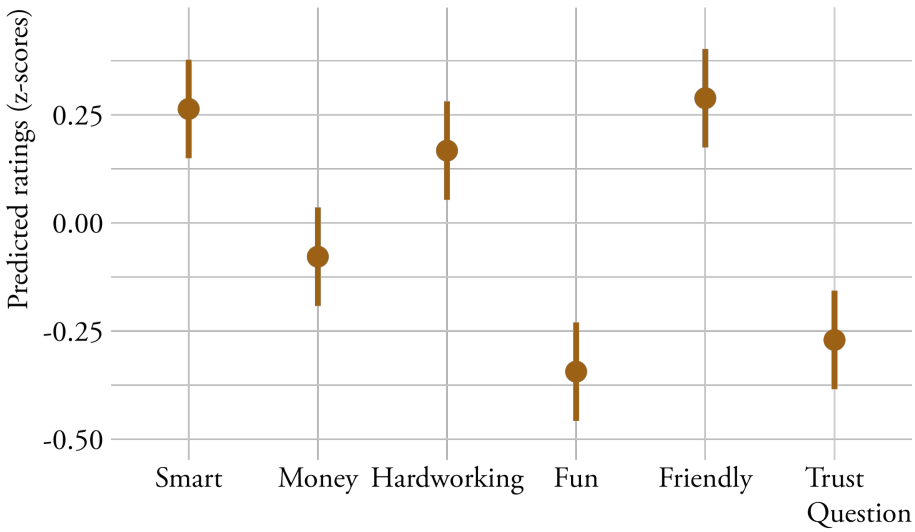


Figure 5.9. Verbal-guise task, by question, without age, Plymouth

Figure 5.9 summarizes the results of the model. Since there were no age effects, it shows the average predicted ratings across all ages for each question type. Pairwise comparisons (see Section 4.5) of the six questions reveal three main statistically significant clusters. First, the SMART (M_{sm} = 0.26; *95% CI*$_{sm}$ = [0.15, 0.37]), HARDWORKING (M_{hw} = 0,17; *95% CI*$_{hw}$ = [0.06. 0.28]) and FRIENDLY (M_{fr} = 0.29; *95% CI*$_{fr}$ = [0.18, 0.40]) questions were rated highest overall and did not differ from each other ($t_{sm\text{-}hw}(3159) = 1.69$, $p = .540$; $t_{sm\text{-}fr}(3159) = -0.44$, $p = .998$; $t_{hw\text{-}fr}(3159) = -2.12$, $p = .275$). Next, the FUN (M_{fn} = -0.34; *95% CI*$_{fn}$ = [-0.45, -0.23]), and TRUST (M_{tr} = -0.27; *95% CI*$_{tr}$ = [-0.38, -0.16]) questions received the overall lowest results while not differing from each other ($t_{fn\text{-}tr}(3159) = -1.28$, $p = .794$). All the variables within each of these two clusters differed significantly from the variables in the opposite group ($t_{sm\text{-}fn}(3159) = 10.63$, $p < .001$; $t_{sm\text{-}tr}(3159) = 9.35$, $p < .001$; $t_{hw\text{-}fn}(3159) = 8.94$, $p < .001$; $t_{hw\text{-}tr}(3159) = 7.66$, $p < .001$; $t_{fr\text{-}fn}(3159) = -11.07$, $p < .001$; $t_{fr\text{-}tr}(3159) = 9.78$,

$p < .001$;). Finally, MONEY (M_{mn} = -0.08; *95% CI*$_{mn}$ = [-0.19, 0.03]) was rated in between the two poles and differed from all other questions ($t_{mn\text{-}sm}(3159) = 5.98$, $p < .001$; $t_{mn\text{-}hw}(3159) = -4.29$, $p < .001$; $t_{mn\text{-}fn}(3159) = 4.65$, $p < .001$; $t_{mn\text{-}fr}(3159) = -6.41$, $p < .001$; $t_{mn\text{-}tr}(3159) = 3.37$, $p = .0099$).

The data presented immediately above shows that the ratings vary by question. That is, children rated unknown speakers higher on the SMART, HARDWORKING and FRIENDLY questions than on the FUN and TRUSTWORTHY ones. As a result, a principal component analysis (PCA) was applied to the data to reduce the number of variables for the analysis by accent. Typically, only factors whose eigenvalue are above one are kept (Jolliffer, 2004, p. 114). In the case of this investigation, only one dimension yielded an eigenvalue with a value higher than one. This dimension included both the SMART and HARDWORKING ratings. This means that all other questions represented different constructs. Therefore, further analyses were carried out on each question individually following the procedure in Dossey et al. (2020). Thus, six different linear mixed-effects models were created, one per question. In each of them, the dependent variable was the ratings participants gave. The independent variables were accent and age. Finally, the random effects of participant ID and speaker were included.

Each of the variables is now discussed one at a time. The questions involving prestige variables (SMART, MONEY and HARDWORKING) are outlined first. They are followed by the questions dealing with the warmth variables (FUN, FRIENDLY and TRUST). The figures in the following six sub-sections should be interpreted the same way. Namely, bold lines show the predicted mean rating for a given age, while ribbons indicate the 95% confidence intervals.

5.3.1 SMART

An ANOVA on the SMART variable reveals a significant accent and age interaction ($F(4,509.24) = 2.51$, $p = .0415$). Figure 5.10 displays the predicted results from this model. The interaction can be understood in the following way. From the age of 7 years to the age of 9 years, children from Plymouth rated all speakers in a similar fashion. That is, there were no signs of accent stereotyping. From this age onwards, however, they started rating speakers differently. To be specific, from the age of 9 years, they rated (i) French-accented English speakers higher than MLE and Plymouth speakers and (ii) SSBE speakers higher than Plymouth speakers. Furthermore, from the age of 9 and a half years, SSBE speakers were also rated higher than MLE speakers. Finally, from the age of 10 years, Chinese-accented English was rated significantly higher than the MLE accent.

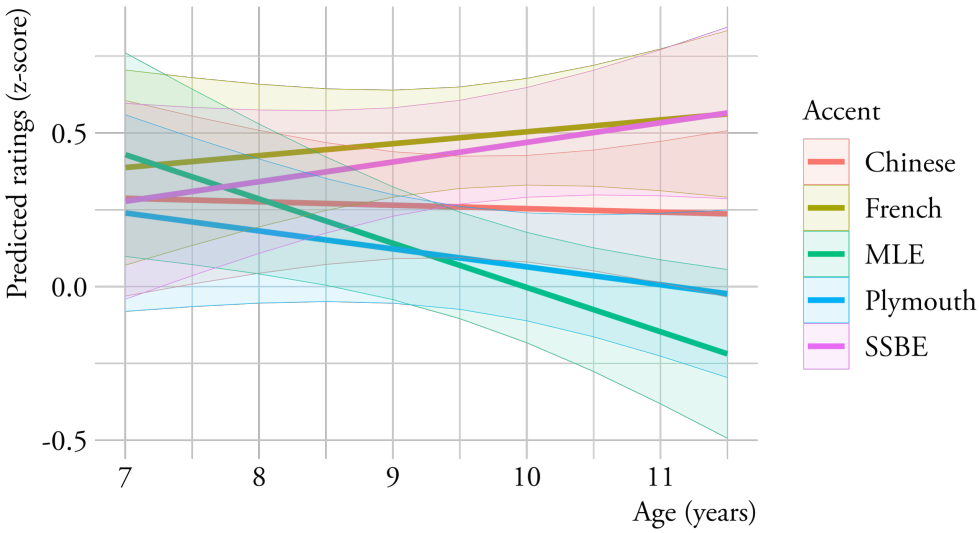


Figure 5.10. Verbal-guise task, SMART, with age, Plymouth

5.3.2 MONEY

An ANOVA on the MONEY question shows a lack of significant effects for accent ($F(4,28.7) = 0.39$, $p = .814$), age ($F(1,107.69) = 0.15$, $p = .703$) and the interaction between the two ($F(4,414.94) = 0.68$, $p = .603$).

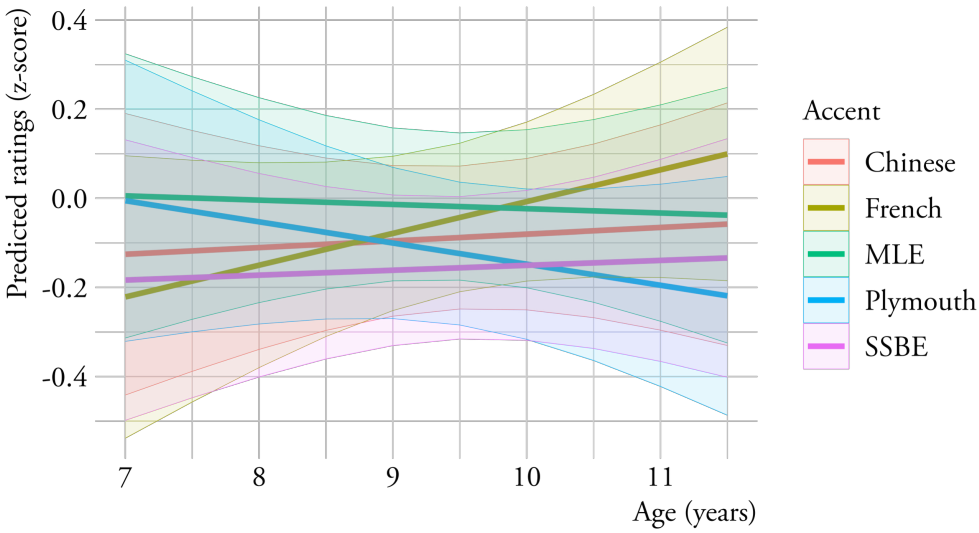


Figure 5.11. Verbal-guise task, MONEY, with age, Plymouth

The results from the model are presented in Figure 5.11. As can be seen, children from Plymouth rated all speakers in the same way throughout the 5-year span, implying that they did not associate people speaking who have different accents with having a higher (or lower) SES.

5.3.3 HARDWORKING

The data from the HARDWORKING question reveals a similar pattern to that of SMART. This was not unexpected, as the PCA mentioned above indicated that these two variables correlate with each other. An ANOVA provides a statistically significant age-accent interaction ($F(4,420.59) = 2.43$, $p = .048$). Figure 5.12 displays this interaction. Younger participants, between the ages of 7 years and 9 and a half years, did not rate accents differently. However, different accents were associated with different levels of HARDWORKING-ness from this age. Namely, both French-accented English and SSBE were rated higher than the MLE and Plymouth accents. Moreover, from the age of 11 years, participants also rated Chinese-accented English higher than the local, Plymouth accent.

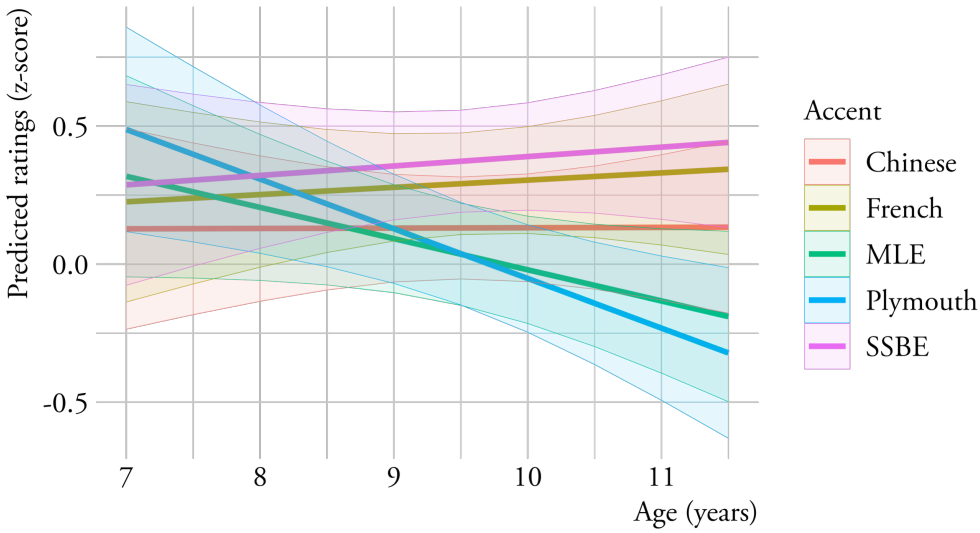


Figure 5.12. Verbal-guise task, HARDWORKING, with age, Plymouth

5.3.4 FUN

The FUN question did not provide any statistically significant effects with regard to accent ($F(4,10.01) = 0.29$, $p = .879$) or age ($F(1,107.51) = 0.61$, $p = .438$). The interaction of these two terms was not significant either ($F(4,423.75) = 1.65$, $p = .161$).

The results from this model can be seen in Figure 5.13. The F-value of the interaction of the FUN question was higher than in the case of MONEY or FRIENDLY. This may be because, despite the lack of overall significant effects, post-hoc analyses showed some significant pairwise comparisons. Older children from the age of 9 and a half years rated the Plymouth accent higher than French- and Chinese-accented English. Furthermore, from the age of 10 and a half, it also received higher scores than SSBE.

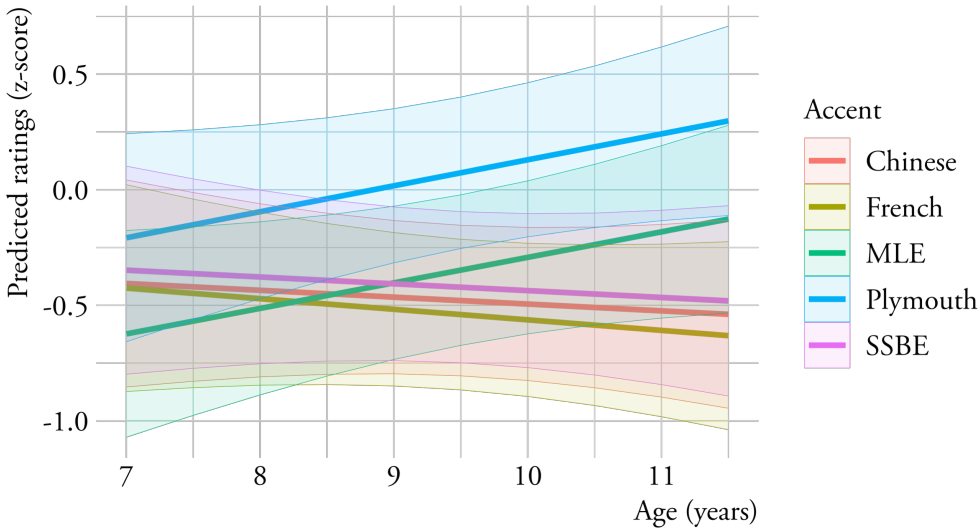


Figure 5.13. Verbal-guise task, FUN, with age, Plymouth

5.3.5 FRIENDLY

The ratings from the FRIENDLY variable were not subject to significant effects of accent ($F(4,27.13) = 1.12$, $p = .369$), age ($F(1,106.82) = 0.30$, $p = .583$) or their interaction ($F(4,415.16) = 0.44$, $p = .780$). See Figure 5.14 for the predicted ratings by the model. As in the case of MONEY above, children from Plymouth across all ages rated speakers in a similar fashion regardless of accent.

5.3.6 TRUST

The final question, TRUST, did not provide significant effects of accent ($F(4,16.03) = 0.44$, $p = .776$) or age ($F(1,107.8) = 1.69$, $p = .195$). However, their interaction almost reached significance ($F(4,424.58) = 2.18$, $p = .0707$); see Figure 5.15. This may be explained by the responses of older children. From the age of 11 years, they rated French-accented English and SSBE higher than the Plymouth accent. Moreover, from the age of 11 and a half years, they also rated these two accents higher than MLE.

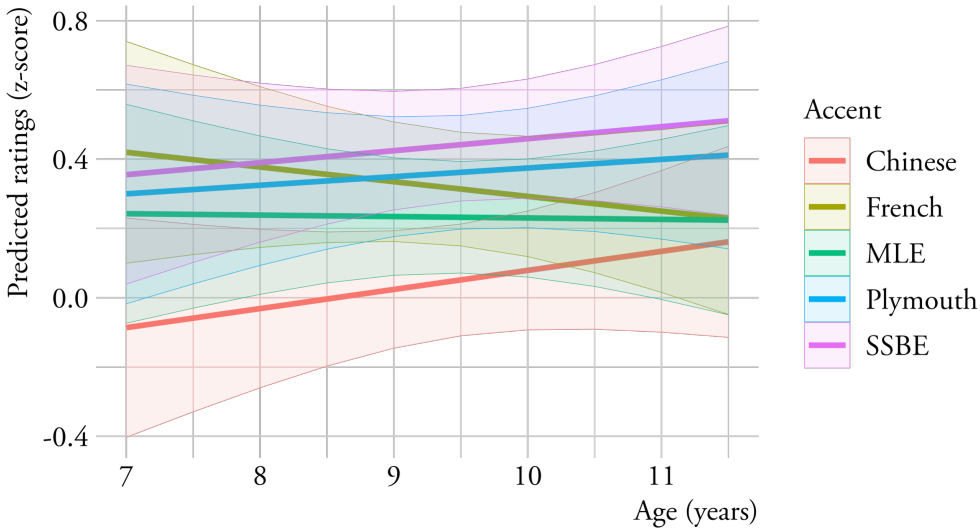


Figure 5.14. Verbal-guise task, FRIENDLY, with age, Plymouth

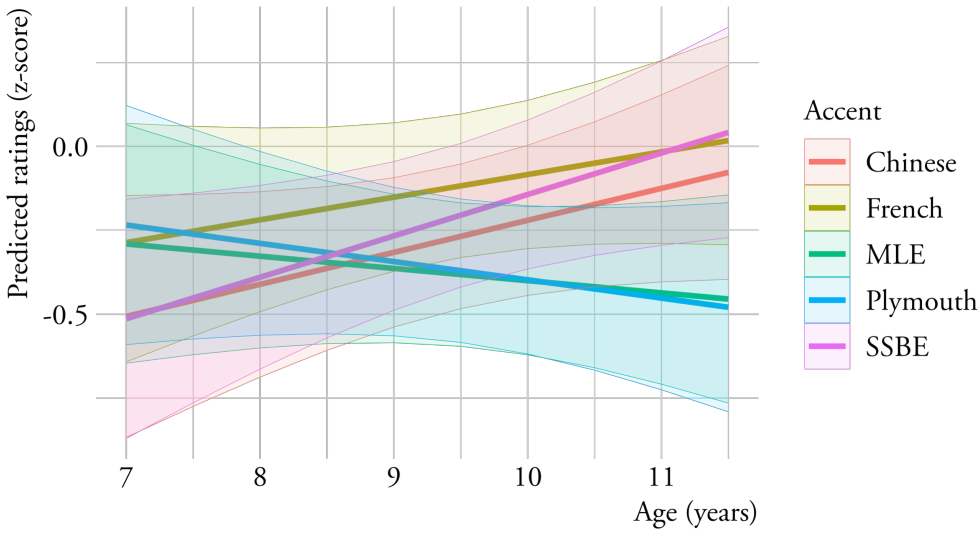


Figure 5.15. Verbal-guise task, TRUST, with age, Plymouth

5.4 Discussion

In this section, the findings from the three previous sections are dealt with. First, each task is discussed individually in the order in which it was presented: the intelligibility

task first, followed by the categorization and, lastly, the main focus of the investigation, namely the verbal-guise task. Afterwards, Section 5.4.1 deals with how the findings from Plymouth inform RQ2 (whether children develop language stereotypes toward L2 accents at an earlier age than toward L1 accents). Finally, Section 5.4.2 sheds light on RQ3 and RQ4, that is, whether there is a relationship between the development of accent stereotypes and the ability to categorize and understand them.

First, Plymouth children had the least difficulty understanding speech-in-noise when it was realized in their local, Plymouth accent and SSBE. This is not surprising. With regard to the local, Plymouth accent, speakers' linguistic processing is tailored to their native accent in the same way it is tailored to their native language. Despite listening to different L1 accents of their native language, speakers have more fine-tuned resources when processing their native accent (Cristia et al., 2012; Cutler, 2012; Dossey et al., 2023). This likely facilitated the processing of speech-in-noise for children from Plymouth when it involved their local accent. As concerns the SSBE accent, its scores may be explained by two factors. First, SSBE has a strong presence in media and education in England. Second, the Plymouth participants' parents likely spoke accents closer to SSBE than to the local Plymouth accent. This is because a significant portion of the participants from Plymouth came from middle- to upper-middle-class families (see Section 4.4) and, in the UK, there is a close relationship between accent and SES (see Section 2.1). These two factors may increase the amount of exposure children from Plymouth receive to SSBE and, as a consequence, their proficiency in it[42].

The accents that scored below the Plymouth and SSBE accents were all nonnative. First, French-accented English received very similar scores to the Plymouth and SSBE accents despite being an L2 accent. In fact, the scores for French-accented English were not distinguishable from those for SSBE or the Plymouth accent[43]. Chinese-accented English and MLE (an L2 and L1 accent respectively) achieved the lowest scores, with MLE also being close to significance to French-accented English. A clear hierarchy of accents was thus present in the data. However, there was not a clear-cut divide in accents when it came to intelligibility. That is, no statistically unequivocal clusters of accents were found.

In a parallel way, the data from the categorization task provided a clear hierarchy of accents without straightforward groups. SSBE scored highest while Chinese-accented English and MLE scored lowest. The main difference between the two tasks was the placement of the Plymouth accent and French-accented English. While the differences between these two accents were not statistically significant in either task, the difference between them in the intelligibility task was greater than in the categorization task.

[42] In fact, the survey data in Section 4.4 reveals that the Plymouth and SSBE accents were the ones the participants from Plymouth listened to the most on a daily basis.

[43] The difference between French-accented English and the Plymouth accent almost reached significance. See Section 5.1.

Furthermore, the average score for the Plymouth accent was higher than that for French-accented English in the intelligibility task, while lower than that for French-accented English in the categorization task. Similarly, French-accented English differed significantly from MLE and Chinese-accented English in the categorization task, but not in the intelligibility task.

Despite these differences, the results from the intelligibility and categorization tasks showed similar overall patterns. Both showed clear developmental and accent effects but no interactions. This is perhaps not unexpected; these two tasks have sometimes been labeled as objective measures of sociolinguistic development (e.g. Dossey et al., 2020). There are two main reasons for this. First, they both have correct answers[44]. That is, when participants heard the word *alligator* in noise, the correct answer was to click on the picture depicting an alligator. Similarly, the correct answers for the categorization tasks involved putting the two speakers of each accent in the same, individual group. Second, they directly rely on linguistic skills. Therefore, it can be posited that, as children's linguistic abilities improve during late childhood (e.g. Nippold, 2004), they are able to perform better in these tasks.

The results of these two tasks displayed similar patterns to those found in the literature. Clopper and Wagner (2019) and Bent et al. (2021) also showed age and accent effects for their intelligibility tasks, without an interaction between the two. In addition, in their free categorization task, Jones et al. (2017) found these same effects with a lack of interactions (see McCullough et al., 2019b for similar results with accent perception tasks). Moreover, Dossey et al. (2020) also reported a gradual improvement in this age range for their intelligibility and classification tasks. In fact, the ranking of the selection of accents in Dossey et al. (2020) (different L1 accents from the US) also differed across tasks. For example, the Midland accents ranked among the highest on their intelligibility task, while it scored among the lowest on their classification task. This matches how the accents in this investigation ranked differently in the intelligibility and categorization tasks.

Of particular relevance in these respects are the results from Jones et al. (2017) and Dossey et al. (2020). Both studies used the same free classification task employed in this thesis, as described in Section 4.1.3. Therefore, the similarity in patterns can be argued to reveal a general developmental effect rather than being an artifact of the specific accents included in these studies[45].

Turning to the results of the verbal-guise task, the findings were more complex. In this case, there were no main age or accent effects. Of the six questions included in the study, two gave rise to significant age and accent interactions, two provided individual significant pairwise comparisons and two generated non-significant results across the board. Chronologically speaking, Plymouth children did not show signs of accent

44 Note that this is not how it was framed for participants; see Section 4.1.

45 But see Section 7.4 for a comparison concerning the London data.

stereotypes between the ages of 7 and 9 years. That is, they rated speakers of different accents similarly within each question. Nevertheless, after this age, accent became a relevant factor when rating speakers on some variables (see Section 7.4 for a visual representation of the interaction between age and accent ratings):

- From the age of 9 years, they rated (i) speakers of French-accented English as smarter than speakers of the MLE and Plymouth accents and (ii) speakers of SSBE as smarter than the Plymouth accent.

- From the age of 9 and a half years, they rated SSBE speakers as smarter than MLE speakers. Furthermore, French-accented English and SSBE speakers were rated as more hardworking than MLE and Plymouth speakers. Finally, Plymouth speakers were rated as more fun than French- and Chinese-accented English speakers.

- From the age of 10 years, Chinese-accented English was rated as smarter than MLE.

- From the age of 10 and a half years, the Plymouth accent was rated as more fun than SSBE.

- From the age of 11, Chinese-accented English was perceived as more hardworking than the Plymouth accent. Moreover, French-accented English and SSBE were considered more trustworthy than the Plymouth accent.

- From the age of 11 and a half years, speakers of French-accented English and SSBE were also rated as more trustworthy than speakers of MLE.

The fact that it was two of the prestige variables — SMART and HARDWORKING — that provided statistically significant interactions is perhaps expected based on the literature. On the one hand, sociolinguistic research demonstrates that, even among adult participants, prestige variables are more robust and consistent across studies. In contrast, results from warmth variables are typically less congruous (see discussions in e.g. Imuta & Spence, 2020; McCullough et al., 2019a; and results from e.g. Nesdale & Rooney, 1996). For instance, while speakers of standard accents are typically rated as being more intelligent than speakers of non-standard accents (see Section 2.3), it is not always the case that speakers of non-standard accents are rated higher in warmth ratings, such as friendliness or likability. As an example, one can compare the results from Ladegaard (1992), where the non-standard Danish accents were rated low in prestige but high in solidarity, to Arthur et al. (1974), in which speakers of GA were rated higher in both sets of variables over non-standard speakers. The fact that variables

involving prestige are more solid in adult populations may make it easier for children to learn these associations. This can explain why these two questions were the ones that showed significant interactions.

On the other hand, SMART is one of the first sociolinguistic variables children consistently learn to associate with accents, as the summary of the literature in Section 3.3.3 showed. Four-year-olds in Dossey et al. (2020) rated speakers differently based on smartness (and friendliness) first. This was also true of the participants in McCullough et al. (2019a) at this age. At the ages of 5 to 6 years, children in Satraki (2015), Santhanagopalan et al. (2021) and Kinzler and DeJesus (2013a, 2013b) also rated speakers differently on intelligence (together with kindness in Santhanagopalan et al. (2021) and niceness in Kinzler and DeJesus (2013b)). Since children appear to be sensitive to the SMART variable from a very early age, it is not surprising it as one of the first variables with clear effects in the data in this investigation.

With regard to HARDWORKING, its similarities in results with SMART imply that children perceived these categories in a similar fashion. Nevertheless, it showed significant results at a later age, when participants were 9 and a half years old. This could be argued to show that SMART is a more basic concept for children. In fact, a similar result could be said to be present in Satraki (2015). In their results, children were sensitive to their SMART variable at an earlier age than for their LAZY variable (LAZY could be said to measure a similar, if opposite, concept to our HARDWORKING). Needless to say, even though these concepts can be construed to be each other's opposites, it is not possible to posit that they are measuring the same construct or that children interpret them in a similar way.

One aspect that distinguishes the results from the verbal-guise task in Section 5.3 and those from the literature is the participants' ages. This thesis includes participants from the age of 7 years. Yet it is only from the age of 9 years when Plymouth participants use SMART for their stereotyped responses to accents. Why participants in this study learned these associations at a later age may be understood by comparing the results to Kinzler and DeJesus (2013b). In their study, 5- to 6-year-olds had accent stereotypes if they were from the northern US, where the standard accent is spoken. When their participants were from the southern US and spoke with an accent low in prestige, they did not have signs of accent stereotypes until the age of 9 to 10 years. Just like the southern accent in the US, the Plymouth accent is a low prestige accent in the UK (see Section 4.2.1). Thus, considering the participants' native accent, the results in this study do in fact match those in Kinzler and DeJesus (2013b): both groups from low-prestige accent regions started rating accents differently on the SMART question from the age of 9 years.

Finally, two further characteristics of these results are worth mentioning. First, the fact that the SMART and HARDWORKING ratings changed through development and were not stable associations also matches results from Dossey et al. (2020) and McCullough et al. (2019a). These two studies showed that the development of

sociolinguistic evaluations is a long process that extends into puberty. This is also present in our data. The SMART ratings became stable at the age of 10 years and those of HARDWORKING at the age of 11. Second, the last prestige variable, MONEY, did not show any significant results. This can be argued to parallel the results found in Dossey et al. (2020). In their study, their RICH variable — which can be argued to be comparable to the MONEY question in the present investigation — did not become significant until the age of 12. Since this age range is outside our participant pool, the lack of significant effects is perhaps not surprising.

The results from the warmth ratings are discussed below. As mentioned above, the less coherent nature of these variables, even in adult populations, makes direct comparison of results across the literature more difficult. FUN and TRUST were the variables that showed some significant pairwise comparisons. Of these two, FUN began being a relevant variable at an earlier age, at the age of 9 and a half years. Just as in the case of prestige variables, FUN becomes significant in the literature at earlier ages. For example, Giles et al. (1983) showed that children between the ages of 7 and 8 years rated accents differently on their FUNNY question[46]. Furthermore, Satraki's (2015) FUNNY variable also provided significant results in this age range. With regard to TRUST, it also became a relevant variable later on in development in McCullough et al. (2019a), with their somehow comparable HONEST category. Nevertheless, in their data, children began rating certain speakers as more honest between the ages of 8 and 9 years. In contrast, it is the 11-year-olds in this study that first rated accents differently on this question.

These age differences may also be the result of our population of participants, as discussed above for the prestige variables. Kinzler and DeJesus (2013b) did not include this variable. Therefore, it is not possible to establish the same kind of comparison at this point. In terms of the last solidarity variable (FRIENDLY), participants rated all speakers equally. Two other studies have included the same question. First, Dossey et al. (2020) showed that children rated accents differently on FRIENDLY from the age of 4 to 5 years. In contrast, the data in McCullough et al. (2019a) did not show children rating different accents as more friendly than others until the age of 8 to 9 years.

5.4.1 Accent types: native, nonnative, L1 and L2

In this section, the value of a categorical divide between L1 and L2 accents for research within developmental sociolinguistics is discussed. The data from the intelligibility and categorization tasks is dealt with first. This is followed by a discussion about whether children develop stereotypes toward L2 accents first and whether the L1-L2 divide is of theoretical importance for understanding children's accent associations.

46 But see footnote 13 in Section 3.3.3

First, the results from the intelligibility task revealed that children's understanding of accents in late childhood is not dependent on whether they were native or nonnative accents. In fact, not even a strong, categorical distinction between L1 and L2 accents seemed to fully explain the data. On the one hand, the accents children had the least difficulty understanding included the native Plymouth accent (the native accent) and SSBE and French-accented English (two nonnative accents). Thus, it cannot be concluded that children performed better with their local, Plymouth accent than with other accents. Arguably, the important role of SSBE in media, education and British society as a whole (see Section 4.2.1) could suggest that it should be considered to be positioned between a native and nonnative accent (see also the discussion in Section 7.4.1). Nevertheless, the difference in scores between SSBE and French-accented English was not statistically significant and that between Plymouth and French-accented English was only close to significant (p = .088). Therefore, a native/nonnative divide does not explain our results regardless of what accent is chosen as the native one.

On the other hand, the data was not explained by a L1/L2 divide. As the previous discussion illustrates, French-accented English and SSBE (an L2 and L1 accent, respectively) obtained comparable results. Similarly, participants also obtained matching results for MLE and Chinese-accented English, L1 and L2 accents, respectively. In fact, the scores for French-accented English were closer to those for MLE and SSBE than to those for Chinese-accented English. This lack of a clear L1-L2 divide matches the results in Bent et al. (2021) (see also Hanulíková, 2020, where an L2 accent scores higher than an L1 accent). They discussed how, in studies like Bent and Holt (2018), Floccia et al. (2006) and Goslin et al. (2012), L1 accents had an advantage over L2 accents in word recognition tasks. However, such patterns could not be found in Evans and Lourido (2019), Floccia, Butler, Goslin, et al. (2009) or Paquette-Smith et al. (2021). Consequently, they carried out a study in which multiple L1 and L2 accents were included. Furthermore, they also performed a modified version of the Levenshtein Distance Algorithm (Levenshtein, 1965), as used in Levy et al. (2019), to quantify how similar the accents in their stimuli were to each other. Their results showed that it is phonetic distance from the native accent, rather than being L1 or L2, that accounted for the difficulty in understanding other accents. For instance, participants performed better with the L2 German-accented English than with the L1 Scottish English.

The results of the categorization task provided a more intricate picture. First, the accuracy scores paralleled those of the intelligibility task. Neither the native/nonnative categories nor the L1/L2 ones account for the participants' accuracy scores. Children were as good at categorizing the Plymouth accent (their native and L1 accent) as French-accented English (a nonnative and L2 accent) and SSBE (a nonnative and L1 accent). In a similar fashion, MLE and Chinese-accented English obtained similar scores that did not differ from each other. What's more, Plymouth's scores were closer to those of Chinese-accented English than to MLE. In comparison, the results from the

co-occurrence matrices painted a different picture. Throughout the 5-year span in the Plymouth data, children were more likely to categorize speakers of L2 and L1 accents separately. That is, when participants grouped speakers of different accents together in one cluster, they were more likely to include, for example, (i) a speaker of Chinese-accented English and French-accented English than (ii) a speaker of Chinese-accented English and MLE. As shown in Figure 5.8, the structure of the branch for L1 accents did change as a result of age. Seven- and 8-year-olds categorized L1 accents as native (Plymouth) versus non-native (MLE and SSBE). In contrast, the older children, 9- to 11-year-olds, made clusters based on standard-ness: SSBE, on the one hand, and the Plymouth and MLE accents, on the other.

There are no studies including both L1 and L2 accents with a free classification task, to the best of my knowledge. Therefore, a direct comparison of these results to the relevant literature is not possible. However, similarities in patterns do appear. Floccia, Butler, Girard, et al. (2009), Girard et al. (2008), Wagner, Clopper, et al. (2014), Evans and Lourido (2019) and Paquette-Smith et al. (2019) showed that children are able to better distinguish their native accent when contrasted with an L2 accent than with an L1 accent. This matches the results of the co-occurrence matrices, seeing that L1 and L2 accents were more likely to cluster independently of each other. These studies used a binary, forced-choice design. Children had to choose whether a person spoke with their native accent or a different one (either an L1 or L2 accent)[47]. Moreover, they only included two accents besides the participants' native one. Therefore, it is not possible to discern between accuracy and group confusion.

The fact that the accuracy scores in Section 5.2.1 did not reflect such a clear-cut divide between native and nonnative accents is also reflected in Wagner, Clopper, et al. (2014), Jones et al. (2017), McCullough et al. (2019b), Dossey et al. (2020) and Paquette-Smith et al. (2019). Besides the contrast just described, these studies also included a condition in which participants had to distinguish either between (i) a nonnative L1 accent and an L2 accent or (ii) multiple nonnative L1 accents. These studies showed that children find it difficult to distinguish between nonnative accents, regardless of whether they are L1 or L2. For example, participants in Wagner, Clopper, et al. (2014) could distinguish between their native accent (a Midland accent from the US) and the L2 Marathi-accented English. In comparison, if the contrast involved (i) their native accent and Lancashire English or (ii) Lancashire English and Marathi-accented English, participants performed below chance.

Likewise, McCullough et al. (2019b) and Dossey et al. (2020) showed that not all nonnative L1 accents are categorized the same way[48]. For instance, in Dossey et al.

47 There are small methodological differences between these studies. However, the overall designs are comparable.

48 McCullough et al. (2019b) and Jeffries (2019) also showed that children can perceptually differentiate L1 accents at an earlier age than they can categorize them.

(2020), participants between the ages of 4 and 9 years could not properly categorize (i) their native, Midlands accent when compared to the Southern US accent or (i) the Southern accent when compared to the Mid-Atlantic US accent. Nonetheless, they were able to distinguish between the Southern and Northern US accents.

Using a different task design, the data in Weatherhead et al. (2019) is also worth comparing to that in Section 5.2. In their two experiments, children aged 4 to 5 years were presented with speakers of multiple L1 and L2 accents. In the first experiment, they were shown a map of the US[49] with a cross on the upper, left corner. They were told that the cross marked where they lived and that different people lived in different places on the map. After listening to one speaker of each accent, children were asked to mark on the map where they thought each person lived. The researchers' hypothesis was that the more different the accent was from the participants' own, the further away from the cross they would place the speaker on the map. They also tested whether L2 accents would consistently be placed further away than L1 accents. Their younger group of participants placed their native accent closer to the cross than all other nonnative accents. However, there were no differences across the different nonnative accents, whether L1 or L2. Their older group placed the native accent closer to the cross too. In contrast, they placed other nonnative accents at different distances. However, it did not reflect a L1-L2 divide.

In their second experiment, Weatherhead et al. (2019) low-pass filtered their stimuli so that only prosodic information remained[50]. In this case, participants of both ages not only placed the native accent closer to the cross compared to the nonnative accents, but they also placed L1 accents significantly closer to the cross than L2 accents. The results of Weatherhead et al. (2019), together with those in Section 5.2, show that children's abilities to categorize accents based on the L1-L2 distinction is not a yes-or-no issue. It is subject to both the design of the task, the stimuli used as well as the specific analysis employed (as the difference in results between accuracy and the confusion matrices shows).

Finally, with respect to the co-occurrence matrices for the L1 accents specifically, the most comparable study is Jones et al.'s (2017) investigation of L1 accents from the US. The results from their younger participants (6- to 7-year-olds) are partially comparable to the findings in Section 5.2.2[51]. Jones et al. (2017) posited that this younger group classified speakers based on familiarity: on the one hand, those speakers the participants had more exposure to and, on the other, those they were less familiar with. This can

49 A map of the US was used, instead of Canada (where the experiment took place), to avoid participants being familiar with it, as young participants were not expected to be familiar with a US map.

50 They also modified the task slightly to accommodate the change in auditory stimuli.

51 Note that a strict one-to-one comparison is not possible since the confusion matrices in Jones et al. (2017) merged 2-year groups: 6- and 7-year-olds, 8- and 9-year-olds, and 10- and-11-year-olds. In contrast, our results are shown by year.

arguably parallel the fact that our younger (7 to 8 years old) participants structured the L1 cluster with the native accent, on the one hand, and nonnative accents, on the other. Furthermore, from the age of 9 years, the results in both Section 5.2.2. and Jones et al. (2017) show that children categorized accents in a more adult-like fashion that matches social norms.

Below, how the results of the verbal-guise task relate to accent types is discussed with a focus on the significant effects described in Section 5.4. The way children rated different accents from the age of 9 years does not seem to rely on either a L1/L2 distinction or a native/nonnative one. The earliest signs of accent stereotypes were shown with the question SMART. At this point, they rated both an L2 accent (French-accented English) and a nonnative L1 accent (SSBE) as smarter than the native accent. Moreover, French-accented English (an L2 accent) was also rated higher than the nonnative L1 accent MLE. A similar structure was found with HARDWORKING, where an L2 accent and a nonnative L1 accent were rated higher than the native, Plymouth accent at the age of 9 and a half years. A comparable pattern appeared for TRUST, with French-accented English and SSBE being rated higher than the MLE and Plymouth accents from the age of 11 and a half years. The only variable that seemed to follow a L1-L2 split is FUN. From the age of 9 and a half years, the native Plymouth accent was rated higher than both L2 accents in the stimuli.

