How is the Social Meaning of Linguistic Variation Stored in Memory?

Hielke Anne David Vriesendorp

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

University of Sheffield

2021

Abstract

This thesis uses experimental methods to investigate the storage of the social meaning of accent variation in memory, focussing on two main aspects:

1. The connection between linguistic variation and its social meaning in memory storage
2. The size and detail of the linguistic representations that are used to access social meaning

The first aspect is explored by testing whether social and linguistic information are joined together within the same representations in memory or not. This would be the case if the episodic memory traces that exemplar theory postulates incorporate social meaning (Docherty & Foulkes 2014). Alternatively third-wave variationist research posits that there is a fluid link between variation and its social meaning (Eckert 2008), implying separate storage in memory. Innovations to exemplar theory allow for both exemplars and more abstract generalisations across exemplars to be used in speech processing. This introduces new types of representations to its theory of speech processing, which may store social and linguistic information separately, although this has not been explored extensively yet. In this PhD I present three highly automatic 'socio-contextual priming' experiments which suggest that it is possible for listeners to process linguistic variation without its social meaning becoming activated as well, suggesting that the two types of information are stored separately.

The second aspect is explored by means of two accent recognition experiments, which tested whether listeners recognise accents more accurately depending on whether the stimulus word is frequent, not frequent or a non-word. These experiments found that listeners are able to recognise accents even in non-words, implying that lexical information is not required for the retrieval of social (accent) information, and that sub-lexical information is used to access social meaning. Accuracy was higher in the experiment which used real words, but there was no difference between accuracy in high-frequency and low-frequency words within the experiment. This suggests that the number or strength of lexical representations that are used to access social meaning does not facilitate access to social meaning, but that the presence of a lexical phonological form is helpful for the contextualisation of sub-lexical information used to access social meaning. Furthermore, the accent recognition experiments found that listeners perceived fewer voices in the experiments than were used in reality, suggesting that the representations listeners used for recognition were only detailed enough for them to distinguish between more general character types rather than individual speakers.

Table of Contents

List of Figures

List of Tables

Acknowledgements

This PhD would not have been possible without the great help and enthusiasm of my experimental participants, as well as all the retweeters and sharers and commenters and boosters and I'll-message-my-family-ers. I would like to thank them. It is a very special type of gratefulness to feel when someone's mum drives down from York to get her clear /l/ and centring FACE vowel on tape, or your friend tells you he's found a group of eight strangers in the queue for a youthful rock band who are now all performing your accent recognition task on the one laptop they have between them.

The project was made possible through the University Prize Scholarship, which allowed me not just to do the project but also to take my ideas into the world at different conferences and summer schools and sharpen them further there. I would like to thank the New Zealand Institute of Language, Brain and Behaviour, in particular Lynn Clark and Andy Gibson, for having me during a research visit which reshaped the PhD as I knew it, and for teaching me the joys of the mixed-effect regression model and RStudio.

At the University of Sheffield, I would like to thank Meesha Warmington for giving me access to Gorilla, the software that radically transformed the project in terms of both scale and quality. I'd like to thank my academic mentor Angie Frost for giving me the tools to work through this long term project steadily and confidently.

I would like to thank my examiners Ranjan Sen and Erez Levon for an enjoyable viva experience and their helpful feedback.

Finally, I want to thank my supervision team for their guidance and invaluable expertise: Kirsten Bartlett, Chris Montgomery, and Emma Moore. I feel incredibly lucky to have been able to work with such a supportive and brilliant group of people. In particular, I'd like to thank Kirsten for her never-ending enthusiasm, Chris for his down-to-earth advice, and Emma for her insights into the perfectionism that I didn't even know I had.

Personal thanks

Writing this PhD has been a cocktail of conflicting experiences. It was isolating and sometimes damaging, it was full of excitement, and dark and anxiety inducing. It also gave me the time to develop, and to enjoy student life for a little longer – and of course it paid for my food, my bills, my travels, my books and my crop tops. It mirrors my experience with moving to Sheffield itself. The three years living here were coloured by newness and loss, frustration and excitement. I'd like to thank the following people for being there in this time.

First, I want to thank my friends Johanna Grace, Nathaniel Dziura, Holly Dann, and Sarah Tasker for their friendship, support, and insight. One of yous said it best: we don't ave much money but we do have a laugh.

I am thankful to my therapist Samantha Montague for seeing straight through the stories I told myself and giving me the space (and worksheets) to start healing.

I'd like to thank Amraj Lally for his outspoken joy, our connection, and for challenging me into becoming a fuller version of myself. I'd like to thank Tara Cohen for bringing me back to lightness and myself, whenever British culture got the best of me.

Thank you to Rory Wilson for your companionship in all its different forms. You've opened up a whole new world to me – of worthiness, love, and being seen.

Thank you to my parents Greet Meijer en Herman Vriesendorp, voor jullie liefde door dik en dun.

1 Introduction

1.1 The storage of social meaning

This thesis investigates the question of how the social meaning of linguistic variation is stored in cognition. This question is at the intersection of two fields in sociolinguistics which have gained a lot of traction in the past 20 years. Firstly it engages with recent work on the social meaning of linguistic variation in the 'third-wave' approach to linguistic variation (Eckert 2012; Hall-Lew, Moore & Podesva 2021). In this approach, the social meaning of variation is seen as the dynamic, flexible inferences that are made about a speaker in interaction on the basis of the variants in their speech. Secondly, it engages with work on the cognitive underpinnings of sociolinguistic perception, in particular work on 'exemplar theory' (Docherty & Foulkes 2014; Hay 2018). This theory posits that speech is processed by matching input to stored detailed episodic memories, or 'exemplars', which include social and contextual information, or to more abstract generalisations across those exemplars.

There are two aspects to the main question about the storage of social meaning which I investigate. The first is whether social meaning is stored in united memory representations, such as exemplars, where linguistic and social information is part of the same memory unit, or whether social and linguistic information are stored separately, but closely linked. The former fits the notion of the exemplar in exemplar theory, in which social and linguistic information are stored in the same episodic memory: a memory of the same 'episode' or event. The latter would be more in line with third-wave approaches which posit flexible links between social meaning and the linguistic variation it is attached to, but is also possible in 'hybrid' exemplar theory. Here, the more abstract generalisations that listeners make on the basis of exemplars may store social and linguistic information separately. I investigate this by using highly automatic socio-contextual priming experiments to test whether the processing of linguistic information always implies the activation of its social meaning, and vice versa. From the experiments I conducted, it seems possible to activate one without the other, providing evidence that the two are stored separately.

The second aspect of the main question is that of what linguistic information social meaning is connected to. Sociolinguistic work on exemplar theory (Foulkes & Docherty 2006; Johnson 2006; Hay 2018) argues that social meaning is stored in word-sized exemplars. However, hybrid exemplar theory has progressed to include the notion that language users form generalisations across these highly detailed exemplars (e.g. a large number of different episodic memories of people saying a word in different ways) and create more abstract linguistic representations (e.g. the phonological forms of that word) (McLennan, Luce & Charles-Luce 2003; Goldinger 2007; Pierrehumbert 2016). These innovations open up the possibility that other potential memory representations can be used to access social meaning.

I investigated which size (lexical or sub-lexical) and how detailed (from phonological form to detailed exemplar) these representations are by conducting two accent recognition tasks with isolated word stimuli of different frequencies: high-frequency, low-frequency, and non-words. These found that participants were not better at recognising accents in highly frequent words than in low-frequency words, and that they were able to recognise accents in non-words. The findings suggest that lexical representations do not seem to play a major role in the retrieval of the social (accent) information, and that sub-lexical representations are able to be used to access social meaning. There was a difference between accuracy in real words and non-words, however, which suggests that having a phonological representation of a word provides useful cues about what variable a socially relevant sublexical variant is a variant of (i.e. 'what sound is meant'). Furthermore, the finding in these experiments that participants were not able to distinguish as many different voices as were present in the tasks suggests that the amount of detail that the linguistic representations used in accent processing have is approximately enough to recognise character types (but not exact speakers).

These findings were generated from research which sought to answer the following project-wide research questions:

1. How are representations of linguistic variation and the social meaning this variation has connected in memory?
    a. Are they stored in separate but closely connected representations?
    b. Or are they stored within the same representations (e.g. episodic memories)?
2. Which linguistic representations are used in the retrieval or interpretation of the social meaning of linguistic variation?
    a. What size are these units?
    b. How detailed are these units?

In the individual chapters of the thesis, more detailed research questions are used for the implementation of the experiments themselves.

1.2 Outline

This thesis consists of eight chapters. After the introduction, I discuss the theoretical framework underpinning my research questions in Chapter 2. Chapters 3 and 4 then describe two socio-contextual priming experiments designed to investigate whether the social and linguistic information of language variation are stored separately or as a part of the same mental representations. Then, Chapters 5 and 6 describe two accent recognition experiments designed to investigate the size of the mental representations that carry the social meaning of linguistic variation. The discussion in Chapter 7 ties together the findings of these chapters in order to answer the two main research questions. The chapters are summarised below.

Chapter 2

In the theoretical framework, I argue that third-wave variation studies and early cognitive work on sociolinguistics from the perspective of exemplar theory could be seen as at odds with each other. Whilst third-wave variation studies find that the link between social meaning and linguistic variation is highly flexible, exemplar theory originally implied that social information was a part of the linguistic representations in memory storage. The innovations of hybridity in more recent exemplar theory however opens up the possibility of separate storage and flexible links, although this has not yet been explored thoroughly. This sets up the first research question: is social meaning flexibly linked to – and therefore separately stored from – linguistic information or are they stored together?

I go on to argue that, with the introduction of more abstract generalisations across exemplars in hybrid exemplar theory, new possibilities have opened up with respect to which linguistic representations are used to access social meaning in the theory – these may be highly detailed exemplars, or more abstract representations. I disentangle two parameters for these representations that are sometimes confounded: the size (e.g. words, segments, utterances) and the detail of those (e.g. from phonological form to highly detailed exemplar). On the basis of this I put the second main research question forward: which size, and how detailed, are the linguistic representations which are used to access social meaning?