Crucially, all these ratings paralleled those of adults. For example, French-accented English and SSBE are often given high prestige scores (Bishop et al., 2005; Coupland & Bishop, 2007; Magne, 2019; Sharma et al., 2022). Similarly, the Plymouth accent has been shown to receive higher responses in warmth categories than in prestige ones (Bishop et al., 2005; Coupland & Bishop, 2007; Sharma et al., 2022). Thus, the fact that a split between native accent and nonnative accents for FUN is present in the data from Plymouth may not be the result of children's use of this divide. Instead, they may simply have learned the adult sociolinguistic norm around them.

The main difference between the children's ratings above and those found in the literature with older populations was that of Chinese-accented English. The current literature shows a trend of Chinese-accented speakers being rated low on both prestige and warmth ratings (Bishop et al., 2005; Magne, 2019; Sharma et al., 2022). This attitudinal shift in children may be the result of the new role of Mandarin in education. The British Council and the Institute of Education at University College London identified Mandarin as one of the key languages for the UK's future (Campbell-Cree, 2017; Institute of Education, 2021). Importantly, this new view of Mandarin has led to a significant increase in the number of primary and secondary schools teaching Mandarin as a foreign language, together with French, Spanish and German (Allan, 2023; Department of Education, 2017; HM Treasury, 2015; Institute of Education, 2021; IOE Digital, 2019; Zhu, 2017). Consequently, the new framing of Chinese as an important language for educational and professional development may have led to children viewing Chinese-accented English as an accent high in prestige. In fact, the

notion that this new view of Chinese may reflect a recent change in society could also explain why the association between Chinese-accented English and SMART and HARDWORKING was acquired later during development. The more robust societal views of SSBE and French-accented English as accents with high prestige may make it easier for children to learn. Hence, children were able to reflect this association at an earlier age.

To conclude this section, the results from Plymouth, together with the literature mentioned in this section, do not suggest that the development of sociolinguistic perception in general and accent stereotypes in particular follow a strict L1/L2 distinction or a native/nonnative one. This supports the view posited in Wagner, Clopper, et al. (2014, p. 1079) that accents "are not simply ordered on a single linear continuum but may be presented more multidimensionally."

5.4.2 Accent stereotypes, intelligibility and categorization

This section discusses whether the development of accent stereotypes is related to accent intelligibility and the ability to categorize accents. The results from the intelligibility task are dealt with first.

Section 5.1 above shows that there was a clear developmental effect on how well children understood different accents. This could be argued to parallel the fact that younger children did not show any signs of linguistic biases toward accents, while older children did. However, SSBE and Plymouth were the accents children understood best throughout development while they were rated highest and lowest in SMART and HARDWORKING, respectively. Consequently, there does not seem to be a clear relationship between intelligibility and accent stereotypes. This is further support to Dossey et al. (2020), who posited that children develop both skills in parallel.

The relationship between categorization and stereotyping is split into the accuracy scores (see Section 5.2.1) and the co-occurrence matrices (see Section 5.2.2). The accuracy scores followed a similar pattern to that of the intelligibility task. In contrast, being able to properly categorize accents can be considered a pre-requisite for children to make social judgements about them (see Wagner, Clopper, et al., 2014 and the discussion in Section 1.1). Thus, it could be hypothesized that the data in this chapter shows this link: as children become better at categorizing accents, they become able to attach different stereotypes to them. However, the Plymouth data is not completely conclusive in this regard. A strong version of this relationship does not appear to be supported. If this were the case, one would expect the first accents children show stereotypes toward to be the accents they can categorize correctly first. For instance, if children were able to properly categorize speakers of SSBE first, one would expect they would also be able to attach different stereotypes to this accent at an earlier age. This does not appear to be the case. In this investigation, the first stereotypical associations involved the accents of Plymouth and MLE being rated lower than SSBE and French-

accented English. However, MLE was the accent all children had the most difficulty categorizing. As a result, one cannot posit that those accents children categorize best are those they first develop accent stereotypes toward during development. Potentially, a weaker version of the relationship between categorization and stereotyping can still be maintained. Namely, that a general level of categorization skills is required for children to be able to judge accents differently. Nonetheless, the data in this chapter does not allow for a clear answer (but see Chapter 7).

The results depicted in the confusion matrices (Section 5.2.2) can be argued to be parallel to those of the verbal-guise data. As mentioned in Section 5.4.1, the accent stereotypes the Plymouth participants had matched the general sociolinguistic norm in England (SSBE and French-accented English rated higher than the Plymouth and MLE accents). In a parallel way, their co-occurrence matrices show that, from the age of 9 years (the same age when they start reproducing accent stereotypes), they categorize L1 accents according to the social structure of accents: the standard SSBE accent, on the one hand, and Plymouth and MLE, on the other. Before this age, they organized this cluster into native (Plymouth) and nonnative (SSBE and MLE) accents. Thus, there is a relationship between children being able to sort accents according to their social status and judging them based on adult sociolinguistic attitudes.

In conclusion, a relationship between intelligibility and accent stereotypes does not find support in this study: the accents children understood best received both positive and negative evaluations. In contrast, a relationship between categorization and stereotyping is more likely. Children did not show stereotypes toward those accents they categorize first, as a strong version of this relationship would imply. It could be argued, however, that simply achieving a certain level of skill for categorizing accents is required for children to be able to reproduce accent stereotypes. Furthermore, the change in the structure of the clusters children made between the ages of 7 and 8 years, on the one hand, and 9, 10 and 11 years, on the other, did match the age at which they started showing signs of stereotypes. Thus, there is reason to believe there is a closer relationship between them than between intelligibility and stereotyping.

5.5 Summary

This chapter describes the results from the data collection in Plymouth (Southwest England). First, the results from the intelligibility task show that children become better at understanding speech in noise the older they get. Moreover, they understand some accents (the Plymouth and SSBE accents as well as French-accented English) better than others (Chinese-accented English and MLE). Next, the data from the categorization task gives rise to a similar pattern. The older participants get, the better they perform on this task. How well they categorize accents does not change with age. However, what speakers they put together in the same groups is subject to

developmental effects. Younger participants group the Plymouth and SSBE accents together more frequently than with MLE, while older participants tend to group MLE together with the Plymouth accent more frequently than with SSBE. Finally, with regard to the verbal-guise task, two questions provide significant interactions. These are the SMART and HARDWORKING questions. Up until the age of 9 years, children from Plymouth do not rate speakers differently on these questions. After this age, they start rating French-accented English, SSBE and, eventually, Chinese-accented English higher than MLE and the Plymouth accent.

6 Accent stereotypes in London, a linguistically heterogeneous city

This chapter describes and discusses how and when children from London develop accent stereotypes and what factors might affect this development. This depiction is based on the data collected using the methods described in Chapter 4. This chapter is therefore parallel to Chapter 5, where the focus was on the data from Plymouth. First, the results of the two auxiliary tasks are described: the intelligibility task in Section 6.1 and the categorization task in Section 6.2. These two are followed by the findings of the thesis's main task, the verbal-guise, in Section 6.3. The chapter closes with a discussion of the results (Section 6.4). First, each task is dealt with individually, followed by how the London data informs the RQs introduced in Section 1.1. The next chapter compares the findings from both Plymouth and London to obtain a deeper insight into the development of stereotypes to accents in childhood.

6.1 Intelligibility task

In this task, participants were asked to listen to a word with background noise and choose a picture that matched the word they heard (see more details in Section 4.1.1). The findings from this task are structured in two ways: by word and by accent. First, the number of correct answers per word is shown in Table 6.1. The percentage column was calculated based on the number of times each word was presented to participants. Each participant heard each word once. Since there are a total of 105 participants from London, the maximum number of correct answers for each word is 105 (one per participant). The results in Table 6.1 provide a clear hierarchy of word accuracy. Some words were chosen correctly most of the time, such as *church* with 96.19% right answers and alligator with 92.38%. On the opposite end, some words were only answered correctly around a third of the time, such as thorn and palm with 33.33% and 30.48% correct answers, respectively. These results show that some words were more difficult to understand than others when masked with noise.

Second, the data was analyzed by merging correct answers across accents. The number of correct answers per accent is shown in Table 6.2. In a similar way to the by-

word results, the percentages were calculated based on the total number of times London participants listened to words in each accent. The maximum possible number of correct answers in London is 420 (105 participants times four different words per accent). A gradient of accents is present. London children had the least difficulty understanding the Plymouth and French-accented English accents, both of which obtained the same number of correct answers. SSBE followed in terms of correct responses, succeeded by MLE and, in last place, Chinese-accented English.

Table 6.1. Intelligibility task, London, correct answers per word
Correct answers (raw and percentage) per word in London. The maximum number of correct answers per word is 105.

Target word	Correct answers	Percentage
church	101	96.19
mouse	98	93.33
alligator	97	92.38
house	96	91.43
wolf	95	90.48
bird	94	89.52
leaf	90	85.71
branch	89	84.76
porridge	86	81.90
farm	85	80.95
foot	82	78.10
salt	81	77.14
key	79	75.24
toy	77	73.33
goat	76	72.38
bowl	65	61.90
face	64	60.95
goose	61	58.10
thorn	35	33.33
palm	32	30.48

Table 6.2. Intelligibility task, London, correct answers per accent
The maximum number of correct answers per accent is 420

Accent	Correct answers	Percentage
Plymouth	339	80.71
French	339	80.71
SSBE	324	77.14
MLE	304	72.38
Chinese	277	65.95

Table 6.2 shows differences across accents. But it does not show whether these differences are statistically significant, whether there are developmental effects or an interaction between age and accent. To test this, a linear mixed-effects model (see

Section 4.5) was employed. The dependent term was the accuracy score per participant per accent. The independent variables were age and accent. These two terms were included with an interaction between them. Moreover, two different random effects were included, individual speaker per accent and participant ID. An analysis of variance provides a significant effect of age ($F(1,104) = 4.44$, $p = .038$) as well as of accent ($F(4,1976) = 45.02$, $p < .001$). The model does not find an interaction between these two ($F(4,1976) = 0.58$, $p = .674$). The predicted values of this model are presented in Figure 6.1. The X-axis depicts age and the Y-axis represents accuracy. Bold lines represent the mean accuracy predicted for a given age; the ribbons show the 95% confidence intervals predicted by the model.

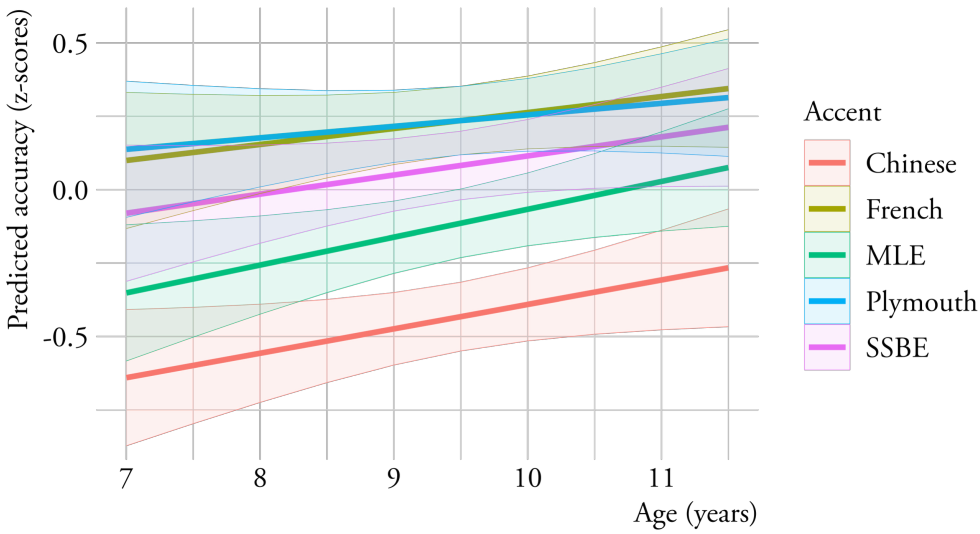


Figure 6.1. Intelligibility results, with age, London

As the statistical significance of the developmental effect demonstrates, the accuracy of younger participants was lower (M_{7yo} = -0.17, 95% CI_{7yo} = [-0.40, 0.07]) than that of older ones (M_{11yo} = 0.10, 95% CI_{11yo} = [-0.10, 0.30]). The effect of accent on accuracy remains consistent throughout development. Consequently, a model was run with age and accent as predictors but without an interaction between them. The other terms in the model were the same as in the model including the interaction. To obtain a clearer overview of how different accents affected children's intelligibility scores, Figure 6.2 depicts the results of this new model. In this case, the predicted average accuracy across all ages is presented on the Y-axis. These scores are shown against the accents included in the stimuli.

The ranking of the accents in Figure 6.2 mirrors that of Table 6.2: French-accented English and Plymouth received the highest scores, followed by SSBE, then MLE and,

finally, Chinese-accented English. To test which individual differences in scores were significant, pairwise comparisons (see Section 4.5) were performed.

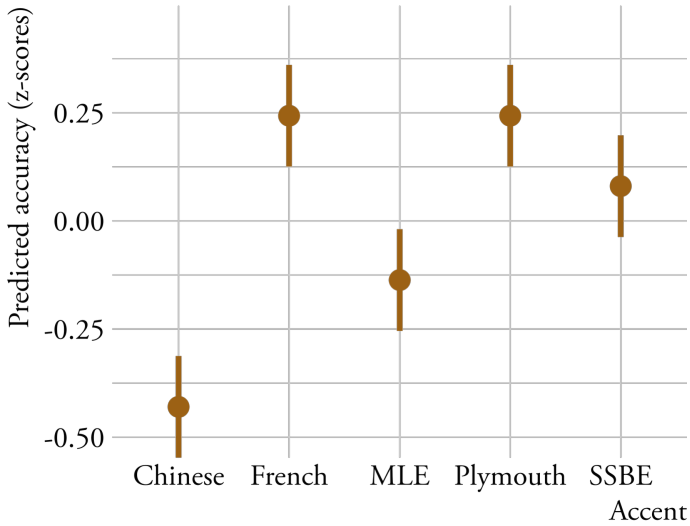


Figure 6.2. Intelligibility results, without age, London

These pairwise comparisons revealed that the ranking of accents can be organized in three main clusters. First, Chinese-accented English (M_{chi}= -0.43; *95% CI*$_{chi}$ = [-0.55, -0.31]) was rated significantly lower than all other accents ($t_{chi\text{-}fre}$(5.02) = -11.31, p < .001; $t_{chi\text{-}ply}$(5.02) = -11.31, p < .001; $t_{chi\text{-}mle}$(5.02) = -4.93, p = .023; $t_{chi\text{-}ssb}$(5.02) = -8.58, p = .002). Second, MLE obtained the second lowest scores (M_{mle} = -0.14; *95% CI*$_{mle}$ = [-0.25, -0.02]). Besides MLE's difference with Chinese-accented English, MLE's score was also significantly different from the Plymouth accent ($t_{mle\text{-}ply}$(5.02) = 6.39, p = .008) and French-accented English ($t_{mle\text{-}fre}$(5.02) = 6.39, p = .008). With respect to SSBE (M_{ssb}= 0.08; *95% CI*$_{ssb}$ = [-0.04, 0.20]), the difference between MLE and SSBE was not significant ($t_{mle\text{-}ssb}$(5.02) = -3.65, p = .070). Nonetheless, the difference between SSBE and MLE was greater than that between SSBE and either the Plymouth accent or French-accented English. Third, Plymouth (M_{ply} = *0.24; 95% CI*$_{ply}$ = [0.13, 0.36]) and French-accented English (M_{fre} = 0.24; *95% CI*$_{fre}$ = [0.13, 0.36]) achieved the highest scores. Their scores did not differ from each other ($t_{ply\text{-}fre}$(5.02) = 0.01, p = .999). Furthermore, neither of them had a statistically significant difference from SSBE ($t_{ply\text{-}ssb}$(5.02) = 2.74, p = .177; $t_{fre\text{-}ssb}$(5.02) = 2.74, p = .177).

Shortly, how well children understood words in noise can be described as involving three categories. They obtained the lowest scores when the stimuli were presented in Chinese-accented English. The second-to-lowest scores were obtained with the MLE stimuli. Finally, the best scores were achieved with the Plymouth, French-accented

English and SSBE accents. Nevertheless, the difference between SSBE and MLE was not significant, though it bordered on significance.

6.2 Categorization task

The categorization task asked participants to listen to two speakers of each accent and place them in groups if they thought the speakers were born in the same city (see Section 4.1.3 for more details). How London participants solved this task is split into two sections. Section 6.2.1 describes their accuracy scores. Section 6.2.2 deals with how participants grouped the individual speakers.

6.2.1 Accuracy scores

In this context, accuracy means how well participants grouped the ten speakers into their five corresponding groups, following the method used in Dossey et al. (2020) (see Section 4.5.1).

The average score by accent is shown in Table 6.3. French- and Chinese-accented English showed the highest overall accuracy. In contrast, participants from London had the most difficulty categorizing MLE speakers. The scores pertaining to the Plymouth and SSBE accents were in between these two poles. The data was analyzed using a linear mixed-effects model (see Section 4.5). The dependent variable was accuracy. The predictors were age and accent with an interaction between them. Finally, a random effect of participant was included. In this case, it was not possible to include a random effect for individual speakers, since the accuracy score was calculated by accent and it generalized over both speakers (see Section 4.5.1).

Table 6.3. Categorization task, London, average scores
The scores span from -100 to 100

Accent	Average score
French	55.60
Chinese	43.10
Plymouth	36.55
SSBE	34.94
MLE	20.54

An ANOVA shows that there was a significant effect of accent ($F(4,936) = 5.61$, $p < .001$), whereas there was no overall developmental effect ($F(1,104) = 1.92$, $p = .168$) or an interaction between the two ($F(4,936) = 0.985$, $p = .414$). The predicted results from the model can be seen in Figure 6.3. The X-axis depicts age, while the Y-axis

represents accuracy. Bold lines represent the predicted mean score for a given age, while ribbons indicate the 95% confidence intervals.

As the lack of a developmental effect above implies, Figure 6.3 shows that the scores for all accents, except Chinese-accented English, did not change significantly across development. In contrast, the scores obtained with Chinese-accented English did improve with age. A post-hoc analysis revealed that this developmental change for Chinese-accented English was statistically significant ($t(413.88) = 2.33$, $p = .021$). This means that, while the overall scores were similar for the younger participants (M_{7yo} = -0.11, *95% CI*$_{7yo}$ = [-0.35, 0.13]) and the older ones (M_{11yo} = 0.08, *95% CI*$_{11yo}$ = [-0.12, 0.28]), the ability to categorize Chinese-accented English was weaker at the age of 7 years (M_{7yo} = -0.18, *95% CI*$_{7yo}$ = [-0.47, 0.10]) than at the age of 11 years (M_{11yo} = 0.29, *95% CI*$_{11yo}$ = [0.08, 0.50]).

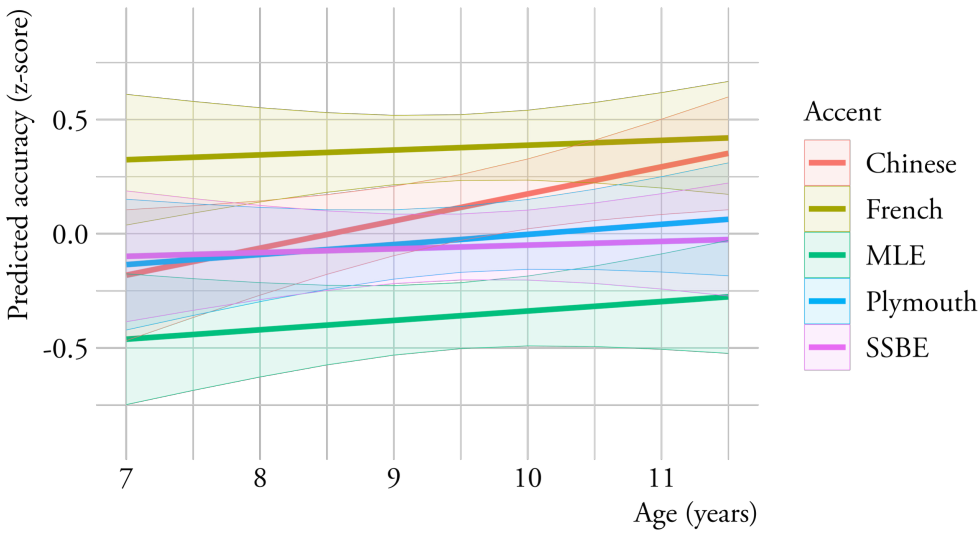


Figure 6.3. Categorization results, with age, London

Put differently, the ranking of Chinese-accented English with respect to the other accents changed with age, while the ranking of the other four accents remained stable throughout development. A summary of developmental changes of Chinese-accented English in relation to the other four accents can be seen in Figure 6.4. The horizontal axis depicts age in years, while the vertical axis shows which accents pairwise comparisons were performed with. Blue-shade regions indicate that the difference between Chinese-accented English and the accent in the relevant row is statistically significant. At the age of 7 years, Chinese-accented English only differed significantly from French-accented English. Between the ages of 7 and a half and 9 years, it differed from both French-accented English and MLE. At the age of 9 and a half, its score was also significantly different from SSBE, together with the previous accents. At the age of

10 years, Chinese-accented English differed from all other accents in a significant way. Finally, from the age of 10 and a half years onwards, the scores for Chinese-accented English were no longer significantly different from French-accented English, while still differing from scores for the MLE, SSBE and Plymouth accents in a statistically significant way.

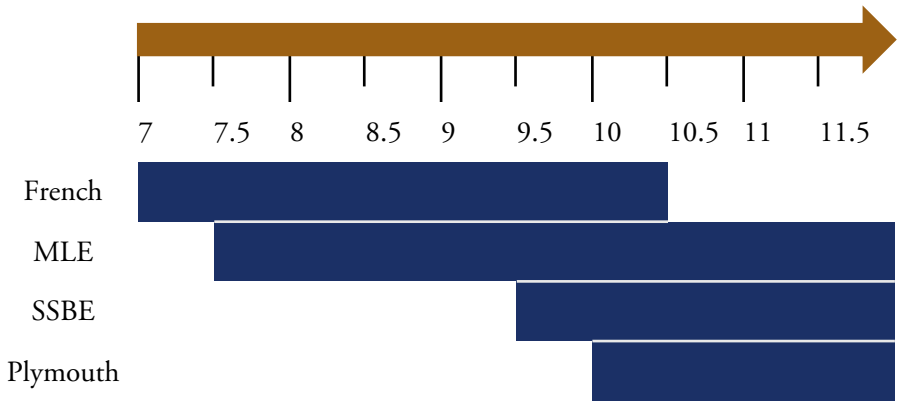


Figure 6.4. Accent differences with Chinese-accented English in the categorization task
Numbers indicate age in years; blue regions indicate that Chinese-accented English is significantly different from the relevant accent.

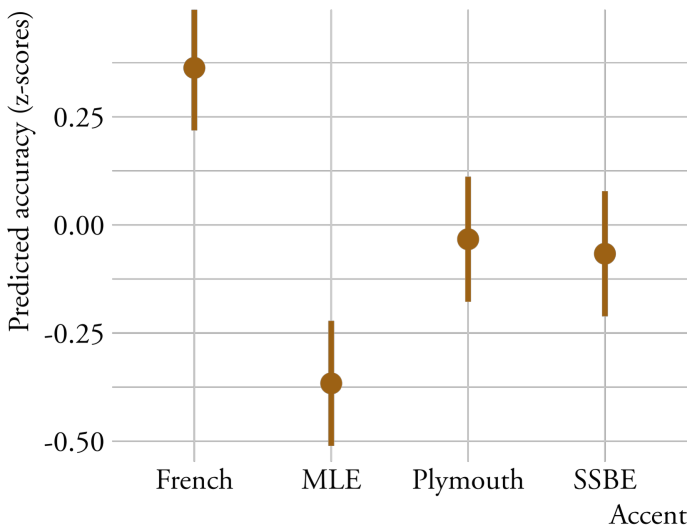


Figure 6.5. Categorization results, without age, London

To analyze the differences across the accents that are not subject to age effects (French-accented English, Plymouth, MLE and SSBE), a model without an interaction between

age and accent was run. The rest of the terms were the same as in the model described earlier in this section. Figure 6.5 shows the predicted average scores across all ages from this model. In this plot, the predicted mean accuracy and 95% confidence intervals are presented against the accents included in the stimuli (without including Chinese-accented English).

The predicted scores in Figure 6.5 were analyzed to test whether the differences were significant. For this purpose, pairwise comparisons (see Section 4.5) of the four accents uncovered two main constellations. First, French-accented English (M_{fre} = 0.36; *95% CI*$_{fre}$ = [0.22, 0.51]) was significantly different from all other accents ($t_{fre\text{-}mle}(5.04) = 8.65$, $p = .002$; $t_{fre\text{-}ply}(5.04) = 4.70$, $p = .027$; $t_{fre\text{-}ssb}(5.04) = 5.10$, $p = .019$). Second, the Plymouth accent (M_{ply} = -0.03; *95% CI*$_{ply}$ = [-0.18, 0.11]) and SSBE (M_{ssb} = -0.07; *95% CI*$_{ssb}$ = [-0.21, 0.08]) did not differ from each other ($t(5.04) = 0.40$, $p = .993$). Finally, the difference in scores between MLE (M_{mle} = -0.37; *95% CI*$_{ssb}$ = [-0.51, -0.22]) and the Plymouth accent was not significant ($t_{mle\text{-}ply}(5.04) = -3.95$, $p = .052$); in a similar way, the difference between MLE and SSBE did not reach significance ($t_{mle\text{-}ssb}(5.04) = -3.56$, $p = .076$). It is important to note that the difference between Plymouth and SSBE was smaller than that between either of these two accents and MLE. In both cases (Plymouth–MLE and SSBE–MLE), the differences almost reached statistical significance.

6.2.2 Co-occurrence matrices

A co-occurrence matrix was created to analyze *how* London participants grouped speakers, not only how accurate the groups were. For this purpose, the number of times an individual speaker was grouped with each of the other speakers was calculated. The result of such a co-occurrence matrix for the London data is available in Figure 6.6. Each cell indicates the percentage of times the accents in the relevant row/column were grouped together. The percentages need to be understood in a by-row fashion. Thus, out of all the times an MLE speaker was placed in a group with other speakers, 5.71% was with a speaker of Chinese-accented English, 4.29% with a French-accented English speaker, 34.29% with another MLE speaker, etc. The distribution of the co-occurrence matrix with the raw values was analyzed using a chi-square test of independence. This test indicates that the pattern in Figure 6.6 was significantly different from chance ($X^2(16) = 905.85$, $p < .001$).

A clearer picture of the hierarchical structure of the groups in Figure 6.6 was produced with the R package *pheatmap* (Kolde, 2019). Its results can be seen in Figure 6.7. This dendrogram illustrates two main groups. One group includes French- and Chinese-accented English, while the other includes SSBE, MLE and the Plymouth accent. The latter group contains a deeper structure. SSBE and MLE were grouped with each other more frequently than with the Plymouth accent.

Section 6.2.1 shows that only Chinese-accented English is subject to developmental effects. Therefore, year-by-year co-occurrence matrices were created to see whether the accents participants included with Chinese-accented English also changed with age. Figure 6.8 depicts these matrices.

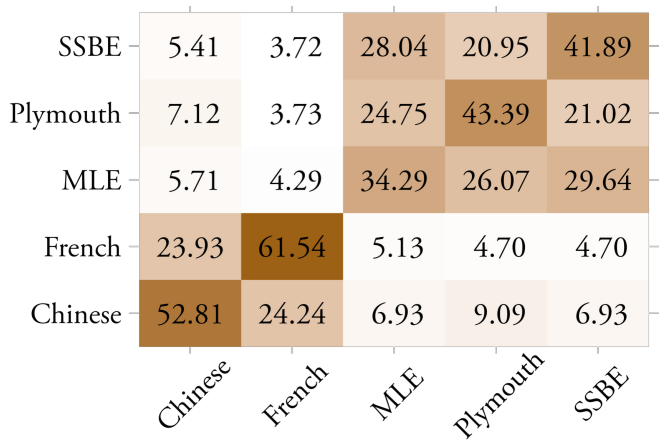


Figure 6.6. Co-occurrence matrix, without age, London
Values in percentage row-wise. The darker the color of a cell, the higher the percentage of groups.

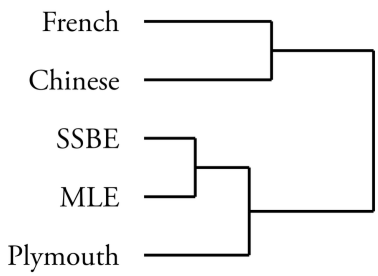


Figure 6.7. Co-occurrence dendrogram, without age, London

All five sub-plots were statistically different from chance after employing chi-square tests of independence on each sub-plot with the raw counts (rather than the percentages in the figures) ($X^2_{7yo}(16) = 88.57$, $p < .001$; $X^2_{8yo}(16) = 212.02$, $p < .001$; $X^2_{9yo}(16) = 272.01$, $p < .001$; $X^2_{10yo}(16) = 201.91$, $p < .001$; $X^2_{11yo}(16) = 205.89$, $p < .001$). They reveal that, despite the change in accuracy through the five-year span, Chinese-accented English was more frequently clustered with French-accented English throughout development.

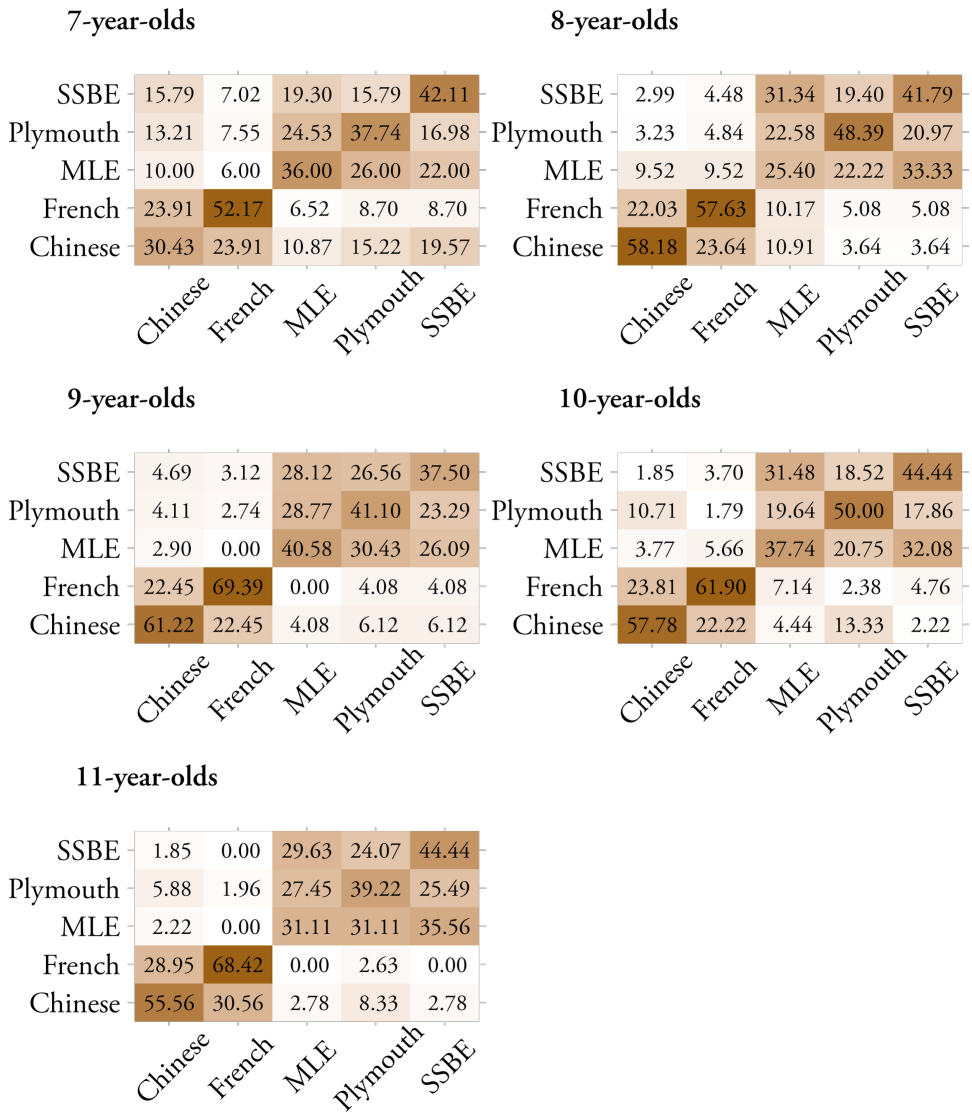


Figure 6.8. Co-occurrence matrices, with age, London
Values in percentage row-wise. The darker the color of a cell, the higher the percentage of groups.

6.3 Verbal-guise task

In this section, the results of the main task of this study, i.e. the verbal-guise task, are described. As a reminder, participants heard one speaker of each accent read the same

story out loud and were asked some questions about them, like *how hardworking do you think this person is?* (see Section 4.1.2 for details). The results of this task are analyzed in two ways. First, one analysis tests whether children rated speakers differently by question. Second, within each question, an analysis of the ratings tested whether participants rated accents differently, parallel to the method used in Dossey et al. (2020).

To answer the first question, a linear mixed-effects model (see Section 4.5) was created. In it, the ratings were the dependent variable. The predictor variables were question type (see Section 4.1.2) and age. Furthermore, two random effects were included: participant ID and individual speaker of each accent. The model was analyzed using an ANOVA, which reports a significant effect of question type ($F(5,2900) = 36.44$, $p < .001$), but no effect of age ($F(1,2878) = 0.17$, $p = .678$) or a question–age interaction ($F(5,2870) = 0.31$, $p = .909$). Consequently, it is possible to see that children rated unknown speakers differently on different questions. Figure 6.9 depicts the predicted results from this model. Since no age effect or age–question interactions were found, the figure only displays the effects of question type on ratings.

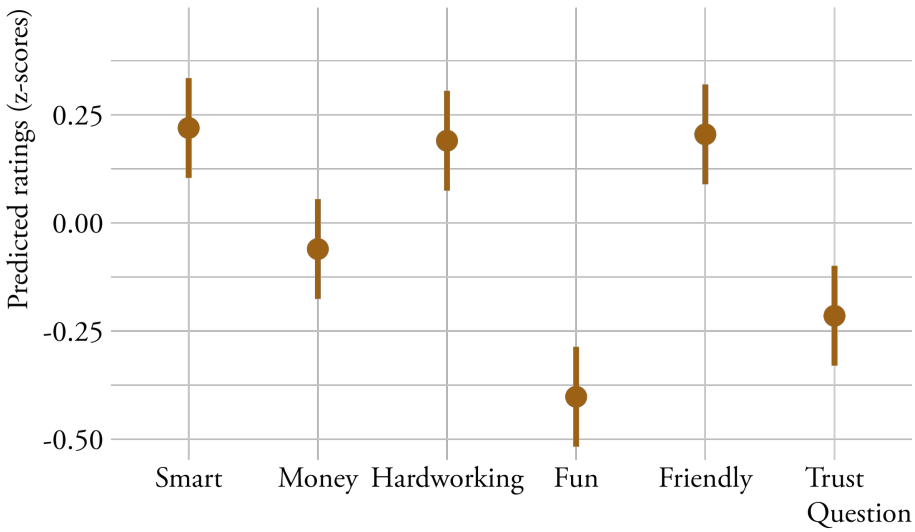


Figure 6.9. Verbal-guise task, by question, without age, London

As Figure 6.9 shows, children from London rated speakers higher on certain variables, such as SMART, than on others, like FUN. In order test whether the differences across question types were significant, pairwise comparisons (see Section 4.5) were implemented. The results reveal three main groups. First, the SMART ($M_{sm} = 0.22$; *95% CI*$_{sm}$ = [0.10, 0.33]), HARDWORKING ($M_{hw} = 0.19$; *95% CI*$_{hw}$ = [0.07. 0.31]), and FRIENDLY ($M_{fr} = 0.20$; *95% CI*$_{fr}$ = [0.09, 0.32]) ratings received the highest scores and did not significantly differ from each other ($t_{sm\text{-}hw}(2811) = 0.49$, $p = .996$; $t_{sm\text{-}}$

$_{fr}(2811) = 0.24$, $p = .999$; $t_{hw\text{-}fr}(2811) = -0.24$, $p = .999$). These three question types were followed by the questions involving MONEY (M_{mn} = -0.06; *95% CI*$_{mn}$ = [-0.18, 0.05]) and TRUST (M_{tr} = -0.21; *95% CI*$_{tr}$ = [-0.33, -0.10]). These two did not differ from each other ($t_{mn\text{-}tr}(2811) = 2.53$, $p = .114$), whereas both of them differed from the first three ($t_{mn\text{-}sm}(2811) = 4.60$, $p < .001$; $t_{mn\text{-}hw}(2811) = -4.11$, $p < .001$; $t_{mn\text{-}fr}(2811) = -4.36$, $p < .001$; $t_{sm\text{-}tr}(2811) = 7.13$, $p < .001$; $t_{hw\text{-}tr}(2811) = 6.65$, $p < .001$; $t_{fr\text{-}tr}(2811) = 6.89$, $p < .001$). Finally, speakers were rated lowest on the FUN variable (M_{fn} = -0.40; *95% CI*$_{fn}$ = [-0.52, -0.29]), which differed from all other questions ($t_{fn\text{-}mn}(2811) = 5.61$, $p < .001$; $t_{sm\text{-}fn}(2811) = 10.21$, $p < .001$; $t_{fn\text{-}tr}(2811) = -3.08$, $p = .026$; $t_{hw\text{-}fn}(2811) = 9.72$, $p < .001$; $t_{fr\text{-}fn}(2811) = -9.97$, $p < .001$). Thus, regardless of accent, London children rated speakers as smart, friendly and hardworking and as having less money, being less trustworthy and, finally, even less fun.

Due to the similarities in scores in the previous paragraph, a PCA was applied to the data to test whether the number of variables in the analysis could be reduced. The standard procedure involves choosing components with an eigenvalue above one (Jolliffer, 2004, p. 114). The data provided only one dimension that yields an eigenvalue higher than one. This component encompassed the variables SMART and HARDWORKING. This means that all other ratings can be taken to measure individual constructs. As a consequence, the analyses described below are conducted individually per question, in line with the analysis of Dossey et al. (2020).

Six different linear mixed-effects models were thus performed, each of them involving one of the six questions included. In each of them, the dependent variable was the ratings participants gave. The predictors were accent and age, with an interaction between the two. Finally, participant ID and individual speaker were included as random effects. All plots in the following six sections should be read in the same fashion: age is indicated on the X-axis and ratings on the Y-axis; moreover, bold lines show the predicted mean rating for a given age, while ribbons indicate the predicted 95% confidence intervals. First, the results from the three questions involving prestige variables are described: SMART, MONEY and HARDWORKING. Lastly, the findings regarding the warmth questions, FUN, FRIENDLY and TRUST are presented.

6.3.1 SMART

The data from the SMART question, analyzed using an ANOVA, revealed a significant interaction between accent and age ($F(4,475) = 2.62$, $p = .0398$). The ratings predicted by the model can be seen in Figure 6.10. This interaction should be understood in the following way (a graphical depiction of the interaction is shown in Section 7.4). From the ages of 7 to 8 and a half years, the accents of Plymouth and SSBE were rated higher than Chinese-accented English. Moreover, from the ages of 7 and a half to 10 years, the Plymouth accent was also rated higher than MLE. SSBE was rated higher than MLE from the age of 8 years onwards. French-accented English was rated higher than

MLE from the age of 8 and a half. In contrast, Chinese-accented English received higher ratings than MLE only at the age of 10 years and higher ratings than the Plymouth accent at the age of 11 and a half years.

In summary, the results can be divided into three stages. First, the Plymouth and SSBE accents obtained higher SMART ratings compared to Chinese-accented English and MLE (at a slightly later age). Second, SSBE, Plymouth and French-accented English were perceived higher than MLE. Finally, SSBE as well as French- and Chinese-accented English were rated higher than MLE, with Plymouth also being rated below Chinese-accented English by the oldest participants.

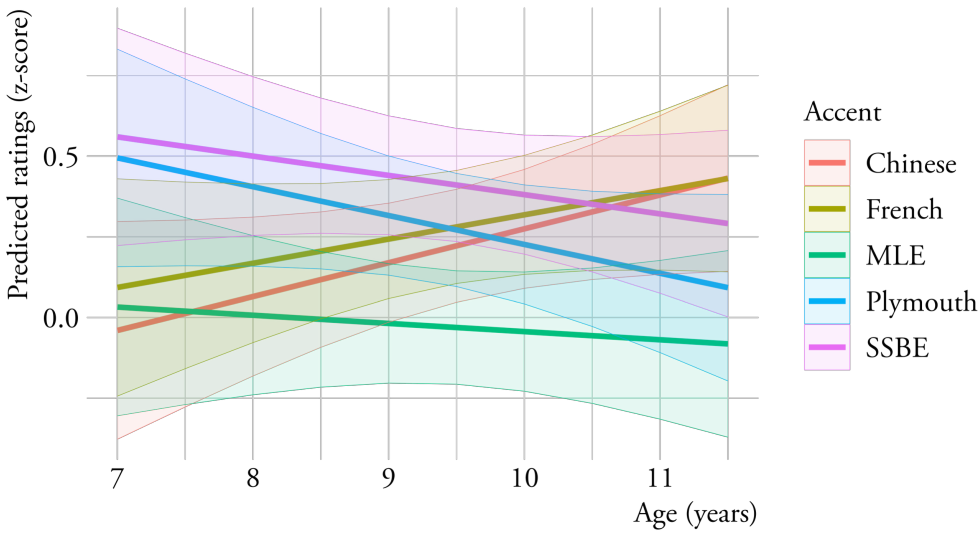


Figure 6.10. Verbal-guise task, SMART, with age, London

6.3.2 MONEY

On the second prestige variable, MONEY, an ANOVA revealed no age ($F(1,95.71) = 0.05$, $p = .831$) or accent effects ($F(4,75.94) = 0.53$, $p = .711$). Moreover, there was no interaction between these two ($F(4,377.05) = 0.90$, $p = .462$). The predicted results from this question are plotted in Figure 6.11. As the lack of significant results indicate, Figure 6.11 shows that children from London rated all speakers, regardless of accent and age, in a similar fashion.

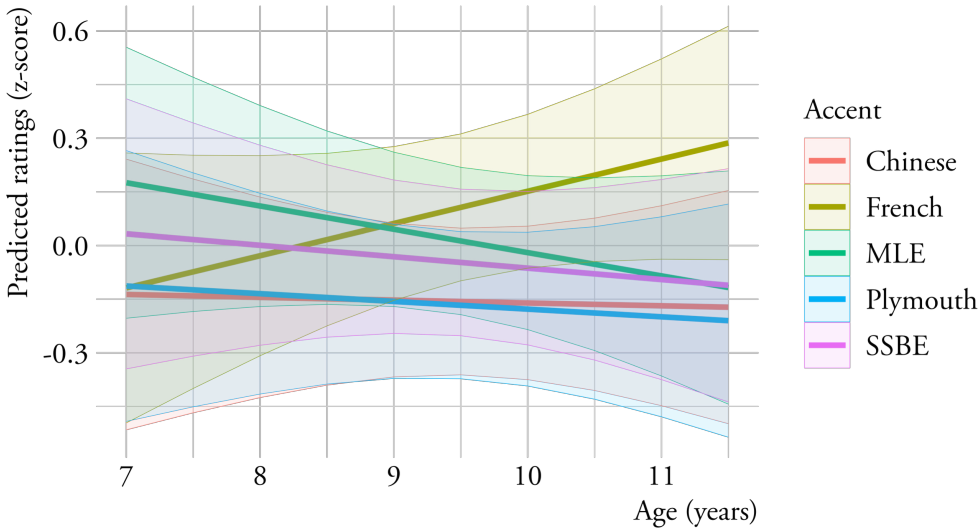


Figure 6.11. Verbal-guise task, MONEY, with age, London

6.3.3 HARDWORKING

As could be expected by the results of the PCA described above, the data from HARDWORKING exhibited a similar pattern to that of SMART. A significant interaction between age and accent is found ($F(4,453.15) = 3.87$, $p = .005$). The predicted values of the model can be seen in Figure 6.12.

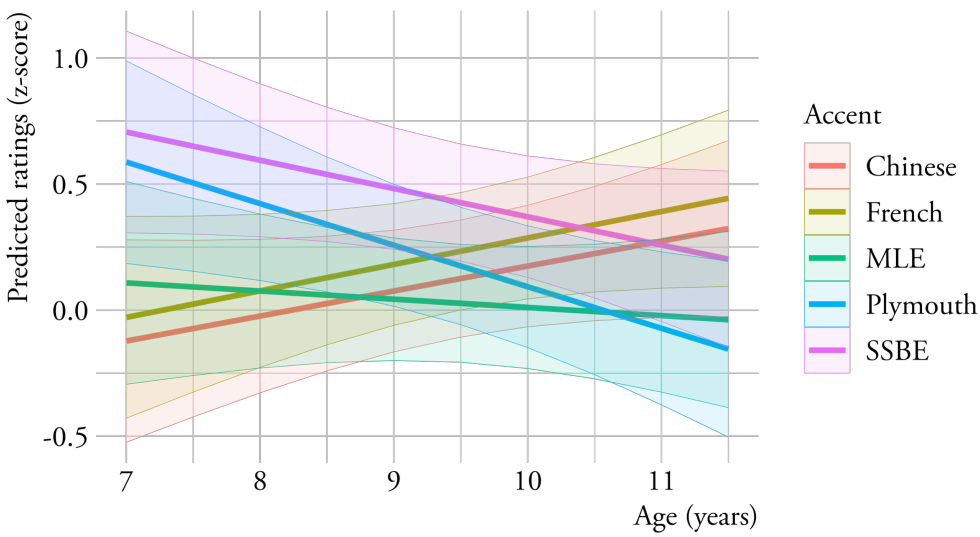


Figure 6.12. Verbal-guise task, HARDWORKING, with age, London

The interaction in this question is due to the following developmental changes (a visual representation of these changes is shown in Section 7.4). Between the ages of 7 and 8 years, London participants rated the SSBE and Plymouth accents higher than all other accents (MLE, French- and Chinese-accented English). At the age of 8 years, SSBE was still rated higher than these three accents. However, the Plymouth accent was only rated above Chinese-accented English. Afterwards, between the ages of 8 and a half and 10 years, SSBE received ratings higher than MLE and Chinese-accented English. This was followed by a brief period in which all accents were rated equally. Finally, from the age of 11 years onwards, French- and Chinese-accented English as well as SSBE were rated higher than the Plymouth accent.

6.3.4 FUN

The FUN question did not give rise to any statistically significant effects with regard to accent ($F(4,384.04) = 0.27$, $p = .898$) or age ($F(1,96.04) = 0.59$, $p = .442$). Similarly, the interaction between these two variables was not significant either ($F(4,384.04) = 0.80$, $p = .525$). The predicted ratings from the model can be seen in Figure 6.13, which depicts how London children rate all accents equally on the FUN question across development.

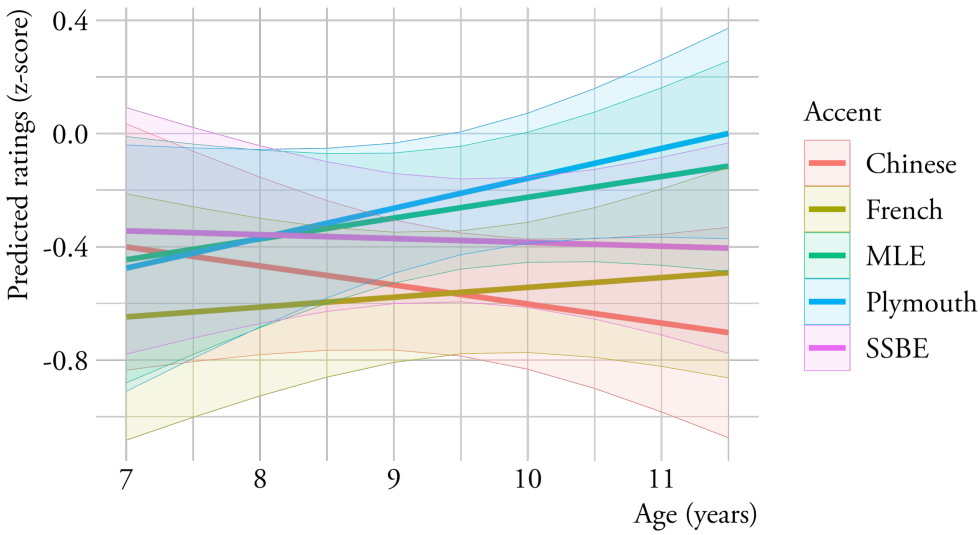


Figure 6.13. Verbal-guise task, FUN, with age, London

6.3.5 FRIENDLY

The ratings from the FRIENDLY variable were not subject to statistically significant effects of accent ($F(4,14.68) = 0.38$, $p = .821$), age ($F(1,451.20) = 0.0005$, $p = .982$) or their interaction ($F(4,450.72) = 0.31$, $p = .872$); see Figure 6.14 for an illustration of the predicted results. Just as with FUN above, the ratings the five speakers received did not vary as a function of accent or age.

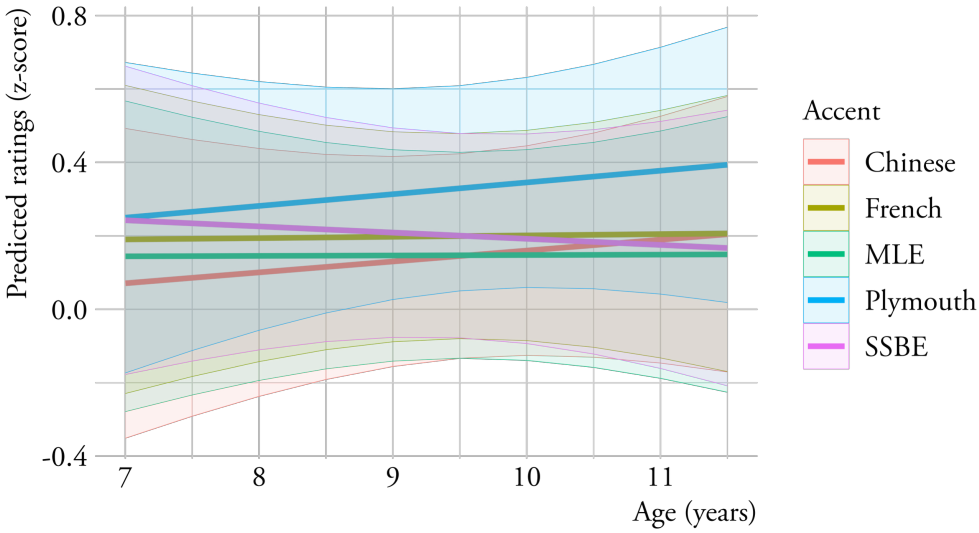


Figure 6.14. Verbal-guise task, FRIENDLY, with age, London

6.3.6 TRUST

Finally, TRUST did not provide significant effects of accent ($F(4,73.44) = 1.16$, $p = .0.333$) or age ($F(1,94.97) = 0.004$, $p = .949$) either. Similarly, their interaction did not reach significance ($F(4,382.56) = 0.56$, $p = .688$); see Figure 6.15. The pattern observable here was similar to that for the two other warmth questions above, with neither accent nor age influencing the ratings of participants from London.

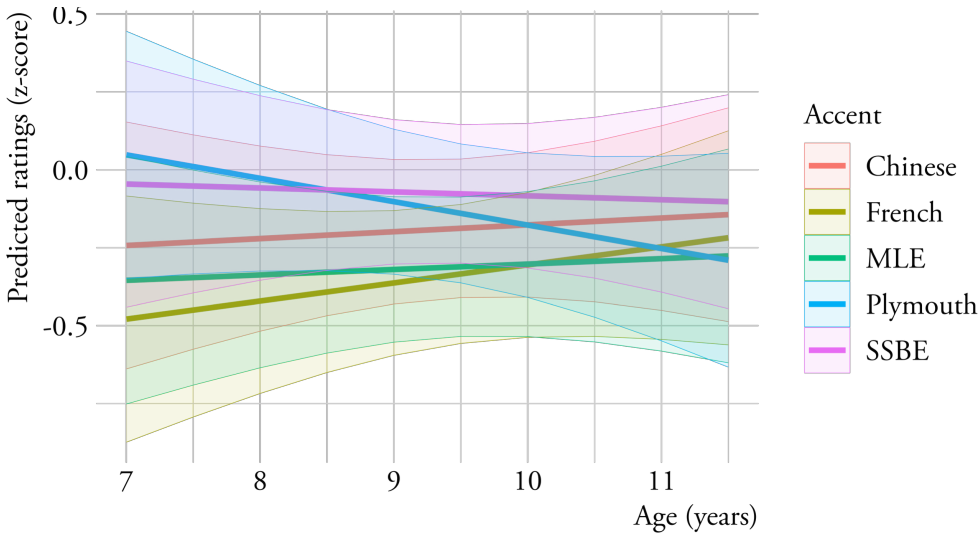


Figure 6.15 Verbal-guise task, TRUST, with age, London

6.4 Discussion

The findings from the three previous sections are discussed below. First, each individual task is commented on in the same order in which they were reported: the intelligibility task first, then the categorization task and, lastly, the verbal-guise task, the main focus of this thesis. Then, how these results inform the RQs in Section 1.1 is dealt with: whether the development of accent stereotypes relies on a typology of accent types, such as L1/L2 accents (Section 6.4.1), and whether children's development of accent stereotypes relies on their ability to understand them and categorize them (Section 6.4.2).

In the intelligibility task, participants from London scored highest when the speech-in-noise included the accents of Plymouth, French-accented English and SSBE. Furthermore, Chine-accented English received the lowest scores, differing significantly from all other accents. Of these, the most expected result was that for SSBE. As described in Section 5.4 and 4.2.1, SSBE plays an important role in media and education. Moreover, Section 2.1 described how accents in England are not only reliant on geography, but also SES. Most of the families in the London sample came from middle- to upper-middle SES (see Section 4.4). Therefore, it is likely that several of the participants' parents spoke accents closer to SSBE than to MLE. This indicates that children had a high exposure to SSBE, as shown in Section 4.4. This higher exposure, in turn, may have made it easier for them to process SSBE in adverse conditions.

In contrast, the fact that the Plymouth accent and French-accented English obtained equally high scores is more unexpected. This is especially true considering that MLE, the local accent, obtained scores below these three accents (although not significantly different from SSBE). Two main factors may explain this pattern. First, London's multicultural and international nature means that many accents are found in the city. This includes both L1 accents tightly associated with London (such as SSBE, RP, Cockney, Estuary English, Popular London and, of course, MLE) as well L2 accents due to immigration (such as accents from speakers from India, Nigeria, Italy, Bangladesh or Poland; Office for National Statistics, 2022).

Second, approximately half of the London participants in the sample had parents who were either non-native speakers of English or had a non-local accent (see Section 4.4). This may have led to a situation in which MLE did not have the same native accent status as, for example, the Plymouth accent had in Plymouth, a city which is much more linguistically homogenous. These factors, together with the relationship between accent and class described above, may suggest that SSBE could potentially be considered closer to the London participants' native accent rather than MLE. This could help explain why their intelligibility scores were higher for stimuli in SSBE than in MLE.

All together, these results show that the stimuli in the set-up show a gradient in scores that improves throughout development. The fact that scores improve with age is hardly unexpected. Children's (socio)linguistic development during late childhood allows them to understand different kinds of speech better in adverse conditions (see e.g. Dossey et al., 2020).

The results of the categorization task are discussed below. On the one hand, the the ranking of accents in terms of accuracy was also subject to a hierarchy without unequivocal groups, as in the intelligibility task. While French-accented English was categorized best by younger participants (below the age of 10 years), it received the same scores as Chinese-accented English from older ones. Similarly, Plymouth and SSBE were categorized better than MLE. However, the differences between (i) the Plymouth accent and MLE and (ii) SSBE and MLE were larger than the difference between the Plymouth and SSBE accents.

Concerning the relative order of accents across both the intelligibility task and the categorization task, French was highest on both and MLE among the lowest. The main differences between the tasks were thus the relative order of the Plymouth accent and SSBE as well as that of Chinese-accented English. The former two accents ranked higher in the intelligibility task, while they ranked below French-accented English in the categorization task. Interestingly, in both tasks, the difference between MLE and SSBE was smaller than that between MLE and Plymouth, which may reflect the higher exposure London participants had to these accents.

The major difference between these two tasks is the lack of age effects. Namely, while children's understanding of words in noise improved throughout development, their

ability to categorize accents remained stable across the age span. The exception to this is Chinese-accented English. As discussed above and in Section 5.4 for the Plymouth data, the age effect on the intelligibility task is an expected finding. It is the lack of developmental effects on the categorization task that requires further investigation.

The literature on children's categorization of accents typically shows age effects, with older children outperforming younger ones (see discussion in Section 5.4). The London results from the categorization task were therefore unexpected. A key element in understanding this pattern may be the fact that participants from London had high levels of exposure to linguistic diversity (see Section 4.4). Research focusing on adults shows that people who are regularly exposed to more varieties of a language are better at sorting accents (e.g. Clopper & Pisoni, 2004; Clopper & Pisoni, 2007; Williams et al., 1999). Parallel results with children were also reported by Floccia, Butler, Girard, et al. (2009). In their accent classification task, children from families where at least one parent spoke with a non-local accent perform better than children whose parents only spoke with the local accent (see similar results in Evans & Lourido, 2019; McCarthy & Evans, 2019). It may be the case that, if younger participants from London were better at the categorization task due to their linguistic exposure, the age effect was levelled out. A direct comparison between these results and those from Plymouth is presented in Chapter 8, since the differences in exposure between the Plymouth and London participants are significant (see Section 4.4).

The results of the verbal-guise task paint a more intricate picture. Instead of main accent or age effects, the data showed significant interactions with two of the variables included (SMART and HARDWORKING). Briefly put, the youngest participants from London rated speakers with a Plymouth and SSBE accent as smarter than speakers of Chinese-accented English and MLE as well as more hardworking than speakers of Chinese- and French-accented English and MLE. With age, this pattern is partially reversed. Specifically, older London children rated the Plymouth accent low on the SMART and HARDWORKING variables and Chinese- and French-accented English were rated higher. In contrast, SSBE remained fairly high and MLE rather low throughout development.

The fact that two of the prestige variables were the only ones to provide significant results is perhaps unsurprising. Research with adults has shown that prestige variables are more robust across studies than warmth variables, as discussed in Section 5.4. This may have made it easier for children to learn associations based on these variables. Furthermore, the review in Chapter 3 depicted SMART as one the earliest variables children assign to accents (see Section 5.4 for a further discussion). This means that these results match the overall pattern found in the literature. Moreover, the fact that the youngest participants from London already rated accents differently on these two variables is also paralleled in the literature showing that even children as young as 4 years old can rate speakers as smart based on their accent. Naturally, it is not possible

to establish *when* London children began having these stereotypes, since data was not collected from younger children.

The results from the solidarity ratings offered a different pattern. For all questions, the results were not significant. This can reflect the less solid nature of these sorts of variables. For example, both McCullough et al. (2019a) and Dossey et al. (2020) used a similar set-up. They both collected data in the US (Columbus, Ohio) from participants in comparable age ranges and included the variables SMART and FRIENDLY. Their results differed regarding when children started rating accents as friendly. In McCullough et al. (2019a), FRIENDLY did not become a relevant variable until the ages of 8 and 9 years. In Dossey et al. (2020), 4-year-olds rated the Mid-Atlantic US accent lower in friendliness than the rest of the accents in their stimuli. Therefore, the results from London can arguably lend support to the view that, even for children, warmth ratings are less stable than prestige ratings.

A final comment on the verbal-guise task follows. The ratings by children from London demonstrate that accent stereotypes do not constitute knowledge children simply learn at one point during development. These social structures are malleable during development, matching the patterns present in McCullough et al. (2019a) and Dossey et al. (2020).

6.4.1 Accent types: native, nonnative, L1 and L2

This section comments on how the results from London inform the need for a strict distinction between native/nonnative or L1/L2 accents if we are to understand how children develop accent stereotypes. It also discusses whether such divides explain how well children understand and categorize accents.

First, the data from the intelligibility task revealed that neither of these distinctions fit the results. With regard to a native/nonnative accent classification, the Plymouth and French-accented English accents were those with the highest scores. In fact, both scored statistically higher than MLE, the local accent. SSBE also obtained results similar to those for the Plymouth accent and French-accented English. Nevertheless, the scores for SSBE were closer to MLE than those of Plymouth and French-accented English were to MLE. This may reflect the important role of SSBE in British culture as well as children's exposure to it, as described in the previous section. Importantly, MLE was significantly different from Plymouth as well as Chinese- and French-accented English. This could be taken as evidence that children were sensitive to the native/nonnative divide. Nonetheless, the fact that it did not score best in the speech-in-noise task but, in fact, second to last, puts this hypothesis into question. If this were a key variable, we would expect the local accent to score highest.

When it comes to the L1/L2 distinction, the findings did not support such a strict classification. French-accented English received scores similar to those for both the Plymouth and SSBE accents. Moreover, its scores were more similar to all the L1

accents than to Chinese-accented English, the other L2 accent included as stimuli. In a parallel fashion, the results from the Plymouth accent, an L1 accent, were closer to French-accented English than to MLE, another L1 accent. Overall, the results suggest that whether an accent increases (or decreases) in intelligibility is not based on it being, for instance, nonnative or L1 (see Section 5.4.1 for further discussion).

A different picture is observed in the findings from the categorization task. The accuracy scores are discussed first, followed by the results from the co-occurrence matrices. Younger participants did not rely on the native/nonnative split. For instance, they were just as good at categorizing the local MLE accent as Chinese-accented English. Furthermore, the differences between the SSBE and Plymouth accents, on the one hand, and MLE, on the other, were not significant[52]. Furthermore, the fact that MLE scored among the lowest in accuracy and, with age, significantly lower than Chinese-accented English is not consistent with previous results in the literature. For example, Floccia, Butler, Girard, et al. (2009) reported that their 7-year-old participants were better than their 5-year-old group at classifying their local, Plymouth accent than a nonnative L1 accent and an L2 accent. Moreover, the fact that it is an L2 accent that was closest in accuracy for younger children also suggests that MLE's differences from all other accents may not be based on the fact that it is the native accent.

Turning to the categorization of accents as L1 or L2, the accuracy scores did not support the need for such a divide in accent categorization. First, as just mentioned, younger London participants' scores for MLE and Chinese-accented English did not differ in a statistically significant manner. Likewise, between the ages of 7 and 9 and a half years, the scores for Chinese-accented English did not differ from those for the Plymouth and SSBE accents. In contrast to these results, the co-occurrence matrices showed a sharp division between L1 accents, on the one hand, and L2 accents, on the other. Throughout all five years, children were more likely to group French- and Chinese-accented English speakers together, separately from the Plymouth, SSBE or MLE speakers, all of which were grouped together more frequently than with the L2 accents. With regard to the structure of the L1 cluster, MLE was more frequently grouped together with SSBE. This can be argued to refute the importance of a native/nonnative division for categorization[53].

The results of the verbal-guise task can be discussed in two steps based on the participants' age. This is due to the crosscut pattern of the ratings, as shown by the interaction between age and accent for the SMART and HARDWORKING questions described in Section 6.3. Younger children rated speakers differently than the older

[52]Note that the differences between Plymouth–MLE and SSBE–MLE almost reach significance.

[53] Nevertheless, if SSBE is to be considered between a native and nonnative accent, as described in the previous sections, it could be argued that the L1 branch does support the differentiation between native accents (SSBE and MLE) and nonnative accents (the Plymouth accent).

children did. The results of the former group are commented on first. These participants rated the Plymouth and SSBE accents higher on SMART than Chinese-accented English and MLE. A similar pattern was found with HARDWORKING. In this case, the younger children rated MLE as well as Chinese- and French-accented English lower than the Plymouth and SSBE accents. The fact that the two L2 accents were seen as less smart and less hardworking, while two of the L1 accents were seen as smarter and more hardworking, might be taken as evidence that the L1/L2 divide has some explanatory power. Nonetheless, MLE (an L1 and local accent) was rated just as low as French- and Chinese-accented English in the HARDWORKING variable and as low as Chinese-accented English in the SMART question. This suggests that a strong version of this relationship is unlikely. That is, the data does not support the notion that children in this age range will rate all L1 accents higher than all L2 accents (or vice versa). It could still be the case that L2 accents are more likely to be rated lower, which would reflect a weaker version of the relationship. In any case, the effect is not absolute.

It is worth noting that the youngest participants already show signs of accent stereotyping. As a consequence, the London data does not allow a test of whether children develop stereotypes toward L2 accents at an earlier age than toward L1 accents (see RQ2 in Section 1.1). It could be the case that London children began rating Chinese- and French-accented English lower at an earlier age than MLE. But the data does not allow this to be tested.

Turning to the native/nonnative split, neither the SMART nor the HARDWORKING variable support the notion that this division is important. For both variables, MLE was rated in a fashion similar to Chinese- and French-accented English.

The accent ratings by the older participants from London showed a different pattern from younger ones. Nonetheless, their data did not follow a native/nonnative split or a L1/L2 one. Older children rated SSBE and both Chinese- and French-accented English higher in the SMART and HARDWORKING questions. Furthermore, MLE and the Plymouth accent were rated as less smart, while only the accent from Plymouth was rated as less hardworking.

The ratings of older participants matched the overall adult ratings on the accents included in our study. They rated SSBE and French-accented English high in variables within the prestige factor and the MLE and Plymouth accents low (e.g. Coupland & Bishop, 2007; Magne, 2019; see Section 4.2.1). The main difference between the responses from the older participants and those of adults was their views on Chinese-accented English. The literature dealing with adult participants shows how Chinese-accented English is typically rated low on prestige variables (e.g. Coupland & Bishop, 2007; Magne, 2019). The fact that children flipped this pattern may be the result of how the Chinese language, and consequently Chinese-accented English, has been framed in education in the recent past. To be specific, Chinese has been described as one of the most important foreign languages for the UK's future. In addition, more and more children have begun taking Chinese as a foreign language at school. This may

have led to children viewing Chinese-accented English as prestigious, in contrast to previous generations (see Section 5.4.1 for further discussion).

In conclusion, the development of accent stereotypes as well as the ability to understand and categorize different types of accents does not entirely rely on whether the accent is native/nonnative or L1/L2. This overall pattern gives further support to studies such as Wagner, Clopper, et al. (2014) and Weatherhead et al. (2019), which suggested that the way children react to accents throughout their development is based on multiple factors and not a strict accent typology.

6.4.2 Accent stereotypes, intelligibility and categorization

This section examines how the results from London inform two of the RQs in Section 1.1, namely, whether there is a relationship between the development of accent stereotypes and being able to understand and categorize accents. The results of the intelligibility task are discussed first, followed by the results of the categorization task.

The results from Sections 6.1 and 6.3 do not appear to support a relationship between children having more difficulty understanding certain accents and forming stereotypes toward them. The intelligibility task did not show an interaction between participants' age and accents. That is, in this investigation, children throughout the age span primarily had difficulty understanding Chinese-accented English. Nevertheless, the way this accent was rated for the SMART and HARDWORKING questions in the verbal-guise task changed with age: from being rated significantly as less smart and hardworking by the younger participants to being rated as smarter and more hardworking by the older ones. One could argue that children's improved understanding of Chinese-accented English is what caused their change in stereotypical views. Nevertheless, the results on the Plymouth accent and French-accented English do not support this argument. Both these two accents were the ones children had the least difficulty understanding in noise. However, they were subject to opposite trends in the verbal-guise task. The Plymouth accent was rated high on both the SMART and HARDWORKING questions by the youngest participants and low by the oldest ones. In contrast, French-accented English was rated as less hardworking by younger children and as more hardworking by the older ones. Therefore, a relationship between intelligibility and accent stereotypes is not supported by the results from London.

Whether the ability to categorize accents influences children's abilities to attach stereotypes to them is discussed below. First, the accuracy scores for the categorization task do not seem to establish a direct relationship with those for the verbal-guise task. As mentioned in the previous paragraph, the SMART and HARDWORKING ratings were subject to developmental and accent-based interactions. Nevertheless, the categorization task only shows accent effects[54]. French-accented English remained the

[54] But see the discussion of Chinese-accented English below.

accent children categorized best, with similar scores by 7-year-olds and 11-year-olds. Nevertheless, it goes from being rated lower for HARDWORKING to being rated highest. Furthermore, the Plymouth accent and SSBE received similar scores in the categorization task across all ages, with the younger participants being as good at categorizing these two accents as the older ones. Nevertheless, SSBE was rated consistently high for SMART and HARDWORKING, while the Plymouth accent changed from being rated high by the younger children to being rated low by the older ones. Therefore, it does not seem possible to establish a strong relationship between the results of these two tasks.

The accuracy scores for Chinese-accented English in the categorization task may support a counter-argument, since they were subject to developmental effects. These scores changed from being similar to those for the Plymouth and SSBE accents and significantly different from French-accented English among younger participants to being comparable to those of French-accented English among older ones. Nevertheless, these changes did not match the ratings in the verbal-guise task. In this case, Chinese-accented English obtained ratings parallel to those of French-accented English in being rated lower for SMART and HARWORKING by younger children but higher by older participants. Therefore, the age-based changes for Chinese-accented English do not support a strong relationship between these two tasks either.

Second, the co-occurrence matrices in Section 6.2.2 do not support a strong relationship between the categorization and verbal-guise tasks. As Figure 6.7 above shows, children from London tended to confuse French- and Chinese-accented speakers with each other more frequently than they did with speakers of Plymouth, SSBE or MLE. Nevertheless, the SMART and HARDWORKING ratings for MLE matched the ratings for the two L2 accents provided by younger children. In the case of older participants, it is SSBE that was rated similarly to the two L2 accents. Hence, the two branches in the co-occurrence matrix do not match the verbal-guise results. Similarly, the internal structure of the L1 branch in Figure 6.7 merged MLE with SSBE separate from the Plymouth accent. Nevertheless, MLE and SSBE were at opposite poles of the SMART and HARDWORKING ratings. MLE is significantly lower than SSBE on these two variables.

In conclusion, the ratings speakers of different accents receive in the verbal-guise change with development. In contrast, the accent scores in the two auxiliary tasks remained stable through the 5-year span in the sample. Therefore, the London data from the three sociolinguistic tasks does not support a strong relationship between intelligibility and categorization, on the one hand, and accent stereotypes, on the other. Nevertheless, the next chapter compares the data from London and Plymouth. This may provide further insights into how different aspects of sociolinguistic development interact during late childhood.

6.5 Summary

This chapter focuses on the results from the London data collection. The intelligibility task gives rise to developmental and accent main effects, but no interaction between the two. In other words, older children from London outperform the younger participants. Overall, London children answered correctly more frequently when listening to French-accented English, and the Plymouth and SSBE accents than when listening to MLE or Chinese-accented English. In the categorization task, their results did not change with age, except for Chinese-accented English, which was subject to developmental effects. In this task, participants from London scored best with French-accented English and worst with MLE, with the Plymouth and SSBE accents scoring in between. Lastly, the verbal-guise task provides significant interactions for the SMART and HARDWORKING questions. Younger participants rated SSBE and the Plymouth accent higher than MLE and French- and Chinese-accented English. In contrast, older participants rated SSBE, together with French- and Chinese-accented English, higher than the Plymouth and MLE accents.

7 Comparing accent stereotypes across Plymouth and London

Chapters 5 and 6 describe how the data collected for this thesis helps us understand how and when children develop accent stereotypes in Plymouth and London, respectively. In contrast, this chapter directly compares the results from these two locations. The goal is to obtain a clearer view of how, during development, a child's environment affects the way they acquire accent stereotypes. Moreover, this comparison makes it possible to unveil whether different factors have different effects on these two populations. To carry out this comparison, the results from each task are compared across cities one at a time[55]: first the auxiliary tasks (the intelligibility task in Section 7.1 and the categorization task in Section 7.2) and, then, the core of this thesis, the verbal-guise task in Section 7.3. These comparisons are followed by a discussion in Section 7.4. A general discussion of each task is followed by one concerning how the comparison of the Plymouth and London data inform the RQs in Section 1.1.

7.1 Intelligibility task

This section deals with the results from the intelligibility task, in which participants heard a word with background noise and had to choose a picture that matched the word they heard (see Section 4.1.1 for details). The results from this task are organized in two ways. First, accuracy scores are summarized by stimulus word. Then, scores are analyzed by accent.

The percentage of correct answers per word across cities is shown in Table 7.1. These percentages were calculated based on the total number of times each word was shown to participants in each city: once per participant. This means that, for each word, the maximum number of correct answers is 105 in London and 119 in Plymouth (the raw number of correct answers for each word can be found in Section 5.1 for Plymouth and Section 6.1 for London). In both cities, there are comparable ranges of accuracy: from

[55] For the sake of simplicity, this chapter only reports results *across* cities. For the results *within* each city, the reader is referred to that city's relevant chapter.

97.48% in Plymouth and 96.19% in London with *church* in both cases, to 27.73% for *thorn* in Plymouth and 30.48% for *palm* in London. Furthermore, the ranking of words across the two cities appears to be similar. For instance, the five words with the lowest accuracy are the same ones in both Plymouth and London: *thorn*, *palm*, *goose*, *face* and *bowl*. An ANOVA on a binomial generalized linear model (see Section 4.5), with accuracy as the dependent variable and both city and target word as predictors, reveals no effect of city ($X^2(1) = 0.89$, $p = .346$) or a city–target word interaction ($X^2(19) = 18.09$, $p = .516$).

Table 7.1. Intelligibility task, across cities, correct answers per word
The maximum number of correct answers per word is 119 for Plymouth and 105 for London.

Plymouth		London	
Target word	Percentage	**Target word**	Percentage
church	97.48	**church**	96.19
branch	92.44	**mouse**	93.33
house	92.44	**alligator**	92.38
alligator	91.60	**house**	91.43
mouse	91.60	**wolf**	90.48
wolf	89.92	**bird**	89.52
foot	87.39	**leaf**	85.71
porridge	87.39	**branch**	84.76
bird	85.71	**porridge**	81.90
farm	82.35	**farm**	80.95
leaf	82.35	**foot**	78.10
key	80.67	**salt**	77.14
toy	78.15	**key**	75.24
salt	73.95	**toy**	73.33
goat	68.91	**goat**	72.38
goose	68.07	**bowl**	61.90
bowl	65.55	**face**	60.95
face	50.42	**goose**	58.10
palm	35.29	**thorn**	33.33
thorn	27.73	**palm**	30.48

Second, the data was analyzed to test for the effect of accent on intelligibility. Table 7.2 shows how many correct answers participants gave for each accent across cities. The percentages in each cell are calculated based on how many words participants heard in each accent. The stimuli included four words per accent. Therefore, the highest number of correct answers is four times the number of participants in each city (476 in Plymouth and 420 in London). The ranking of accents is comparable across cities. Participants answered correctly most frequently when the word was uttered with a Plymouth accent. At the opposite end, Chinese-accented English received the lowest number of correct answers, followed by MLE. The main difference across the two cities was the ranking of SSBE and French-accented English.

Table 7.2. Intelligibility task, across cities, correct answers per accent
The maximum number of correct answers per word is 476 for Plymouth and 420 for London.

Plymouth		London	
Accent	Percentage	**Accent**	Percentage
Plymouth	81.72	**Plymouth**	80.71
SSBE	80.46	**French**	80.71
French	77.52	**SSBE**	77.14
MLE	72.69	**MLE**	72.38
Chinese	69.96	**Chinese**	65.95

The data in Table 7.2 does not enable conclusions regarding the significance of these differences or their interactions with age. Therefore, a linear mixed-effects model (see Section 4.5) was created to answer these questions. The dependent variable was accuracy scores (see Section 4.5). The predictors were age, accent and city. These three terms were included with a three-way interaction between them. Finally, participant ID and individual speaker were included as random effects. An analysis of variance reveals a significant interaction between accent type and city ($F(4,4256) = 5.06$, $p < .001$). The predicted scores and 95% confidence intervals (bold lines and ribbons respectively) are shown in Figure 7.1. It shows that participants in both cities become better at understanding speech-in-noise with age at comparable rates. Similarly, the ranking of accents remained relatively stable in both cities regardless of age. As the significant interaction in this paragraph indicates, the main difference between the two cities concerns how well participants understood different accents.

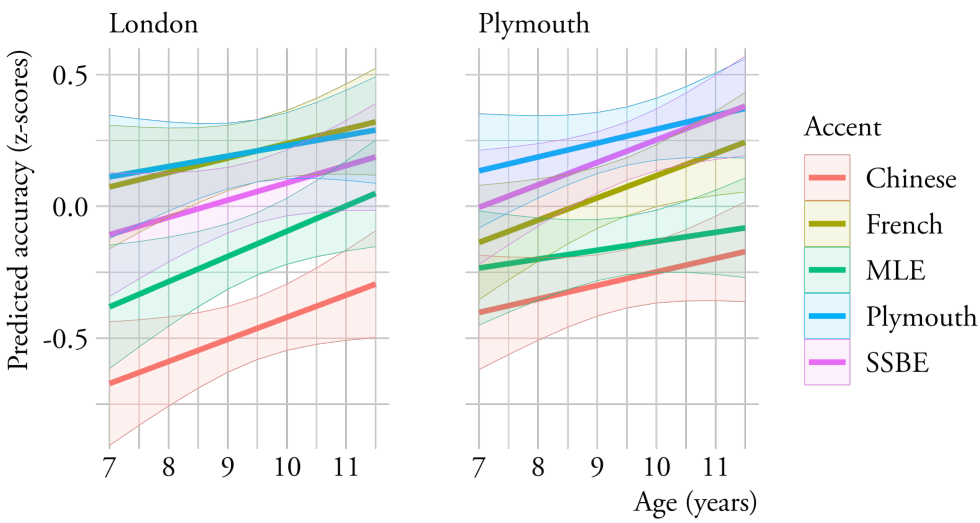


Figure 7.1. Intelligibility results, with age, across cities

The ranking of accents in the intelligibility task across cities is discussed below. For a clearer view of how good participants were at different accents in each city, a new model was run that excluded the interaction between (i) age, on the one hand, and (ii) accent and city, on the other. The rest of the terms were the same as in the model above. The predicted values of this model are shown in Figure 7.2. Predicted average scores across all ages (Y-axis) are plotted against accent (X-axis) and city. Figure 7.2 shows that the results from both cities can be split into two groups. On the one hand, Chinese-accented English and MLE are the accents children had the most difficulty understanding, with the former obtaining the lowest scores in both cities. On the other hand, the Plymouth accent, SSBE and French-accented English are the accents children scored highest in regardless of city.