Chapter 3

In the first experimental chapter, I describe an experiment which aimed to test whether it is possible to process the social meaning of linguistic variation (e.g. ‘American’) without activating the linguistic representation carrying that social meaning (e.g. the regional lexical variant ‘CANDY’) and vice versa. This would indicate separate representations in storage. The study used a variation on a semantic priming experiment, a ‘socio-contextual priming experiment’, to see if exposure to a first (prime) stimulus influenced responses to a second (target) stimulus – which would imply spreading or joint activation. When the prime stimulus was the regional lexical variant *y'all* this, under certain circumstances, resulted in inhibitory priming effects (slower response times) for British participants when they responded to target stimuli which were accent stereotypes of 'Americanness' (e.g. 'GUNS'). However, prime stimuli that were the names of certain accents (e.g. American) under other circumstances, resulted in facilitatory priming effects to target stimuli which were regional lexical variants (e.g. 'CANDY'). The difference in these priming effects suggests that social and linguistic information are linked in different ways, resulting in different types of associative processes depending on the direction of processing. This suggests a separation of social and linguistic information in memory storage, as it would be expected that if they are a part of the same representations, priming would always happen in the same way in the same direction.

Chapter 4

Chapter 4 describes a pair of follow-up studies, which streamlined the socio-contextual priming experiment of Chapter 3 and made a number of methodological improvements. In this chapter's socio-contextual experiments, the primes were auditory and contained an accent variant. The paired experiments differed in the duration of the inter-stimulus interval between prime and target but no priming effects were found in either experiment. These results again suggest that it may be possible to process and activate linguistic accent information without activating corresponding social meaning, and that the two are stored separately in memory.

Chapter 5

Chapter 5 describes an explicit accent recognition task, which aimed to uncover which linguistic representations carry social (in this case accent) information and, specifically, what size they are. It tested whether word-sized representations (such as the ones described in exemplar theory) were central to this by investigating whether participants were better at recognising accents in isolated high-frequency words than in isolated low-frequency words, all else being equal. This experiment found no difference between the two, suggesting that having a large number of lexical exemplars, or often activated word-sized representations at any level of linguistic detail, does not help recognition. At the same time the experiment found anecdotal evidence that participants distinguished fewer different voices than were used in the experiments, opening up the possibility that the representations used for recognition were somewhat less detailed than would be necessary to distinguish between different speakers.

Chapter 6

Chapter 6 describes a second accent recognition task which investigated whether participants were able to recognise accents in isolated non-word stimuli, in order to see if sub-lexical representations allowed participants to accurately assess the accent of these stimuli. This was found to be the case, providing evidence that sub-lexical representations can be used to access social meaning. At the same time, the performance in this task was considerably worse than in the word-based recognition task of Chapter 5. This suggests that some phonological information about a word is helpful to social meaning processing, because it allows for the contextualisation of sublexical cues in the input. The experiment also asked participants to self-report the number of different voices they thought they heard, which was again much lower than the real number of voices. This again suggests that the linguistic representations that are used to access social meaning may be less detailed than necessary to distinguish between different speakers and that it may be closer to the levels of detail required to distinguish between character types.

Chapter 7

The discussion in Chapter 7 combines findings from the previous four experimental chapters to answer the main two research questions. On the basis of the data from Chapters 3 and 4, it seems most likely that social information and linguistic information are stored separately, which is at odds with the idea that social meaning is carried by exemplars as united memory representations which include both the linguistic and the socio-contextual detail of an utterance. It is in line with third-wave variationist approaches to social meaning which argue that social meaning is flexibly linked to variation, and provides evidence within hybrid exemplar theory that social meaning is not necessarily stored within the exemplar. On the basis of the data from Chapters 5 and 6, it is most likely that the linguistic representations that are used to access social meaning are sub-lexical. The findings also suggest that that sub-lexical representations benefit from being embedded in phonological information about the word they occur in. The fact that participants tended to hear fewer speakers than were used in both the experiments suggests that the amount of detail included in the sub-lexical representations is approximately enough for participants to recognise character types, but not speakers.

2 Theoretical framework

2.1 Introduction

Whilst there is a long tradition of sociolinguistic research on the way linguistic variation can indicate a wide range of potential characteristics of and information about a speaker – including class (Campbell-Kibler 2005), gender (Strand 1999), sexuality (Levon 2007), regional origin (Clopper 2004), local group identity (Bucholtz 1999), cultural orientation (Vriesendorp & Rutten 2017), a persona (Podesva 2007), and interactional stance (Moore & Podesva 2009) – a new question has come to the fore in some recent sociolinguistic perception research: how is this information accessed cognitively?

Taking a cognitive approach to this question, many sociolinguistic perception researchers use a theory of speech processing called 'exemplar theory' (see Docherty & Foulkes 2014). In this theory a speech utterance is processed by matching a new input to detailed memories of other utterances. These detailed memories, or 'exemplars', contain information about the linguistic variation contained in the utterance as well as information about its speaker and its further context. This seems at odds with sociolinguistic approaches to social meaning which suggest that there is a fluid inference-based link between variation and its meaning (Eckert 2008). At the same time, more modern versions of exemplar models (Goldinger 2007; McLennan & Luce 2005) allow for speakers to generalise across these exemplars and form more abstract representations. This is to say they form representations which include less detail than an exemplar and therefore match a larger number of linguistic realisations (e.g. the phoneme /p/ matching both [p] and [p^h] in English]). However, here, which of these representations (exemplars or abstract generalisations) are used to access social meaning has not been extensively explored.

In the following chapter I describe the theoretical background to the research questions asked in Section 1.1 above. In Section 2.2, I describe sociolinguistic research on the way that the social meaning of variation is used and constructed in interaction, and how it suggests that social meaning and linguistic variation are stored separately and linked in practice. I also describe how in the exemplars that exemplar models describe, the two are implied to be united in one in the mental representations that listeners have of linguistic variation, and how in hybrid exemplar theory both separate storage and united storage would be possible in more abstract representations of linguistic variation. In Section 2.3, I explore the research which has been conducted to locate which linguistic representations are used in speech processing, with respect to both the size of these representations and the detail that these representations have, in order to explore which of them are used to access social meaning.

2.2 Connecting linguistic variation and social meaning

2.2.1 Introduction

Different types of social meaning

As briefly mentioned above, sociolinguistic research has found that linguistic variation can reflect and be used to construct a large range of different types of social information, from broad social demographic information to local interactional stances. One of the main broad demographic information types is regional origin. Plichta & Preston (2005), for example, find that American listeners are able to correctly classify realisations of the PRICE vowel in US accents on a continuum of locations from the North to the South of the US. Other examples of linguistic variation carrying broader social-demographic information are class (Campbell-Kibler 2005), gender (Strand 1999) and age (Walker & Hay 2011).

At a more local level, Bucholtz (1999) finds in a group of high-school 'nerd girls', linguistic variation is used to construct a locally relevant identity: here, girls were using hypercorrect forms and fully released /t/ to indicate membership of their group as a 'nerd'. Bucholtz (1999) also shows that this membership, and the speakers' identities, are constructed in interaction. In this way, linguistic variation carries the social meaning of locally constructed identities. Similarly, Vriesendorp & Rutten (2017) find similar construction of local identities and stances in an instant messaging corpus, where loose network of Dutch gay adolescents use codeswitching to English to align with 'gay celebratory' attitudes and identities.

Furthermore, research has found that linguistic variation can index personas (Podesva 2007) and character types (Starr 2021) – socially recognised figures, such as 'valley girls', 'cholos', 'cowboys', 'jocks', and 'burnouts' (Eckert 2008: 455–6). Podesva (2007) shows how one speaker uses different phonation types and pitch ranges to construct different personas of himself – ranging from this speaker's professional persona to his 'gay diva' persona.

When these figures are recognised beyond local communities, and understood macro-socially, Starr (2021: 249) argues that these become 'character types' rather than the personas of one person. An example of this is the California 'valley girl' stereotype. D'Onofrio (2016) finds in a range of perception experiments that back realisations of the TRAP vowel carry the meaning 'valley girl'. Such character types float somewhat between highly local social meanings (e.g. membership of a 'nerd' group) and broader macro-social groups (e.g. regional origin). D'Onofrio (2016) herself uses the term 'persona' to talk about these findings, but in this PhD I follow Starr's (2021: 249) categorisation in which more widely understood personas are called character types.

Defining social meaning

With this wide range of possible social meanings of linguistic variation, how exactly should social meaning be conceptualised theoretically? This is at the core of many third-wave variation studies (Eckert 2012; Levon 2017): what does it mean for a variant or a variety to have a 'social meaning'? In this section, I discuss how third-wave approaches and exemplar theory approaches differ in how they describe and define social meaning. In the third-wave approach I follow Hall-Lew, Moore & Podesva's (2021: 5) definition of social meaning as 'the set of inferences that can be drawn on the basis of how language is used in a specific interaction'. This implies that social meaning is part of an inferential process and is flexible, as it depends on its use in interaction.

At the same time, sociolinguistic research on exemplar theory argues that exemplar theory is able to explain the processing of socio-indexical meaning elegantly because exemplars are memory traces which contain social information as a part of the memory of the linguistic utterance (Docherty & Foulkes 2014; Foulkes & Docherty 2006). This suggests that there is no need for inference to reach a social meaning, and that this is automatically activated as a part of the relevant exemplar memories themselves. At the same time, hybrid exemplar innovations argue that listeners make generalisations across exemplars. If such generalisations happen separately for social and linguistic information across exemplars and listeners create representations of linguistic information and social information separately, this would allow for the indexical inferred links that third-wave approaches argue for.

The two subsections below aim to address both perspectives on this question: is social meaning stored in memory representations separate from linguistic information, connected through indexical links, or are they a part of the same memory representations?

2.2.2 Social meaning in the third wave: flexible and inferred

Mutability

In the 'Third Wave of Variation Studies' (Eckert 2012), social meaning is not seen as the static distributional information about linguistic variation across social categories but rather as the dynamic inferences that can be made about speakers and interactions on the basis of how this variation is used in a specific interaction (Hall-Lew, Moore & Podesva 2021: 5). These are shown in third wave variation studies to be agentively and dynamically used by speakers to position themselves in the social landscape (Eckert 2012).

Bucholtz's (1999) study mentioned in Subsection 2.2.1 is an example of this, showing how the group of nerd girls described agentively used super-standard and hypercorrect phonological and syntactic forms associated with intelligence to construct their in-group nerd girl identity. Similarly, Podesva (2007) describes a speaker using the expressive connotations of falsetto phonation, and higher and wider pitch ranges in specific contexts to construct a 'diva' persona.