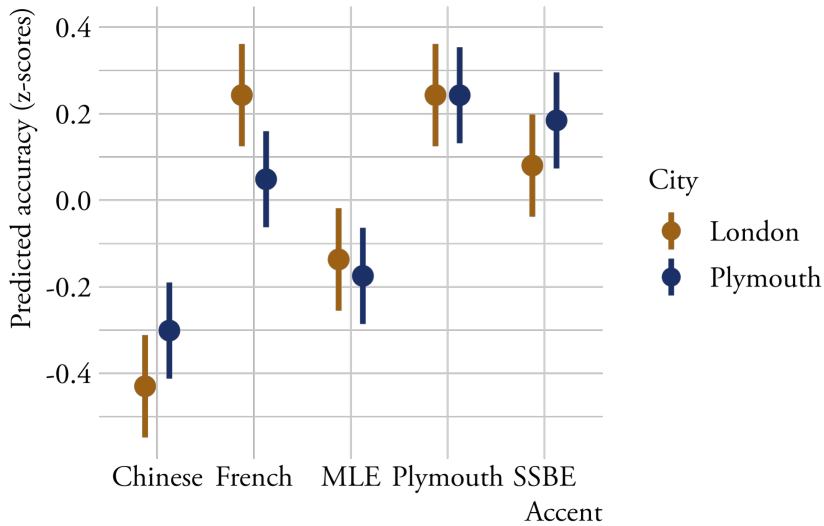


Figure 7.2. Intelligibility results, without age, across cities

Further analyses were carried out to determine what differences in scores across cities are significant. For this purpose, pairwise comparisons (see Section 4.5) of the accents across cities were performed[56]. The difference for French-accented English across cities is the only significant pairwise comparison ($t_{fre}(594.40) = 2.35$, $p = .019$). The other four pairwise comparisons do not reach significance ($t_{chi}(594.40) = -1.55$, $p = .120$; $t_{mle}(594.40) = 0.46$, $p = .645$; $t_{ply}(594.40) = 0.003$, $p = .997$; $t_{ssb}(594.40) = -1.26$, $p = .208$). Furthermore, a post-hoc analysis reveals that the rankings of French-accented English, the Plymouth accent and SSBE differ across cities. These differences are to be understood in the following way. In the city of Plymouth, French-accented English

[56] For the differences between accents *within* cities, see Section 5.1 and 6.1 for Plymouth and London, respectively.

scores below the Plymouth accent and SSBE. In London, however, it is SSBE that scores below French-accented English and the Plymouth accent. These differences in rankings are statistically significant (the difference between French-accented English and Plymouth across cities $t(4248) = 2.36$, $p = .0183$; and the difference between French-accented English and SSBE across cities $t(4248) = 3.631$, $p < .001$).

7.2 Categorization task

The results from the categorization task are the topic of this section. In this task, participants listened to two speakers of each accent and were asked to place in groups the speakers they thought were born in the same city (see Section 4.1.3 for details). How children performed on this task is split into two sections. Their accuracy scores are compared first. Next, the focus is on the content of the groups they created, which is done by looking at each city's co-occurrence matrices.

7.2.1 Accuracy scores

Accuracy in this context means how correct the clusters participants made are. This was quantified using the analysis conducted in Dossey et al. (2020) (see Section 4.5.1 for details).

The average scores by accent across cities are shown in Table 7.3. These figures reveal two main differences. On the one hand, the order of accents is different across cities. SSBE and French-accented English scored best in Plymouth, while they placed fourth and first, respectively, in London. Similarly, Chinese-accented English scored second-to-last in Plymouth and second best in London. The main similarities across cities are the ranking of MLE, last in both cities, and Plymouth, third in both cities. On the other hand, the average scores in London appear to be slightly higher than in Plymouth. For instance, the highest average score in Plymouth is 41.21, whereas in London it is 55.60.

Table 7.3. Categorization task, across cities, average scores
Potential scores span from -100 to 100.

Plymouth		London	
Accent	Average score	**Accent**	Average score
SSBE	41.21	**French**	55.60
French	40.20	**Chinese**	43.10
Plymouth	36.28	**Plymouth**	36.55
Chinese	24.79	**SSBE**	34.94
MLE	19.39	**MLE**	20.54

A linear mixed-effects model (see Section 4.5) was run on these results to test whether these differences were significant and whether there were interactions of any kind. The dependent variable was accuracy, while the independent variables were age, accent and city. These three terms were included with a three-way interaction between them. Moreover, the random effect of participant ID was also added. It was not possible to include individual speakers for each accent as a random effect since the accuracy score merges both speakers. An ANOVA uncovers a city–age interaction ($F(1,222) = 8.64$, $p = .004$) as well as a city–accent interaction ($F(4,2007) = 7.17$, $p < .001$). The scores from this model can be seen in Figure 7.3. It depicts the predicted scores (bold lines) and 95% confidence intervals (ribbons) for a given age and accent.

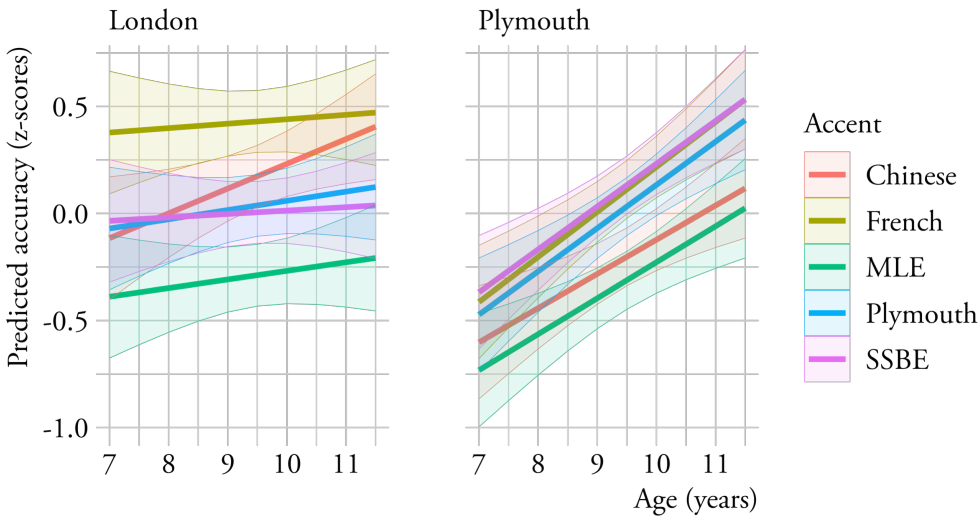


Figure 7.3. Categorization results, with age, across cities

The two interactions from the previous paragraph are discussed in turn. The city–age interaction reflects the fact that children in Plymouth obtained better scores with age (see Section 5.2.1). In contrast, children from London performed the task equally well regardless of age (see Section 6.2.1). One implication of this difference is that the youngest children from Plymouth performed significantly worse than the youngest children from London, as revealed by a post-hoc test ($t(217.99) = -3.40$, $p < .001$). In contrast, the overall scores from the oldest participants did not differ across cities ($t(217.99) = 0.91$, $p = .366$). The age at which the scores from the Plymouth participants begin to match those from their London peers was between the ages of 9 and 9 and a half years. While their scores were significantly different at the age of 9 ($t(218) = -6.57$, $p = .012$), they were not at the ages of 9 and a half years ($t(218) = 2.87$, $p = .091$).

As described in Section 6.2.1, Chinese-accented English was subject to developmental effects in London. A post-hoc test reveals that this improvement with age in London was not significantly different from the improvement with age exhibited by children from Plymouth ($t(839) = 0.64$, $p = .521$). However, the general scores for Chinese-accented English in London were higher than in Plymouth, which leads us to the second interaction of the model described above: that between city and accent.

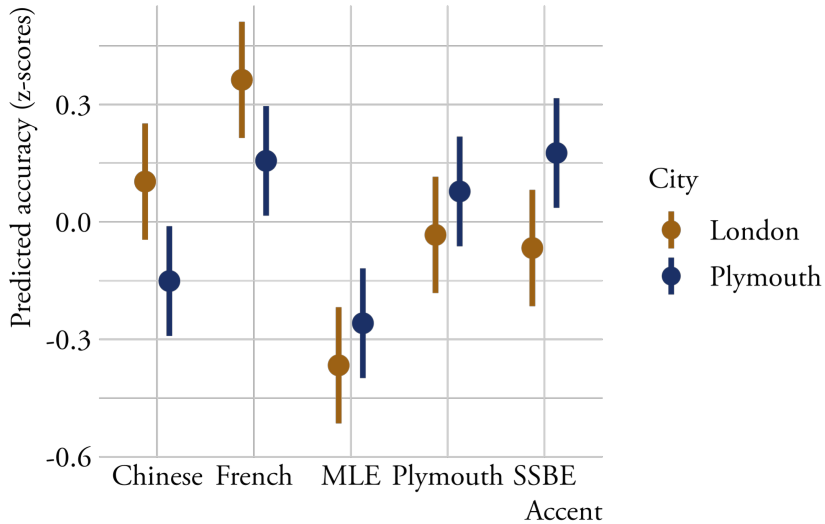


Figure 7.4. Categorization results, without age, across cities

Figure 7.4 depicts the average scores across all ages for each accent in each city to more easily see the differences in accent rankings. Two main patterns are apparent. Children from London obtained higher scores than their Plymouth peers when categorizing the L2 accents (French- and Chinese-accented English). In contrast, children from London achieved lower accuracy scores than those from Plymouth when categorizing the L1 accents in the task (MLE, SSBE and the Plymouth accent).

These differences in means across sites were analyzed using pairwise comparisons (see Section 4.5). Of all five accents, three are significantly different: Chinese-accented English ($t(738.78) = 2.44$, $p = .015$), French-accented English ($t(738.78) = 1.99$, $p = .0472$) and SSBE ($t(738.78) = -2.33$, $p = .021$). In contrast, the scores for MLE ($t(738.78) = -1.03$, $p = .302$) and the Plymouth accent ($t(738.78) = -1.07$, $p = .286$) do not significantly differ across cities. Therefore, participants from London were better than their Plymouth peers at correctly categorizing Chinese- and French-accented English. In contrast, Plymouth children scored higher on SSBE than did children from London. Finally, despite the higher means from Plymouth for the MLE and Plymouth accents, the differences were not significant across cities.

7.2.2 Co-occurrence matrices

Section 7.2.1 demonstrates that children from Plymouth and London differed in their accuracy scores. Nevertheless, it doesn't allow inferences to be made about which speakers they grouped together. For this purpose, co-occurrence matrices were created. Figure 7.5 below depicts Figure 5.5 and Figure 6.6 again. Each cell in Figure 7.5 represents the percentage of times accents were clustered together. The percentages are to be read row-wise. That is, in London, out of all the times SSBE was put in a group with other speakers, 5.41% was with a Chinese-accented English speakers, 28.04% with an MLE speaker, etc. Both subplots show a clear split between accents. Regardless of city, children tended to group Chinese- and French-accented English speakers together more frequently than with the L1 accents. Conversely, they tended to cluster speakers of MLE, Plymouth and SSBE in groups of their own more frequently than with the L2 accents.

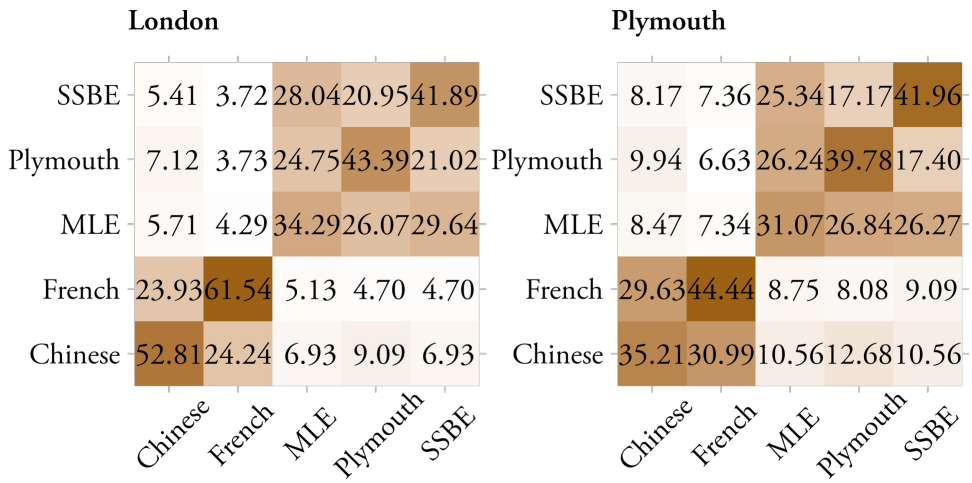


Figure 7.5. Co-occurrence matrix, across cities
Values in percentage row-wise. The darker the color of a cell, the higher the percentage of groups.

A clearer view of the inner structure of the L1 accents is shown in Figure 7.6. As is apparent in Figure 7.5, Figure 7.6 shows that the two L2 accents were grouped together more frequently across cities. Nevertheless, there were differences in how the branch including the L1 accents was organized. In London, children grouped MLE and SSBE together more frequently than either of them was grouped with the Plymouth accent. In contrast, participants from Plymouth grouped the Plymouth and MLE speakers in clusters more frequently than they did with speakers of SSBE. Children from Plymouth appeared to group L1 accents based on whether they were standard (SSBE) or not (the Plymouth and MLE accents), whereas children from London categorized accents based

on those they were exposed to more frequently (SSBE and MLE; see Section 4.4) and those they listened to less (the Plymouth accent).

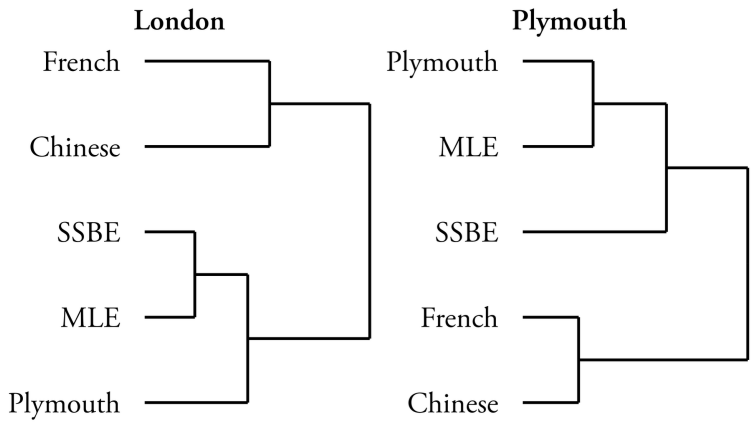


Figure 7.6 Co-occurrence dendrogram, across cities

7.3 Verbal-guise task

This section reports on the results from the main task in this thesis: the verbal-guise. Participants listened to one speaker per accent and were asked different questions about them, such as *how friendly do you think this person is?* (see Section 4.1.2 for details). The data from this task is described in two ways. First, whether participants gave different ratings for different questions is tested. Second, how accent affected these ratings is accounted for.

A linear mixed-effects model (see Section 4.5) was created to examine the first question. The dependent variable was the ratings participants gave. The independent variables were age, question type and city with a three-way interaction. Furthermore, participant ID and individual speaker of each accent were included as random effects. This model was analyzed using an ANOVA. It revealed a general effect of question ($F(5,6139.80) = 4.40$, $p = .0006$), but no effect of city ($F(1,6145) = 0.12$, $p = .732$) or city–question interaction ($F(5,6139.80) = 0.25$, $p = .938$). Figure 7.7 below depicts the predicted results of this model. Since there are no developmental effects (see Section 5.3 and Section 6.3), the plot shows the predicted values from a model that only included the interaction of question type and city as predictor variables. The lack of effects involving where children were from corroborates what can be seen visually in Figure 7.7. Regardless of location, children rated speakers higher on SMART,

HARDWORKING and FRIENDLY. In contrast, they also rated them lower on FUN and TRUSTWORTHY[57].

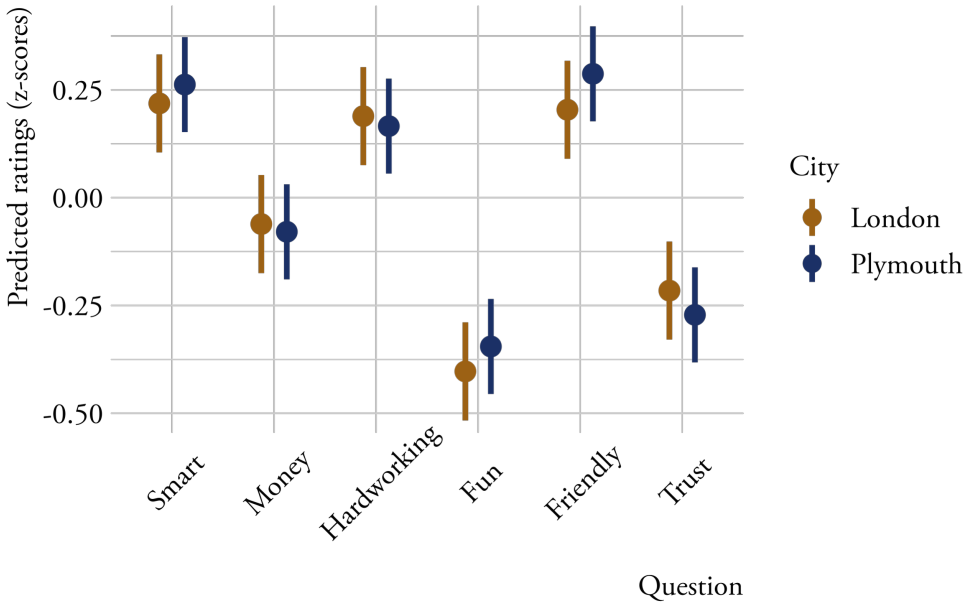


Figure 7.7 Verbal-guise task, by question, without age, across cities

The next step in the analysis involved investigating whether, within each question, ratings differed based on accent and the location of the data collection. With this goal in mind, six different analyses were carried out[58]. In each of them, the response variable was the ratings provided by participants. The predictors were age, accent and city with a three-way interaction between them. Finally, participant ID and individual speaker were included as random effects. In the next six sub-sections, the plots should be understood in the same way. Bold lines indicate predicted ratings for a particular accent at a particular age. Ribbons correspond to the 95% confidence intervals. The three question types involving prestige variables (SMART, MONEY and HARDWORKING) are dealt with first. These are followed by the warmth variables (FUN, FRIENDLY and TRUST).

[57] For the pairwise comparisons *within* each city, see Section 5.3 for Plymouth and Section 6.3 for London.

[58] Since the PCA for Plymouth (Section 5.3) and London (Section 6.3) did not reveal more than one component individually, the analysis with individual variables is the starting point.

7.3.1 SMART

The first variable discussed is SMART. An ANOVA on the data from this question reveals a significant interaction involving accent, age and city ($F(4,815.35) = 4.25$, $p = .003$). Figure 7.8 depicts the predicted data from this model organized by city. The significant interaction is the result of the different developmental trajectories in the two cities. In Plymouth, children did not rate accents differently on the SMART variable until the age of 9 years. From this age onwards, French-accented English, SSBE and Chinese-accented English began being rated as smarter than the Plymouth and MLE accents. In contrast, children in London associated different accents with different degrees of smartness by the age of 7 years. The way younger participants rated accents, however, is different from the way older children did. The younger group rated the accents of Plymouth and SSBE higher than Chinese-accented English and MLE; the older group considered Chinese- and French-accented English and SSBE to be smarter than MLE and Plymouth.

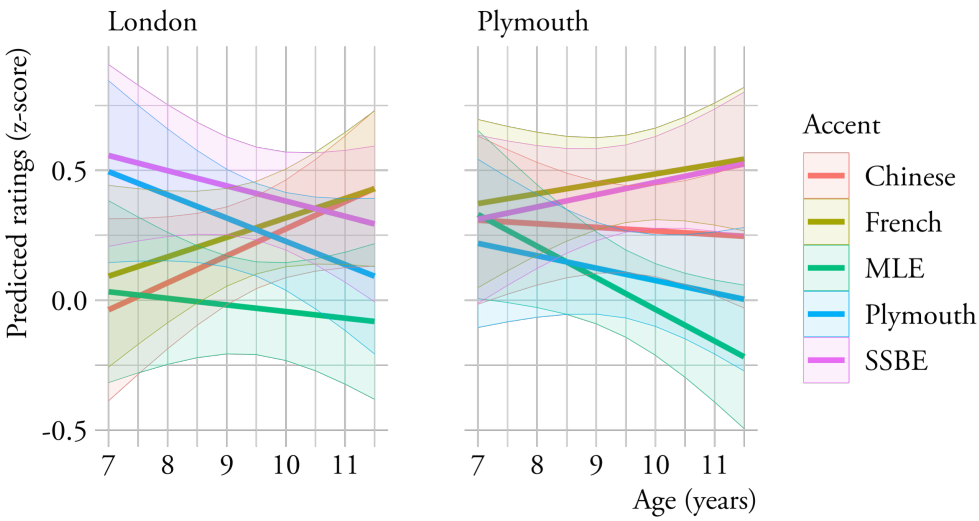


Figure 7.8 Verbal-guise task, SMART, with age, across cities

The main difference between the two cities is therefore how younger participants rated speakers and how these ratings changed throughout development. In fact, none of the pairwise comparisons across cities for the older participants were significant ($t_{chi}(1014.91) = 1.43$, $p = .152$; $t_{fre}(1014) = 1.16$, $p = .246$; $t_{ply}(1009) = -1.14$, $p = .253$; $t_{mle}(1025) = 1.23$, $p = .218$; $t_{ssb}(1016) = -1.02$, $p = .308$).

7.3.2 MONEY

The second question is perceived SES, here labeled MONEY. This variable did not give rise to significant effects for the accent–city ($F(4,810.24) = 0.35$, $p = .842$), city–age ($F(1,203.07) = 0.001$, $p = .973$), or accent–city–age ($F(4,808.83) = 0.23$, $p = .920$) interactions, based on an ANOVA. The predicted results from the model can be seen in Figure 7.9. These results can be interpreted as evidence that children from either city did not associate different accents with different amounts of MONEY.

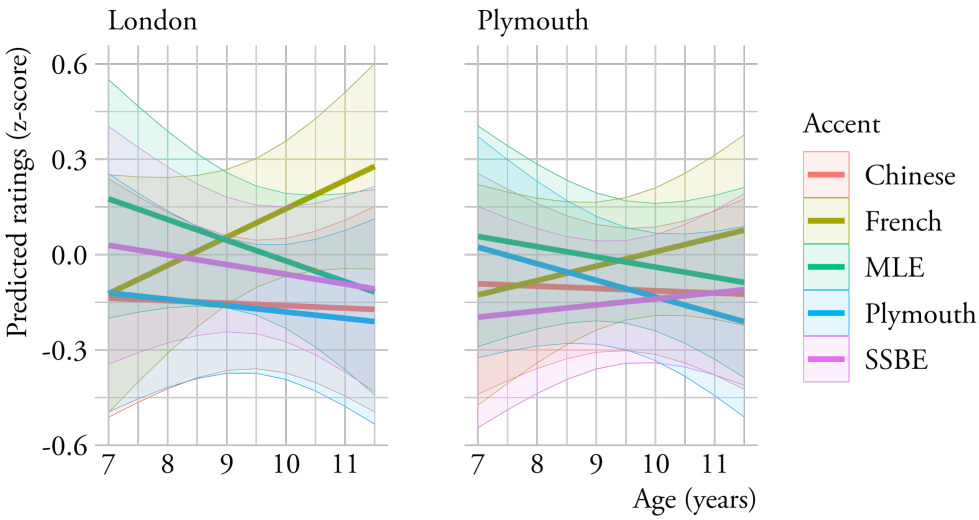


Figure 7.9. Verbal-guise task, MONEY, with age, across cities

7.3.3 HARDWORKING

The last prestige variable is HARDWORKING. The analysis of the data from this question shows a pattern parallel to that of SMART. An ANOVA reveals a significant effect of the age–accent–city interaction ($F(4,801.71) = 4.30$, $p = .002$). Figure 7.10 displays this interaction with the predicted scores from the model.

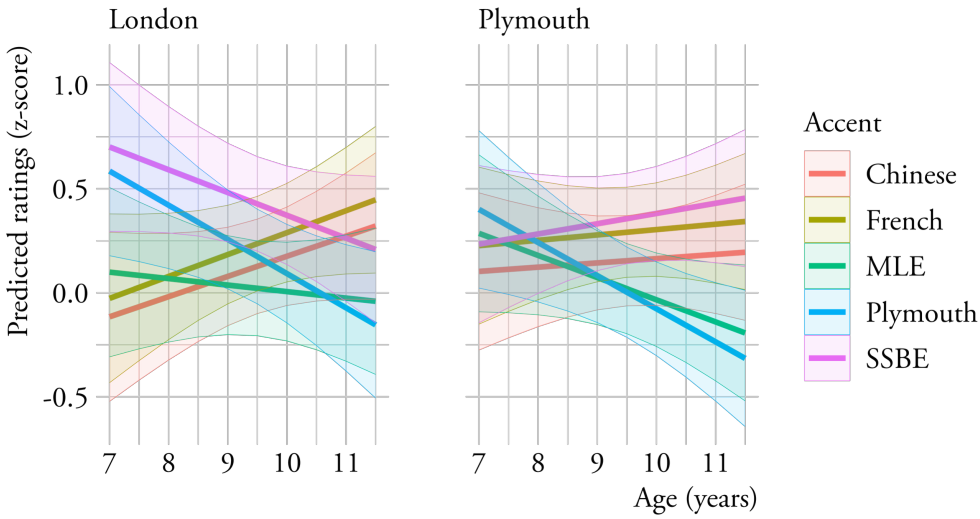


Figure 7.10. Verbal-guise task, HARDWORKING, with age, across cities

The significant interaction in the previous paragraph is the result of differences in developmental trajectories, parallel to 7.3.1 for SMART. In Plymouth, participants between the ages of 7 and 9 and a half years rated speakers of different accents the same way. From this age onwards, they began rating SSBE and French-accented English (and, eventually, Chinese-accented English) higher than MLE and the Plymouth accent. In contrast, London children already rated some speakers as more hardworking than others by the age of 7 years. Namely, they rated the SSBE and Plymouth accents higher than Chinese- and French-accented English and MLE at this age. As they grew older, their accent associations changed. By the age of 11 years, the pattern partially swaps: Chinese- and French-accented English as well as SSBE were rated higher than the Plymouth accent. By this age, none of the pairwise comparisons across cities were significant among the older participants ($t_{chi}(916.96) = -0.46$, $p = .648$; $t_{fre}(915.58) = 1.08$, $p = .280$; $t_{ply}(927.08) = -0.77$, $p = .441$; $t_{mle}(924.80) = 1.08$, $p = .279$; $t_{ssb}(916.24) = -1.42$, $p = .155$). This means that by the age of 11 years, children in both cities gave similar HARDWORKING ratings to the five different accents in the study.

7.3.4 FUN

The first warmth variable reported is FUN. None of its interactions gives rise to statistically significant effects in an analysis of variance (accent–city $F(4,814.22) = 0.40$, $p = .808$; city–age $F(1,205.13) = 0.06$, $p = .802$; accent–age–city $F(4,813.11) = 0.27$, $p = .896$). Figure 7.11 depicts the predicted values from the model by city. It shows how, regardless of city or age, participants rated speakers in a similar fashion.

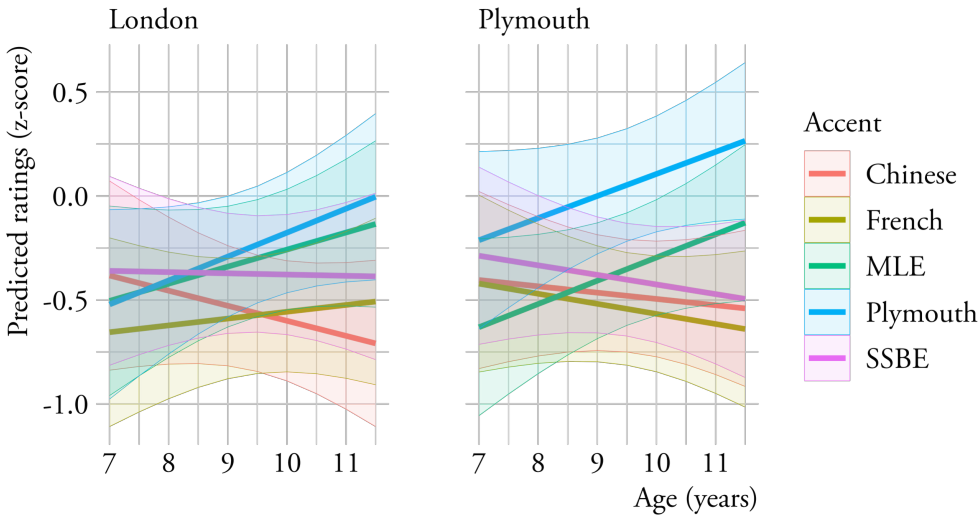


Figure 7.11. Verbal-guise task, FUN, with age, across cities

7.3.5 FRIENDLY

The second question within the warmth block is FRIENDLY. Its ratings are not subject to significant interactions between accent–city ($F(4,1018) = 0.29$, $p = .878$), age–city ($F(1,1018) = 0.03$, $p = .859$) or accent–age–city ($F(4,1017) = 0.15$, $p = .963$). See Figure 7.12 for an illustration of the predicted data from the model across the two locations. Just as with FUN, the ratings do not vary as a function of the speakers' accents or the participants' age and location.

7.3.6 TRUST

The last question discussed is TRUST. Like the previous two variables, it does not give rise to significant interactions involving accent–city ($F(4,815) = 1.08$, $p = .365$), age–city ($F(1,205.49) = 0.28$, $p = .594$) or accent–age–city ($F(4,814) = 0.684$, $p = .603$). The predicted values from the model can be seen in Figure 7.13 arranged by city.

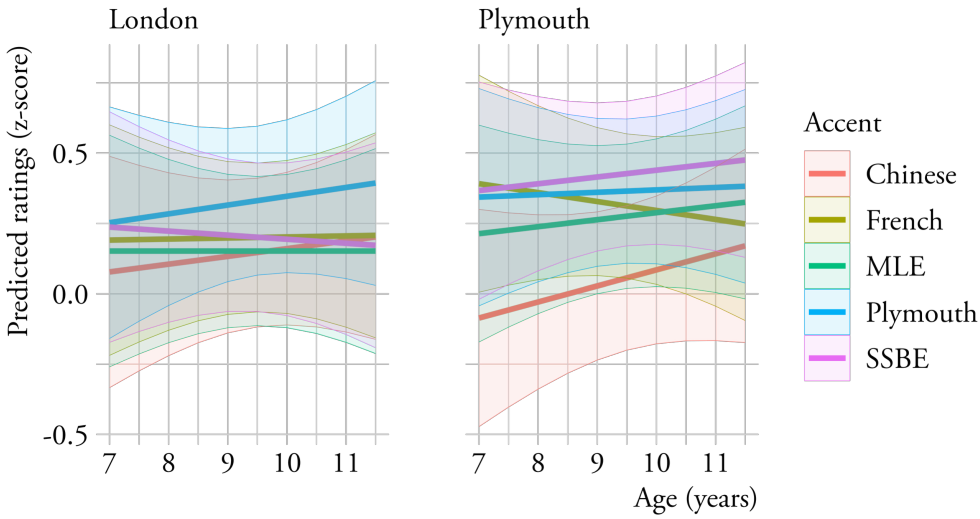


Figure 7.12. Verbal-guise task, FRIENDLY, with age, across cities

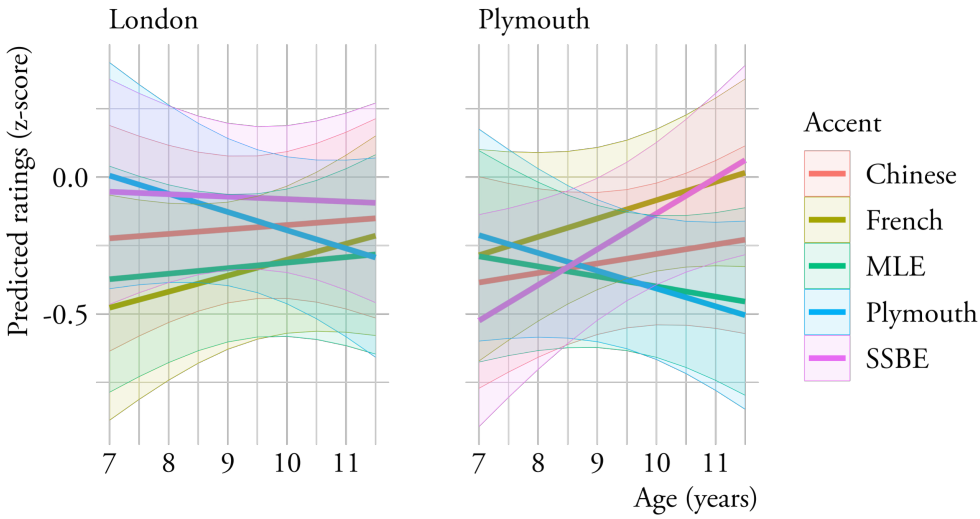


Figure 7.13. Verbal-guise task, TRUST, with age, across cities

7.4 Discussion

This section deals with how a direct comparison between the Plymouth and London data enables a better understanding of how children develop accent stereotypes as well

as the different factors underlying this development[59]. With this purpose in mind, the results of each individual task are first discussed. This section closes with a reflection on how these comparisons inform the RQs in Section 1.1.

The intelligibility task provides very similar results across sites. Children in both cities in Southern England became better at understanding speech-in-noise with age. Furthermore, this improvement took place at similar rates. This is perhaps unsurprising. Without considering the local accents at this point (but see Section 7.4.1 below), children in both locations had similar levels of exposure to the accents included in the experimental set-up. Therefore, one can expect their ability to hear "through the noise" to be comparable in this task.

One question may be raised regarding the greater exposure to linguistic diversity among children from London shown in Section 4.4. Specifically, it is reasonable to expect that these participants will perform better on this task due to their life experience listening to different accents. In fact, exposure to accents has been shown to be important for accent intelligibility (Cristia et al., 2012; Maye et al., 2008; Norris et al., 2003). However, a key factor for the effect of exposure on intelligibility is that it involves *specific* exposure to a relevant accent to improve children's ability to understand that accent. In other words, having exposure to nonnative accent *A* helps children understand that specific accent, but it does not appear to improve intelligibility of nonnative accent *B*. For example, Evans and Lourido (2019) showed how children's ability to understand nonnative accents did not differ between bilingual children, who encountered more linguistic diversity in their everyday life, and monolingual children, who were exposed to less variation (see also McCarthy & Evans, 2019; Paquette-Smith et al., 2021). In our case, children in London had more exposure to linguistic diversity *in general* (see Section 4.4), but not to the accents used as stimuli. This might explain the similar scores across cities.

Finally, the general age effects in both cities can be argued to be the result of the fact that language acquisition does not stop after the preschool years (e.g. Nippold, 2004). Children's ability to understand speech improves throughout late childhood and into adulthood, even the building blocks are acquired at an earlier age (Goldman & Fristoe, 2000).

The main difference between the two locations is the significant difference in scores for French-accented English and its relationship to the Plymouth and SSBE accents. Briefly put, children from London were better at French-accented English than were children from Plymouth. Moreover, while children from London scored higher on French-accented English and the Plymouth accent than on SSBE, participants in Plymouth obtained higher scores for the Plymouth accent and SSBE than for French-accented English. The higher scores for SSBE in Plymouth may be the result of their

[59] For discussions concerning each city individually, see Section 5.4 for Plymouth and Section 6.4 for London.

higher degree of direct exposure to this specific accent (see Section 4.4). However, this reasoning cannot be applied to French-accented English, as children in both locations heard this specific accent a comparable number of hours. Therefore, it is currently not possible to explain this difference.

The categorization task reveals greater differences across locations. With regard to the accuracy scores presented in Section 7.2.1, the most noticeable contrast is that of developmental trajectories, or lack thereof. In London, younger participants and older participants were equally good at categorizing the speakers in the task. In contrast, in Plymouth, older participants outperformed younger ones. Interestingly, the comparison of both cities demonstrates that the differences in developmental patterns between the two cities was the result of the younger group from London being better at the task than the corresponding group from Plymouth. In fact, the scores of the older groups across sites did not differ from each other.

This difference in categorizing skills can be attributed to the main difference between the participants from London and Plymouth: linguistic exposure. As Section 4.4 described, the main demographic variables differentiating these groups were those involving exposure to linguistic diversity. It may therefore be argued that it is increased linguistic exposure that enabled younger children from London to outperform their Plymouth peers.

The relationship between exposure to linguistic diversity and the ability to categorize accents is supported by findings in the relevant literature. As described in Section 6.4, people who regularly hear different accents — both adults and children — perform better at categorizing accents than people who are not exposed to such diversity (e.g. Clopper & Pisoni, 2004, 2007; Evans & Lourido, 2019; Floccia, Butler, Girard, et al., 2009; Jeffries, 2019; Williams et al., 1999). Importantly, this need not be specific exposure to the accents included in the stimuli, unlike the case with intelligibility above. The following two studies illustrate how *general* linguistic exposure affects children's ability to categorize accents[60]. First, Floccia, Butler, Girard, et al. (2009) divided participants into monodialectal, if both of the participant's parents spoke with the local accent, or bidialectal, if at least one of the parents spoke with a non-local accent. In a similar way, Evans and Lourido (2019) studied monolingual and both bi- and multilingual children from London. Both studies included accents in their corresponding stimuli that were somewhat unfamiliar to participants. Despite this, children who had experienced more linguistic diversity were better at sorting accents into the correct groups.

The older group of participants in both Plymouth and London achieved comparable results in the categorization task. This may reflect the fact that developing the skill to categorize accents is subject to more drastic changes during early and mid-childhood

[60] Alcorn et al. (2020, p. 668) further proposed that it is exposure to linguistic diversity during early years that may create better representations of nonnative accents.

(Dossey et al., 2020; Jones et al., 2017). While adult-like results are not achieved in this task until the age of 16 years, the learning curve becomes less steep around the age of 12 years (Jones et al., 2017, p. 26). If the older participants in our sample, who span from 11 years to 11 years and 11 months, were approaching a flatter developmental stage, the differences across groups may be diluted. Finally, even the oldest participants in the sample obtained far from perfect scores, as shown in Table 7.3. This should not be surprising. Jones et al. (2017, p. 25) showed that even adults fail to correctly categorize different accents using the same type of task.

Regarding the accents' rankings across cities, children from London were better at categorizing the two L2s in the stimuli: Chinese- and French-accented English. This may perhaps reflect the fact that the London participants had more exposure to L2 accents in general (see Section 4.4). Not only did they have a higher percentage of parents who spoke with L2 accents than did participants from Plymouth; they also interacted more with people from outside the family unit who spoke with L2 accents than did participants from Plymouth. Nevertheless, exposure cannot explain the similarities involving the local accents (the MLE and Plymouth accents) in the two cities. Specifically, children from both London and Plymouth were equally good at categorizing the Plymouth accent and MLE. Because the Plymouth children had significantly higher exposure to their local accent and London children had much more exposure to MLE, exposure cannot be the only factor at play. This is further discussed in Section 7.4.1. below.

The findings from the verbal-guise task also reveal differences across the two locations. The main differences concerned the developmental trajectories for the questions SMART and HARDWORKING. Visual representations of the developmental differences for these two question types are shown below. Figure 7.14 depicts the developmental trajectory in the SMART variable in both cities. Figure 7.15 does the same for HARDWORKING.

Figures 7.14 and 7.15 can be read in the same way. In each figure, the table to the left shows the developmental trajectory for Plymouth, while the table to the right depicts the developmental trend in London. In each table, the horizontal axis depicts age in years at 6-month intervals. The vertical axis illustrates pairwise comparisons. These are to be interpreted in a two-step fashion. First, the left-most accent labels indicate which accent the labels to the right are being compared to. Each horizontal bold line delimits which accent is used in these comparisons. For example, the first four rows of Figure 7.14 perform pairwise comparisons between Chinese-accented English, on the one hand, and French-accented English, MLE, SSBE and the Plymouth accent, on the other.

Second, colored regions indicate that the pairwise comparison between the two accents is significant in the relevant age range. Pink shaded areas illustrate that the left-most accent was rated significantly lower than the accent in the relevant row. Green shaded areas show that the left-most accent was rated significantly higher than the

accent in the specific row. For example, Figure 7.14 shows that, between the ages of 10 and 12 years in London, Chinese-accented English is rated as smarter than MLE. Similarly, it also shows that, in Plymouth, MLE is rated lower than SSBE between the ages of 9 and a half and 12 years.

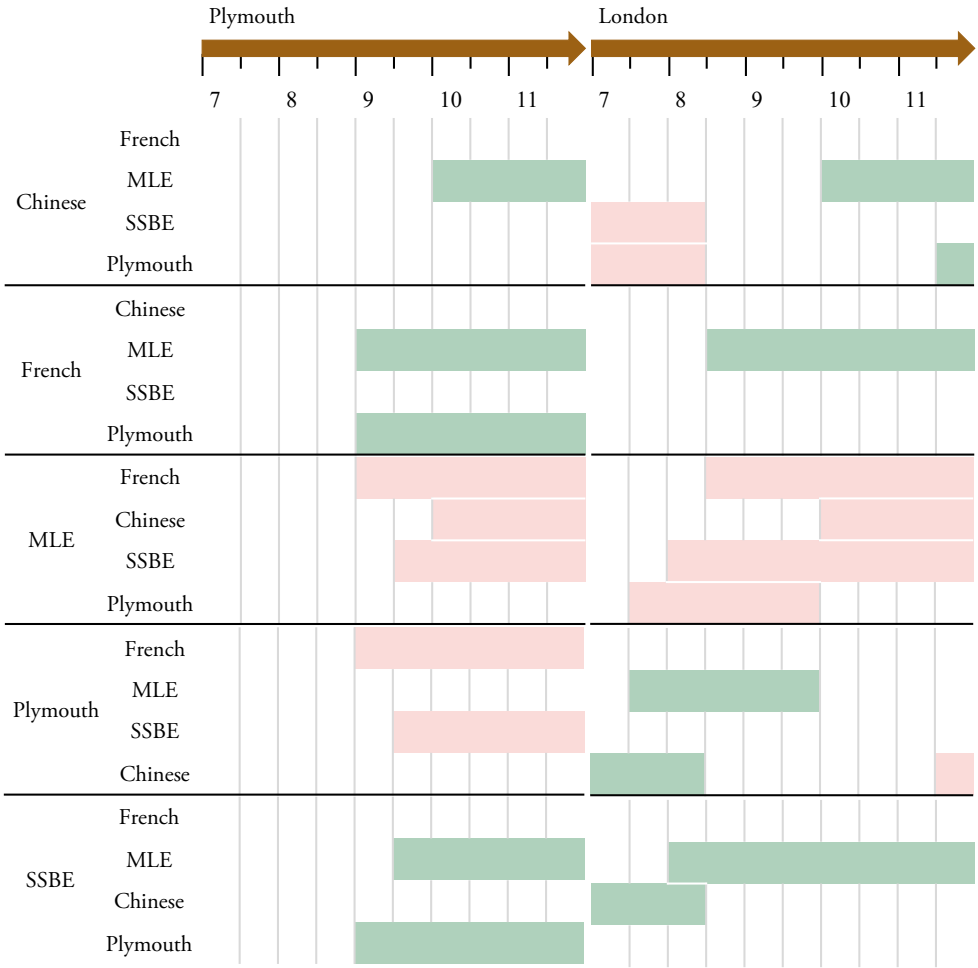


Figure 7.14. Developmental changes, SMART, across cities
Pink regions indicate that the left-most accent was rated lower than the accent in the relevant row. Green regions indicate that the left-most accent was rated higher than the accent in the relevant row.

Figures 7.14 and 7.15 show that the developmental patterns for SMART and HARDWORKING were similar. Therefore, they are commented on together. The trajectories are quite distinct between the two cities. In London, children's accent stereotypes changed with age: the accents younger participants rated lowest were rated

highest by older participants. Likewise, the accents that obtained the highest ratings by younger participants received the lowest scores by older ones. The main exceptions are SSBE, which was generally rated positively, and MLE, which received overall negative scores. In contrast, Plymouth children did not show signs of accent stereotyping until the age of 9 years. From there onwards, they slowly developed an adult-like pattern (see Section 5.4). Importantly, the ratings by older participants across the two cities did not differ from each other. The main difference between London and Plymouth is therefore how children reach that point.

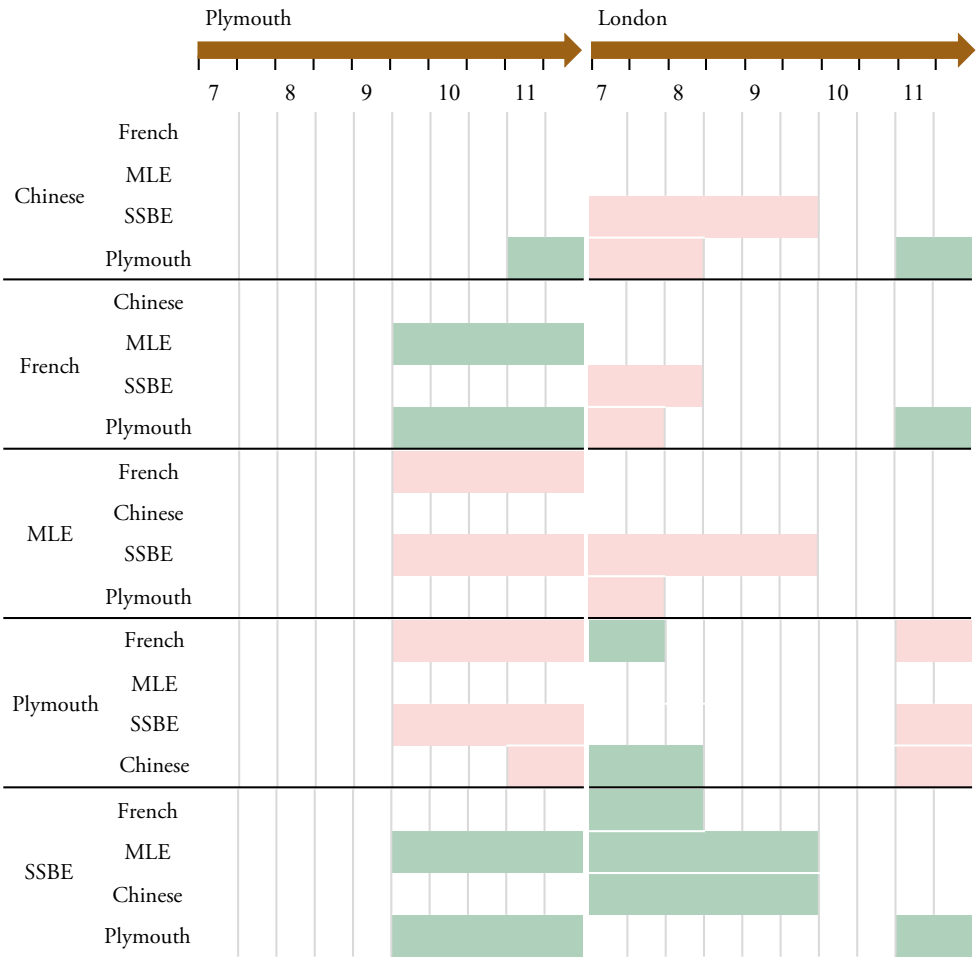


Figure 7.15. Developmental changes, HARDWORKING, across cities
Pink regions indicate that the left-most accent was rated lower than the accent in the relevant row. Green regions indicate that the left-most accent was rated higher than the accent in the relevant row.

The fact that children learning the same language in different regions of the same country appear to develop accent stereotypes in different ways is not completely unexpected. As mentioned in Section 3.3.3, Kinzler and DeJesus (2013b) collected data from 5- and 6-year-olds as well as 9- and 10-year-olds in both the northern and southern US. While their younger group from the north linked different traits to different accents, their southern peers did not until the age of 9 and 10 years. In fact, one can argue that there are certain similarities between Kinzler and DeJesus (2013b) and the present thesis. In both studies, participants from a region in which a non-prestigious accent is spoken developed accent stereotypes around the age of 9 years. In contrast, participants from more prestigious regions (London in this study and Chicago in Kinzler & DeJesus, 2013b) showed associations between accents and personality traits at a younger age.

Similarly, the later development of accent stereotypes in children from Plymouth can be said to match the results in Cohen and Haun (2013). In their study of linguistic prejudice, they found that children speaking Brazilian Portuguese did not rate speakers of different accents differently until the ages of 9 and 10 years. Cohen and Haun (2013) pointed out that this was a deviation from previous studies, which have shown signs of linguistic prejudice much earlier. In Section 3.3.1, it was posited that the reason may be that Brazilian Portuguese is considered less prestigious than its European counterpart. The lack of prestige associated with this variety could potentially have affected how children develop language attitudes toward different accents. Such "delayed" signs of linguistic prejudice, as shown in Cohen and Haun (2013), can be argued to match the finding in this investigation that children from Plymouth did not judge speakers differently until approximately the same age, between 9 and 10 years. More generally, the fact that sociocultural context during language development affects how children develop accent stereotypes and language attitudes is also found in Kinzler, Shutts, et al. (2012). In their study, they focused on linguistic prejudice. They showed that whether a child grows up in a rural or urban community, even within the same city (Cape Town, South Africa, in the case of Kinzler, Shutts, et al., 2012), also affects what views children have regarding different languages. Specifically, children growing up in a rural setting did not prefer speakers of either Xhosa (the local language) or English (the *lingua franca*). In contrast, children who attended school in Cape Town chose English speakers over Xhosa (see also Papazachariou et al., 2018 for how children growing up with nonstandard varieties seem to develop language attitudes at later ages).

Finally, only prestige variables (SMART and HARDWORKING) gave rise to statistically significant interactions. None of the questions involving the warmth factor (FUN, FRIENDLY and TRUST) were subject to significant interactions. In Section 5.4, this was related to two lines of research. On the one hand, there is sociolinguistic research with adults showing that questions dealing with the prestige factor are more consistent across studies. In contrast, warmth variables are less stable. On the other hand, the summary of the literature in Section 3.3.3 established that children appear

to learn associations between code and personal traits involving prestige variables at an earlier age than warmth variables. The comparison of the results from both cities in this study seems to give further support to both of these phenomena.

There may well be a further nuance to the statistical significance of the SMART and HARDWORKING questions. As mentioned in Section 2.2.1 for attitudes in general and Section 2.3 for language attitudes specifically, people's reactions to an attitudinal object are highly dependent on context. In the case of language, two classic examples include Callan et al. (1983) and Carranza and Ryan (1975). The former investigated whether Greek-accented English and the local, Australian accent receive different ratings based on, among other things, context. Participants were told that speakers were recorded talking at a bus stop, at home or at school. The effect of accent was strongest in the school context and weakest in the bus stop context. Furthermore, Carranza and Ryan (1975) carried out a similar study including GA and Mexican Spanish in the US. Their contrast of context also included both school and home settings. In their results, the differences in ratings for the two accents were greater in the school setting than in the home setting[61].

The way the current investigation was framed to participants was with a school setting in mind (see more details in Section 4.1). In the verbal-guise, participants were told that the speakers are "5 students read[ing] a text in class" and that their answers will help us "fill out their [the speakers'] reports." This setting may have led to the SMART and HARDWORKING variables being more salient, perhaps to the detriment of the variables related to the warmth factor. This may have had an effect in two ways. On the one hand, both the SMART and HARDWORKING questions are arguably quite relevant in a school setting. For HARDWORKING, anecdotal evidence can be found in the pilot phase of this investigation. When two 7-year-old pilot participants were asked whether they knew what *hardworking* meant, they said it means that (i) they study hard at school and that (ii) they do all their homework (both paraphrased). On the other hand, the content of the stimulus text (a polar bear hunting animals and feeding its cubs; see Section 4.2.2.) may also have made the warmth questions less pertinent due to its topic.

Before dealing with how the results of this chapter inform the RQs in Chapter 1, one further issue is discussed. The ages between 9 and 10 years appear to be particularly important for the results in this investigation. As mentioned above, children from Plymouth began developing accent stereotypes in this age bracket. It is also at this point when their categorization skills matched those from their London peers. Furthermore, in this age range children from London began showing accent stereotypes matching the adult norm (see above and Section 6.4), which contrasts with the associations they had

[61] Incidentally, in both Callan et al. (1983) and Carranza and Ryan (1975), prestige variables were subject to stronger effects than warmth variables, further supporting the discussion in the previous paragraph.

at a younger age. All these changes were also in accordance with the summary of the literature in Chapter 3, where it was concluded that adult-like attitudes begin emerging around the age of 10 years.

The data collection in this thesis does not allow determination of what these changes are the result of. However, the literature on late child development offers some tentative and potential answers that are briefly discussed below. First, the literature from late language development is discussed, followed by literature on social and cognitive development. This should not be understood as implying that these forms of development are independent of each other. Both influence each other (Nippold, 2004, 2007). They are dealt with separately for ease of exposition.

From a linguistic point of view, different important developmental changes take place around the ages of 9 and 10 years (see a summary in Nippold, 2007, pp. 361-368). First, children begin making use of complex meaning structures and understanding that language can "mean more than it says" (Dockrell & Messer, 2004; Peskin & Olson, 2004; Sullivan, 2006). Furthermore, their ability to infer meaning from context is subject to considerable improvement during this age range (Dockrell & Messer, 2004; Nippold, 2002). Second, children begin showing an increased focus on not only what is said, but also how it is said. This is evident by their improved use and understanding of style, register and politeness (Halmari & Smith, 1994; Jisa, 2004; Khorounjaia & Tolchinsky, 2004; Peskin & Olson, 2004). Finally, the way they interact with others also changes. They are more precise in their exchanges and aware of potential misunderstandings (Lloyd, 1991).

With regard to general socio-cognitive development in this age range, further important changes take place. First, children's understanding of taxonomies and transitive properties drastically improves (García Coni et al., 2019). Second, their long-term and working memories become better, partially due to their enhanced cognitive schemata (Abbott & Burkitt, 2023). Third, their self-image also changes. From describing themselves based on physical attributes, they begin to use more abstract and psychological concepts and take into account how others see them (Harter, 1999, 2006). Fourth, they are also better at knowing when and how to display emotions as well as at reading the emotions of others (Manstead, 1995) and at recognizing faces in a wholistic fashion, rather than as multiple individual features (Baenninger, 1994).

Lastly, there are also significant changes in peer relations during this age range. Peer relationships begin becoming more relevant and highly valued than those with family members (Gummerum & Keller, 2008; Rubin et al., 2006). Peers are seen not only as "children I play with," but also as people with shared interests, who help each other and who have desirable attributes (Gummerum & Keller, 2008; McDougall & Hymel, 2007; Youniss, 1980). Children are also better at describing peers using more abstract terms such as their beliefs, values, and personal traits (just like their self-image; Gummerum & Keller, 2008).

The previous three paragraphs briefly highlight results from different areas of late child development. These results have one thing in common. Around the ages of 9 and 10 years, children begin gaining a deeper insight into their social world and the intricacies of language use. This new focus may enable them to tune their attention to how language is used and what others' use of language says about them. Therefore, the fact that children in this age range begin showing accent stereotypes similar to those of adults (and, in the case of Plymouth, to better categorize speakers) may be the result of these developmental changes in some way. Naturally, this investigation has not gathered data that allows direct testing of this relationship. Further research is therefore needed to determine whether this is the case and, if so, what specific socio-cognitive factors affect the developmental of sociolinguistic perception and evaluation (see also Section 8.3).

7.4.1 Accent types: native, nonnative, L1 and L2

This section discusses how the findings from the two cities inform RQ2 from Section 1.1, namely, whether children develop stereotypes toward L2 accents at an earlier age than toward L1 accents. Before this, it deals with whether this distinction between L1 and L2 accents is relevant for explaining the results of the two auxiliary tasks in this investigation.

As discussed in Sections 5.4.1 and 6.4.1, how well children understand accents in noise does not seem to depend on whether the accent is L2 or L1. For example, in the case of children from Plymouth, scores for French-accented English were just as high as for SSBE and, in the case of London, French-accented English scored as well as the Plymouth accent. Furthermore, the results were also argued not to support the need for a strict native-nonnative split. For instance, in Plymouth, children's scores did not differ between the local, Plymouth accent and French-accented English. Nevertheless, the direct comparison between the data from Plymouth and London raises a question regarding this last point. In Section 5.4, it was argued that the reason why the Plymouth accent scored highest among the children from Plymouth is because it is the local accent. This was not the case for London children, however. MLE ranked second-to-last among London participants. This is an unexpected result, since MLE was expected to be the local, native accent. A key to understanding this disparity of the effects of local accents might be revealed by comparing the scores in the two cities. Figure 7.2 in Section 7.1 above shows that the scores for the Plymouth and MLE accents did not differ across cities. This may imply that the higher scores for the Plymouth accent in Plymouth may not be the result of it being the local accent. Why the scores for the presumably local accents did not differ across cities may be explained by the participant pool in this study.

Section 4.4 indicated that most of the children in this investigation came from middle- to upper-middle-class families. This is a key demographic variable. Accent in

the UK is strongly correlated not only with geography, but also with class: the higher a person's SES, the greater the chances of them speaking an accent closer to SSBE (see Section 2.1). Therefore, the families in this study can be expected to speak with standard accents, rather than with local ones. This is perhaps corroborated by the data from the exposure questionnaire in Section 4.4. It showed that SSBE is the accent children, regardless of city, had most direct contact with, followed by the local accent (the Plymouth accent in Plymouth and MLE in London). This may suggest that SSBE could be considered the native accent of our participants (as hinted at in Sections 5.4.1 and 6.4.1). This does not explain why the scores for MLE and Plymouth did not differ between cities. Specifically, it doesn't account for why the Plymouth accent scored significantly higher than MLE in both locations.

One potential reason may be that the Plymouth accent (specifically the speakers included in the stimuli of this thesis) is phonetically closer to SSBE than MLE is to SSBE. As mentioned in Section 5.4.1, phonetic distance across accents has been argued to explain why people have difficulty understanding them. In other words: accents phonetically closer to one's native accent are easier to process, while accents phonetically more different are harder to understand (e.g. Bent et al., 2021). If the Plymouth accent is phonetically closer to SSBE, an accent that may be considered the native accent of our participants, it would explain why it scores higher. However, without quantifying the phonetic differences between the speakers of the SSBE, MLE and Plymouth accents in this study, as in Bent et al. (2021), it is impossible to be certain whether these results reflect a correlation between phonetic distance and intelligibility[62].

A similar discussion applies to the categorization task. In terms of the accuracy scores (Section 7.2.1 above), the results from both cities show that a L1–L2 divide does not explain the data. For example, in Plymouth, French-accented English scored just as high as the Plymouth accent. Comparably, in London, the scores for Chinese-accented English did not differ from the SSBE or Plymouth accents up until the ages of 9 and a half and 10 years, respectively. It is the local accents that require further looking into. Just as above with intelligibility, the scores for the MLE and Plymouth accents did not differ across cities. Children in the two cities were equally good at categorizing these two accents. Moreover, it is true of both cities that the Plymouth accent and SSBE are easier to categorize than MLE, although the difference between Plymouth and MLE in London only bordered on significance (p = .077 for Plymouth–MLE and p = .052 for SSBE–MLE). The fact that both the Plymouth accent and MLE behave in a similar fashion across cities might provide further support for seeing SSBE as closer to the participants' native accent than perhaps the accents of Plymouth and MLE, as discussed in the previous paragraphs. The scores may therefore show that the ability to categorize

[62] While this study does not include such measures, the results from the preparatory study (see Section 4.3.2) showed that the strength ratings of the Plymouth and MLE speakers included in the stimuli did not differ significantly from each other, whereas both of them differed from SSBE.

accents depends on phonetic distances between accents as well. However, there is no literature that directly correlates categorization skills with phonetic differences between accents (the way Bent et al., 2021 do for intelligibility). Therefore, it can only be posited as a tentative explanation.

In contrast to the accuracy scores, the co-occurrence matrices (Section 7.2.2) show a clear division between L1 and L2 accents in the two cities. In both cases, participants were more likely to confuse French-accent English with Chinese-accented English and, conversely, Plymouth, MLE and SSBE with each other. The main difference between participants from the two locations was the structure of the branch containing the L1 accents. In Plymouth, the overall structure was that of the standard accent (SSBE), on the one hand, and the non-standard accents (Plymouth and MLE), on the other. In contrast, children from London grouped the accents they had been exposed to most (MLE and SSBE) more frequently than any of them with the Plymouth accent. This difference may be the result of the specific phrasing of the task. As described in Section 4.1.3, participants were told to "put together the people that were born in the same city." SSBE is frequently said to be an accent that cannot be associated with a particular geographical region (see Section 4.2.1). Nonetheless, both SSBE and MLE are accents associated with the Southeast of England in general and London specifically (e.g. Cole & Strycharczuk, 2024). Therefore, participants from London may simply have assumed that they were "born in the same city," that is, in London.

The verbal-guise task is discussed below. A priori, the comparison of the SMART and HARDWORKING findings suggests that children did not necessarily develop stereotypes toward L2 accents at an earlier age than toward L1 accents. In Plymouth, for example, both French-accented English and SSBE were rated as smarter than the Plymouth accent and MLE when the first signs of accent stereotypes appeared at the age of 9 years. Similarly, the youngest participants from London rated French- and Chinese-accented English as low as MLE. Therefore, a strong divide between L1 and L2 accents does not account for the results. However, the data from London did not show *when* children from London began developing accent stereotypes, as the data from Plymouth did. The younger participants from Plymouth rated all accents in a similar way, which allows us to see *when* and *how* stereotypes appear. Therefore, it is not possible to know whether children in London would have first developed negative associations toward French- and Chinese-accented English and, at a later age, toward MLE.

Interestingly, the stereotype ratings by younger London participants depart from adult norms, as described in Section 6.4: the Plymouth accent was rated high on both SMART and HARDWORKING and French-accented English low on these same questions. This may imply that they were not replicating societal norms. Their responses may be, at least partially, based on a binary divide between "accents similar to mine" and "accents different from mine." However, the fact that MLE, the local accent, was rated low matches the adult cultural norm (see Section 4.2.1). Therefore,

this pattern may reflect an amalgamation between children acquiring sociolinguistic norms and a general differentiation between L2–L1. However, since the data did not show when London participants began showing accent stereotypes, this cannot be confirmed. In contrast to this pattern, when children from Plymouth started showing signs of accent stereotypes, they adopted the adult sociolinguistic norms (see discussion in Section 5.4). Therefore, *when* children acquire stereotypes seems to also affect *how* they do it.

Finally, with regard to the native–nonnative split, such a strict divide does not seem to explain what accent stereotypes children have. The specific kind of evidence for this argument is dependent on what accent(s) should be considered the native accent of our participants (see the discussion at the beginning of this Section). If the Plymouth and MLE accents are considered the native accents, it is not the case that Plymouth children started favoring their local accent over the other four accents or that the London children rated MLE above the rest. In fact, in both cases, the local accent was rated significantly lower. If SSBE were to be taken as the native accent of the participants in this investigation, it did get higher ratings in the two cities. Nonetheless, it was not the case that, at any point in the developmental trajectories described above, SSBE was the only accent rated highest. For example, London participants rated SSBE as high as the Plymouth accent when they were younger and as high as French-accented English when they were older. In a parallel fashion, Plymouth children in this investigation rated SSBE as high as French-accented English between the ages of 9 and 11 years[63]. Therefore, SSBE never received special status by being the native accent, as would be expected if children were to treat their native accent differently from all nonnative accents.

7.4.2 Accent stereotypes, intelligibility and categorization

RQ3 and RQ4 in Section 1.1 concern whether the development of accent stereotypes relies on other sociolinguistic skills: being able to categorize and understand other accents. This section discusses how the data from the two cities inform these RQs. The results of the verbal-guise and the intelligibility task are discussed first. The relationship between the verbal-guise and the categorization task is discussed immediately thereafter.

The data from both cities supports the idea that there is not a relationship between intelligibility and accent stereotypes (whether positive or negative). For example, London's accuracy scores for all accents in the intelligibility task improved in parallel throughout development. In contrast, how these accents were rated in the verbal-guise task changed with age. French-accented English went from being rated among the

[63] Just as it was pointed out regarding the discussion of L1–L2 accents, the London data does not show *when* children started having accent stereotypes. It could be the case that they would have rated SSBE above the Plymouth accent if data from younger children had been collected.

lowest to receiving some of the highest scores. Conversely, the Plymouth accent was rated high by the younger group but was rated lower by older participants. If intelligibility had influenced the ratings in the verbal-guise task, one would expect the accents obtaining lower scores in the intelligibility task to have been rated lower. This is not the case. The same pattern was found in Plymouth. The SSBE and Plymouth accents were the ones children understood best throughout development. However, while SSBE obtained high ratings in the SMART and HARDWORKING questions, the Plymouth accent was rated significantly lower than SSBE.

In contrast, an association between the categorization task and the verbal-guise task is more likely. Children from Plymouth only showed signs of accent stereotypes after the age of 9 years. In contrast, participants from London attached different traits to different accents already by the age of 7 years. This disparity was paralleled in their categorization results. Children from Plymouth improved their scores in the categorization task with age, while, in London, participants were equally good throughout the 5-year span. This means that younger children from London were better at the categorization task than younger children from Plymouth. Importantly, the age at which the scores of children from Plymouth started matching those of their London peers was 9 years, the same age at which they started showing the first signs of accent stereotyping.

Therefore, the data seems to suggest that being able to categorize accents is a pre-requisite for having stereotypes toward them. This would provide a positive answer to RQ3. As mentioned in Section 5.4.2, this relationship seems to be somehow indirect. A strong relationship between the ability to categorize accents and accent stereotypes would imply that the accents children categorize best are those they develop stereotypes toward first. This is not the case. For instance, in Plymouth, children had the most difficulty categorizing the MLE accent. Nonetheless, MLE was one of the first accents children began associating with low ratings in the SMART and HARDWORKING questions. The data from this study supports a weaker version of this relationship. Namely, a certain skill level in categorizing accents is required for children to develop stereotypes toward them.

Finally, the ability to categorize accents was correlated with the degree of exposure children had to linguistic diversity (see the beginning of Section 7.4). Hence, a further, indirect link may be posited between linguistic exposure during development and accent stereotyping. To be specific, the more exposure a child has to linguistic diversity while growing up, the earlier they can associate people speaking with different accents with different traits. As a consequence of this finding, RQ1 in Chapter 1 (whether linguistic diversity in a child's environment while growing up shapes how or when they develop accent stereotypes) receives some support as well. However, this relationship seems to be contingent on children's ability to categorize accents.

This (indirect) association between exposure and language attitudes differs from the results in Paquette-Smith et al. (2019). They collected data from 5- to 6-year-olds in

the Greater Toronto Area. Their task had children listen to two speakers and decide who they would like to become friends with. The stimuli included the local Canadian accent, a British accent and Korean-accented English. They categorized participants into three groups based on how much exposure they had to different accents. Their results suggest that there is no relationship between exposure and how often children would choose the local speaker as a potential friend.

Three potential factors may explain why their data did not show an effect of exposure. First, following the classification of the literature in Chapter 3, Paquette-Smith et al.'s (2019) study investigated linguistic discrimination, while the current study deals with accent stereotyping. It could be the case that exposure to linguistic diversity affects different aspects of language attitudes differently. This would be parallel to the different effects of exposure discussed above with regard to intelligibility and categorization (while exposure to a specific accent only improves intelligibility in said accent, general exposure to linguistic diversity sems to improve children's categorization skills). Therefore, it may be the case that linguistic exposure affects different aspects of language attitudes differently.

Second, how exposure was categorized in Paquette-Smith et al. (2019) may also help explain the disparities in the results. Just like in this thesis, Paquette-Smith et al. (2019) categorized participants based on *direct* exposure to accents (see Section 4.4). Nevertheless, their participants may still receive significant *indirect* exposure. The Greater Toronto Area is "heralded as being one of the most culturally and linguistically diverse regions in the world" (Paquette-Smith et al., 2019, p. 811). Therefore, even if children's parents did not report high levels of direct exposure to multiple accents, children may still be exposed to more incidental linguistic diversity than our group from Plymouth. Incidental linguistic exposure has been shown to be important for language attitudes in children. As described in Section 3.3.2, Howard et al. (2014) reported that even incidental exposure to different accents can affect how willing 19-month-olds were to imitate different speakers (a study that can also be classified under linguistic discrimination; see Chapter 3). Incidental exposure in Toronto may hence have influenced Paquette-Smith et al.'s (2019) participants' behavior regardless of their direct exposure to accents.

Third, the potential importance of general sociocultural factors cannot be neglected. The present investigation collected data in England, while the participants in Paquette-Smith et al. (2019) were in Toronto, Canada. As Section 3.5 argued, it cannot be assumed a priori that children growing up in different countries will develop language attitudes in a parallel fashion. As the results from Kinzler and DeJesus (2013b), Kinzler, Shutts, et al. (2012) and the verbal-guise task in this study corroborate, even children growing up in the same country can experience different developmental trajectories. One can only expect that greater social and cultural differences (such as across two countries) might affect how children react to linguistic diversity. For example, sociolinguists have sometimes noted that speakers from England are very sensitive to

accent differences when it comes to social judgements, perhaps more so than in other English-speaking countries (see, e.g., Adams, 2018; Cruttenden, 1994, 2014; Giles & Powesland, 1975; Gimson, 1970, p. 83; Upton, 2008). If this is the case, children raised in England could potentially grow up learning that linguistic diversity is a much more important social cue. This, in turn, may affect how their development of linguistic attitudes in general progresses.

7.5 Summary

This chapter compares the results from Plymouth and London. Its aim is to reveal how a child's environment during upbringing affects when or how they develop accent stereotypes. The results of the intelligibility task show that children performed equally well across cities. This applies to both their scores per word (Table 7.1) and by accent (Figure 7.1). The main difference between the two cities is that children from London were better at understanding French-accented English than their Plymouth peers. With regard to the categorization task, younger children from London outperformed younger children from Plymouth. In contrast, older participants' scores did not differ across cities. Furthermore, children from London were better than children from Plymouth at correctly categorizing French- and Chinese-accented English. Finally, two questions form the verbal-guise task give rise to significant interactions: SMART and HARDWORKING. Participants from Plymouth only rated speakers differently on these two questions from the ages of 9 years and 9 and a half years, respectively. In contrast, participants from London already showed signs of accent stereotypes at the age of 7 years. Their specific associations changed with age. Younger participants from London have therefore accent stereotypes different from those of their older London peers.

8 Conclusion

This chapter closes this thesis on the development of accent stereotypes. It first recaps the main takeaway messages (Section 8.1). Then, in Section 8.2, it summarizes the main findings of the investigation with reference to the RQs posed in Section 1.1 and the general field of the development of language attitudes. Finally, directions for future work are outlined in Section 8.3.

8.1 Summary

The superordinate goal of this thesis is to investigate the factors that drive the development of accent stereotypes in late childhood. With this aim in mind, Chapters 2 and 3 provide the theoretical background of this investigation. With regard to the former, Chapter 2 encapsulates how the terms *accent* and *stereotype* are understood in this thesis. The goal of doing so is twofold. On the one hand, it is to bridge the gap between sociolinguistics and (developmental) psychology, the two main academic branches that have studied the development of language attitudes. On the other, it is to problematize how these (and related) terms have been used in the literature.

Chapter 2 defines *accent* as the regularized segmental and suprasegmental differences across language varieties. Moreover, it highlights two common misunderstandings associated with the linguistic concept of *accent*. First, Chapter 2 stresses that every speaker of a spoken language has an accent. Second, it is emphasized that accents are not clear-cut linguistic entities that can be unequivocally described. The difficulty in establishing an unambiguous accent as well as the concept of accent strength entail that two speakers of the same accent may speak in significantly different ways. With regard to *stereotype*, Chapter 2 takes a social psychological approach and includes it under the umbrella term *attitude*. Stereotypes are defined as beliefs about the traits and behaviors of a given social group. These beliefs have implications for how individuals treat others and how they interpret their behaviors. Chapter 2 further indicates that stereotypes, like accents, are not stable entities. Chapter 2 notes that this flexibility of stereotypes does not reflect different *kinds* of stereotypes, but different ways in which they are *evoked*. Chapter 2 finishes by introducing the field of language attitudes, noting how it has been partially neglected in the field of social cognition and, finally, how the study of language attitudes can benefit from being anchored in Cognitive Linguistics.

Chapter 3 offers an overview of the literature dealing with how children develop language attitudes. Research in this area is rather fragmented, which makes it challenging to create generalizations to understand the processes underlying this development. Chapter 3 breaks down this field into three broad categories. These include the development of (i) the Standard Language Ideology, (ii) the association between codes and speakers' geographical background and (iii) attitudes toward speakers of different codes (including prejudice, discrimination and stereotypes). Chapter 3 illustrates how children use language as a social tool from a very young age: even newborns as young as 0 to 5 days old prefer to listen to the language their mother spoke during pregnancy or rhythmically similar languages. However, the specific timeline of how language attitudes develop from birth is influenced by whether a child speaks other languages, their exposure to linguistic diversity, SES and what specific code the child speaks. Chapter 3 also emphasizes the difficulties associated with reaching these conclusions. The vast differences in methods and analyses across studies pose serious barriers to the field.

Chapter 4–7 constitute the experimental side of this investigation. Chapter 4 describes and motivates the methods used for the data collection. It does so against the backdrop of the importance of methodological considerations in developmental psychology in general and developmental sociolinguistics specifically. It describes how the methodological choices were made to create a set-up appropriate for the youngest participants as well as to allow comparison of the results with the current literature. The results of the data collection are introduced in the following three chapters. Chapter 5 deals with the results from Plymouth, while Chapter 6 focuses on the data from London. Chapter 7 then compares the results across the two cities in an effort to better understand how a child's environment during upbringing affects their development of accent stereotypes.

A brief description of the results follows (see Section 8.2 for a discussion of how these results relate to the study's RQs). First, participants in both cities became better with age at understanding different accents. The main difference between the cities of Plymouth and London was how the accents in the stimuli ranked across locations. In London, participants obtained better scores with French-accented English and the Plymouth accent than with SSBE. In Plymouth, participants were better at the local Plymouth accent and SSBE than with French-accented English. Second, the categorization task provided more considerable differences. Children from Plymouth became better with age at categorizing accents. In contrast, there was no developmental effect on the London data. This difference between the two cities was due to the younger participants from London being better at the task than the younger participants from Plymouth. Furthermore, participants from London were, overall, better at categorizing the L2 accents in the stimuli than were their peers from Plymouth.

The verbal-guise task provided significant results for two out of the six variables included (SMART and HARDWORKING). The developmental trajectories for these

two variables were different across the two cities. In Plymouth, children rated speakers the same way between the ages of 7 and 9 years. From this age onwards, they started showing accent stereotypes, rating French-accented English and SSBE higher than the Plymouth accent and MLE. In contrast, London children already showed signs of accent stereotypes by the age of 7 years. Their development followed a cross-like pattern. The younger London participants rated the SSBE and Plymouth accents highest among all other accents, while the older participants rated SSBE and French- and Chinese-accented English highest. Importantly, the ratings of older participants did not differ across cities.

8.2 Contributions

The overarching goal of this thesis is to answer the twofold question of *how English children develop accent stereotypes in late childhood and what factors influence this process*. As already mentioned in Section 1.1, it is not possible for this thesis to answer this question due to the multifaceted nature of accent stereotypes. Nonetheless, the results of this investigation enable some insights to be gained, thereby contributing to the topic. Specifically, there does not appear to be one single way in which children develop accent stereotypes. How a child learns to associate someone's accent with their personal qualities depends on their cognitive abilities as well as their cultural and social context. This is not completely unexpected, as shown by two lines of research. On the one hand, Section 7.4 discussed how the development of language attitudes in general seems to be dependent on sociocultural factors. On the other, Section 2.4 framed this thesis within the theory of Cognitive Linguistics. This theoretical approach sees language as embodied in a cultural, physical and psychological sense. Therefore, the fact that how children learn accent stereotypes is not a one-size-fits-all process can be argued to be a given.

The superordinate question in the previous paragraph was narrowed down to four RQs in Section 1.1. The results from Chapter 5, 6 and 7 (summarized in the previous section) allow some answers to be given and inform a general discussion within the development of language attitudes. This is the topic of this section, where each RQ is discussed in turn.

RQ1 concerns whether exposure to linguistic diversity during a child's development affects when or how they show accent stereotypes. The answer to this question would seem to be *yes*. The children from London in this study were exposed to significantly more linguistic diversity in their daily life than were the Plymouth children. Moreover, children from London showed accent stereotypes at an earlier age. Therefore, exposure to linguistic diversity can be correlated with an earlier onset of stereotypes. Whether this relationship is direct or indirect is not certain. The fact that young participants from London had accent stereotypes is argued to relate to the fact that they were better

able to categorize the accents in the stimuli (see below). In Sections 6.4 and 7.4, the fact that the young London participants were better at the categorization task was linked to their greater exposure to different accents and languages. Therefore, it is possible that the effect of linguistic diversity on accent stereotypes is mediated by the ability to categorize accents. Being exposed to linguistic diversity enables children to categorize accents at an earlier age, which, in turn, makes it possible to develop accent stereotypes at an earlier age.