As Eckert (2008) argues, the exact meaning of variation in practice depends on the interaction itself. She describes how the full release of the /t/ plosive can carry a multitude of social meanings in American English, including a gay diva persona and preciseness (Podesva 2004), nerdiness or intelligence (Bucholtz 1999), and being Orthodox Jewish (Benor 2001). However, these do not all arise at once. She argues that a linguistic variant has a range or 'field' of potential meanings, depending on the context it occurs in and how it is used in interaction. Pharao et al. (2014) find further evidence for this in a study which showed that [s+] is interpreted differently by Danish listeners depending on whether it occurs in 'modern Copenhagen speech' (where it is perceived as feminine or gay) or 'street language' (where it did not change perceptions that way).

Moore & Podesva (2009) discuss how the different social meanings of the use of tag questions at a northern English secondary school can be 'activated' in different contexts by different speakers. As they argue, the most immediate meaning of tag questions was conduciveness. However, different groups of speakers used them for different interactional purposes: some used them to conduce evaluative stances towards others, and to criticize behaviour, whilst others used them to conduce a shared viewpoint. Other groups then used them to construct authority or to conduce mutual involvement and friendships. Beyond these interactional stances, some groups used them to index local identity, such as belonging to their own 'cool' group. At an even more macro level, the use of the tag questions could be interpreted as feminine and working class. All of this, according to Moore & Podesva (2009) is mediated by whether the interpreter has ideological links between the variation in question and social meanings such as conduciveness, coolness, and femininity.

This fits with Agha's (2003, 2005) theoretical discussion of the concept of 'enregisterment'. This is the process in which a linguistic repertoire becomes socially recognised and differentiable. Johnstone (2009), discussing the enregisterment of 'Pittsburghese', argues that linguistic variation often goes unnoticed unless variants are linked with an ideological scheme. If two groups of people are expected to act differently from each other (e.g. men and women; Ochs 1992) differences in variant usage can come to become meaningful. These ideological ideas can then be shared both explicitly and implicitly, for example by overtly talking about the variation or by connecting the variation to social stereotypes. Cognitively this fits well with Sumner et al.'s (2014) study in which they found that for words with different pronunciation variants the infrequent, atypical but socially idealised variants were remembered better long term than the frequent typically occurring variants. They argue that the encoding of speech happens in a weighted way, so that rarer forms because of extra attention paid to them during processing can take a central place in memory. This would allow for ideology to influence (socio)linguistic processing through weighting at the cognitive level.

Inference

In third-wave approaches to social meaning, this dynamic conceptualisation has implications for how social information and linguistic information are linked to each other: if social meaning is contextual, can change, and needs ideological links to allow for social meaning to be interpreted, then this meaning must be separate from the variation it is connected to. This is evident from third-wave theories of how linguistic and social information are linked: these links are conceptualised as inferred (Acton 2021).

Hall-Lew et al.'s (2021: 5) definition of social meaning mentioned above puts inferences at the centre of social meaning: 'the set of *inferences* that can be drawn on the basis of how language is used in a specific interaction' (italicisation my own). This is developed in detail by Acton (2021), who argues that people tend to form links between two things co-occurring, so that perceiving one thing indexes another. For sociolinguistic variation this means that perceiving a linguistic variant that co-occurs often with speakers who have certain characteristics can lead a listener to infer that a speaker using this same variant is likely to share these same characteristics. Between the social and the linguistic information, therefore, there is a process of interpretation that is mediated by ideology and context. Acton (2021) argues that this inference can become conventionalised, but holds that for social meaning to arise at all this process of inferences is the first step.

As Acton (2021) argues, the process of inference allows for a number of aspects of social meaning that have been found in third wave research, and are discussed above: the purposeful agentive use of variation to position oneself in the social landscape (Bucholtz 1999; Podesva 2007), the underspecified nature of social meaning which depends on context to be fully realised (Eckert 2008), and its connection to ideology (Moore & Podesva 2009).

Indexicality

The way social meaning and linguistic variation are connected in the third wave is through 'indexicality' (Ochs 1992). As D'Onofrio (2021) argues, social meaning is part of a sociolinguistic sign within Peirce's (1991) semiotic system.

Peirce (1991) distinguishes between three types of signs: symbolic, iconic, and indexical signs. A symbol in this classification is a sign where the sign vehicle (the thing that means something) and object (what it means) are related to each other in a way that is completely arbitrary, and an icon is a sign where the two resemble each other. The sociolinguistic sign is an indexical sign: one where the sign vehicle 'points at' the object. The two are related in reality without needed in resemble each other. In the case of social meaning it is related to linguistic variation based on co-occurrence and relatedness in ideology. For example, some words may only be used by educated wine connoisseurs, and therefore only co-occur with speakers who are those connoisseurs, and language users may comment or hold beliefs about using those words. This way using those terms in interaction may come to indicate that someone who uses them is a member of that group (Silverstein 2003).

For this meaning to arise, it is not just necessary that the linguistic variation and social characteristics cooccur, but also that language users interpret the links between the two as meaningful and connected (D'Onofrio 2021: 123). This is the third element of the sign: what Peirce (1991) calls the interpretant. D'Onofrio (2021) argues that cognitively, the interpretant consists of the processes involved in linking form and meaning. However, as she argues, the interpretant can also become the cognitive representation that is the result of those processes (D'Onofrio 2021: 123). This is similar to Acton's (2021) argument that the inferred link between variation and social meaning can conventionalise. This opens up the possibility that there are mental representations where variation and social meaning are part of one unit, although this is secondary to the original interpretative processes.

In summary, third wave approaches to social meaning have a clear separation between representations of social meaning and representations of linguistic variation which are indexically linked to each other. This is summarised schematically in the Figure 1 below. It shows that variation in rhoticity and the openness of LOT vowels, as well as, for example, a certain realisation of the word *water*, form their own representations. These variants are then linked to a common social meaning, i.e. that a speaker is likely American, through indexical links (based on co-occurrence and ideology).

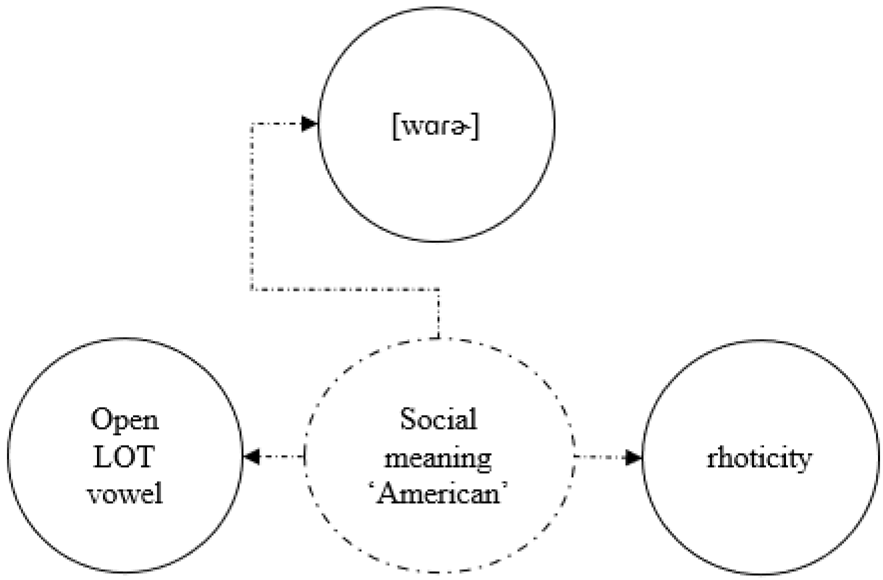

Figure 1. Schematic representation of social meaning and linguistic variation as separate cognitive units connected through indexical links.

Cognitive evidence for separate storage

Recent work by Campbell-Kibler (2021) provides evidence that social information and linguistic information may be stored separately, by testing three ways in which social meanings of masculinity and variation in /s/ are connected: the influence of speaker gender on /s/ production, the influence of /s/ variation on perceptions of masculinity, and the influence of masculinity cues on the perception boundaries of /s/ (on a continuum with /ʃ/). She found that there were no correlations between how strong the effects were of these influences within participants. This suggests that these three ways that social and linguistic information influence each other are not caused by one and the same link (or togetherness of the two within representations), but by multiple links between the two types of information.

2.2.3 'Pure' exemplar theory: social meaning as a part of the exemplar

Interestingly, exemplar theory, as another (relatively) recent account of the connection between language and social meaning, does not necessarily follow the third wave's conceptualisation of social meaning as flexibly linked. Here new linguistic input is matched to episodic memory traces of earlier input. In this theory listeners have memories containing detailed information about the entire 'episode' of an event in their memory storage: this is to say the meaning of the utterance, the realisation of the sounds, information about the speaker, the context, the location, the emotions the interactants had, et cetera.

This is in contrast with the hypothesis of 'speech normalisation' (Joos 1948 cited by Goldinger 1998: 252). This is the idea that any information that is not relevant to phoneme or word recognition is filtered out in the process of perception. The idea was often used in word recognition models (e.g. McClelland & Elman 1986) to explain how listeners are able to understand different and new speakers despite the wide variety in speakers' phonetic differences (such as voice quality, speech rate, accent, and pitch). When a number of studies from approximately the 1990s onwards provided evidence against the normalisation hypothesis this filtering process had to be reconsidered.

A number of studies found that phonemically irrelevant acoustic details are stored in the brain: Nygaard, Sommers & Pisoni (1994) found that listeners who hear a word twice are better at recognising it the second time if the word was pronounced by the same voice as the first time, than when the word was pronounced by a different voice. Furthermore, Church and Schacter (1994) found that acoustic details such as a speaker's voice, intonation and fundamental frequency affect how quickly listeners respond to a word they have already heard. In their experiment listeners performed a recognition task for words that they had already heard in an earlier phase of the experiment, and were found to be quicker to respond to those words if they were pronounced by the same voice or with the same intonation than when they were pronounced by a different voice or with different intonation. These sorts of results were found in a large number of studies in the 1990s and 2000s (Goldinger, Pisoni & Logan 1991; Goldinger 1996; Bradlow, Nygaard & Pisoni 1999; Goldinger 2007; McLennan & Luce 2005).