RQ2 concerns whether children develop stereotypes toward L2 accents at an earlier age than toward L1 accents. The data from Plymouth shows that the answer to this RQ is *no*. When Plymouth children began showing signs of accent stereotypes at the age of 9 years, they rated French-accented English and SSBE higher than the Plymouth accent and MLE. That is, an L2 and an L1 accent obtained higher ratings than two L1s. In contrast, the data from London did not allow for clear conclusions. The youngest participants in the dataset already rated speakers differently based on accent. Therefore, it was not possible to establish *when* they developed accent stereotypes. The younger participants from London rated French- and Chinese-accented English (the two L2s in the study) as low as MLE (an L1 accent). Therefore, it can be concluded that, at least between the ages of 7 and 11 years, the L2-L1 divide is not relevant for what accent stereotypes children have. Finally, although not directly related to RQ2, the L2-L1 split did not fully account for the results in either the intelligibility task or the categorization task. The only instance where children's responses matched a distinction between L2 and L1 accents was in the confusion matrices of the categorization task. Children were not necessarily better (or worse) at categorizing one kind of accent. Nevertheless, when they confused accents, French-accented speakers tended to be mixed with Chinese-accented speakers. Similarly, speakers with a Plymouth, MLE or SSBE accent tended to be grouped together more frequently than with L2 accents.

RQ3 asks whether there is a relationship between the ability to categorize accents and accent stereotypes. The comparison of the data across the two cities reveals that the answer is *yes, likely*. The younger participants from London scored better on the categorization task than the younger participants from Plymouth did. The age at which the Plymouth children started obtaining the same scores as their London peers was between 9 and 9 and a half years. Importantly, it was also at this age the Plymouth children showed the first signs of accent stereotypes, whereas the youngest children from London already rated accents differently. It is thus possible to conclude that the ability to categorize accent is necessary for children to have stereotypes toward them. Nevertheless, this statement needs to be further nuanced. Specifically, a strong version of the relationship between accent categorization and stereotypes was not supported. It was not the case that the accents children categorized best were those they developed stereotypes toward first. For instance, MLE was the accent children had the most difficulty categorizing in both cities. However, it was also one of the accents children rated negatively early on during development. Therefore, the relationship between

categorization and stereotypes should be taken in a weaker form: a general skill in categorizing accents is required to develop stereotypes toward them.

Lastly, RQ4 asks whether the development of accent stereotypes correlates with the ability to understand different accents. The data from the two cities suggests that the answer is *no*. If there had been a relationship between intelligibility and accent stereotypes, the expectation would have been that lower scores on the intelligibility task would lead to more negative ratings. This was not the case. Children from both London and Plymouth obtained similar scores on the intelligibility task and, moreover, were subject to comparable developmental effects. In contrast, children from Plymouth did not show accent stereotypes until the age of 9 and a half years, while the participants form London already rated speakers differently based on accent by the age of 7 years. Similarly, the accents children understood best, such as SSBE or the Plymouth accent, were not those they rated most positively. For instance, the Plymouth accent was rated lowest in SMART and HARWORKING by the older participants in both cities.

The results of this investigation can also cast light on issues concerning the development of language attitudes more generally. Specifically, there is a discussion about to what extent accent sensitivity and, as a result, accent-based social decisions are an innate aspect of social cognition or the result of a learned behavior due to exposure (a view often labeled as *protracted learning*).

On the one hand, some researchers posit that humans are predisposed to use language variation as a cue for social information (e.g. Kinzler et al., 2010; Kinzler & Dautel, 2012; Pietraszewski & Schwartz, 2014a, 2014b; Rakić et al., 2011a). To support this, they refer to research showing that infants and newborns choose speakers of their own native language and their own native accent before they are able to understand language or cultural norms. Kinzler et al. (2010, pp. 624-625), for instance, proposed that accent differences may have played a more important role than race in human evolution. On the other hand, some researchers have defended the notion that humans' accent sensitivity is the result of exposure to the linguistic patterns in their environment during development (e.g. Creel, 2018; Creel & Jimenez, 2012; Creel & Seubert, 2015; Dossey et al., 2020; Jones et al., 2017; Wagner, Clopper, et al., 2014)[64]. The evidence for this viewpoint stems from results showing that not all tasks elicit accent biases in young children. Furthermore, infants' and children's ability to differentiate accents seems to follow a U-shape development curve. For example, Butler et al. (2011) reported that 5-month-old infants were able to distinguish between L1 accents, while the 5-year-old participants in Floccia, Butler, Girard, et al. (2009) could not. These results call into question the "innate" nature of accent sensitivity. As Creel (2018, p. 9) put it, "if effects only appear in a limited range of tasks, it seems less likely

[64] Exemplar theory (e.g. Docherty et al., 2018; Docherty & Foulkes, 2014; Docherty et al., 2013; Foulkes & Docherty, 2006) can also be said to maintain such a view.

that accent is a salient social signal than that accent is one of a variety of social signals that are available to children."

The results from the present thesis can be interpreted to support the importance of exposure for the development of language attitudes, that is, the view that accent sensitivity and the use of accent for social decisions are subject to exposure-based effects. The results of both the categorization task and the verbal-guise task can be argued to support this view. With regard to the categorization task, the amount of exposure younger children from London received was argued to greatly influence their better scores compared to their Plymouth peers. Concerning the verbal-guise task, children from the two cities developed accent stereotypes in different ways. In London, the youngest participants already rated SSBE and the Plymouth accent above other accents, with this pattern turning into adult-like norms with age. In contrast, Plymouth children only began showing signs of accent stereotypes (matching the social norms of adults) from the age of 9 years. If accents were supposed to be a predetermined source of social information, one would expect children to first judge the speakers of their native accent — their ingroup — positively and all other speakers — the outgroups — negatively. This was not the case in this investigation. Children in different socio-cultural contexts appear to develop accent stereotypes differently, which is difficult to account for from a nativist perspective. This is not to be understood as meaning that language variation is not important for infants and children. The literature clearly demonstrates that it is. However, how language variation interacts with other aspects of (social) cognition and the sociocultural context of a child appears to be subject to more complex interactions than a nativist account would imply.

8.3 Future research

Chapter 1 highlights that the field of development of language attitudes has not received a great deal of attention. Therefore, there are many ways in which the results from this thesis could inform future investigations. This section focuses on some of them and does not aim at being comprehensive.

First, a future line of research could compare accent stereotypes like those obtained using the verbal-guise task above to those obtained using more implicit methods. The matched-guise technique (and consequently the verbal-guise) is argued in Section 2.3 to be an indirect method that elicits explicit attitudes. Therefore, it does not allow us to tap into the cognitive processes underlying children's stereotypical responses. Returning to the terminology from the Associative-Propositional Evaluation model in Section 2.2.2, more implicit methods would be better suited to testing whether children's accent stereotypes are the result of "the *activation* of mental associations in memory" (Gawronski & Bodenhausen, 2011, p. 61; emphasis in original) or of propositional processes. For this purpose, different methodologies could be used. For

example, Campbell-Kibler (2012) and Pantos (2014) demonstrated how the IAT[65] (Greenwald et al., 1998) can be used to study accent stereotypes. Moreover, the IAT has also been adapted to study the development of non-linguistic attitudes in children (e.g. Baron & Banaji, 2006; Cvencek et al., 2011; Storage et al., 2020). Therefore, a sociolinguistic version of the IAT adapted to younger participants would allow us to see the cognitive processes behind children's accent stereotypes and how they acquire them.

Second, a further direction for future research could involve how accent strength affects when and how children develop accent stereotypes. The strength of an accent is important in the sociolinguistic research (see Section 3.5). On the one hand, speakers considered to have stronger accents receive more negative ratings than speakers of the same, but weaker, accent. On the other, children can differentiate a nonnative accent from their own if it is stronger. A question then arises whether children will show stereotypes at different ages for the same accent based on whether it is strong or weak. Similarly, it could be the case that children show different stereotypes altogether based on whether an accent is stronger or weaker. Therefore, employing a set-up similar to the one described in Chapter 4, in which the stimuli include the "same" accent with different degrees of strength, could help answering these questions. This would allow us to gain a deeper insight into the nature of children's accent stereotypes. For example, if children rate a French-accented speaker the same way regardless of whether the speaker has a strong or a weak accent, it may be that they are basing their responses on the fact that it is a French speaker. However, if they rate a French-accented speaker with a weak accent differently than one with a stronger accent, it may reveal that children do not only rely on the social category of "French speaker," but also on phonetic distance, like adults do.

Third, a different research direction could investigate to what extent the development of accent stereotypes is dependent on the child's general social, linguistic and cognitive development. As mentioned in Section 7.4, the development of accent stereotypes likely relies on children's general socio-cognitive skills. Therefore, the field would benefit from knowing how this development relates to other cognitive processes. One potential line of research could focus on the relationship between the development of accent stereotypes and children's phonological development. Accent stereotypes are arguably a phonetic and phonological task: listeners need to notice that there are phonetic differences across stimuli and categorize these differences based on metalinguistic knowledge. In this context, Nathan et al. (1998) showed relevant results. In their study, they asked 4- and 7-year-olds from London to repeat and define words produced with their local London accent or with a Glaswegian accent. The younger participants provided more *phonetic* repetitions. That is, if they heard the word *bird* in

[65] Of course, there are a plethora of methods that could be utilized to study these implicit associations, such as EEG, lexical-decision tasks, eye-tracking, etc. See Section 2.2.2 and Section 2.4.

a Glaswegian accent with [eɪ], they would pronounce it that way rather than with their native [ɜː] phone. Furthermore, the younger participants also provided more incorrect definitions. In contrast, the older participants repeated the word in their own accent and provided more correct definitions. Nathan et al. (1998) argued that the responses of younger participants were the result of them having fewer abstract phonological representations, in contrast to the older participants. These results show that a child's phonological development affects their perception of accent variation. Therefore, it could be expected that children whose phonological development is faster may be more likely to show accent stereotypes at an earlier age. For this purpose, a future study could include children of the same age who have different scores on direct measures of phonological development (e.g. Dodd et al., 2003).

Last, but certainly not least, the importance of studying how language attitudes develop in children learning other languages (and varieties of the same language) cannot be overstated. Research on language acquisition in general is highly skewed toward focusing on English or other major European languages (e.g. Kidd & Garcia, 2022). The same applies to linguistics as a whole (e.g. Majid and Levinson 2010) and psychology/cognitive science (Blasi et al., 2022; Henrich et al., 2010). This state of affairs was encapsulated by Henrich et al. (2010), who noted that most behavioral research is based on WEIRD societies: *W*estern, *E*ducated, *I*ndustrialized, *R*ich, and *D*emocratic. The same appears to apply to the study of how language attitudes develop. Most of the research presented in Chapter 3 is based on these same populations. If researchers wish to understand the factors that influence how children develop accent stereotypes (or language attitudes more broadly), data is needed from a significant number of cultures and sub-cultures. For instance, this thesis and Kinzler and DeJesus (2013b) have shown that children within the same WEIRD nation can exhibit different developmental trajectories. Therefore, the field can only benefit from investigating how children from a significant portion of the 7,000 existing languages (and the accents and dialects within them) (Eberhard et al., 2024) learn to use language as a tool for social judgements.

9 References

Abbott, R., & Burkitt, E. (2023). *Child development and the brain. From embryo to adolescence*. Bristol University Press.

Adams, Z. (2018). *The persuasiveness of British accents: Enhancing parental self-efficacy to manage children's oral health behaviours* [PhD Dissertation, Queen Mary University of London]. London.

Adorno, T. W., Frenkel-Brunswik, E., Levinson, D., & Sanford, N. (1950). *The authoritarian personality*. Harper.

Agha, A. (2003). The social life of cultural value. *Language & Communication*, *23*(3), 231-273.

Ahrens, F. (2004, 2nd August). Accent on higher tv ratings. *The Washington Post*, https://www.washingtonpost.com/archive/politics/2004/08/02/accent-on-higher-tv-ratings/9ade18bd-6685-43f5-bf2c-a38344337a29/.

Alcorn, S., Meemann, K., Clopper, C. G., & Smiljanic, R. (2020). Acoustic cues and linguistic experience as factors in regional dialect classification. *Journal of the Acoustical Society of America*, *147*(1), 657-670.

Algeo, J. (1998). America is ruining the English language. In L. Bauer & P. Trudgill (Eds.), *Language myths* (pp. 176-182). Penguin Books.

Allan, C. (2023). *Chinese: A growing rival to French in UK schools?* Which School Advisor. Retrieved 16th of August 2024 from https://whichschooladvisor.com/uk/school-news/chinese-a-growing-rival-to-french-in-uk-schools.

Allport, G. (1935). Attitudes. In C. Murchison (Ed.), *Handbook of social psychology* (pp. 798-844). Clark University Press.

Allport, G. (1954). *The nature of prejudice*. Addison-Wesley.

Altendorf, U., & Watt, D. (2008). The dialects in the South of England: Phonology. In B. Kortmann & C. Upton (Eds.), *Varieties of English. The British Isles* (pp. 194-222). Mouton de Gruyter.

Álvarez-Mosquera, P., & Marín-Gutiérrez, A. (2018). Implicit language attitudes toward historically white accents in the South African context. *Journal of Language & Social Psychology*, *37*(2), 238-248.

Ambrazaitis, G., & Bruce, G. (2006). Perception of South Swedish word accents. *Working Papers in Linguistics: Lund University*, *52*, 5-8.

Amdisen, P. (2019, 27th May). Læserbrev: The ice is melting at the poles. *Århus Stiftstidende*, https://stiften.dk/debat/laeserbrev-the-ice-is-melting-at-the-poles.

Anisfeld, E., & Lambert, W. (1964). Evaluational reactions of bilingual and monolingual children to spoken languages. *Journal of Abnormal and Social Psychology*, *69*(1), 89-97.

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*(1), 388-407.

Arredondo, M. M., & Gelman, S. A. (2019). Do varieties of Spanish influence U.S. Spanish–English bilingual children's friendship judgments? *Child Development*, *90*(2), 655-671.

Arthur, B., Farrar, D., & Bradford, G. (1974). Evaluation reactions of college students to dialect differences in the English of Mexican-Americans. *Language and Speech*, *17*(3), 255-270.

Ashmore, R., & Del Boca, F. (1981). Conceptual approaches to stereotypes and stereotyping. In D. Hamilton (Ed.), *Cognitive processes in stereotyping and intergroup behavior* (pp. 15-50). Psychology Press.

Bäckström, M., & Björklund, F. (2007). Structural modeling of generalized prejudice: The role of social dominance, authoritarianism, and empathy. *Journal of Individual Differences*, *28*(1), 10-17.

Baenninger, M. (1994). The development of face recognition: Featural or configurational processing? *Journal of Experimental Child Psychology*, *57*(3), 377-396.

Baetens, K., der Cruyssen, L. V., Achtziger, A., Vandekerckhove, M., & Van Overwalle, F. (2011). N400 and LPP in spontaneous trait inferences. *Brain Research*, *1418*, 83-92.

Bagno, M. (2003). *A norma oculta: Língua e poder na sociedade brasileira*. Parábola.

Baillargeon, R. (1987a). Object permanence in 3½- and 4½-month-old infants. *Developmental Psychology*, *23*(5), 655-664.

Baillargeon, R. (1987b). Young infants' reasoning about the physical and spatial properties of a hidden object. *Cognitive Development*, *2*(3), 179-200.

Baillargeon, R. (1993). The object concept revisited: New directions in the investigation of infants' physical knowledge. In C. Granrud (Ed.), *Visual perception and cognition in infancy* (pp. 265–315). Erlbaum.

Baker, P., & Eversley, J. (2000). *Multilingual capital: The languages of London's school children and their relevance to economic, social and educational policies*. Battlebridge Publications.

Ball, P. (1983). Stereotypes of anglo-saxon and non-anglo-saxon accents: Some exploratory Australian studies with the matched guise technique. *Language Sciences*, *5*(2), 163-183.

Barbu, S., Nardy, A., Chevrot, J.-P., & Juhel, J. (2013). Language evaluation and use during early childhood: Adhesion to social norms or integration of environmental regularities? *Linguistics*, *51*(2), 381-411.

Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Handbook of social cognition. Volume I. Basic processes* (pp. 1-40). Lawrence Erlbaum Associates, Inc.

Barnes-Holmes, D., Barnes-Holmes, Y., Hayden, E., Milne, R., & Stewart, I. (2006). Do you really know what you believe? Developing the implicit relational assessment procedure (IRAP) as a direct measure of implicit beliefs. *The Irish Psychologist*, *32*(7), 169-177.

Baron, A. S., & Banaji, M. R. (2006). The development of implicit attitudes: Evidence of race evaluations from ages 6 and 10 and adulthood. *Psychological Science*, *17*(1), 53-58.

Barrett, M. (2007). *Children's knowledge, beliefs and feelings about nations and national groups*. Psychology Press.

Barrett, M., & Short, J. (1992). Images of European people in a group of 5–10-year-old English schoolchildren. *British Journal of Developmental Psychology*, *10*(4), 339-363.

Barrett, M., Wilson, H., & Lyons, E. (2003). The development of national in-group bias: English children's attributions of characteristics to English, American and German people. *British Journal of Developmental Psychology*, *21*(2), 193-220.

Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory & Cognition*, *10*(1), 82-93.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617-645.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1-48.

Baugh, J. (2000). Racial identification by speech. *American Speech*, *75*(4), 362-364.

Bauman, R., & Briggs, C. (2003). *Voices of modernity: Language ideologies and the politics of inequality*. Cambridge University Press.

Bayard, D., & Green, J. A. (2005). Evaluating English accents worldwide. *Te Reo*, *48*, 21-28.

Beck, E. (2014). *The role of socio-indexical information in regional accent perception by five to seven year old children* [PhD Dissertation, The University of Michigan]. Michigan.

Bell, A. (2001). Back in style: Reworking audience design. In P. Eckert & J. Rickford (Eds.), *Style & sociolinguistic variation* (pp. 139-169). Cambridge University Press.

Bell, A. (2007). Designing and testing questionnaires for children. *Journal of Research in Nursing*, *12*(5), 461-469.

Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, *42*, 155-162.

Bent, T., Brown, E., Holt, R. F., Jamsek, I. A., Arzbecker, L. J., Liang, L., & Van Engen, K. J. (2021). How pronunciation distance impacts word recognition in children and adults. *Journal of the Acoustical Society of America*, *150*(6), 4103-4117.

Bent, T., & Holt, R. F. (2018). Shhh… I need quiet! Children's understanding of American, British, and Japanese-accented English speakers. *Language & Speech*, *61*(4), 657-673.

Bentin, S., McCarthy, G., & Wood, C. C. (1985). Event-related potentials, lexical decision and semantic priming. *Electroencephalography and Clinical Neurophysiology*, *60*(4), 343-355.

Berthele, R. (2010). Investigations into the folk's mental models of linguistic varieties. In D. Geeraerts, G. Kristiansen, & Y. Peirsman (Eds.), *Advances in cognitive sociolinguistics* (pp. 265-290). De Gruyter Mouton.

Bettens, K., Alighieri, C., Bruneel, L., De Meulemeester, L., & Van Lierde, K. (2020). Peer attitudes toward children with cleft (lip and) palate related to speech intelligibility, hypernasality and articulation. *Journal of Communication Disorders*, *85*, Article 105991.

Bevan, N. (2023, 27th of January). Merthyr presenter Katie Owen mocked for her Valleys accent at glitzy London awards bash. *WalesOnline*, https://www.walesonline.co.uk/lifestyle/tv/presenter-katie-owen-mocked-accent-26090622.

Bijvoet, E. (2020). *Attityder till spår av andra språk i svenskan: En forskningsöversikt*. Institutet för språk och folkminnen.

Birney, M. E., Rabinovich, A., Morton, T. A., Heath, H., & Ashcroft, S. (2019). When speaking English is not enough: The consequences of language-based stigma for nonnative speakers. *Journal of Language and Social Psychology*, *39*(1), 67-86.

Bishop, H., Coupland, N., & Garrett, P. (2005). Conceptual accent evaluation: Thirty years of accent prejudice in the UK. *Acta Linguistica Hafniensia*, *37*(1), 131-154.

Blasi, D. E., Henrich, J., Adamou, E., Kemmerer, D., & Majid, A. (2022). Over-reliance on English hinders cognitive science. *Trends in Cognitive Sciences*, *26*(12), 1153-1170.

Boberg, C. (1999). The attitudinal component of variation in American English foreign (a) nativization. *Journal of Language and Social Psychology*, *18*(1), 49-61.

Bodenhausen, G. V., Schwarz, N., Bless, H., & Wänke, M. (1995). Effects of atypical exemplars on racial beliefs: Enlightened racism or generalized appraisals? *Journal of Experimental Social Psychology*, *31*(1), 48-63.

Bogardus, E. S. (1927). Race friendliness and social distance. *Journal of Applied Sociology*, *11*, 272-287.

Bolinger, D. (1965). The atomization of meaning. *Language*, *41*(4), 555-573.

Borovsky, A., & Creel, S. C. (2014). Children and adults integrate talker and verb information in online processing. *Developmental Psychology*, *50*(5), 1600–1613.

Bosch, L., & Sebastián-Gallés, N. (1997). Native-language recognition abilities in 4-month-old infants from monolingual and bilingual environments. *Cognition*, *65*(1), 33-69.

Bouchard, E. (1969). Psycholinguistic attitude study. *Studies in Language and Language Behavior*, *8*, 437-450.

Bouchon, C., Floccia, C., Fux, T., Adda-Decker, M., & Nazzi, T. (2015). Call me Alix, not Elix: Vowels are more important than consonants in own-name recognition at 5 months. *Developmental Science*, *18*(4), 587-598.

Boughton, Z. (2006). When perception isn't reality: Accent identification and perceptual dialectology in French. *Journal of French Language Studies*, *16*(3), 277-304.

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

Bozzola, E., Spina, G., Agostiniani, R., Barni, S., Russo, R., Scarpato, E., Di Mauro, A., Di Stefano, A. V., Caruso, C., Corsello, G., & Staiano, A. (2022). The use of social media in children and adolescents: Scoping review on the potential risks. *International Journal of Environmental Research and Public Health*, *19*(16), Article 9960.

Braber, N., Ching, J., Jarman, J., Pautz, N., Robson, J., & Stevens, O. (2024). 'Is he a barrister or not?' a study on perceived and actual accentism at the bar of England and Wales. *International Journal of Speech, Language and the Law*, *31*(1), 1-22.

Bradac, J. J. (1990). Language attitudes and impression formation. In H. Giles & W. Robinson (Eds.), *Handbook of language and social psychology* (pp. 387-412). John Wiley and Sons.

Bradac, J. J., Mulac, A., & House, A. (1988). Lexical diversity and magnitude of convergent versus divergent style shifting: Perceptual and evaluative consequences. *Language & Communication*, *8*(3-4), 213-228.

Bratož, S., Pirih, A., & Štemberger, T. (2021). Identifying children's attitudes towards languages: Construction and validation of the LANGattMini scale. *Journal of Multilingual and Multicultural Development*, *42*(3), 234-246.

Brennan, E. M., & Brennan, J. S. (1981a). Accent scaling and language attitudes: Reactions to Mexican American English speech. *Language and Speech*, *24*(3), 207-221.

Brennan, E. M., & Brennan, J. S. (1981b). Measurements of accent and attitude toward Mexican-American speech. *Journal of Psycholinguistic Research*, *10*(5), 487-501.

Brewer, M. B., Dull, V., & Lui, L. (1981). Perceptions of the elderly: Stereotypes as prototypes. *Journal of Personality and Social Psychology*, *41*(4), 656-670.

Brown, B. L. (1980). Effects of speech rate on personality attributions and competency evaluations. In H. Giles, P. Smith, & W. P. Robinson (Eds.), *Language: Social psychological perspectives.* (pp. 293-300). Pergamon.

Brown, B. L., Giles, H., & Thakerar, J. N. (1985). Speaker evaluations as a function of speech rate, accent and context. *Language & Communication*, *5*(3), 207-220.

Brown, C. (2011). American elementary school children's attitudes about immigrants, immigration, and being an American. *Journal of Applied Developmental Psychology*, *32*(3), 109-117.

Bruce, G. (2005). Intonational prominence in varieties of Swedish revisited. In S.-A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 410–429). Oxford University Press.

Bugel, T. (2009). Explicit attitudes in Brazil towards varieties of Portuguese. *Studies in Hispanic and Lusophone Linguistics*, *2*(2), 275-304.

Burgess, M. C. R., Burgess, S. R., Dill, K. E., Stermer, S. P., & Brown, B. P. (2011). Playing with prejudice: The prevalence and consequences of racial stereotypes in video games. *Media Psychology*, *14*(3), 289-311.

Buson, L., & Billiez, J. (2013). Representations of stylistic variation in 9- to 11-year-olds: Cognitive processes and salience. *Linguistics*, *51*(2), 325-354.

Butler, J., Floccia, C., Goslin, J., & Panneton, R. (2011). Infants' discrimination of familiar and unfamiliar accents in speech. *Infancy*, *16*(4), 392-417.

Byers-Heinlein, K., Burns, T. C., & Werker, J. F. (2010). The roots of bilingualism in newborns. *Psychological Science*, *21*(3), 343-348.

Byers-Heinlein, K., & Garcia, B. (2015). Bilingualism changes children's beliefs about what is innate. *Developmental Science*, *18*(2), 344-350.

Cabrera, J. C. M. (2011). «unifica, limpia y fija.» la RAE y los mitos del nacionalismo lingüístico español. *El dardo en la Academia*, *1*, 157-314.

Callan, V., Gallois, C., & Forbes, P. (1983). Evaluative reactions to accented English. Ethnicity, sex role and context. *Journal of Cross-Cultural Psychology*, *14*(4), 407-426.

Cameron, C. D., Brown-Iannuzzi, J. L., & Payne, B. K. (2012). Sequential priming measures of implicit social cognition: A meta-analysis of associations with behavior and explicit attitudes. *Personality and Social Psychology Review*, *16*(4), 330-350.

Campbell-Cree, A. (2017). *Which foreign languages will be most important for the UK post-brexit?* British Council. Retrieved 16th of August 2024 from https://www.britishcouncil.org/research-insight/foreign-languages-post-brexit.

Campbell-Kibler, K. (2012). The implicit association test and sociolinguistic meaning. *Lingua*, *122*(7), 753-763.

Cargile, A. C., Giles, H., Ryan, E. B., & Bradac, J. J. (1994). Language attitudes as a social process: A conceptual model and new directions. *Language & Communication*, *14*(3), 211-236.

Carranza, M. A., & Ryan, E. B. (1975). Evaluative reactions of bilingual Anglo and Mexican American adolescents toward speakers of English and Spanish. *International Journal of the Sociology of Language*, *6*, 83-104.

Cattani, A., Abbot-Smith, K., Farag, R., Krott, A., Arreckx, F., Dennis, I., & Floccia, C. (2014). How much exposure to English is necessary for a bilingual toddler to perform like a monolingual peer in language tests? *International Journal of Language & Communication Disorders*, *49*(6), 649-671.

CENSUS 2021. (2023a). *How life has changed in Plymouth: Census 2021*. Office for National Statistics. Retrieved 1st of October 2024 from https://www.ons.gov.uk/visualisations/censusareachanges/E06000026/.

CENSUS 2021. (2023b). *How life has changed in the city of London: Census 2021*. Office for National Statistics. Retrieved 1st of October 2024 from https://www.ons.gov.uk/visualisations/censusareachanges/E09000001/.

Chaiken, S., & Trope, Y. (1999). *Dual-process theories in social psychology*. Guilford Press.

Chambers, C. T., & Johnston, C. (2002). Developmental differences in children's use of rating scales. *Journal of Pediatric Psychology*, *27*(1), 27-36.

Chambers, J. K. (2003). *Sociolinguistic theory: Linguistic variation and its social significance*. Blackwell.

Chambers, J. K., & Trudgill, P. (1980). *Dialectology*. Cambridge University Press.

Chan, A., & Li, D. (2000). English and Cantonese phonology in contrast: Explaining Cantonese ESL learners' English pronunciation problems. *Language, Culture and Curriculum*, *13*(1), 67 – 85.

Chen, Y., Michael, R., Gilbert, H., & Lerman, J. (2001). Vowel production by Mandarin speakers of English. *Clinical Linguistics and Phonetics*, *15*(6), 427-440.

Cheshire, J., Fox, S., Kerswill, P., & Torgersen, E. (2008). Ethnicity, friendship network and social practices as the motor of dialect change: Linguistic innovation in London. *Sociolinguistica*, *22*(1), 1-23.

Cheshire, J., Kerswill, P., Fox, S., & Torgersen, E. (2011). Contact, the feature pool and the speech community: The emergence of Multicultural London English. *Journal of Sociolinguistics*, *15*(2), 151-196.

Chesire, J. (1998). Double negatives are illogical. In L. Bauer & P. Trudgill (Eds.), *Language myths* (pp. 113-122). Penguin Books.

Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press.

Chomsky, N. (1986). *Knowledge of language: Its nature, origin and use*. Praeger.

Christakis, D., & Hale, L. (2025). Introduction. In D. Christakis & L. Hale (Eds.), *Handbook of children and screens: Digital media, development and well-being from birth through adolescence*. Springer.

Cichocki, W. (2012). An overview of the phonetics and phonology of Acadian French spoken in northeastern New Brunswick (Canada). In R. Gess, C. Lyche, & T. Meisenburg (Eds.), *Phonological variation in French. Illustrations from three continents* (pp. 211-234). John Benjamins Publishing Company.

Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. *The Journal of the Acoustical Society of America*, *116*(6), 3647-3658.

Clopper, C., & Wagner, L. (2019). Regional dialect intelligibility across the lifespan. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th international congress of phonetic sciences* (pp. 2324-2328). Australasian Speech Science and Technology Association Inc., and International Phonetic Association.

Clopper, C. G., & Pisoni, D. B. (2004). Homebodies and army brats: Some effects of early linguistic experience and residential history on dialect categorization. *Language Variation and Change*, *16*(1), 31-48.

Clopper, C. G., & Pisoni, D. B. (2007). Free classification of regional dialects of American English. *Journal of Phonetics*, *35*(3), 421-438.

Clyne, M. (2000). Lingua franca and ethnolects in Europe and beyond. *Sociolinguistica*, *14*(1), 83-89.

Cohen, E., & Haun, D. (2013). The development of tag-based cooperation via a socially acquired trait. *Evolution and Human Behavior*, *34*(3), 230-235.

Cole, A. (2021). Disambiguating language attitudes held towards sociodemographic groups and geographic areas in South East England. *Journal of Linguistic Geography*, *9*(1), 13-27.

Cole, A., & Strycharczuk, P. (2024). The search for linguistically coherent accents. *English World-Wide*, *45*(1), 1-29.

Conrey, F. R., & Smith, E. R. (2007). Attitude representation: Attitudes as patterns in a distributed, connectionist representational system. *Social Cognition*, *25*(5), 718-735.

Corning, A. F. (2002). Self-esteem as a moderator between perceived discrimination and psychological distress among women. *Journal of Counseling Psychology*, *49*(1), 117-126.

Corriveau, K. H., Kinzler, K. D., & Harris, P. L. (2013). Accuracy trumps accent in children's endorsement of object labels. *Developmental Psychology*, *49*(3), 470-479.

Costa, D., & Serra, R. (2022). Rhoticity in English, a journey over time through social class: A narrative review. *Frontiers in Sociology*, *7*, Article 902213.

Coupland, N. (1985). 'Hark, hark, the lark': Social motivations for phonological style-shifting. *Language & Communication*, *5*(3), 153-171.

Coupland, N. (2007). *Style: Language variation and identity*. Cambridge University Press.

Coupland, N., & Bishop, H. (2007). Ideologised values for British accents. *Journal of Sociolinguistics*, *11*(1), 74-93.

Coupland, N., Williams, A., & Garrett, P. (1994). The social meanings of Welsh English: Teachers' stereotyped judgements. *Journal of Multilingual and Multicultural Development*, *15*(6), 471-489.

Cox, W. T. L., Abramson, L. Y., Devine, P. G., & Hollon, S. D. (2012). Stereotypes, prejudice, and depression: The integrated perspective. *Perspectives on Psychological Science*, *7*(5), 427-449.

Creel, S. C. (2018). Accent detection and social cognition: Evidence of protracted learning. *Developmental Science*, *21*(2), Article 21:e12524.

Creel, S. C., & Jimenez, S. R. (2012). Differences in talker recognition by preschoolers and adults. *Journal of Experimental Child Psychology*, *113*(4), 487-509.

Creel, S. C., & Seubert, E. (2015). Children's early perceptual and late-emerging social sensitivity to accented speech. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *37*, Article 19v7z5v8.

Cremona, C., & Bates, E. (1977). The development of attitudes toward dialect in Italian children. *Journal of Psycholinguistic Research*, *6*(3), 223-232.

Cristia, A., Seidi, A., Vaughn, C., Schmale, R., Badlow, A., & Floccia, C. (2012). Linguistic processing of accented speech across the lifespan. *Frontiers in Psychology*, *3*, Article 479.

Croft, W. l., & Cruse, D. A. (2004). *Cognitive linguistics*. Cambridge University Press.

Cruttenden, A. (1994). *Gimson's pronunciation of English* (5th ed.). Arnold.

Cruttenden, A. (1997). *Intonation*. Cambridge University Press.

Cruttenden, A. (2014). *Gimson's pronunciation of English* (8th ed.). Routledge.

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2007). The bias map: Behaviors from intergroup affect and stereotypes. *Journal of Personality and Social Psychology*, *92*(4), 631-648.

Cutler, A. (1984). Stress and accent in language understanding and production. In D. Gibbon & H. Richter (Eds.), *Intonation, accent and rhythm* (pp. 77-90). Mouton de Gruyter.

Cutler, A. (2012). *Native listening: Language experience and the recognition of spoken words*. MIT Press.

Cutler, A., Sebastián-Gallés, N., Soler-Vilageliu, O., & Ooijen, B. v. (2000). Constraints of vowels and consonants on lexical selection: Cross-linguistic comparisons. *Memory & Cognition*, *28*(5), 746-755.

Cutler, C., & Røyneland, U. (2015). Where the fuck am I from? Hip-hop youth and the (re)negotiation of language and identity in Norway and the US. In J. Nortier & B. A. Svendsen (Eds.), *Language, youth and identity in the 21st century: Linguistic practices across urban spaces* (pp. 139-164). Cambridge University Press.

Cvencek, D., Greenwald, A. G., & Meltzoff, A. N. (2011). Measuring implicit attitudes of 4-year-olds: The preschool implicit association test. *Journal of Experimental Child Psychology*, *109*(2), 187-200.

Dabrowska, E., & Divjak, D. (2015). Introduction. In E. Dabrowska & D. Divjak (Eds.), *Handbook of cognitive linguistics* (pp. 1-9). De Gruyter Mouton.

Dautel, J. B., & Kinzler, K. D. (2018). Once a French speaker, always a French speaker? Bilingual children's thinking about the stability of language. *Cognitive Science*, *42*, 287-302.

Davies, A. (2003). *The native speaker: Myth and reality*. Multilingual Matters.

Day, R. (1982). Children's attitudes towards language. In E. Bouchard & H. Giles (Eds.), *Attitudes towards language variation: Social and applied contexts* (pp. 116-131). Edward Arnold.

Day, R. R. (1980). The development of linguistic attitudes and preferences. *TESOL Quarterly*, *14*(1), 27-37.

De Houwer, J. (2003). The extrinsic affective simon task. *Experimental Psychology*, *50*(2), 77-85.

de Vogelaer, G., & Toye, J. (2017). Acquiring attitudes towards varieties of Dutch: A quantitative perspective. In G. D. Vogelaer & M. Katerbow (Eds.), *Acquiring sociolinguistic variation* (pp. 117-154). John Benjamins Publishing Company.

Degner, J., & Dalege, J. (2013). The apple does not fall far from the tree, or does it? A meta-analysis of parent–child similarity in intergroup attitudes. *Psychological Bulletin*, *139*(6), 1270-1304.

DeJesus, J., Liberman, Z., & Kinzler, K. (2019). Early origins of identity. Infants' and children's thinking about language and culture. *Zero to Three*, *39*(3), 10-16.

DeJesus, J. M., Hwang, H. G., Dautel, J. B., & Kinzler, K. D. (2017). Bilingual children's social preferences hinge on accent. *Journal of Experimental Child Psychology*, *164*, 178-191.

DeJesus, J. M., Hwang, H. G., Dautel, J. B., & Kinzler, K. D. (2018). "American = English speaker" before "American = white": The development of children's reasoning about nationality. *Child Development*, *89*(5), 1752-1767.

Dennhag, I., Steinvall, A., Hakelind, C., & Deutschmann, M. (2019). Exploring gender stereotypes about interpersonal behavior and personality factors using digital matched-guise techniques. *Social behavior and personality*, *47*(8), 1-13.

Department of Education. (2017). *Hundreds of pupils on track towards fluency in Mandarin*. GOV.UK. Retrieved 16th of August 2024 from https://www.gov.uk/government/news/hundreds-of-pupils-on-track-towards-fluency-in-mandarin.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*(1), 5-18.

Devine, P. G., & Sharp, L. B. (2009). Automaticity and control in stereotyping and prejudice. In T. D. Nelson (Ed.), *Handbook of prejudice, stereotyping, and discrimination.* (pp. 61-87). Psychology Press.

Diamond, A. (1985). Development of the ability to use recall to guide action, as indicated by infants' performance on AB. *Child Development*, *56*(4), 868-883.

Dobrow, J. R., & Gidney, C. L. (1998). The good, the bad, and the foreign: The use of dialect in children's animated television. *The Annals of the American Academy of Political and Social Science*, *557*, 105-119.

Docherty, G., Foulkes, P., Gonzalez, S., & Mitchell, N. (2018). Missed connections at the junction of sociolinguistics and speech processing. *Topics in Cognitive Science*, *10*(4), 759-774.

Docherty, G. J., & Foulkes, P. (2014). An evaluation of usage-based approaches to the modelling of sociophonetic variability. *Lingua*, *142*, 42-56.

Docherty, G. J., Langstrof, C., & Foulkes, P. (2013). Listener evaluation of sociophonetic variability: Probing constraints and capabilities. *Linguistics*, *51*(2), 355-380.

Dockrell, J., & Messer, D. (2004). Lexical acquisition in the early school years. In R. A. Berman (Ed.), *Language development across childhood and adolescence* (pp. 35-52). John Benjamins Publishing Company.

Dodd, B., Holm, A., Hua, Z. H. U., & Crosbie, S. (2003). Phonological development: A normative study of British English-speaking children. *Clinical Linguistics & Phonetics*, *17*(8), 617-643.

Doss, R. C., & Gross, A. M. (1994). The effects of black English and code-switching on intraracial perceptions. *Journal of Black Psychology*, *20*(3), 282-293.

Dossey, E., Clopper, C. G., & Wagner, L. (2020). The development of sociolinguistic competence across the lifespan: Three domains of regional dialect perception. *Language Learning and Development*, *16*(4), 330-350.

Dossey, E., Jones, Z., & Clopper, C. G. (2023). Relative contributions of social, contextual, and lexical factors in speech processing. *Language and Speech*, *66*(2), 322-353.

Dovidio, J. F. (2001). On the nature of contemporary prejudice: The third wave. *Journal of Social Issues*, *57*(4), 829-849.

Dovidio, J. F., Brigham, J., Johnson, B., & Gaertner, S. (1996). Stereotyping, prejudice, and discrimination: Another look. In C. N. Macrae, C. Stangor, & M. Hewstone (Eds.), *Stereotypes and stereotyping* (pp. 276-322). Guilford.

Dovidio, J. F., & Gluszek, A. (2012). Accents, nonverbal behavior, and intergroup bias. In H. Giles (Ed.), *Handbook of interpersonal communication* (pp. 87-99). Routledge.

Dragojevic, M., & Giles, H. (2016). I don't like you because you're hard to understand: The role of processing fluency in the language attitudes process. *Human Communication Research*, *42*(3), 396-420.

Dragojevic, M., Giles, H., Goatley-Soan, S., & Dayton, Z. A. (2024). Americans' attitudes toward British accents: The role of social categorisation, perceived group prototypicality, and processing fluency. *Journal of Multilingual and Multicultural Development*, 1-18.

Dragojevic, M., Giles, H., & Watson, B. (2013). Language ideologies and language attitudes: A foundational framework. In H. Giles & B. Watson (Eds.), *The social meanings of language, dialect and accent. International perspectives on speech styles.* (pp. 1-25). Peter Lang.

Dragojevic, M., & Goatley-Soan, S. (2022a). Americans' attitudes toward foreign accents: Evaluative hierarchies and underlying processes. *Journal of Multilingual and Multicultural Development*, *43*(2), 167-181.

Dragojevic, M., & Goatley-Soan, S. (2022b). The verbal-guise technique. In R. Kircher & L. Zipp (Eds.), *Research methods in language attitudes* (pp. 203-218). Cambridge University Press.

Dragojevic, M., Mastro, D., Giles, H., & Sink, A. (2016). Silencing nonstandard speakers: A content analysis of accent portrayals on American primetime television. *Language in Society*, *45*(1), 59-85.

Dunn, L. M., Dunn, D. M., Sewell, K., Styles, B., Brzyska, B., Shamsan, Y., & Burge, B. (2009). *British picture vocabulary scale* (3rd ed.). GL Assessment Limited.

Durante, F., Fiske, S. T., Kervyn, N., Cuddy, A. J. C., Akande, A., Adetoun, B. E., Adewuyi, M. F., Tserere, M. M., Ramiah, A. A., Mastor, K. A., Barlow, F. K., Bonn, G., Tafarodi, R. W., Bosak, J., Cairns, E., Doherty, C., Capozza, D., Chandran, A., Chryssochoou, X., & Iatridis, T. (2013). Nations' income inequality predicts ambivalence in stereotype content: How societies mind the gap. *British Journal of Social Psychology*, *52*(4), 726-746.

Durkin, K., & Judge, J. (2001). Effects of language and social behaviour on children's reactions to foreign people in television. *British Journal of Developmental Psychology*, *19*(4), 597-612.

Durrant, S., Luche, C. D., Cattani, A., & Floccia, C. (2015). Monodialectal and multidialectal infants' representation of familiar words. *Journal of Child Language*, *42*(2), 447-465.

Eberhard, D., Simons, G., & Fenning, C. (2024). *Ethnologue: Languages of the world* (27th ed.). SIL International.

Eckert, P. (1989). *Jocks and burnouts: Social categories and identity in the high school*. Teachers college press.

Eckert, P. (2012). Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology*, *41*, 87-100.

Edwards, J. (1982). Language attitudes and their implications among English speakers. In E. B. Ryan & H. Giles (Eds.), *Attitudes towards language variation: Social and applied contexts* (pp. 20-33). E. Arnold.

Edwards, J. (2012). *Multilingualism: Understanding linguistic diversity*. Continuum.

Esling, J. (1998). Everyone has an accent except me. In L. Bauer & P. Trudgill (Eds.), *Language myths* (pp. 169-175). Penguin Books.

Evans, B. G., & Lourido, G. T. (2019). Effects of language background on the development of sociolinguistic awareness: The perception of accent variation in monolingual and multilingual 5-to 7-year-old children. *PHONETICA*, *76*(2-3), 142-162.

Fabricius, A. (2002). Weak vowels in modern RP: An acoustic study of happy-tensing and kit/schwa shift. *Language Variation and Change*, *14*(2), 211-237.

Fabricius, A. (2018). Social change, linguistic change and sociolinguistic change in Received Pronunciation. In N. Braber & S. Jansen (Eds.), *Sociolinguistics in England* (pp. 35-66). Palgrave Macmillan.

Fabricius, A. (2019). The anticlockwise checked vowel chain shift in modern RP in the twentieth century. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th international congress of phonetic sciences* (pp. 592-596). Australasian Speech Science and Technology Association Inc., and International Phonetic Association.

Fagyal, Z., Kibbee, D., & Jenkins, F. (2006). *French: A linguistic introduction*. Cambridge University Press.

Fantini, A. (2012). Sociolinguistic development and gender identification in childhood: A bilingual case study (birth to age 10). *The bilingual Review*, *31*(2), 99-124.

Fazio, R. H. (1995). Attitudes as object-evaluation associations: Determinants, consequences, and correlates of attitude accessibility. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences.* (pp. 247-282). Lawrence Erlbaum Associates, Inc.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*(2), 229-238.

Fiedler, S., keller, C., & Hanulíková, A. (2019). Social expectations and intelligibility of Arabic-accented speech in noise. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th international congress of phonetic sciences* (pp. 3085-3089). Australasian Speech Science and Technology Association Inc., and International Phonetic Association.

Fielding, G., & Evered, C. (1980). The influence of patients' speech upon doctors: The diagnostic interview. In R. Clair & H. Giles (Eds.), *The social and psychological contexts of language* (pp. 51-72). Erlbaum.

Fillmore, C. (1975). An alternative to checklist theories of meaning. *Proceedings of the First Annual Meeting of the Berkeley Linguistics Society*, *1*, 123-131.

Fiske, S., Harris, L., Lee, T., & Russell, A. (2016). The future of research on prejudice, stereotyping and discrimination. In T. Nelson (Ed.), *Handbook of prejudice, stereotyping, and discrimination* (pp. 487–498). Psychology Press.

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, *11*(2), 77-83.

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*(6), 878-902.

Fiske, S. T., & Taylor, S. E. (2017). *Social cognition: From brains to culture* (3rd ed.). SAGE.

Flege, J. E. (1984). The detection of French accent by American listeners. *Journal of the Acoustical Society of America*, *76*(3), 692-707.

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

Floccia, C., Butler, J., Girard, F., & Goslin, J. (2009). Categorization of regional and foreign accent in 5- to 7-year-old British children. *International Journal of Behavioral Development*, *33*(4), 366-375.

Floccia, C., Butler, J., Goslin, J., & Ellis, L. (2009). Regional and foreign accent processing in English: Can listeners adapt? *Journal of Psycholinguistic Research*, *38*(4), 379-412.

Floccia, C., Goslin, J., & Girard, F. (2006). Does a regional accent perturb speech processing? *Journal of Experimental Psychology: Human Perception and Performance*, *32*(5), 1276-1293.

Floccia, C., Sambrook, T. D., Luche, C. D., Kwok, R., Goslin, J., White, L., Cattani, A., Sullivan, E., Abbot-Smith, K., Krott, A., Mills, D., Rowland, C., Gervain, J., Plunkett, K., Hoff, E., & Core, C. (2018). Vocabulary of 2-year-olds learning English and an additional language: Norms and effects of linguistic distance. *Monographs of the Society for Research in Child Development*, *83*(1), 1-134.

Formanowicz, M., & Suitner, C. (2019). Sounding strange(r): Origins, consequences, and boundary conditions of sociophonetic discrimination. *Journal of Language and Social Psychology*, *39*(1), 4-21.

Fotiou, C., & Ayiomamitou, I. (2021). "We are in Cyprus, we have to use our language, don't we?" pupils' and their parents' attitudes towards two proximal linguistic varieties. *Linguistics and Education*, *63*, Article 100931.

Foulkes, P., & Docherty, G. (2006). The social life of phonetics and phonology. *Journal of Phonetics*, *34*(4), 409-438.

Fox, S. (2015). *The new Cockney: New ethnicities and adolescent speech in the traditional East End of London*. Palgrave Macmillan.

Fuertes, J. N., Gottdiener, W. H., Martin, H., Gilbert, T. C., & Giles, H. (2012). A meta-analysis of the effects of speakers' accents on interpersonal evaluations. *European Journal of Social Psychology*, *42*(1), 120-133.

Fuller, J. M. (2018). Ideologies of language, bilingualism, and monolingualism. In A. De Houwer & L. Ortega (Eds.), *The Cambridge handbook of bilingualism* (pp. 119-134). Cambridge University Press.

Funk, F., & Todorov, A. (2013). Criminal stereotypes in the courtroom: Facial tattoos affect guilt and punishment differently. *Psychology, Public Policy, and Law*, *19*(4), 466-478.

García Coni, A., Ison, M., & Vivas, J. (2019). Conceptual flexibility in school children: Switching between taxonomic and thematic relations. *Cognitive Development*, *52*, Article 100827.

Garner, T., & Rubin, D. L. (1986). Middle class blacks' perceptions of dialect and style shifting: The case of southern attorneys. *Journal of Language and Social Psychology*, *5*(1), 33-48.

Garrett, P. (2007). Language attitudes. In C. Llamas, L. Mullany, & P. Stockwell (Eds.), *The Routledge companion to sociolinguistics* (pp. 116-121). Routledge.

Garrett, P. (2010). *Attitudes to language*. Cambridge University Press.

Garrett, P., & Coupland, N. (1999). Evaluating dialect in discourse: Teachers' and teenagers' responses to young English speakers in Wales. *Language in Society*, *28*(3), 321-354.

Garrett, P., Coupland, N., & Williams, A. (2003). *Investigating language attitudes: Social meanings of dialect, ethnicity and performance*. Cambridge University Press.

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin*, *132*(5), 692-731.

Gawronski, B., & Bodenhausen, G. V. (2011). The associative-propositional evaluation model: Theory, evidence, and open questions. *Advances in experimental social psychology*, *44*, 59-127.

Gawronski, B., & Creighton, L. A. (2013). Dual process theories. In D. Carlston (Ed.), *The Oxford handbook of social cognition* (pp. 282-312). Oxford University Press.

Gawronski, B., Hofmann, W., & Wilbur, C. J. (2006). Are "implicit" attitudes unconscious? *Consciousness and Cognition*, *15*(3), 485-499.

Geeraerts, D. (2018). *Ten lectures on cognitive sociolinguistics*. Brill.

Geeraerts, D., & Cuyckens, H. (2007). Introducing cognitive linguistics. In G. Dirk & C. Herbert (Eds.), *Oxford handbook of cognitive linguistics* (pp. 3 - 21). Oxford University Press.

Geeraerts, D., Kristiansen, G., & Peirsman, Y. (2010a). *Advances in cognitive sociolinguistics*. Mouton de Gruyter.

Geeraerts, D., Kristiansen, G., & Peirsman, Y. (2010b). Introduction: Advances in cognitive sociolinguistics. In D. Geeraerts, G. Kristiansen, & Y. Peirsman (Eds.), *Advances in cognitive sociolinguistics* (pp. 1-20). De Gruyter Mouton.

Gelman, S. A., & Wellman, H. M. (1991). Insides and essences: Early understandings of the non-obvious. *Cognition*, *38*(3), 213-244.

Ghavami, N., & Peplau, L. A. (2013). An intersectional analysis of gender and ethnic stereotypes: Testing three hypotheses. *Psychology of Women Quarterly*, *37*(1), 113-127.

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

Giles, H. (1971). Patterns of evaluation to R.P., South Welsh and Somerset accented speech. *British Journal of Social and Clinical Psychology*, *10*(3), 280-281.

Giles, H. (1972). The effect of stimulus mildness-broadness in the evaluation of accents. *Language and Speech*, *15*(3), 262-269.

Giles, H. (2012). *The handbook of intergroup communication*. Routledge.

Giles, H., Baker, S., & Fielding, G. U. Y. (1975). Communication length as a behavioural index of accent prejudice. *International Journal of the Sociology of Language*, *1975*(6), 73-82.

Giles, H., Harrison, C., Creber, C., Smith, P., & Freeman, N. (1983). Developmental and contextual aspects of children's language attitudes. *Language & Communication*, *3*, 141-146.

Giles, H., & Marlow, M. L. (2011). Theorizing language attitudes existing frameworks, an integrative model, and new directions. *Annals of the International Communication Association*, *35*(1), 161-197.

Giles, H., & Powesland, P. (1975). *Speech style and social evaluation*. Academic Press.

Giles, H., & Watson, B. M. (2013). *The social meanings of language, dialect and accent: International perspectives on speech styles*. Peter Lang.

Gill, M. (2016, 27th of October). Angela Rayner called 'thick as mince' in abusive emails about her accent. *HuffingtonPost*, https://www.huffingtonpost.co.uk/entry/angela-rayner-hits-back-at-abusive-emails-over-her-accent_uk_5811d990e4b04660a438156a.

Gimson, A. (1970). *An introduction to the pronunciation of English* (2nd ed.). Edward Arnold.

Girard, F., Floccia, C., & Goslin, J. (2008). Perception and awareness of accents in young children. *British Journal of Developmental Psychology*, *26*(3), 409-433.

Gluszek, A., & Dovidio, J. F. (2010a). Speaking with a nonnative accent: Perceptions of bias, communication difficulties, and belonging in the United States. *Journal of Language and Social Psychology*, *29*(2), 224-234.

Gluszek, A., & Dovidio, J. F. (2010b). The way they speak: A social psychological perspective on the stigma of nonnative accents in communication. *Personality and Social Psychology Review*, *14*(2), 214-237.

Gluszek, A., & Hansen, K. (2013). Language attitudes in the Americas. In H. Giles & B. M. Watson (Eds.), *The social meanings of language, dialect and accent: International perspectives on speech styles* (pp. 26-63). Peter Lang.

Goldman, R., & Fristoe, M. (2000). *Goldman-fristoe test of articulation* (2nd ed.). Pearson Assessments.

Goorimoorthee, T., Csipo, A., Carleton, S., & Ensslin, A. (2019). Language ideologies in videogame discourse: Forms of sociophonetic othering in accented character speech. In A. Ensslin & I. Balteiro (Eds.), *Approaches to videogame discourse: Lexis, interaction, textuality* (pp. 269-287). Bloomsbury Academic.

Goslin, J., Duffy, H., & Floccia, C. (2012). An ERP investigation of regional and foreign accent processing. *Brain and Language*, *122*(2), 92-102.

Gover, A. R., Harper, S. B., & Langton, L. (2020). Anti-Asian hate crime during the covid-19 pandemic: Exploring the reproduction of inequality. *American Journal of Criminal Justice: The Journal of the Southern Criminal Justice Association*, *45*(4), 647-667.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*, 1464-1480.

Gries, S. (2013). Sources of variability relevant to the cognitive sociolinguist, and corpus- as well as the psycholinguistic methods and notions to handle them. *Journal of Pragmatics*, *52*, 5-16.

Grondelaers, S., van Hout, R., & van Gent, P. (2018). Re-evaluating the prestige of regional accents in Netherlandic Standard Dutch: The role of accent strength and speaker gender. *Journal of Language and Social Psychology*, *38*(2), 215-236.

Gummerum, M., & Keller, M. (2008). Affection, virtue, pleasure, and profit: Developing an understanding of friendship closeness and intimacy in western and Asian societies. *International Journal of Behavioral Development*, *32*(3), 218-231.

Guo, R., Cheng, F., & Evans, B. (2023). What you see is (not) what you hear? The role of social expectations on speech intelligibility in a British context. In R. Skarnitzl & J. Volín (Eds.), *Proceedings of the 20th international congress of phonetic sciences – ICPhS 2023* (pp. 3542-3546). International Phonetic Association.

Hackman, D. A., & Farah, M. J. (2009). Socioeconomic status and the developing brain. *Trends in Cognitive Sciences*, *13*(2), 65-73.

Hackman, D. A., Farah, M. J., & Meaney, M. J. (2010). Socioeconomic status and the brain: Mechanistic insights from human and animal research. *Nature Reviews Neuroscience*, *11*(9), 651-659.

Haiman, J. (1980). Dictionaries and encyclopedias. *Lingua*, *50*, 329-357.

Halmari, H., & Smith, W. (1994). Code-switching and register shift: Evidence from finnish-English child bilingual conversation. *Journal of Pragmatics*, *21*(4), 427-445.

Hannisdal, B. (2006). *Variability and change in Received Pronunciation: A study of six phonological variables in the speech of television newsreaders* [PhD Dissertation, University of Bergen]. Bergen.

Hansen, K., Steffens, M. C., Rakić, T., & Wiese, H. W. (2017). When appearance does not match accent: Neural correlates of ethnicity-related expectancy violations. *Social Cognitive & Affective Neuroscience*, *12*(3), 507-515.

Hanulíková, A. (2018). The effect of perceived ethnicity on spoken text comprehension under clear and adverse listening conditions. *Linguistics Vanguard*, *4*(1), Article 20170029.

Hanulíková, A. (2020). Effects of talker identity on speech intelligibility: A lifespan perspective. In L. Taschenberger (Ed.), *Book of abstracts: 2nd workshop on speech perception and production across the lifespan (sppl2020)* (pp. 114-115). Zenodo.

Hanulíková, A. (2024). Navigating accent bias in German: Children's social preferences for a second-language accent over a first-language regional accent. *Frontiers in Language Sciences*, *3*, Article 3:1357682.

Harari, H., & McDavid, J. W. (1973). Name stereotypes and teachers' expectations. *Journal of Educational Psychology*, *65*, 222-225.

Harrington, J., Palethorpe, S., & Watson, C. (2000). Monophthongal vowel changes in Received Pronunciation: An acoustic analysis of the Queen's christmas broadcasts. *Journal of the International Phonetic Association*, *30*(1-2), 63-78.

Harter, S. (1999). *The construction of the self: A developmental perspective*. Guilford Press.

Harter, S. (2006). The self. In *Handbook of child psychology: Social, emotional, and personality development* (Vol. 3, pp. 505-570). John Wiley & Sons, Inc.

Hay, J., & Drager, K. (2010). Stuffed toys and speech perception. *Linguistics*, *48*(4), 865-892.

Hay, J., Nolan, A., & Drager, K. (2006). From fush to feesh: Exemplar priming in speech perception. *The Linguistic Review*, *23*(3), 351-379.

Heaton, H., & Nygaard, L. C. (2011). Charm or harm: Effect of passage content on listener attitudes toward American English accents. *Journal of Language & Social Psychology*, *30*(2), 202-211.

Heeringa, W., & Nerbonne, J. (2001). Dialect areas and dialect continua. *Language Variation and Change*, *13*(3), 375-400.

Hendriks, B., Meurs, F., & Groot, E. (2017). The effects of degrees of Dutch accentedness in ELF and in French, German and Spanish. *International Journal of Applied Linguistics*, *27*(1), 44-66.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2-3), 61-83.

Hernandez, R. (1993, 2nd of March). When an accent becomes an issue; immigrants turn to speech classes to reduce sting of bias. *The New York Times*, https://www.nytimes.com/1993/03/02/nyregion/when-accent-becomes-issue-immigrants-turn-speech-classes-reduce-sting-bias.html.

Hill, J. H. (2008). *The everyday language of white racism*. Wiley Blackwell.

Hilton, J. L., & von Hippel, W. (1996). Stereotypes. *Annual Review of Psychology*, *47*(1), 237-271.

Hinrichs, L., & White-Sustaíta, J. (2011). Global Englishes and the sociolinguistics of spelling. *English World-Wide*, *32*(1), 46-73.

Hirschfeld, L. A., & Gelman, S. A. (1997). What young children think about the relationship between language variation and social difference. *Cognitive Development*, *12*(2), 213-238.

HM Treasury. (2015). *Boost to Mandarin teaching in schools*. GOV.UK. Retrieved 16th of August 2024 from https://www.gov.uk/government/news/chancellor-announces-boost-to-mandarin-teaching-in-schools.

Hodson, J. (2014). *Dialect in film and literature*. Palgrave Macmillan.

Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, *26*(1), 55-88.

Holm, L. A. (2018). *Krabber*. Straarup & Co.

Holmes-Elliott, S., & Levon, E. (2024). The jet set: Modern RP and the (re)creation of social distinction. *Language Variation and Change*, *36*(2), 1-22.

Honey, J. (2017). Sociophonology. In F. Coulmas (Ed.), *The handbook of sociolinguistics* (pp. 92-106). Blackwell Publishers.

Hosoda, M., & Stone-Romero, E. (2010). The effects of foreign accents on employment-related decisions. *Journal of Managerial Psychology*, *25*(2), 113-132.

Howard, L. H., Carrazza, C., & Woodward, A. L. (2014). Neighborhood linguistic diversity predicts infants' social learning. *Cognition*, *133*(2), 474-479.

Hualde, J. I. (2005). *The sounds of Spanish*. Cambridge University Press.

Hughes, A., Trudgill, P., & Watt, D. (2012). *English accents and dialects. An introduction to social and regional varieties of English in the British Isles*. Routledge.

Hwang, H. G., & Markson, L. (2018). Locals don't have accents: Children weigh phonological proficiency over syntactic or semantic proficiency when categorizing individuals. *Journal of Child Language*, *45*(4), 1018-1034.

Hymes, D. (1972). On communicative competence. In J. Pride & J. Holmes (Eds.), *Sociolinguistics* (pp. 269-293). Penguin.

Imuta, K., & Spence, J. L. (2020). Developments in the social meaning underlying accent- and dialect-based social preferences. *Child Development Perspectives*, *14*(3), 135-141.

Institute of Education. (2021). *Mandarin excellence programme*. IOE Confucious Institute for Schools. Retrieved 16th of August 2024 from https://ci.ioe.ac.uk/mandarin-excellence-programme/.

Ioannidou, E. (2004). On language and ethnic identity among Greek cypriot students. *Cyprus Review*, *16*, 29-52.

IOE Digital. (2019). *More British children are learning Mandarin Chinese – but an increase in qualified teachers is urgently needed*. Institute of Education Blog. Retrieved 16th of August 2024 from https://blogs.ucl.ac.uk/ioe/2019/02/11/more-british-children-are-learning-mandarin-chinese-but-an-increase-in-qualified-teachers-is-urgently-needed/.

Ito, T. A., & Urland, G. R. (2003). Race and gender on the brain: Electrocortical measures of attention to the race and gender of multiply categorizable individuals. *Journal of Personality and Social Psychology*, *85*(4), 616-626.

Jeffries, E. (2019). Preschool children's categorization of speakers by regional accent. *Language Variation and Change*, *31*(3), 329-352.

Jisa, H. (2004). Growing into academic French. In R. A. Berman (Ed.), *Language development across childhood and adolescence* (pp. 135-162). John Benjamins Publishing Company.

Johnson, F. L. (2012). Rhetorical positioning of US policy statements about multilingual education—with reference to the EU. *Language, Culture and Curriculum*, *25*(1), 73-87.

Jolliffer, I. (2004). *Principal component analysis*. Springer.

Jones, D. (1909). *The pronunciation of English*. Cambridge Universitgy Press.

Jones, D. (1917). *An English pronouncing dictionary*. Dent.

Jones, Z., Yan, Q., Wagner, L., & Clopper, C. G. (2017). The development of dialect classification across the lifespan. *Journal of Phonetics*, *60*, 20-37.

Jørgensen, J. N., & Quist, P. (2001). Native speakers' judgments of second language Danish. *Language Awareness*, *10*(1), 41-56.

Jost, J. T., Glaser, J., Kruglanski, A. W., & Sulloway, F. J. (2003). Political conservatism as motivated social cognition. *Psychological Bulletin*, *129*(3), 339-375.

Kaiser, I., & Kasberger, G. (2021). Children's sociolinguistic preferences: The acquisition of language attitudes within the Austrian standard-dialect continuum. In A. Ghimenton, A. Nardy, & J.-P. Chevrot (Eds.), *Sociolinguistic variation and language acquisition across the lifespan* (pp. 129-160). John Benjamins Publishing Company.

Kang, O., & Rubin, D. L. (2009). Reverse linguistic stereotyping: Measuring the effect of listener expectations on speech evaluation. *Journal of Language and Social Psychology*, *28*(4), 441-456.

Katz, D., & Braly, K. (1933). Racial stereotypes of one hundred college students. *The Journal of Abnormal and Social Psychology*, *28*(3), 280-290.

Kavas, A., & Kavas, A. (2008). An exploratory study of undergraduate college students' perceptions and attitudes toward foreign accented faculty. *College Student Journal*, *42*(3), 879-890.

Kean, S. (2012). *The violinist's thumb and other lost tales of love, war, and genius as written in our genetic code*. Little, Brown and Company.

Keesing, R. M. (1979). Linguistic knowledge and cultural knowledge: Some doubts and speculations. *American Anthropologist*, *81*(1), 14-36.

Kehoe, M., Patrucco-Nanchen, T., Friend, M., & Zesiger, P. (2018). The relation between phonological and lexical development in French-speaking children. *Clinical Linguistics & Phonetics*, *32*(12), 1103-1125.

Kerswill, P. (2014). The objectification of 'Jafaican'. The discoursal embedding of Multicultural London English in the British media. In J. Androutsopoulos (Ed.), *Mediatization and sociolinguistic change* (pp. 427-456). De Gruyter.

Kerswill, P., Torgersen, E. N., & Fox, S. (2008). Reversing "drift": Innovation and diffusion in the London diphthong system. *Language Variation and Change*, *20*(3), 451-491.

Khorounjaia, E., & Tolchinsky, L. (2004). Discursive constraints on the lexical realization of arguments in Spanish. In R. A. Berman (Ed.), *Language development across childhood and adolescence* (pp. 83-109). John Benjamins Publishing Company.

Kidd, E., & Garcia, R. (2022). How diverse is child language acquisition research? *First Language*, *42*(6), 703-735.

Kiesner, J., Maass, A., Cadinu, M., & Vallese, I. (2003). Risk factors for ethnic prejudice during early adolescence. *Social Development*, *12*(2), 288-308.

Kinzler, K., Shutts, K., DeJesus, J., & Spelke, E. (2010). Accent trumps race in guiding children's social preferences. *Social Cognition*, *27*(4), 623-634.

Kinzler, K. D. (2013). The development of language as a social category. In M. R. Banaji & S. A. Gelman (Eds.), *Navigating the social world: What infants, children, and other species can teach us.* (pp. 314-317). Oxford University Press.

Kinzler, K. D. (2021). Language as a social cue. *Annual Review of Psychology*, *72*, 241-264.

Kinzler, K. D., Corriveau, K. H., & Harris, P. L. (2011). Children's selective trust in native-accented speakers. *Developmental Science*, *14*(1), 106-111.

Kinzler, K. D., & Dautel, J. B. (2012). Children's essentialist reasoning about language and race. *Developmental Science*, *15*(1), 131-138.

Kinzler, K. D., & DeJesus, J. M. (2013a). Children's sociolinguistic evaluations of nice foreigners and mean Americans. *Developmental Psychology*, *49*(4), 655-664.

Kinzler, K. D., & DeJesus, J. M. (2013b). Northern = smart and Southern = nice: The development of accent attitudes in the United States. *The Quarterly Journal of Experimental Psychology*, *66*(6), 1146-1158.

Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2007). The native language of social cognition. *PNAS*, *104*(30), 12577-12580.

Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2012). "Native" objects and collaborators: Infants' object choices and acts of giving reflect favor for native over foreign speakers. *Journal of Cognition and Development*, *13*(1), 67-81.

Kinzler, K. D., Shutts, K., & Spelke, E. S. (2012). Language-based social preferences among children in South Africa. *Language Learning and Development*, *8*(3), 215-232.

Kircher, R., & Fox, S. (2019). Attitudes towards Multicultural London English: Implications for attitude theory and language planning. *Journal of Multilingual and Multicultural Development*, *40*(10), 847-864.

Kitamura, C., Panneton, R., & Best, C. T. (2013). The development of language constancy: Attention to native versus nonnative accents. *Child Development*, *84*(5), 1686-1700.

Kite, M. E., Whitley, B. E., & Wagner, L. S. (2023). *Psychology of prejudice and discrimination.* (4th ed.). Routledge.

Kjempff, M. (2022, 29th of August). Selam måtte tage et dansk navn for at få drømmejobbet. *TV2 Kosmopol*, https://www.tv2kosmopol.dk/koebenhavn/selam-maatte-tage-et-dansk-navn-for-at-faa-droemmejobbet.

Klarskov, K., & Henriksen, M. (2011, 14th of May). Sferne grinte af Søvndals engelsk. *Berlingske*, https://www.berlingske.dk/politik/sferne-grinte-af-soevndals-engelsk.

Klink, A., & Wagner, U. (1999). Discrimination against ethnic minorities in Germany: Going back to the field. *Journal of Applied Social Psychology*, *29*(2), 402-423.

Koch, L. M., & Gross, A. M. (1997). Children's perceptions of black English as a variable in intraracial perception. *Journal of Black Psychology*, *23*(3), 215-226.

Kolde, R. (2019). *Pheatmap: Pretty heatmaps*. In CRAN, https://CRAN.R-project.org/package=pheatmap.

Kotsinas, U. B. (1988). Immigrant children's Swedish — a new variety? *Journal of Multilingual and Multicultural Development*, *9*(1-2), 129-140.

Kristiansen, G. (2003). How to do things with allophones: Linguistic stereotypes as cognitive reference points in social cognition. In R. Dirven, R. Frank, & M. Pütz (Eds.), *Cognitive models in language and thought* (pp. 69-120). Walter de Gruyter.

Kristiansen, G. (2009). Style-shifting and shifting styles: A socio-cognitive approach to lectal variation. In G. Kristiansen & R. Dirven (Eds.), *Cognitive sociolinguistics. Language variation, cultural models, social systems* (pp. 45-90). Walter de Gruyter.

Kristiansen, G. (2010). Lectal acquisition and linguistic stereotype formation. In G. Dirk, K. Gitte, & P. Yves (Eds.), *Advances in cognitive sociolinguistics* (pp. 223-264). De Gruyter Mouton.

Kristiansen, G., & Dirven, R. (2008). *Cognitive sociolinguistics: Language variation, cultural models, social systems*. Mouton de Gruyter.

Kristiansen, G., Franco, K., De Pascale, S., Rosseel, L., & Weiwi, Z. (2022). *Cognitive sociolinguistics revisited*. De Gruyter Mouton.

Kristiansen, T. (2001). Two standards: One for the media and one for the school. *Language Awareness*, *10*(1), 9-24.

Kristiansen, T. (2003). Language attitudes and language politics in Denmark. *International Journal of the Sociology of Language*, *159*, 57-71.

Kristiansen, T. (2009). The macro-level social meanings of late-modern Danish accents. *Acta Linguistica Hafniensia*, *41*(1), 167-192.

Kristiansen, T. (2010). Attitudes, ideology and awareness. In R. Wodak, B. Johnstone, & P. Kerswill (Eds.), *The SAGE handbook of sociolinguistics* (pp. 265-278). Sage.

Kristiansen, T. (2013). Language attitudes in the nordic countries. In H. Giles & B. Watson (Eds.), *The social meanings of language, dialect and accent. International perspectives on speech styles.* Peter Lang.

Kristiansen, T., & Grondelaers, S. (2013). *Language (de)standardisation in late modern Europe: Experimental studies*. Novus Forlag.

Kryger, J. (2017, 11th of May). Ibrahim og Hakan: Vi hedder Peter og Jens på arbejdet. *TV2 Østjylland*, https://www.tv2ostjylland.dk/aarhus/ibrahim-og-hakan-vi-hedder-peter-og-jens-pa-arbejdet.

Kubota, J. T., & Ito, T. A. (2007). Multiple cues in social perception: The time course of processing race and facial expression. *Journal of Experimental Social Psychology*, *43*(5), 738-752.

Kubota, J. T., Li, J., Bar-David, E., Banaji, M. R., & Phelps, E. A. (2013). The price of racial bias: Intergroup negotiations in the ultimatum game. *Psychological Science*, *24*(12), 2498-2504.

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, *44*(4), 978-990.

Kustatscher, M. (2017). Young children's social class identities in everyday life at primary school: The importance of naming and challenging complex inequalities. *Childhood*, *24*(3), 381-395.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*(12), 463-470.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, *62*, 621-647.

Kutas, M., & Hillyard, S. A. (1980a). Event-related brain potentials to semantically inappropriate and surprisingly large words. *Biological Psychology*, *11*(2), 99-116-116.

Kutas, M., & Hillyard, S. A. (1980b). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*, 203-205.

Kutlu, E. (2020). Now you see me, now you mishear me: Raciolinguistic accounts of speech perception in different English varieties. *Journal of Multilingual and Multicultural Development*, *44*(6), 511-525.

Labov, W. (1964). Stages in the acquisition of standard English. In R. Shuy, A. Davis, & R. Hogan (Eds.), *Social dialects and language learning* (pp. 77-104). National Council of Teachers of English.

Labov, W. (1966a). The linguistic variable as a structural unit. *Washington Linguistics Review*, *3*, 4-22.

Labov, W. (1966b). *The social stratification of English in New York city*. Center for Applied Linguistics.

Labov, W. (1972). *Sociolinguistic patterns*. University of Pennsylvania Press.

Labov, W. (2002). The anatomy of style-shifting. In P. Eckert & J. R. Rickford (Eds.), *Style and sociolinguistic variation* (pp. 85-108). Cambridge University Press.

Labov, W. (2006). *The social stratification of English in New York city* (2nd ed. ed.). Cambridge University Press.

Ladegaard, H. J. (1992). Sprogholdninger i Danmark: En socialpsykologisk analyse af stereotype holdninger til sprog og sprogbrugere i fire danske lokaliteter. *Nordisk Psykologi*, *44*, 173-189.

Ladegaard, H. J. (2000). Language attitudes and sociolinguistic behaviour: Exploring attitude-behaviour relations in language. *Journal of Sociolinguistics*, *4*(2), 214-233.

Lakoff, G. (1971). Presupposition and relative well-formedness. In D. Steinberg & L. Jakobovits (Eds.), *Semantics: An interdisciplinary reader in philosophy, linguistics and psychology* (pp. 329-340). Cambridge University Press.

Lakoff, G. (1977). Linguistic gestalts. *Chicago linguistic society*, *13*, 236-287.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. University of Chicago Press.

Lambert, W. E., Frankle, H., & Tucker, G. R. (1966). Judging personality through speech: A French-Canadian example. *Journal of Communication*, *16*(4), 305-321.

Lambert, W. E., Hodgson, R. C., Gardner, R. C., & Fillenbaum, S. (1960). Evaluational reactions to spoken languages. *Journal of Abnormal & Social Psychology*, *60*(1), 44-51.

Lamberton, D. (2002). *The economics of language*. Edward Elgar.

Langacker, R. W. (1979). Grammar as image. *Linguistic Notes from La Jolla*, *6*, 88-126.

Langacker, R. W. (1987). *Foundations of cognitive grammar: Theoretical prerequisites*. Stanford University Press.

Langacker, R. W. (2008). *Cognitive grammar: A basic introduction*. Oxford University Press.

Lass, N. J., Ruscello, D. M., & Lakawicz, J. A. (1988). Listeners' perceptions of nonspeech characteristics of normal and dysarthric children. *Journal of Communication Disorders*, *21*(5), 385-391.

Lass, N. J., Ruscello, D. M., Stout, L. L., & Hoffman, F. M. (1991). Peer perceptions of normal and voice-disordered children. *Folia Phoniatrica*, *43*(1), 29-35.

Lavandera, B. R. (1978). Where does the sociolinguistic variable stop? *Language in Society*, *7*(2), 171-182.

Lavigne, A., Christy, L., Spock, S., & Edwards, G. (2002). Sk8er boi. On *Let Go*. Arista.

Lenth, R. (2025). *Emmeans: Estimated marginal means, aka least-squares means*. In CRAN, https://CRAN.R-project.org/package=emmeans.

Levenshtein, V. I. (1965). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet physics. Doklady*, *10*, 707-710.

Levi, S. V. (2021). Perception of indexical properties of speech by children. In J. Pardo, L. C. Nygaard, R. Remez, & D. B. Pisoni (Eds.), *The handbook of speech perception* (pp. 465-483).

Levi, S. V., Winters, S. J., & Pisoni, D. B. (2007). Speaker-independent factors affecting the perception of foreign accent in a second language. *Journal of the Acoustical Society of America*, *121*(4), 2327-2338.

Levin, H., Giles, H., & Garrett, P. (1994). The effects of lexical formality and accent on trait attributions. *Language & Communication*, *14*(3), 265-274.

Levon, E., Sharma, D., & Ilbury, C. (2022). *Speaking up: Accents and social mobility*. The Sutton Trust.

Levon, E., Sharma, D., Watt, D. J. L., Cardoso, A., & Ye, Y. (2021). Accent bias and perceptions of professional competence in England. *Journal of English Linguistics*, *49*(4), 355-388.

Levy, H., Konieczny, L., & Hanulíková, A. (2019). Processing of unfamiliar accents in monolingual and bilingual children: Effects of type and amount of accent experience. *Journal of Child Language*, *46*(2), 368-392.

Lewis, G. (1975). Attitude to language among bilingual children and adults in Wales. *International Journal of the Sociology of Language*, *4*, 103-121.

Lewis, G. J., & Bates, T. C. (2010). Genetic evidence for multiple biological mechanisms underlying in-group favoritism. *Psychological Science*, *21*(11), 1623-1628.

Lewis, G. J., Kandler, C., & Riemann, R. (2014). Distinct heritable influences underpin in-group love and out-group derogation. *Social Psychological and Personality Science*, *5*(4), 407-413.

Liberman, Z., Woodward, A. L., & Kinzler, K. D. (2017). Preverbal infants infer third-party social relationships based on language. *Cognitive Science*, *41*(3), 622-634.

Lindemann, S. (2003). Koreans, Chinese, or Indians? Attitudes and ideologies about non-native English speakers in the United States. *Journal of Sociolinguistics*, *7*(3), 348-364.

Lindsey, G. (2019). *English after RP. Standard British pronunciation today*. Palgrace Macmillan.

Lippi-Green, R. (2012). *English with an accent. Language, ideology and discrimination in the United States*. Routledge.

Lippmann, W. (1922). *Public opinion*. Macmillan.

Lipski, J. (2012). Geographical and social varieties of Spanish: An overview. In J. I. Hualde, A. Olarrea, & E. O'Rourke (Eds.), *The handbook of hispanic linguistics* (pp. 1-26). Wiley-Blackwell.

Lipski, J. (2018). Dialects of Spanish and Portuguese. In C. Boberg, J. Nerbonne, & D. Watt (Eds.), *The handbook of dialectology* (pp. 498-509). Wiley-Blackwell.

Lloyd, P. (1991). Strategies used to communicate route directions by telephone: A comparison of the performance of 7-year-olds, 10-year-olds and adults. *Journal of Child Language*, *18*(1), 171-189.

Loureiro-Rodriguez, V., & Acar, E. (2022). The matched-guise technique. In R. Kircher & L. Zipp (Eds.), *Research methods in language attitudes* (pp. 185–202). Cambridge University Press.

Loureiro-Rodriguez, V., Boggess, M. M., & Goldsmith, A. (2013). Language attitudes in Galicia: Using the matched-guise test among high school students. *Journal of Multilingual and Multicultural Development*, *34*(2), 136-153.

Maass, A., Arcuri, L., & Suitner, C. (2014). Shaping intergroup relations through language. In T. Holtgraces (Ed.), *The Oxford handbook of language and social psychology* (pp. 157-176). Oxford University Press.

Macaulay, R. K. S. (2007). *Talk that counts: Age, gender, and social class differences in discourse*. Oxford University Press.