On the basis of these findings, Goldinger (1998) proposed an 'episodic lexicon': he argues that words in a speaker's lexicon consist of a large number of memory traces they have of hearing that word at earlier points in time. When someone listens to a word, the new trace of that word is matched with the existing memory traces in long-term memory, which are activated according to how similar they are. This creates an echo, consisting of all the similar, more or less, activated traces. This echo contains information not necessarily present in the sound input, on the basis of the information present in the echoed traces (such as the use and meaning of the word, but possibly also speaker information and other contextual information). This means that the stored exemplars are not abstracted to underlying phonological forms, and that the only abstraction present is the combined echo of (separate) activated memory traces – specifically of words. This version of exemplar theory is now often referred to as 'pure' exemplar theory, as it does not include any more abstract representations than exemplars to account for speech processing. In Subsection 2.2.4 below I discuss 'hybrid' innovations on pure exemplar theory which do include abstract generalisations across exemplars.

Social and linguistic information as a unit

As Foulkes & Docherty (2006) argue, the way (early) exemplar theory works means that the way social meaning is perceived is elegantly incorporated into speech perception more generally: when a listener processes new input it is matched to the other memory traces, not just on the basis of the linguistic information, but also on the basis of the social and contextual aspects of the input and the matched memory traces (Johnson 2006; Johnson 2007).

As is shown in Figure 2, in Foulkes & Docherty's (2006) and Docherty & Foulkes' (2014) exemplar model, words are stored as exemplars of previous utterances of that word, which contain acoustic detail (see levels A and B in the figure), as well as social information (level C). The clustering of exemplars based on this detail then produces knowledge of sociolinguistic variation. Crucially, this means that social information about the speaker of the utterance is an integral *part* of the linguistic representations used to process speech, meaning that there is no need for the creation of links between social meaning and linguistic variation.

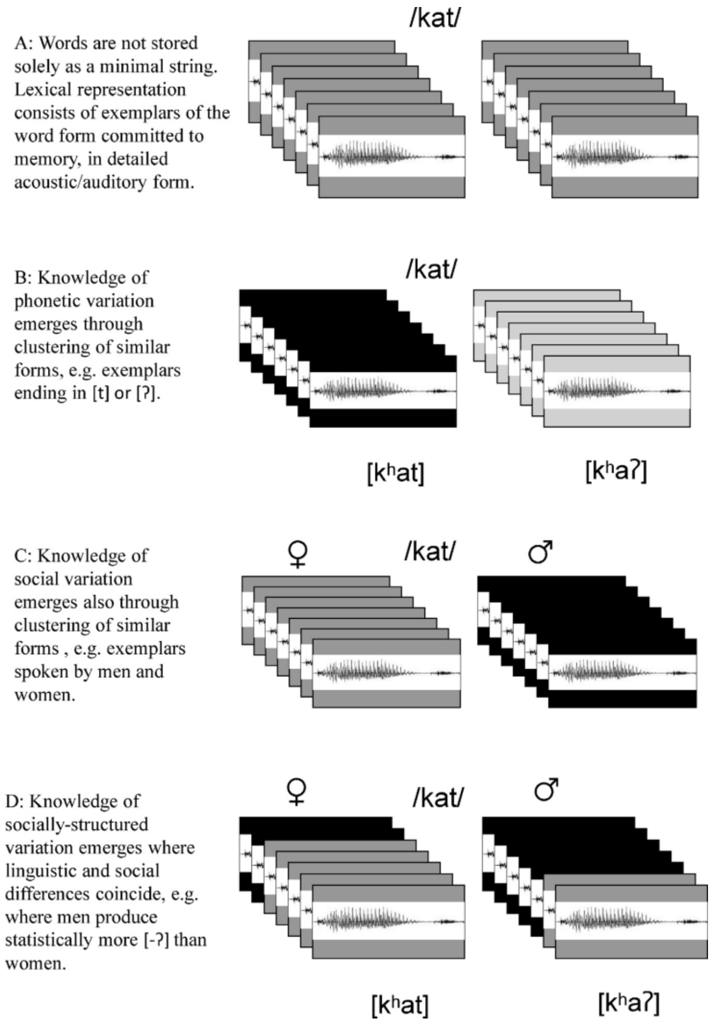

Figure 2. Schematic illustration of socio-phonetic variability within an exemplar-model (Docherty & Foulkes 2014: 44).[1]

Whilst work on exemplar theory in sociolinguistics often takes macro-social categories as an example (e.g. the gender example in Figure 2 above, but also age in Walker & Hay (2011)), exemplar theory would, in its own way, also be able to account for the contextual nature and mutability of social meaning identified in third wave research: in an exemplar model, new input is not just matched with linguistic memories, but entire episodic memories including their social context. This means the flexibility of different social meanings associated with linguistic variants in different contexts would not be underspecified, but rather hyperspecified: exemplars would contain so much social, contextual, and linguistic detail that new input can be matched with exemplars which do not just share a linguistic variant, but also enough social, contextual, and other linguistic information. This way a highly specific, contextual meaning can be retrieved.

For example, new input of a man at a party using fully realised /t/ (Podesva 2004) would match other memories of men using fully realised /t/ (perhaps also in party or other casual settings), and certainly more than it would memories of school girls using fully realised /t/ (Bucholtz 1999). This way, it follows elegantly that fully realised /t/ in the context of a party in Podesva's (2004) study is understood as gay diva-esque, rather than as part of a nerd girl identity, even if it can carry that meaning in a different

[1] Reprinted with permission from Elsevier.

setting. This way, the exemplar approach can be extended to more localised, complex social meanings, such as local identity construction, stances, and persona and character type meanings.

This scenario also allows for the possibility that inference does play a role in the exemplar theoretical approach to social meaning. If there are multiple possible meanings of /t/ release, and multiple cues (linguistic and non-linguistic) pointing towards different social meanings, there is room for uncertainty. As Kleinschmidt, Weatherholtz & Jaeger (2018) argue, a listener would then have to infer which of the possibilities is the most likely (even within an exemplar approach). However, at the level of the exemplar itself, social and linguistic information would still be jointly stored together.

In summary, early exemplar theory approaches to social meaning suggest that social and linguistic information are united in mental representations, more specifically in the exemplar, which is an episodic memory containing both linguistic and social information. These representations would still be connected as being a part of the same variety, and be clustered together (Docherty & Foulkes 2014), potentially with indexical links (the co-occurrence of certain forms together), as well as iconic links (with the same linguistic variants, and social characteristics connecting different exemplars).

Support for united storage of social and linguistic information

Cognitively oriented sociolinguistic research provides further evidence that social information is encoded within the same representation as linguistic detail. One key paper in this argument is by Walker & Hay (2011), who find that lexical access is easier when socio-phonetic detail and the social distribution of lexemes match. In their study, they found that listeners are quicker to recognise words which are more often used by older people than by younger people, when those words are pronounced in an 'old' sounding voice. This can be accounted for in pure exemplar approaches by positing that listeners have more and stronger episodic memories of older sounding people pronouncing those 'old' words than younger sounding people. This then makes accessing these easier. In addition to Walker & Hay's (2011) findings for New Zealand English, Kim (2016) finds similar findings in Korean.

In addition to this, some perception experiments have found that listener perceptions of linguistic variation can be primed with social cues, without listeners realising this. Niedzielski (1999) finds that listeners hear more Canadian raising (a raised and fronted MOUTH vowel) in the same speaker if they had been told the speaker was Canadian as opposed to being from Detroit, despite them listening to the same recording. In exemplar theory, this can be accounted for as the result of the social information that a speaker was from Canada activating exemplars of Canadians pronouncing words with Canadian raising, making it more likely that someone perceived new input as similar to those activated exemplars.

Similarly, Hay, Nolan & Drager's (2006) study of perception of the KIT vowel in a New Zealand English speaker found that listeners hear more Australian English (higher and fronter) vowels for the same recordings, with the same acoustic input, if they were given an answer sheet which had the word "Australian" written on it, than if they were given one with "New Zealander" written on it. This was the case even though nearly all participants were aware they were listening to a New Zealander. Again, in an exemplar approach this can be accounted for as the result of social cues activating certain exemplars because of their matching social information, skewing perceptions.

2.2.4 Hybrid exemplar theory: joint or separate storage?

Modern iterations of exemplar theory have introduced nuancing innovations to the idea of exemplars in processing. Following criticism of the lack of abstract representations in pure exemplar theory, more recent, so-called 'hybrid models' have introduced the notion that listeners generalise over the exemplars they have available to them, and create more abstract representations on the basis of those (Pierrehumbert 2002; McLennan, Luce & Charles-Luce 2003; Goldinger 2007). This may also allow for generalisations across the social information included in an exemplar. These generalisations may either be structured in the same way as exemplars themselves (uniting social and linguistic information) or create separately stored social and linguistic representations. The latter would align with the flexible links between social and linguistic information in third-wave sociolinguistic approaches.

Criticism of 'pure' exemplar theory

Goldinger's (1998) episodic lexicon faced strong criticism because it did not include any abstract representations in accounting for speech processing. Firstly, Cutler (2008) argued that abstractions play an undeniable role in speech recognition, both at a prelexical and at a lexical level. At the phoneme level it has been found that listeners who hear an idiosyncratic pronunciation of a certain segment for a speaker (e.g. /ʃ/ for /s/) generalise this to all instances of that segment, and expect to hear that idiosyncrasy even in words that they have not heard that speaker pronounce before and cannot have exemplars of (see for example Norris, McQueen & Cutler 2003; Kraljic & Samuel 2006; Eisner & McQueen 2006). Similar evidence for language users' ability to create abstract generalisations was found by German, Carlson & Pierrehumbert (2013) who taught American students to learn a Glaswegian accent - more specifically to use [tʰ] for /t/ in word-medial positions and to use [ɾ] for /r/. Learning occurred after being exposed to only 24 examples and was generalised over words outside the training set. This evidence taken together suggests that listeners use abstract, generalised representations of linguistic information smaller than words (sub-lexical units such as allophones) in speech processing.

Another argument made by Cutler (2008) is that speech recognition is influenced by factors outside episodic memory of speech sounds. Weber and Cutler (2004) and Escudero, Hayes-Harb & Mitterer (2008) show that the way words are spelled influences how they are perceived. And, as Cutler (2008) notes, language users often know words which they have learnt from reading but have never heard, and yet they do have an idea of what they sound like. This connection between orthography and pronunciation implies that there must be some abstract lexical and sub-lexical representations stored in the brain. This would be how language users have an idea what <l> and <larpst> sound like, without having encountered them as spoken words.

Ernestus (2014) provides support for both the presence of episodic traces and of underlying abstract representations of words, from the perspective of acoustic reduction. Acoustically reduced word forms in connected speech are a very common phenomenon in spoken language, and exemplar theory can elegantly explain processing these by positing that they have their own episodic traces stored in the lexicon. This way people are able to easily recognise those forms even when they are very different from their unreduced variants (for example [jeʃei] for *yesterday*¸ or Dutch [ɛik] for *eigenlijk* ('actually')). This gives exemplar theory an advantage over abstractionist models where those reduced forms would be mapped onto the full phonological forms of those words as a variant ([ɛik] is much closer to many other Dutch words than to the full form [ɛixənlək], including the word *eik* [ɛik] ('oak'), but can still be recognised as *eigenlijk*).

However, there are also some problems for exemplar models in findings about reduced word processing, as Ernestus (2014) points out: highly reduced forms are often reconstructed in perception to the extent that people perceive sounds that were not actually there: Dutch listeners who heard a highly reduced version of the suffix /-lək/, i.e. [-k], still reported hearing a [l] (Kemps et al. 2004). This implies that a more unreduced phonological form takes precedence over more specific, detailed episodic traces (although orthography might also play a role). Furthermore, Ernestus, Baayen & Schreuder (2002) have found that the recognition of highly reduced word forms is very low in isolation (only 52%), which seems to indicate that these forms are not stored lexically, even though exemplar theory would require that they would be, just as any other perceived word form.

Taking this evidence together, it has become clear that exemplar theory can account for a number of phenomena that a speech normalisation account cannot explain (Church & Schacter 1994; Ranbom & Connine 2007; Ernestus 2014), but also that it needs to incorporate at least some more abstract representations into its model of speech processing to account for phenomena like the restoration effect

(Kemps et al. 2004) and the generalisability of allophonic patterns (German, Carlson & Pierrehumbert 2013).

Hybrid models

Pure exemplar theory is often used to elucidate the potential problems of episodic models of language processing (e.g. Cutler 2008; Ernestus 2014; Pierrehumbert 2016). However, this representation of exemplar theory has become somewhat of a strawman: most if not all more recent proponents of the episodic approach argue for hybrid models of speech perception in which both abstractions and episodic traces play a role (see McLennan, Luce & Charles-Luce 2003; Hawkins 2003; Hay, Pierrehumbert & Beckman 2004; Johnson 2006; Goldinger 2007; Wedel 2012; German, Carlson & Pierrehumbert 2013; Pierrehumbert 2016). Even Goldinger (1998: 265) points out that it would be possible for exemplar models to incorporate more abstract representations.

Complementary Learning Systems

Goldinger (2007) proposes a hybrid model, called the Complementary Learning System. In this system, acoustic input is first processed by a stable long-term unit (in the 'cortical complex') which allows for prior knowledge (e.g. prior contextual factors or knowledge about how to unitise the heard soundwave into phonetic units) to help processing. It is then combined with new contextual information and voice specificity in episodic traces and combined into a fast learning unit (in the 'hippocampal complex') which allows for rapid memorisation of specific events. The stable cortical complex then is able to slowly learn statistical regularities from the hippocampal cortex (creating more abstract knowledge).

For example, when a listener encounter a particular speaker who uses [ʃ] for /s/, they need to know in which words /s/ occurs, to be able to learn this pattern. This prior knowledge allows for [ʃ] occurring in such positions to be interpreted as an /s/, which is then fed into the fast learning hippocampal complex which is able to rapidly determine that all recent instances of /s/ have been [ʃ] (instead of being overruled by the stable knowledge that [ʃ] means /ʃ/). If then this keeps occurring, the stable network will slowly establish that this speaker's /s/ is always realised as [ʃ].

Adaptive Resonance Theory

McLennan, Luce & Charles-Luce's (2003) model, based on Grossberg and Stone's (1986) Adaptive Resonance Theory, goes into more detail about the types of linguistic representations that may be used in processing in a hybrid exemplar model. In the model, acoustic-phonetic input causes activation not with pre-set levels of linguistic units (phonetic features, phonological forms, and lexemes), but with more fluid 'chunks': these can be any sound representations from very specific, episodic memories of words to maximally abstract phonological representations. Importantly they can also be representations at more intermediate levels of abstractness, such as allophones. If those chunks are more strongly represented in the brain (for example they are more frequent) they are activated more quickly than weaker representations. This means that abstract phonological forms have higher resonance and are processed quicker than more specific (and therefore less frequent) representations.

McLennan, Luce, Charles-Luce (2003) find differences in what 'chunks' get activated first in speech processing, in a series of repetition priming experiments – experiments in which it is tested whether exposure to one stimulus influences participants' response to a second stimulus. When listeners in their experiment heard allophonically specific input (e.g. AmE [æɾəm], *atom*) before hearing allophonically different but phonemically the same input (e.g. [ætəm]), this helped them recognise the second stimulus faster. However, faster recognition only occurred if the experimental conditions made sure processing was slow, ambiguous, or effortful. This implies that the underlying, more abstract phoneme representation /t/ was activated in slow and effortful processing (for example in a difficult lexical decision task), influencing response times then, but that it was not activated in quick processing (in an easy lexical decision task).

In McLennan, Luce, Charles-Luce's (2003) model this can be explained by the high frequency of the allophone [ɾ] for /t/ in American English, which would lead to quicker resonance of that chunk than the underlying form (as [ætəm] as the form closest to /ætəm/ is less frequent than [æɾəm] in American English). A nuance to this is that the processing of the fully abstract phonological form /t/ is still more frequent than the also very frequent allophonically specific form [ɾ], as even the allophonic form [ɾ] would activate the phonemic form /r/. This suggests that not only frequency but also the amount of detail shared between phonetic input and representations in memory plays a role.

In addition to this McLennan and Luce's (2005) find that priming effects of acoustic specificity (i.e. effects of using for example the same voice and its acoustic specifics twice) only occur if processing is delayed or effortful. Matching talker identity or speech rate of both the prime and the target words only increased priming effects in a shadowing task (repeating a word that has just been played) if the participants' reactions were delayed (when they had to wait for a cue before shadowing instead of shadowing immediately), or if processing was slower (when lexical decision tasks were more difficult). This still fits their model well: exemplars of one specific voice are of course less frequently processed than their more abstract representations and would gain resonance even later. Taken together these studies suggest that the allophonic form [ɾ] is the most immediate in processing, then the underlying form /t/ and then further voice detail.

The structure of abstract representations in exemplar models

Because hybrid exemplar theory is open to the use of all different types of linguistic representations from abstract to specific and large to small, it is not straightforward to disprove or prove it as a theory as a whole. However, in this thesis it is used as a framework within which to investigate which possible representations are used most importantly and centrally in the processing of social meaning, rather than testing the theory as a whole.

For example, hybrid exemplar theory opens up new possible ways that social meaning could either be stored jointly with or separately from linguistic representations. The abstractions that listeners create by generalising across exemplars may either incorporate social and linguistic information within the same representations, or they may be stored separately as socio-contextual generalisations and as linguistic generalisations.

If the former is the case, listeners would have a representations of linguistic variants as spoken by certain speakers (or groups, or by speakers with a certain stance, or emotions et cetera). A listener might then have a representation of 'Speaker A's rhoticity' for example. This is more abstract than an exemplar, which would be an instance of Speaker A's rhoticity (i.e. in a specific word at a specific moment), but still structured in the same way as an exemplar.

More abstractly, they would have representations of 'Americans using rhoticity', et cetera, where the linguistic and the social are part of the same memory unit. Such representations would be connected to each other through indexical and iconic networks (as per Peirce 1991): indexical through the co-occurrence of linguistic variation in certain speakers or speaker groups, and iconic through the resemblance of the social information across these representations. This is summarised in Figure 3 below.

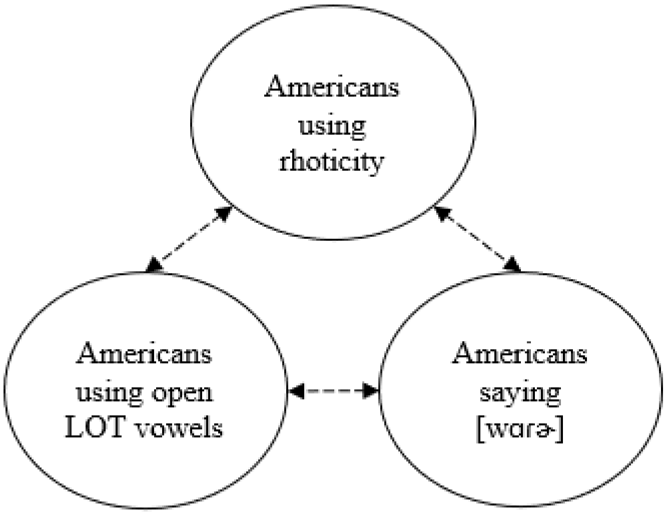

Figure 3. Abstract representations of social meaning and linguistic information where social and linguistic information are united in the same representations, whilst being less detailed than exemplars. The representations are linked through iconicity and indexicality.

However, it is also possible that listeners create generalisations over different aspects of episodic memories. This way more abstract representations of different types of information would be stored separately in different representations. For example, linguistically they might establish a representation for a dialect variant, like rhoticity, or open LOT vowels, across the exemplars where they have heard speakers use rhoticity or those open LOT vowels. Separately, they may generalise across episodic memories of one speaker, or a group of similar speakers, to establish a social meaning representation, such as 'Speaker A's accent' or 'American' (but also social class meanings, femininity, local identities, stances, and character typed for example). These generalisations then do not necessarily include linguistic information, but are closely linked to them due to ideology and co-occurrence, creating the sociolinguistic sign (D'Onofrio 2021). This would fit the third-wave variationist conceptualisation of social meaning and linguistic information being separate and flexibly linked.