Macrae, C. N., & Bodenhausen, G. V. (2000). Social cognition: Thinking categorically about others. *Annual Review of Psychology*, *51*(1), 93-120.

Macrae, C. N., Milne, A. B., & Bodenhausen, G. V. (1994). Stereotypes as energy-saving devices: A peek inside the cognitive toolbox. *Journal of Personality & Social Psychology*, *66*(1), 37-47.

Maegaard, M., Jensen, T. J., Kristiansen, T., & Jørgensen, J. N. (2013). Diffusion of language change: Accommodation to a moving target. *Journal of Sociolinguistics*, *17*(1), 3-36.

Maegaard, M., & Pharao, N. (2021). Features, meanings, and indexical fields. In L. Hall-Lew, E. Moore, & R. J. Podesva (Eds.), *Social meaning and linguistic variation: Theorizing the third wave* (pp. 203-221). Cambridge University Press.

Magne, V. (2019). Students' attitudes towards accents of English. *New Vistas*, *5*(1), 32-37.

Manstead, A. S. R. (1995). Children's understanding of emotion. In J. A. Russell, J.-M. Fernández-Dols, A. S. R. Manstead, & J. C. Wellenkamp (Eds.), *Everyday conceptions of emotion: An introduction to the psychology, anthropology and linguistics of emotion* (pp. 315-331). Springer Netherlands.

Markel, N. N., Eisler, R. M., & Reese, H. W. (1967). Judging personality from dialect. *Journal of Verbal Learning and Verbal Behavior*, *6*(1), 33-35.

Martin, N. G., Eaves, L. J., Heath, A. C., Jardine, R., Feingold, L. M., & Eysenck, H. J. (1986). Transmission of social attitudes. *Proceedings of the National Academy of Sciences of the United States of America*, *83*(12), 4364-4368.

Marvasti, A. (2005). U.S. Academic institutions and perceived effectiveness of foreign-born faculty. *Journal of Economic Issues*, *39*(1), 151-176.

Matsuda, M. J. (1991). Voices of America: Accent, antidiscrimination law, and a jurisprudence for the last reconstruction. *The Yale Law Journal*, *100*(5), 1329-1407.

Maye, J., Aslin, R. N., & Tanenhaus, M. K. (2008). The weckud wetch of the wast: Lexical adaptation to a novel accent. *Cognitive Science*, *32*(3), 543-562.

McCarthy, K., & Evans, B. (2019). The perception of familiar and unfamiliar accents by bilingual and monolingual children. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th international congress of phonetic sciences* (pp. 2203-2207). Australasian Speech Science and Technology Association Inc., and International Phonetic Association.

McClelland, J., & Rumelhart, D. (1986). A distributed model of human learning and memory. In J. McClelland & D. Rumelhart (Eds.), *Parallel distributed processing* (pp. 170 - 215). Massachusetts Institute of Technology Press.

McCullough, E. A., Clopper, C. G., & Wagner, L. (2019a). The development of regional dialect locality judgments and language attitudes across the life span. *Child Development*, *90*(4), 1080-1096.

McCullough, E. A., Clopper, C. G., & Wagner, L. (2019b). Regional dialect perception across the lifespan: Identification and discrimination. *Language and Speech*, *90*(4), 1080-1096.

McDougall, P., & Hymel, S. (2007). Same-gender versus cross-gender friendship conceptions: Similar or different? *Merrill-Palmer Quarterly*, *53*(3), 347-380.

McKenzie, R. M., & Carrie, E. (2018). Implicit-explicit attitudinal discrepancy and the investigation of language attitude change in progress. *Journal of Multilingual and Multicultural Development*, *39*(9), 830-844.

McKenzie, R. M., & Gilmore, A. (2017). 'The people who are out of "right" English': Japanese university students' social evaluations of English language diversity and the internationalisation of Japanese higher education. *International Journal of Applied Linguistics*, *27*(1), 152-175.

Medin, D., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 179-196). Cambridge University Press.

Mellor, D., & Moore, K. A. (2014). The use of likert scales with children. *Journal of Pediatric Psychology*, *39*(3), 369-379.

Milroy, J., & Milroy, L. (1999). *Authority in language: Investigating standard English* (3. ed.). Routledge.

Milroy, L., & McClenaghan, P. (1977). Stereotyped reactions to four educated accents in Ulster. *Belfast Working Papers in Language and Linguistics*, *2*(4), 1-10.

Monteith, M. J., Parker, L. R., & Burns, M. D. (2016). The self-regulation of prejudice. In T. D. Nelson (Ed.), *Handbook of prejudice, stereotyping, and discrimination* (pp. 409-432). Psychology Press.

Moon, C., Cooper, R. P., & Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant Behavior and Development*, *16*(4), 495-500.

Moreno Fernández, F. (2016). *A framework for cognitive sociolinguistics*. Routledge.

Morling, B., & Masuda, T. (2012). Social cognition in real worlds: Cultural psychology and social cognition. In S. Fiske & C. N. Macrae (Eds.), *The SAGE handbook of social cognition* (pp. 429-450). Sage.

Morrish, J. (1999, 21st of March). The accent that dare not speak its name. *The Independent on Sunday*, https://www.phon.ucl.ac.uk/home/estuary/morrish.htm.

Moskowitz, G. B., Gollwitzer, P. M., Wasel, W., & Schaal, B. (1999). Preconscious control of stereotype activation through chronic egalitarian goals. *Journal of Personality & Social Psychology*, *77*(1), 167-184.

Mufwene, S. (2004). *The ecology of language evolution*. Cambridge University Press.

Mugglestone, L. (2007). *Talking proper: The rise of accent as social symbol*. Oxford University Press.

Mulac, A. (1975). Evaluation of the speech dialect attitudinal scale. *Speech Monographs*, *42*, 184-189.

Mulac, A., Hanley, T. D., & Prigge, D. Y. (1974). Effects of phonological speech foreignness upon three dimensions of attitude of selected American listeners. *The Quarterly Journal of Speech*, *60*, 411-420.

Munro, M. J. (2003). A primer on accent discrimination in the Canadian context. *TESL Canada Journal*, *20*(2), 38-51.

Munro, M. J., & Derwing, T. M. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language & Speech*, *38*(3), 289-306.

Murphy, M. L. (2003). *Semantic relations and the lexicon: Antonymy, synonymy, and other paradigms*. Cambridge University Press.

Nagy, N., & Irwin, P. (2010). Boston (r): Neighbo(r)s nea(r) and fa(r). *Language Variation and Change*, *22*(2), 241-278.

Nardy, A., Chevrot, J.-P., & Barbu, S. (2013). The acquisition of sociolinguistic variation: Looking back and thinking ahead. *Linguistics*, *51*(2), 255-284.

Nathan, L. I. Z., Wells, B., & Donlan, C. (1998). Children's comprehension of unfamiliar regional accents: A preliminary investigation. *Journal of Child Language*, *25*(2), 343-365.

Nayakankuppam, D., Priester, J. R., Kwon, J. H., Donovan, L. A. N., & Petty, R. E. (2018). Construction and retrieval of evaluative judgments: The attitude strength moderation model. *Journal of Experimental Social Psychology*, *76*, 54-66.

Nazzi, T. (2005). Use of phonetic specificity during the acquisition of new words: Differences between consonants and vowels. *Cognition*, *98*(1), 13-30.

Neergaard, K. D., & Huang, C.-R. (2022). Mandarin Chinese syllable structure and phonological similarity. Perception and production studies. In C.-R. Huang, Y.-H. Lin, I.-H. Chen, & Y.-Y. Hsu (Eds.), *The Cambridge handbook of Chinese linguistics* (pp. 245-274). Cambridge University Press.

Nerlich, B., & Clarke, D. D. (2007). Cognitive linguistics and the history of linguistics. In D. Geeraerts & H. Cuyckens (Eds.), *The Oxford handbook of cognitive linguistics*. Oxford University Press.

Nesdale, D., & Rooney, R. (1990). Effect of children's ethnic accents on adults' evaluations and stereotyping. *Australian Journal of Psychology*, *42*(3), 309-319.

Nesdale, D., & Rooney, R. (1996). Evaluations and stereotyping of accented speakers by pre-adolescent children. *Journal of Language and Social Psychology*, *15*(2), 133-154.

Nespor, M., Peña, M., & Mehler, J. (2003). On the different roles of vowels and consonants in speech processing and language acquisition. *Lingue e Linguaggio*, *2*, 203-229.

Newman, D. M. (2006). *Identities and inequalities: Exploring the intersections of race, class, gender, and sexuality*. McGraw-Hill.

Ng, S. H. (2007). Language-based discrimination: Blatant and subtle forms. *Journal of Language and Social Psychology*, *26*(2), 106-122.

Nguyen, B. B.-D. (1993). Accent discrimination and the test of spoken English: A call for an objective assessment of the comprehensibility of nonnative speakers. *California Law Review*, *81*(5), 1325-1362.

Niedzielski, N. (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology*, *18*(1), 62-85.

Nier, J. A. (2005). How dissociated are implicit and explicit racial attitudes? A bogus pipeline approach. *Group Processes & Intergroup Relations*, *8*(1), 39-52.

Nippold, M. A. (2002). Lexical learning in school-age children, adolescents, and adults: A process where language and literacy converge. *Journal of Child Language*, *29*(2), 449-488.

Nippold, M. A. (2004). Research on later language development: International perspectives. In R. A. Berman (Ed.), *Language development across childhood and adolescence* (pp. 1-8). John Benjamins Publishing Company.

Nippold, M. A. (2007). *Later language development: School-age children, adolescents, and young adults*. PRO-ED.

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology*, *47*(2), 204-238.

Nortier, J. (2001). Street language in the Netherlands. In A.-B. Stenström, U. B. Kotsinas, & E. Drange (Eds.), *Ungdommers språkmøter* (pp. 29-141). Nordisk Ministerråd.

Nortier, J., & Dorleijn, M. (2008). A Moroccan accent in Dutch: A sociocultural style restricted to the Moroccan community? *International Journal of Bilingualism*, *12*(1-2), 125-142.

Nosek, B. A., & Banaji, M. R. (2001). The go/no go association task. *Social Cognition*, *19*(6), 625-666.

Nyvold, M. (2024, 10th of June). »hvis nogen fik og stadig får et godt grin af det, så er det jo nok sundt«. *Klimamonitor*, https://klimamonitor.dk/nyheder/art9938763/Hvis-nogen-fik-og-stadig-får-et-godt-grin-af-det-så-er-det-jo-nok-sundt.

O'Shannessy, C. (2013). Researching children's acquisition of sociolinguistic competence. In J. Holmes & K. Hazen (Eds.), *Research methods in sociolinguistics* (pp. 304-324). John Wiley and Sons.

Oakhill, J., Garnham, A., & Reynolds, D. (2005). Immediate activation of stereotypical gender information. *Memory & Cognition*, *33*(6), 972-983.

Odato, C. (2010). *Children's development of knowledge and beliefs about English like(s)* [PhD Dissertation, The University of Michigan]. Ann Arbor.

Office for National Statistics. (2022). *Population by country of birth*. London Datastore. Retrieved 17th of March 2025 from https://data.london.gov.uk/dataset/country-of-birth.

Orey, B. D. A., & Park, H. (2012). Nature, nurture, and ethnocentrism in the Minnesota twin study. *Twin Research and Human Genetics*, *15*(1), 71-73.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. University of Illinois Press.

Pace, A., Luo, R., Hirsh-Pasek, K., & Golinkoff, R. M. (2017). Identifying pathways between socioeconomic status and language development. *Annual Review of Linguistics*, *3*, 285-308.

Pantos, A. (2019). Implicitness, automaticity, and consciousness in language attitudes research. Are they related and how do we characterize them? *Linguistics Vanguard*, *5*(s1), Article 20180007.

Pantos, A. J. (2014). Defining the cognitive mechanisms underlying reactions to foreign accented speech: An experimental approach. In M. Pütz, J. A. Robinson, & M. Reif (Eds.), *Cognitive sociolinguistics: Social and cultural variation in cognition and language use* (pp. 187-212). John Benjamins Publishing Company.

Pantos, A. J., & Perkins, A. W. (2013). Measuring implicit and explicit attitudes toward foreign accented speech. *Journal of Language and Social Psychology*, *32*(1), 3-20.

Papapavlou, A., & Sophocleous, A. (2009). Relational social deixis and the linguistic construction of identity. *International Journal of Multilingualism*, *6*(1), 1-16.

Papazachariou, D., Fterniati, A., Archakis, A., & Tsami, V. (2018). Metapragmatic stereotypes about geographical diversity in Greece. *Journal of Greek Linguistics*, *18*(1), 124-154.

Paquette-Smith, M., Buckler, H., White, K. S., Choi, J., & Johnson, E. K. (2019). The effect of accent exposure on children's sociolinguistic evaluation of peers. *Dev Psychol*, *55*(4), 809-822.

Paquette-Smith, M., Cooper, A., & Johnson, E. K. (2021). Targeted adaptation in infants following live exposure to an accented talker. *Journal of Child Language*, *48*(2), 325-349.

Paradis, C. (2003). Is the notion of linguistic competence relevant in cognitive linguistics? *Annual Review of Cognitive Linguistics*, *1*(1), 207-231.

Paradis, C. (2008). Configurations, construals and change: Expressions of degree. *English Language and Linguistics*, *12*(2), 317-343.

Paradis, C. (2015). Meanings of words: Theory and application. In U. Haß & P. Storjohann (Eds.), *Handbuch wort und wortschatz* (pp. 274-294). De Gruyter.

Patterson, M. L., & Werker, J. F. (2002). Infants' ability to match dynamic phonetic and gender information in the face and voice. *Journal of Experimental Child Psychology*, *81*(1), 93-115.

Payne, B. K., & Gawronski, B. (2010). A history of implicit social cognition: Where is it coming from? Where is it now? Where is it going? In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications.* (pp. 1-15). The Guilford Press.

Pear, T. (1931). *Voice and personality*. Chapman and Hall Ltd.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). Psychopy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195-203.

Pélissier, M., & Ferragne, E. (2022). The N400 reveals implicit accent-induced prejudice. *Speech Communication*, *137*, 114-126.

Perris, E. E., & Clifton, R. K. (1988). Reaching in the dark toward sound as a measure of auditory localization in infants. *Infant Behavior and Development*, *11*(4), 473-491.

Peskin, J., & Olson, D. (2004). On reading poetry. Expert and novice knowledge. In R. A. Berman (Ed.), *Language development across childhood and adolescence* (pp. 211-232). John Benjamins Publishing Company.

Petty, R. E., DeMarree, K. C., & Briñol, P. (2007). The meta-cognitive model (MCM) of attitudes: Implications for attitude measurement, change, and strength. *Social Cognition*, *25*(5), 657-686.

Pharao, N., & Kristiansen, T. (2019). Reflections on the relation between direct/indirect methods and explicit/implicit attitudes. *Linguistics Vanguard*, *5*(s1), Article 20180010.

Piaget, J. (1954). *The construction of reality in the child*. Basic.

Pietraszewski, D., & Schwartz, A. (2014a). Evidence that accent is a dedicated dimension of social categorization, not a byproduct of coalitional categorization. *Evolution and Human Behavior*, *35*(1), 51-57.

Pietraszewski, D., & Schwartz, A. (2014b). Evidence that accent is a dimension of social categorization, not a byproduct of perceptual salience, familiarity, or ease-of-processing. *Evolution and Human Behavior*, *35*(1), 43-50.

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology*, *75*(3), 811-832.

Plant, E. A., Kling, K. C., & Smith, G. L. (2004). The influence of gender and social role on the interpretation of facial expressions. *Sex Roles: A Journal of Research*, *51*, 187-196.

Podesva, R. (2006). *Phonetic detail in sociolinguistic variation: Its linguistic significance and role in the construction of social meaning* [PhD Dissertation, Stanford University]. Stanford.

Põldvere, N., Felice, R. D., & Paradis, C. (2022). *Advice in conversation: Corpus pragmatics meets mixed methods*. Cambridge University Press.

Põldvere, N., Johansson, V., & Paradis, C. (2021). Resonance in dialogue: The interplay between intersubjective motivations and cognitive facilitation. *Language and Cognition*, *13*(4), 643-669.

Posit team. (2024). *Rstudio: Integrated development environment for r*. In Posit Software, PBC, http://www.posit.co/.

Poteat, P. (2007). Peer group socialization of homophobic attitudes and behavior during adolescence. *Child Development*, *78*(6), 1830-1842.

Powlishta, K. K., Serbin, L., Doyle, A.-B., & White, D. (1994). Gender, ethnic, and body type biases: The generality of prejudice in childhood. *Developmental Psychology*, *30*(4), 526-536.

Preston, D. (1993). Folk dialectology. In P. Dennis (Ed.), *American dialect research* (pp. 333-378). John Benjamins Publishing Company.

Preston, D. (1996). Whaddayaknow? The modes of folk linguistic awareness. *Language Awareness*, *5*(1), 40-74.

Preston, D. (1998). They speak bad English in the South & New York city. In L. Bauer & P. Trudgill (Eds.), *Language myths* (pp. 139-149). Penguin Books.

Preston, D. (1999a). Introduction. In D. Preston (Ed.), *Handbook of perceptual dialectology* (pp. xxiii-xxxix). Benjamins.

Preston, D. (1999b). A language attitude approach to the perception of regional variety. In D. Preston (Ed.), *Handbook of perceptual dialectology* (pp. 359-373). Benjamins.

Price, S., Fluck, M., & Giles, H. (1983). The effects of language of testing on bilingual pre-adolescents' attitudes towards Welsh and varieties of English. *Journal of Multilingual and Multicultural Development*, *4*(2-3), 149-161.

Proverbio, A. M., Alberio, A., & De Benedetto, F. (2018). Neural correlates of automatic beliefs about gender stereotypes: Males are more prejudicial. *Brain and Language*, *186*, 8-16.

Proverbio, A. M., Orlandi, A., & Bianchi, E. (2017). Electrophysiological markers of prejudice related to sexual gender. *Neuroscience*, *358*, 1-12.

Pütz, M., Robinson, J. A., & Reif, M. (2014). *Cognitive sociolinguistics. Social and cultural variation in cognition and language use*. John Benjamins Publishing Company.

Quine, W. v. O. (1951). Two dogmas of empiricism. *Philosophical Review*, *60*, 20-43.

R Core Team. (2024). *R: A language and environment for statistical computing*. In R Foundation for Statistical Computing, https://www.R-project.org/.

Rakić, T., Steffens, M. C., & Mummendey, A. (2011a). Blinded by the accent! The minor role of looks in ethnic categorization. *Journal of Personality and Social Psychology*, *100*(1), 16-29-29.

Rakić, T., Steffens, M. C., & Mummendey, A. (2011b). When it matters how you pronounce it: The influence of regional accents on job interview outcome. *British Journal of Psychology*, *102*(4), 868-883.

Ramirez, A., Arce-Torres, E., & Politzer, R. (1976). *Language attitudes and the achievement of bilingual pupils*. Stanford Center for Research and Development in Teaching.

Rampton, B. (2011). From 'multi-ethnic adolescent heteroglossia' to 'contemporary urban vernaculars'. *Language & Communication*, *31*(4), 276-294.

Redinger, D. (2010). *Language attitudes and code-switching behaviour in a multilingual educational context: The case of Luxembourg* [PhD Dissertation, University of York]. York.

Reid, P. T. (1979). Racial stereotyping on television: A comparison of the behavior of both black and white television characters. *Journal of Applied Psychology*, *64*(5), 465-471.

Reinard, J. C. (2013). Persuasion in the legal setting. In J. Dillard & L. Shen (Eds.), *The SAGE handbook of persuasion: Developments in theory and practice* (pp. 331-353). Sage.

Remennick, L. I. (1999). 'Women with a Russian accent' in Israel - on the gender aspects of immigration. *European Journal of Women's studies*, *6*, 441-461.

Riach, P. A., & Rich, J. (2004). Fishing for discrimination. *Review of Social Economy*, *62*(4), 465-486.

Ridge, T. (2002). *Childhood poverty and social exclusion. From a child's perspective*. The Policy Press.

Rodríguez-García, J.-M., & Wagner, U. (2009). Learning to be prejudiced: A test of unidirectional and bidirectional models of parent–offspring socialization. *International Journal of Intercultural Relations*, *33*(6), 516-523.

Rodríguez-Gómez, P., Romero-Ferreiro, V., Pozo, M. A., Hinojosa, J. A., & Moreno, E. M. (2020). Facing stereotypes: ERP responses to male and female faces after gender-stereotyped statements. *Social Cognitive & Affective Neuroscience*, *15*(9), 928-940.

Roets, A., & Van Hiel, A. (2011). Allport's prejudiced personality today: Need for closure as the motivated cognitive basis of prejudice. *Current Directions in Psychological Science*, *20*(6), 349-354.

Rosenthal, M. S. (1974). The magic boxes: Pre-school children's attitudes toward black and standard English. *The Florida FL Reporter*, *12*, 55-62.

Rowatt, W. C., Tsang, J.-A., Kelly, J., LaMartina, B., McCullers, M., & McKinley, A. (2006). Associations between religious personality dimensions and implicit homosexual prejudice. *Journal for the Scientific Study of Religion*, *45*(3), 397-406.

Royeen, C. B. (1985). Adaptation of likert scaling for use with children. *The Occupational Therapy Journal of Research*, *5*(1), 59-69.

Rubin, D., Coles, V. B., & Barnett, J. T. (2016). Linguistic stereotyping in older adults' perceptions of health care aides. *Health Communication*, *31*(7), 911-916.

Rubin, D. L. (1992). Nonlanguage factors affecting undergraduates' judgments of nonnative English-speaking teaching assistants. *Research in Higher Education*, *33*(4), 511-531.

Rubin, D. L. (2012). The power of prejudice in accent perception: Reverse linguistic stereotyping and its impact on listener judgments and decisions. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 3rd pronunciation in second language learning and teaching conference* (pp. 11-17). Iowa State University.

Rubin, D. L., Healy, P., Gardiner, T. C., Zath, R. C., & Moore, C. P. (1997). Nonnative physicians as message sources: Effects of accent and ethnicity on patients' responses to AIDS prevention counseling. *Health Communication*, *9*(4), 351-368.

Rubin, D. L., & Smith, K. A. (1990). Effects of accent, ethnicity, and lecture topic on undergraduates' perceptions of nonnative English-speaking teaching assistants. *International Journal of Intercultural Relations*, *14*(3), 337-353.

Rubin, K. H., Bukowski, W. M., & Parker, J. G. (2006). Peer interactions, relationships, and groups. In N. Eisenberg, W. Damon, & R. Lerner (Eds.), *Handbook of child psychology: Social, emotional, and personality development* (pp. 571-645). John Wiley & Sons, Inc.

Rudzka-Ostyn, B. (1988). *Topics in cognitive linguistics*. John Benjamins.

Russell, B. (1969). *An inquiry into meaning and truth*. Penguin.

Ryan, E. B., Carranza, M. A., & Moffie, R. W. (1977). Reactions toward varying degrees of accentedness in the speech of Spanish–English bilinguals. *Language & Speech*, *20*(3), 267-273.

Ryan, E. B., & Giles, H. (1982). *Attitudes towards language variation: Social and applied contexts*. E. Arnold.

Santhanagopalan, R., DeJesus, J. M., Moorthy, R. S., & Kinzler, K. D. (2021). Nationality cognition in India: Social category information impacts children's judgments of people and their national identity. *Cognitive Development*, *57*, Article 100990.

Satraki, M. (2015). *The development of children's attitudes towards linguistic variation: A study on a dialect/register continuum* [PhD Dissertation, University of Cyprus]. Nicosia.

Saussure, F. d. (2015). *Course in general linguistics* (W. Baskin, Trans.; C. Bally & A. Sechehaye, Eds.). Forgotten Books.

Scharinger, M., Monahan, P. J., & Idsardi, W. J. (2011). You had me at "hello": Rapid extraction of dialect information from spoken words. *NeuroImage*, *56*(4), 2329-2338.

Schmidt, R. (2002). Racialization and language policy: The case of the U.S.A. *Multilingua: Journal of Cross-Cultural and Interlanguage Communication*, *21*(2-3), 141-161.

Schneiderman, E. (1982). Sex differences in the development of children's ethnic and language attitudes. *International Journal of the Sociology of Language*, *38*, 31-44.

Schwarz, N., & Bohner, G. (2001). The construction of attitudes. In A. Tesser & N. Schwarz (Eds.), *Blackwell handbook of social psychology: Intraindividual processes* (pp. 436-457). Wiley Blackwell.

Seggie, I. (1983). Attribution of guilt as a function of ethnic accent and type of crime. *Journal of Multilingual and Multicultural Development*, *4*(2-3), 197-206.

Seligman, C. R., Tucker, G. R., & Lambert, W. E. (1973). The effects of speech style and other attributes on teachers' attitudes toward pupils. *Language in Society*, *2*(1), 131-142.

Sharma, D., Levon, E., & Ye, Y. (2022). 50 years of British accent bias: Stability and lifespan change in attitudes to accents. *English World-Wide*, *43*(2), 135-166.

Shuck, G. (2006). Racializing the nonnative English speaker. *Journal of Language, Identity & Education*, *5*(4), 259-276.

Shutts, K., Kinzler, K., McKee, C., & Spelke, E. (2009). Social information guides infants' selection of foods. *Journal of Cognition & Development*, *10*(1/2), 1-17.

Sidanius, J., & Pratto, F. (2001). *Social dominance: An intergroup theory of social hierarchy and oppression*. Cambridge University Press.

Singh, A. (2002). "We are not the enemy" hate crimes against Arabs, Muslims, and those perceived to be Arab or Muslim after September 11. *Human Rights Watch*, *14*(6), 1-41.

Smith, E. R. (1996). What do connectionism and social psychology offer each other? *Journal of Personality & Social Psychology*, *70*(5), 893-912.

Smith, P. T., Turner, J. E., Brown, P. A., & Henry, L. A. (2006). The distinct contributions of age of acquisition and word frequency in auditory word perception. *The Quarterly Journal of Experimental Psychology*, *59*(12), 2121-2134.

Soley, G., & Sebastián-Gallés, N. (2015). Infants prefer tunes previously introduced by speakers of their native language. *Child Development*, *86*(6), 1685-1692.

Søndergaard, B. (2011, 7th of September). Socialisten, der vil magten. *Kristeligt Dagblad*, https://www.kristeligt-dagblad.dk/mennesker/socialisten-der-vil-magten.

Soukup, B. (2012). Current issues in the social psychological study of 'language attitudes': Constructionism, context, and the attitude-behavior link. *Linguistics and Language Compass*, *6*(4), 212-224.

Souza, A., Byers-Heinlein, K., & Poulin-Dubois, D. (2013). Bilingual and monolingual children prefer native-accented speakers. *Frontiers in Psychology*, *4*, Article 953.

Speelman, D., Spruyt, A., Impe, L., & Geeraerts, D. (2013). Language attitudes revisited: Auditory affective priming. *Journal of Pragmatics*, *52*, 83-92.

Spence, J. L., Hornsey, M. J., & Imuta, K. (2021). Something about the way you speak: A meta-analysis on children's linguistic-based social preferences. *Child Development*, *92*(2), 517-535.

Spence, J. L., & Imuta, K. (2020). Age-related changes in children's accent-based resource distribution. *Journal of Experimental Child Psychology*, *193*, Article 104807.

Spence, M., Rollins Pamela, R., & Jerger, S. (2002). Children's recognition of cartoon voices. *Journal of Speech, Language, and Hearing Research*, *45*(1), 214-222.

Stack, D. M., Muir, D. W., Sherriff, F., & Roman, J. (1989). Development of infant reaching in the dark to luminous objects and 'invisible sounds'. *Perception*, *18*(1), 69-82.

Stamou, A. G., Maroniti, K., & Griva, E. (2015). Young children talk about their popular cartoon and tv heroes' speech styles: Media reception and language attitudes. *Language Awareness*, *24*(3), 216-232.

Stangor, C. (2016). The study of stereotyping, prejudice and discrimination within social psychology: A quick history of theory and research. In T. Nelson (Ed.), *Handbook of prejudice, stereotyping, and discrimination* (pp. 3–27). Psychology Press.

Stewart, M. A., Ryan, E. B., & Giles, H. (1985). Accent and social class effects on status and solidarity evaluations. *Personality and Social Psychology Bulletin*, *11*(1), 98-105.

Stoel-Gammon, C. (2011). Relationships between lexical and phonological development in young children. *Journal of Child Language*, *38*(1), 1-34.

Storage, D., Charlesworth, T. E. S., Banaji, M. R., & Cimpian, A. (2020). Adults and children implicitly associate brilliance with men more than women. *Journal of Experimental Social Psychology*, *90*, Article 104020.

Sue, D. W. (2010). *Microaggressions in everyday life: Race, gender, and sexual orientation*. John Wiley & Sons Inc.

Sullivan, J. (2006). *Developing knowledge of polysemous vocabulary* [PhD Dissertation, University of Waterloo]. Waterloo.

Sumner, W. (1906). *Folkways*. Ginn and Company.

Sutton, L., Smith, N., Dearden, C., & Middleton, S. (2007). *A child's-eye view of social difference*. Joseph Rowntree Foundation.

Swann, J., Deumert, A., Lillis, T., & Mestherie, R. (2004). *A dictionary of sociolinguistics*. Edinburgh University Press.

Tabak, I., & Reiser, B. (2022). Scaffolding. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 53-71). Cambridge University Press.

Tajfel, H. (1969). Cognitive aspects of prejudice. *Journal of Social Issues*, *25*(4), 79-97.

Tajfel, H. (1978). The achievement of group differentiation. In H. Tajfel (Ed.), *Differentiation between social groups: Studies in the social psychology of intergroup relations* (pp. 77-98). Academic Press.

Tajfel, H., & Wilkes, A. L. (1963). Classification and quantitative judgment. *British Journal of Psychology*, *54*(2), 101-114.

Talmy, L. (1978). The relation of grammar to cognition - a synopsis. *Proceedings of TINLAP*, *2*, 14-24.

Talmy, L. (2000). *Toward a cognitive semantics*. MIT Press.

Talmy, L. (2015). Relating language to other cognitive systems: An overview. *Cognitive Semantics*, *1*(1), 1-44.

Taylor, J. (2003). *Linguistic categorization*. Oxford University Press.

Theilgaard, C. (2024, 19th of May). For 15 år siden blev Villy Søvndal latterliggjort for et særligt udtryk – nu er det hans valgslogan. *B.T.*, https://www.bt.dk/politik/for-15-aar-siden-blev-villy-soevndal-latterliggjort-for-et-saerligt-udtryk-nu-er.

Thøgersen, J., & Pharao, N. (2020). Holdningers og mediers rolle i (de)standardisering. *Danske talesprog*, *20*, 103–122.

Thurstone, L. (1928). Attitudes can be measured. *American Journal of Sociology*, *33*, 529-554.

Torgersen, E. N., & Szakay, A. (2012). An investigation of speech rhythm in London English. *Lingua*, *122*(7), 822-840.

Trowell, M. (2007). *A test of the effects of linguistic stereotypes in children's animated film: A language attitude study* [Master of Arts Thesis, University of Northern Texas]. Denton.

Trudgill, P. (1998). The meanings of words should not be allowed to vary or change. In L. Bauer & P. Trudgill (Eds.), *Language myths* (pp. 176-182). Penguin Books.

Trudgill, P. (1999). *The dialects of England*. Blackwell.

Trudgill, P., & Hannah, J. (2017). *International English: A guide to varieties of English around the world*. Routledge.

Tsalikis, J., Ortiz, hyphen, Buonafina, M., & LaTour, M. S. (1992). The role of accent on the credibility and effectiveness of the international business person: The case of Guatemala. *International Marketing Review*, *9*(4), 57-57.

Tucker, G. R., & Lambert, W. E. (1969). White and negro listeners' reactions to various American-English dialects. *Social Forces*, *47*(4), 463-468.

Ubalde, J., Alarcón, A., & Lapresta, C. (2017). Evolution and determinants of language attitudes among Catalan adolescents. *International Journal of Intercultural Relations*, *60*, 92-103.

Upton, C. (2008). Received Pronunciation. In B. Kortmann & C. Upton (Eds.), *Varieties of English. The British Isles* (pp. 237-252). Mouton de Gruyter.

US Department of Justice. (2023). *2020 FBI hate crimes statistics*. Community Relations Service, U.S. Department of Justice. Retrieved 1st of July 2024 from https://www.justice.gov/crs/highlights/2020-hate-crimes-statistics.

Uttley, L., de Boisferon, A. H., Dupierrix, E., Lee, K., Quinn, P. C., Slater, A. M., & Pascalis, O. (2013). Six-month-old infants match other-race faces with a non-native language. *International Journal of Behavioral Development*, *37*(2), 84-89.

van Avermaet, E., & McClintock, C. G. (1988). Intergroup fairness and bias in children. *European Journal of Social Psychology*, *18*(5), 407-427.

van Berkum, J. J. A., van den Brink, D., Tesink, C. M. J., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of Cognitive Neuroscience*, *20*(4), 580-591.

van Bezooijen, R. (1994). Aesthetic evaluation of Dutch language varieties. *Language & Communication*, *14*(3), 253-263.

van der Hulst, H. (2011). Pitch accent systems. In v. M. Oostendorp, C. Ewen, E. Hume, & K. Rice (Eds.), *The Blackwell companion to phonology* (pp. 1003-1026). John Wiley and Sonds.

van Ooijen, B. (1996). Vowel mutability and lexical selection in English: Evidence from a word reconstruction task. *Memory & Cognition*, *24*(5), 573-583.

Viterbori, P., Zanobini, M., & Cozzani, F. (2018). Phonological development in children with different lexical skills. *First Language*, *38*(5), 538-559.

Wagner, L., Clopper, C. G., & Pate, J. K. (2014). Children's perception of dialect variation. *Journal of Child Language*, *41*(5), 1062-1084.

Wagner, L., Dunfield, K. A., & Rohrbeck, K. L. (2014). Children's use of social cues when learning conventions. *Journal of Cognition and Development*, *15*(3), 479-494.

Wakelin, M. (1986). *The Southwest of England*. Benjamins.

Wated, G., & Sanchez, J. I. (2006). The role of accent as a work stressor on attitudinal and health-related work outcomes. *International Journal of Stress Management*, *13*(3), 329-350.

Weatherhead, D. (2017). *The role of social information in infants' and toddlers' linguistic representations* [PhD Dissertation, University of Waterloo]. Waterloo.

Weatherhead, D., Friedman, O., & White, K. S. (2018). Accent, language, and race: 4-6-year-old children's inferences differ by speaker cue. *Child Development*, *89*(5), 1613-1624.

Weatherhead, D., Friedman, O., & White, K. S. (2019). Preschoolers are sensitive to accent distance. *Journal of Child Language*, *46*(6), 1058-1072.

Weatherhead, D., & White, K. S. (2018). And then I saw her race: Race-based expectations affect infants' word processing. *Cognition*, *177*, 87-97.

Weatherhead, D., White, K. S., & Friedman, O. (2016). Where are you from? Preschoolers infer background from accent. *Journal of Experimental Child Psychology*, *143*, 171-178.

Weatherhead, D., White, K. S., & Friedman, O. (2018). Children's accent-based inferences depend on geographic background. *Journal of Experimental Child Psychology*, *175*, 108-116.

Weber, R., & Crocker, J. (1983). Cognitive processes in the revision of stereotypic beliefs. *Journal of Personality and Social Psychology*, *45*(5), 961-977.

Weinger, S. (2000). Economic status: Middle class and poor children's views. *Children & Society*, *14*(2), 135-146.

Weisbuch, M., Pauker, K., & Ambady, N. (2009). The subtle transmission of race bias via televised nonverbal behavior. *Science*, *326*(5960), 1711-1714.

Wells, J. C. (1982). *Accents of English*. Cambridge University Press.

White, K. R., Crites Jr, S. L., Taylor, J. H., & Corral, G. (2009). Wait, what? Assessing stereotype incongruities using the N400 ERP component. *Social Cognitive & Affective Neuroscience*, *4*(2), 191-198.

Wiese, H. (2009). Grammatical innovation in multiethnic urban Europe: New linguistic practices among adolescents. *Lingua*, *119*(5), 782-806.

Williams, A., Garrett, P., & Coupland, N. (1999). Dialect recognition. In D. Preston (Ed.), *Handbook of perceptual dialectology* (pp. 345-358). Benjamins.

Williams, D. R. (1999). Race, socioeconomic status, and health: The added effects of racism and discrimination. In N. E. Adler, M. Marmot, B. S. McEwen, & J. Stewart (Eds.), *Socioeconomic status and health in industrial nations: Social, psychological, and biological pathways.* (pp. 173-188). New York Academy of Sciences.

Williams, D. R., & Rucker, T. D. (2000). Understanding and addressing racial disparities in health care. *Health Care Financing Review*, *21*(4), 75-91.

Williams, D. R., & Williams-Morris, R. (2000). Racism and mental health: The African American experience. *Ethnicity & Health*, *5*(3/4), 243-268.

Wittenbrink, B., Judd, C. M., & Park, B. (2001). Spontaneous prejudice in context: Variability in automatically activated attitudes. *Journal of Personality and Social Psychology*, *81*(5), 815-827.

Wittgenstein, L. (1986). *Philosophical investigations* (G. E. M. Anscombe, Trans.; 3rd ed.). Basil Blackwell.

Wolfram. (2017). Dialect in society. In F. Coulmas (Ed.), *The handbook of sociolinguistics* (pp. 107-126). Blackwell Publishers.

Yi, H. G., Phelps, J. E., Smiljanic, R., & Chandrasekaran, B. (2013). Reduced efficiency of audiovisual integration for nonnative speech. *The Journal of the Acoustical Society of America*, *134*(5), El387-393.

Youniss, J. (1980). *Parents and peers in social development. A sullivan-piaget perspective*.

Zachariassen, D. (2022). *"Sån det der" – nogle prosodiske, syntaktiske og leksikale træk i en urban dialekt i Aarhus* [PhD Dissertation, Aarhus University]. Aarhus.

Zahn, C. J., & Hopper, R. (1985). Measuring language attitudes: The speech evaluation instrument. *Journal of Language and Social Psychology*, *4*(2), 113-123.

Zanna, M. P., & Rempel, J. K. (1998). Attitudes: A new look at an old concept. In D. Bar-Tal & A. W. Kruglanski (Eds.), *The social psychology of knowledge* (pp. 7-15). Cambridge University Press.

Zhang, Y., Nissen, S., & Francis, A. (2008). Acoustic characteristics of English lexical stress produced by native Mandarin speakers. *Journal of the Acoustical Society of America*, *123*(6), 4498-4513.

Zhao, B., Ondrich, J., & Yinger, J. (2006). Why do real estate brokers continue to discriminate? Evidence from the 2000 housing discrimination study. *Journal of Urban Economics*, *59*(3), 394-419.

Zheng, Y., & Samuel, A. G. (2017). Does seeing an Asian face make speech sound more accented? *Attention, Perception, and Psychophysics*, *79*(6), 1841-1859.

Zhu, Z. (2017). *How Mandarin Chinese is growing in schools in England*. British Council. Retrieved 16th of August 2024 from https://www.britishcouncil.org/voices-magazine/how-mandarin-chinese-growing-england-schools.

Zuckerman, D. M., Singer, D. G., & Singer, J. L. (1980). Children's television viewing, racial and sex-role attitudes. *Journal of Applied Social Psychology*, *10*(4), 281-294.