In this case, the evidence for the co-presence of social and linguistic information in the same mental representations used for linguistic processing would have to be reconceptualised within sociolinguistic work on exemplar theory. The finding that it is easier to recognise words which are more typically used by old people when they are pronounced with older sounding voices (Walker & Hay 2011; Kim 2016) may then be interpreted as being due to very close indexical connections between the social meaning of those words (i.e. old age) and the socio-phonetic detail they are pronounced in. In this case, the socio-phonetic detail of the old sounding stimuli activates 'old words' because both are very closely connected to the more general social meaning of 'oldness'. Similarly the findings that listeners are more likely to hear Canadian raising when they are told they are listening to a Canadian (Niedzielski 1999) can be explained as the social cue of knowing that someone is Canadian activating this social meaning which is closely related to Canadian English forms, such as Canadian raising, making it more likely that such a representation becomes active as well.

In summary, third-wave variationist approaches to the storage of social meaning suggest social meaning and linguistic information are stored separately in memory, with flexible and fluid links between the two. Hybrid exemplar approaches to the processing of social meaning leave two options open: it is both possible that the two are stored as a unit within exemplar memories, or more abstract units with the same uniting structure, and that the generalisations listeners create on the basis of exemplars are separate for social and for linguistic aspects of the exemplars, allowing for flexible links.

In Chapter 3 and 4 I investigate these two possible conceptualisations further, by investigating if it is possible to activate one aspect (linguistic variation) without the other (social meaning), by means of a methodology which can find highly automatic activation processes: a variation on semantic priming experiments. This tested whether perceiving linguistic variants of a variety (e.g. the regional lexical variant *y'all* associated with American Englishes to British listeners) implied the automatic activation of related stereotypes (e.g. GUNS) or lexical variants of the same variety (e.g. SIDEWALK).

2.3 The linguistic representations used to access social meaning

2.3.1 Introduction

The second main research question addressed in this thesis' investigation of how social meaning is stored cognitively is the question of which mental representations are used to access social meaning. This question is tied into this thesis's first main research question of joint and separate storage: if social meaning is accessed by the activation of exemplars which contain both social and linguistic detail, the possible range of linguistic representations providing access to social meaning would be limited to theses exemplars (often hypothesised to be lexical; see Johnson 2006). However, if this is not the case, there is a much wider range of possibilities. This is explored below (and in Chapters 5 and 6).

In many third-wave variation studies, social meaning is talked about as being attached to relatively abstract features across instantiations (e.g. allophones). For example, when Eckert (2008) talks about the indexical field of possible social meanings of linguistic variation, she uses the full realisation of /t/ as an example of linguistic information carrying social meaning. This is a pattern of similar realisations of a phoneme, rather than each of its more narrow phonetic instantiations. In socio-phonetic research, the linguistic variation that is looked at is often that of the allophonic realisation (e.g. realisations of (-ing) (Campbell-Kibler 2005), rhoticisation of word-final vowels (Zhang 2008), and backed TRAP vowels (D'Onofrio 2015)).

Perhaps somewhat at odds with this is some of the sociolinguistic research engaging with exemplar theory. Here researchers are arguing for a model in which socio-phonetic variation and its social meaning are stored in word-sized representations, specifically lexical exemplars (Walker & Hay 2011; Hay 2018). In those word-sized representations multiple features of a variety or style would be united in the same representation, for example multiple allophones, as well as pitch and prosody, together carrying their social meaning.[2]

These differing approaches to which features of a variety or style are used to access social meaning invite the question: what do the cognitive representations which are used to access social meaning look like? Are they the size of words or sounds or something else? Do they contain as much linguistic and socio-contextual detail as exemplars, or can they be more abstract?

In order to investigate this further, I use the theoretical framework of hybrid exemplar theory. As described above, this theory posits that listeners do not just use highly detailed lexical exemplars to process speech but may also generalise across these to establish more abstract mental representations (see Subsection 2.2.4). This means that social meaning would not necessarily be carried within episodic memory traces but may also be attached to or be a part of more abstract representations such as phonetic variants.

However, much is still unknown about how the connection of social meaning and linguistic detail would be connected in a hybrid exemplar model. Docherty and Foulkes (2014: 53) briefly mention the need for hybrid models to be refined further (to understand how they are formed, how they evolve and how

[2] Still, not each singular word would include all phonetic aspects of a style, and certainly not (morpho-) syntactic aspects, so multiple aspects of a variety or style would still combine across multiple cognitive representations.

different representations weigh up against each other), but not necessarily to consider the place and role of socio-indexical variation in a hybrid model. This is one of the main research questions in this PhD thesis: what representations allow for the processing of social meaning? Specifically I investigate what size they are (e.g. words, utterances, syllables, morae or segments) and how specific they are (i.e. how much linguistic and social detail is included).

McLennan, Luce, Charles-Luce's (2003) model is particularly suitable to incorporate socio-indexical meaning into the model as its 'chunks' go beyond the dichotomy of episodic word memory and phonological representations. In particular, the inclusion and study of the processing of allophones in McLennan, Luce, Charles-Luce (2003) is helpful, as the storage of accent variation might be at a similar level of abstractness (see sociolinguistic studies focussed on allophones such as Campbell-Kibler 2005; Pharao et al. 2014; D'Onofrio 2015). Most priming experiments that have found evidence for the storage of acoustic specificity use voices as the variable of specificity (cf. Schacter and Church 1992, Church and Schacter 1994, Goldinger 1996, McLennan and Luce 2005), but socio-phonetic variation is often at a different level of specificity: it is not always as specific as only one speaker's exemplars, but may include the exemplars of a number of different speakers sharing some social characteristics or interactions with similar social context. In a hybrid model it would be expected that this could well involve abstractions similar to the allophones discussed in McLennan, Luce, Charles-Luce (2003).

2.3.2 Disentangling the parameters of representation size and detail

I follow McLennan, Luce, Charles-Luce's (2003) notion that sociolinguistic processing may involve the activation of 'chunks' of different sizes and different levels of detail, rather than strictly defined types of representations such as abstract phonemes or fully detailed lexical exemplars. This makes it possible to disentangle two parameters in the dichotomy between abstract phonemes and detailed lexical exemplars. As even the difference between phonemes and allophones (both segment-sized, but at different levels of abstraction and detail) shows, these do not always coincide. I propose to disentangle the following two parameters:

1. The **detail** parameter (ranging from the acoustic signal to the underlying phonemic form)
2. The **size** parameter (ranging from segments to utterances)

The detail parameter consists of the amount of (socio)phonetic detail included in a mental representation: this can be very high for very specific mental representations, including for example voice quality, intonation and pitch (Church & Schacter 1994). This can also be very low in the case of phonemes or phonological forms of words (Cutler 2008), and it can be at different levels between these extremes, in the case of, for example, regional allophones (German, Carlson & Pierrehumbert 2013).

The size parameter then consists of the size of these representations. This includes of course the segment and the word, but they can also be larger than words (see Bybee 2002). Walsh et al.'s (2010) multilevel exemplar theory model demonstrates the possibility of having differently sized units in an exemplar model: in their model of exemplar perception and production they distinguish between a 'unit exemplar database' and a 'constituent exemplar database'. In this model, input is both processed as a unit, and parsed into the constituents of the unit. For example, when a listener processes a syllable this activates exemplars of the syllable in the unit exemplar database, and activates exemplars of the parsed individual constituents of the syllable, phonemes, in the constituent database. Depending on whether the activation of the larger-size unit exemplar reaches a threshold, either the unit is perceived or the constituents are combined and perceived. This model was effective for linguistic data of different sizes (not just the example of the syllables, but also syntactic data with constructions as the unit and words as the constituents). This can be extended to mean that exemplar theory can work with differently sized linguistic representations, and not just word-sized ones.

I have cross-tabulated the size and detail of memory representations to provide an overview of potential memory representations in a hybrid exemplar model in Table 1 below.

Table 1. Schematic representation of intersections of the detail parameter and size parameter of specificity and abstraction.

Size → **Detail↓**	**Utterance**	**Word**	**Syllable**	**Segment**
Very low (underlying form)	/teɪk ðə məʊtəweɪ/	/məʊtəweɪ/	/weɪ/	/eɪ/
Potential social meanings	*British English - quaint - posh - direct*	*British English – neutral – quaint – posh*		
Low (widest sociolinguistic category)	[teːk ðə moːʔəweː] (and similar realisations)	[moːʔəweː] (and similar realisations)	[weː] (and similar realisations)	[eː] (and similar realisations)
Potential social meanings	*North of England – friendly – authentic – direct*	*North of England – friendly – authentic*	*North of England – friendly – authentic*	
Some (medium-width sociolinguistic category)	[tɪːk ðə moːʔtəwɪː] (and very similar realisations)	[moːʔtəwɪː] (and very similar realisations)	[wɪː] (and very similar realisations)	[ɪː] (and very similar realisations)
Potential social meanings	*North-East of England – funny – friendly – authentic – direct*	*North-East of England – funny – friendly – authentic*	*North-East of England – funny – friendly – authentic*	
Character type (narrow sociolinguistic category)	[tɪək ðə mʊəʔtəwɪə] + accent wide articulatory setting	[mʊəʔtəwɪə] + accent wide articulatory setting	[wɪə] + accent wide articulatory setting	[ɪə] + accent wide articulatory setting
Potential social meanings	*Very strong Mackem accent – authentic – friendly – working class – masculine – direct*	*Very strong Mackem accent – authentic – friendly – working class – masculine*	*Mackem lad – authentic – friendly – working class – masculine*	
High (speaker)	[tɪəʔk ə mʊəʔtəwɪə] + voice quality (abstracted)	[mʊəʔtəwɪə] + voice quality (abstracted)	[wɪə] + voice quality (abstracted)	[ɪə] + voice quality (abstracted)
Potential social meanings	*George – Mackem – neighbour – old – grumpy – working class – masculine – direct*	*George – Mackem – neighbour – old – grumpy – working class – masculine*		
Very high (speaker- specific persona)	[tɪəʔk ə mʊəʔtəwɪə] + voice quality (somewhat abstracted)	[mʊəʔtəwɪə] + voice quality (somewhat abstracted)	[wɪə] + voice quality (somewhat abstracted)	[ɪə] + voice quality (somewhat abstracted)
Potential social meanings	*George – Mackem – neighbour – old – working class – masculine – when he's relaxed and friendly*			
All that is processed (exemplars)	[tɪəʔk ə moəʔtəwɪə] + voice quality + speech rate + pitch + volume maximally (but not fully) specific	[mʊəʔtəwɪə] + voice quality + speech rate + pitch + volume maximally (but not fully) specific	[wɪə] + voice quality + speech rate + pitch + volume maximally (but not fully) specific	[ɪə] + voice quality + speech rate + pitch + volume maximally (but not fully) specific
Potential social meanings	*George, my old, grumpy, masculine, working-class neighbour from Sunderland, when he told me to take the motorway as we're having a relaxed friendly cup of tea in his nice and quiet Sunderland home, with a low pitch and relatively slow speech rate and average volume.*			
All (acoustic signal)				
Potential social meanings				

In this table, an example is given of what types of mental representations in memory a listener could *conceivably* have stored from their old neighbour George telling them to 'take the motorway' with his strong Mackem (Sunderland English) accent, with different levels of detail, and of different sizes. Below I discuss the cross-tabulated linguistic mental representations and what social meaning they would be able to carry. Where possible I provide research to support what representations have been found already, however, this is not always the case, and the table should be taken more as an exploration of the conceivable, possible representations involved in sociolinguistic processing.

The first parameter, detail, ranges from maximally detailed phonetic input, the acoustic signal, to abstract representations with as little linguistic detail as possible, such as underlying word forms, syllables, or phonemes (Cutler 2008), but conceivably also larger utterances (for example, listeners would need to have some abstract knowledge of the phrase 'it's raining cats and dogs' beyond its word forms to be able to process its meaning correctly). The second parameter, size, conceivably ranges from utterance to segmental (or even subsegmental). As I argue below, at all levels but the acoustic signal, it is conceivable for listeners to have mental representations of all different sizes stored in memory, and to use them in the recognition of social meaning.

The acoustic signal

At the most detailed, most specific level at the bottom of Table 1 is the acoustic signal. At this point no social meaning is attached and it is not possible to segment yet, as this is the signal before it is processed by the listener. The spectrogram shown in the table is that of someone saying 'take the motorway' in a Mackem accent, but the cut-off points are in a sense arbitrary, as acoustic input itself in the world does not come in separate chunks, but rather as a continuous stream of sound.

Segmentation at the exemplar level

At a slightly more abstract, and slightly less detailed level is the acoustic input as it is processed and stored in maximal detail in the moment. This is the level of the exemplar. As Pierrehumbert (2001) argues, the acoustic signal and the exemplar are not identical. The exemplar is a memory trace stored in memory (Goldinger 1998; Johnson 2006), which implies it has had to have been processed by the human ear and brain. This can be influenced by a range of factors such as established linguistic representations (Goldinger 1998), social cues (Drager 2010) and contextual factors (Kraljic, Brennan & Samuel 2008). This means that even exemplars (of any size) are to some degree abstractions of the acoustic input, losing some of the detail of the acoustic signal.

The exemplar level is also the first level where the utterance may be segmented into smaller parts. In many exemplar models this segmentation is focussed on the word (Goldinger 1998; Johnson 2006; Hay 2018). As a practical choice Johnson (2006: 492) picks the word as the unit of the exemplar in his model, as he argues it is the smallest linguistic unit which still carries meaning. However, this is a somewhat arbitrary distinction: smaller linguistic units like morphemes can carry denotational meaning by themselves outside of existing word exemplars (like 'un-' in 'unbreak my heart') – this is particularly true in more agglutinating languages than English. More importantly, it is not clear why carrying referential meaning would be a decisive criterion for exemplar segmentation.

In the discussion of hybrid exemplar theory (e.g. Pierrehumbert 2016), it is often implied that sub-lexical units must require abstraction beyond the lexical exemplar, on the basis of the argument that they can only be identified by detecting patterns across multiple utterances (e.g. finding [p^h] as a pattern in *pea, pod, ping, pan, pen*, etc.) and generalising across these. However, this is equally true for words, which would also need to be isolated from longer utterances in most cases. Similarly, utterances need to be segmented as one utterance in a continuous flow of auditory input, as separate from earlier or later utterances, to be processed properly, or other sounds, or even silence, to be processed as language. If segmentation is possible at the exemplar level for words then, it must also be possible for segments at the exemplar level.

The level of the exemplar is also where the first connections to social meaning are possible. The new exemplar of your old, grumpy, masculine, working-class, Mackem neighbour George saying [tɪəʔk ə moəʔtəwɪə] with the specific intonation, pitch, voice quality, speech rate, volume, etc. when you're having a relaxed and friendly cup of tea with him, means that the social information of that interaction (the type of person George is, the type of conversation you are having, et cetera) is joined, in the memory, with the linguistic information (e.g. the voice quality, the intonation, the realisation of the FACE vowel, the glottal reinforcement, et cetera). This is how exemplar theory elegantly includes social meaning into speech processing.

Speaker generalisations

Slightly more abstract than the exemplar would be a more generalised idea of how a specific person sounds, either in a specific context (what neighbour George sounds like when he's happy or when he is screaming or whispering, or for example a persona he adopts, et cetera) or more generally (what George sounds like). Evidence for some sort of speaker-level abstractions is found by Lavan, Knight & McGettigan (2019) who find that listeners are better at recognising a voice, when it is an average of what the voice sounded like in earlier instances, than they are at recognising those instances themselves. This was true even though they had never heard the average before. At this level of abstraction, not only is the linguistic detail abstracted, but also the social meaning: the social meaning of this more abstract idea of what someone sounds like does not include specific contextual detail anymore. The meaning becomes 'what my old, grumpy etc. neighbour George sounds like when he is being friendly' or at a slightly higher level of abstraction, 'what my old, grumpy, masculine etc. neighbour George sounds like', rather than 'what my old etc. neighbour George sounded like in this specific instance'.

Furthermore, work by Cai et al. (2017) makes an argument that it is speaker perception that is central to sociolinguistic processing, rather than episodic memories. They found that words with different meanings in American English and British English (e.g. *bonnet*) were semantically interpreted based on the accent they were pronounced in. However, they did not find an effect of how recognisable the accents were in these words: stimuli which participants found less recognisable as American or British were still semantically interpreted on the basis of the accent they were pronounced in. This suggests that the level of socio-phonetic detail in the presented forms did not play much of a role. Instead, Cai et al. (2017) propose that listeners identified speakers rapidly and used this information to determine for example whether that speaker would use the word *bonnet* to mean a hat or the part of a car. Their visualisation is shown in Figure 4 on the next page.

Whilst Cai et al. (2017) use this as an argument against exemplar based models, it is possible to combine the two approaches: in a hybrid exemplar model, it would be possible to say that speaker level abstractions are much more dominant than exemplars, which would mean that effects of speaker-level abstractions, such as the ones found here, eclipse effects at other levels of specificity (whilst still allowing for both to potentially exist and be used).

Generalisations across speakers

The next level of abstraction is to generalise across different speakers which reduces detail further, removing speaker specific voice-quality (although not all voice quality if there are articulatory setting patterns across accents (Wieling & Tiede 2017)). This higher level of abstraction then reaches, in this example, the more abstract concepts of character types, regional identities and associations, although they could also be more general social identities like gender, race, or more local group memberships (Eckert 2012), or social attitudes, stances, or styles (Moore & Podesva 2009). In the example in Table 1 the realisation [mʊəʔtəwɪə], with a certain articulatory setting, might carry the meaning of a strong Mackem accent for those who are very familiar with it. Those with slightly less experience with it may use a wider, less detailed representation where it is recognised as either Geordie or North-Eastern, or even Northern English.

Evidence for the use of more abstract representations of social meaning is found by Hay et al. (2019) who find that abstract social categories facilitate word processing in combined lexical decision and implicit association tasks (for more about these techniques see Subsection 3.2.3). They find that participants sorting old and young faces and words more often used by old and young people subconsciously associate that two. They found the same when participants sorted gendered items and words more often used by men or women.

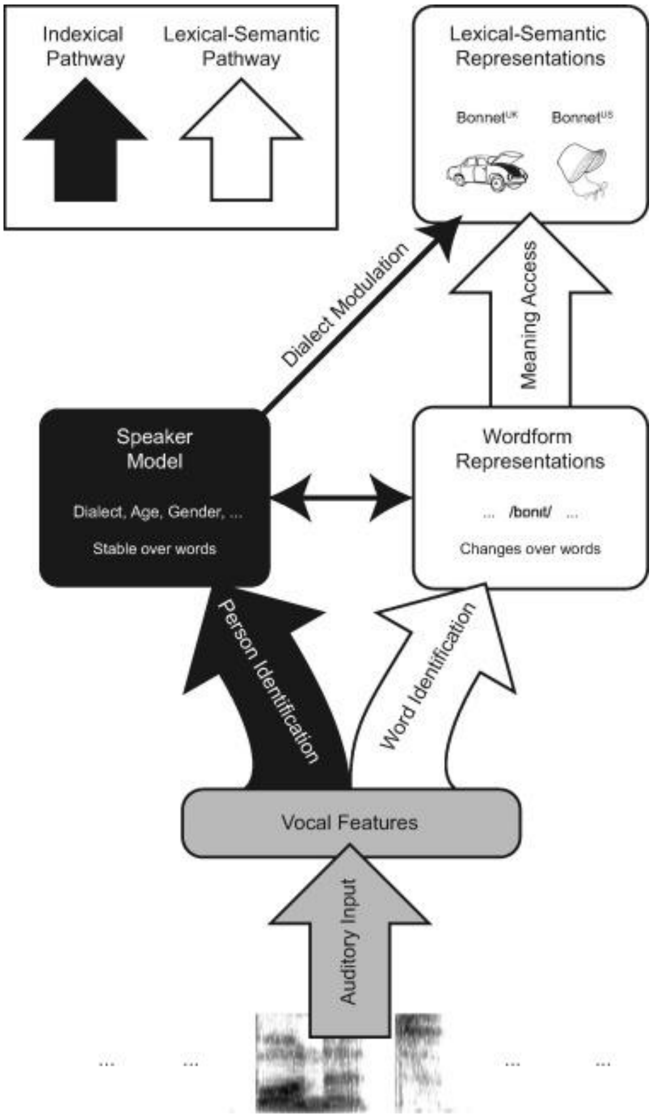

Figure 4. Cai et al.'s (2017)'s speaker model account of spoken word processing (taken from Cai et al. 2017).[3]

Furthermore, neuroscientific research by Dobs et al. (2019) on the perception of images of faces has found that more general (cross-person) characteristics such as gender and age are perceived slightly before more person-specific information about a person's identity. This suggests that social categories

[3] Reprinted under the terms of the Creative Commons CC-BY license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

are central in person perception and may be more primary or central to sociolinguistics than speaker-level representations.

In terms of the size of the linguistic units involved in processing generalisations across speakers, Staum Casasanto (2009) finds influences of social cues (showing participants either a White or a Black face), influences processing of t/d-deletion for American listeners. Interestingly, she finds that this effect was present both in word stimuli and in non-word stimuli. The latter suggests that some sublexical units can be involved in processing sociolinguistic variation.

Above the accent-level even, the lexical underlying form *motorway* would indicate that the speaker is British, which might carry – for non-British people most likely – some social meanings such as posh-ness, quaintness, and Britishness. Alternatively for British speakers it may carry the social meaning 'normal', or no strong meaning at all.

This does not mean that lexical, underlying or non-phonetic, variants always carry the most abstract, social meaning. For example, using the lexical 'underlying form' *howay* may carry the more specific social meaning of being Geordie or Mackem, just as a lexical variant. In fact, there is evidence that lexical and morphosyntactic variation coincides or 'works together' (Snell 2010: 647) with phonetic variation, including for quite specific local social meanings (Moore & Podesva 2009; Snell 2010). If the example utterance used in the table had been 'Howay, man!' the lexical component would have come in at a more specific level. At the same time, the table does not show the possibly more rare but still conceivable case that an underlying form at the syllable or phoneme level carries broad social meaning. Someone having a /ɬ/ phoneme and using it distinctively from /l/ may, for example, indicate that they are Welsh.

Are abstractions necessary?

One further question beyond this is whether the proposed different levels of abstraction of social meaning actually exist and are used. One of the claims of early exemplar theory is that social meaning is processed on the basis of the activation of 'clouds' of memory traces (Foulkes & Docherty 2006) – for example hearing a woman's voice say the word *cat* causes activation of a lot of memory traces of other women (and fewer people of other genders) saying *cat* in a similar way, which all have the social information of the speaker being a woman. This then would, even if all the traces themselves are more detailed with more specific speaker information, cause the overall activation of 'female speaker'.

However, Lavan, Knight & McGettigan (2019) found that listeners create abstractions of speaker voices, which they recognise better than instances of those voices (see 'Speaker generalisations' above). This shows at least one data point in support of abstraction beyond exemplars (at the level of the speaker voice), although there is a clear gap in the research here.

One of the ways this question is addressed in this thesis is by investigating the link between social meaning and linguistic information, as discussed in Section 2.2 above. Chapters 3 and 4 use a variation on semantic priming methodologies to ascertain whether at a highly automatic level, socio-phonetically detailed input always activates the social meaning associated with it. If this is not the case the two may be more likely to be separate. This would imply that exemplars are not the main carriers of social meaning in sociolinguistic processing. Abstractions then would be a necessary part in sociolinguistic processing, as they have been found to be in speech processing more generally (Cutler 2008; Ernestus 2014).

Two parameters

At all the different levels of detail and abstraction discussed above, it is conceivable that the representations used in the processing of sociolinguistic variation are of different sizes, from utterances to segmental representations. This is in contrast with accounts of (hybrid) exemplar theory which assume that lexical effects automatically indicate specificity in mental representations (Ernestus 2014;

Pierrehumbert 2016), and segmental effects automatically indicate abstraction in mental representations (German, Carlson & Pierrehumbert 2013; Pierrehumbert 2016). Instead, I argue that it is more helpful to study these parameters separately. Chapters 5 and 6 focus on the question of what size the representations are that are used to access social meaning, by comparing accent recognition in high-frequency words, low-frequency words, and non-words. However, these experiments also found evidence engaging with separate questions of how more abstract representations of character type related social meaning play a role in sociolinguistic processing.

2.4 Summary and predictions

The framework outlined in this chapter describes different approaches to the main research questions of this thesis regarding the joint or separate storage of social meaning and linguistic variation in memory (research question 1) and the size and detail of the linguistic representations that are used to access social meaning (research question 2) (see the research questions detailed in Section 1.1).

As discussed in Section 2.2 above, third-wave approaches to social meaning, especially Acton (2021), conceive of social meaning as a set of inferences and an interpretative link between linguistic variation and the social information it indexes. In this approach social and linguistic variation information must be stored separately in cognition, and are closely but flexibly connected. Pure exemplar theory expects social meaning and linguistic variation to be united in the same representations (the exemplars), and hybrid exemplar theory allows for both options: the two could either united in both exemplar and more abstract representations, forming a unit of linguistic and social information, or they could be separate in more abstract representations. These possibilities are investigated in Chapters 3 and 4.

Those predictions also influence predictions of the memory representations used in the retrieval or interpretation of the social meaning of linguistic variation: if the two are stored separately, as third-wave approaches predict, then exemplar representations cannot be the meaning-carrying representations (as they unite both types of information). Hybrid exemplar theory would allow for either or both exemplars and higher-level abstractions to be used in sociolinguistic processing. Furthermore, there is a wide range of different potential linguistic representations which may be used to access social meaning, of different sizes and different levels of specificity, beyond the lexical exemplar. Those possibilities are tested in Chapters 5 and 6.

3 Socio-contextual priming between regional lexical variants and accent names and stereotypes

3.1 Introduction

In this chapter, I investigate whether social meaning and linguistic variation are cognitively stored within the same representation or are stored separately. In order to do this I use a method which tests highly automatic processes – a variation on a semantic priming task that I am calling a 'socio-contextual' priming task. This aimed to test whether it was possible for linguistic information to be activated in memory without its corresponding social information being activated as well, and vice versa, before more conscious evaluation processes start. If it is not possible to do this this would support the idea that social meaning and linguistic detail are stored within the same representations, for example in exemplars (see Chapter 2). However, as it was possible to do this in this experiment, the two types of information seem likely to be stored separately, in line with third-wave approaches to social meaning which posit flexible links between the two.

The experimental technique of semantic priming measures whether a participant's exposure to one stimulus influences their response to a second, semantically related, stimulus. The first stimulus is called the prime stimulus, and the second the target stimulus. In semantic priming responses to a target stimulus are compared between two conditions: in one the target stimulus follows a prime related to the target, and in one the target follows an unrelated prime. For example, someone might respond more quickly to reading the target word DOG after having read the related prime word *cat* than after reading the unrelated prime word *idea*. Such a priming effect implies that whilst the representation of the word *cat* becomes active in the brain when someone reads and processes it, the activation of this representation spreads activation to other, related representations as well, such as the word DOG. This happens without people being consciously aware of it. These types of priming effects have widely been found in priming studies in psychology, as I discuss and illustrate in Section 3.2 below.

The current experiment uses not semantic relatedness, like 'cat – DOG', but socio-contextual relatedness: in this case the relatedness between an accent (e.g, *American*) and its regional lexical variants (e.g. CANDY) and vice versa. If there were consistent priming effects in this experiment this would indicate joint or spreading activation between linguistic variants and their social meaning. This would be evidence that the two are either part of the same representation, or very closely connected.

However, the experiment in this chapter finds mixed evidence of automatic activation between linguistic variant and social meaning. It shows that being repeatedly exposed to the regional lexical variant *y'all* results in participants responding more slowly to words relating to Americanness (e.g. GUNS), than words that are unrelated. The results also show that participants responded more quickly to regional lexical variants (e.g. CANDY) when they had been exposed to accent names (e.g. *American*) than an unrelated prime, although only in the absence of prime repetition. The different directionalities of the priming effects (slower response times for one, quicker for another) found in the different conditions suggest that there is a variable link between linguistic representations and their social meaning, when as many evaluative processes as possible are taken out of the equation. This variable linking suggests the separate storage of social and linguistic information.

3.2 Measuring automatic activation

3.2.1 Introduction

Sociolinguistic perception research has come to be more and more interested in below-awareness processing (Drager & Kirtley 2016; Campbell-Kibler 2016), and a range of methodologies have been used to uncover automatic and subconscious processes. Differences in the findings of these different methodologies (Campbell-Kibler 2012; D'Onofrio 2016) suggest that there is a connection between linguistic variation and social meaning that is flexible: in some tasks some linguistic variants are shown to carry different social meaning than in other tasks, depending on how automatic and subconscious the processing was in these tasks.

In order to further investigate this, I used a new highly automatic methodology, a 'socio-contextual priming' experiment – a variation on a masked semantic priming task – to explore whether activation of linguistic variation always implies the activation of relevant social meaning, before any more conscious evaluative processes have started.

3.2.2 Models of awareness in sociolinguistic processing

With the new interest in below-awareness processing in sociolinguistic research, new models of awareness in sociolinguistic processing have been developed. Firstly, Campbell-Kibler (2016) proposes a dual-route model for sociolinguistic processing. Here, there are two routes for processing: one quick, automatic, and below-consciousness route and one slow, introspective one with conscious control. Evaluations based on these different routes, may differ: Campbell-Kibler (2012) finds differences in the perception of /t/ release between more explicit, conscious tasks (direct questions and social evaluations) and the implicit association task (see also Subsection 3.2.3 below)).

D'Onofrio (2016) complicates this dichotomous model of processing by introducing the notion of gradience in awareness. She investigates how one linguistic feature, backing of the TRAP vowel in American English, is linked to three associated social meanings (business professional, 'valley girl', and Californian). She does this by means of four different methods, each inviting a different level of explicit or implicit social evaluations. In an eye-tracking study – measuring processing before conscious awareness – she finds that icons for two out of the three social meanings had effects on how listeners processed backed TRAP vowels (California and 'valley girl'). However, she does not find this for an icon of the third meaning (business persona). That is to say an icon of the state of California or 'valley girl' (shopping bags) influenced whether listeners' eye movements were more likely to go to the option SOCK or the option SACK on the experimental task screen, but that this did not happen with an icon of a briefcase. However, all three social meanings influenced perceptions in a phoneme categorisation task, as a less automatic processing task. Also, in an even less automatic, explicit memory task, it was only the business professional meaning which influenced participant accuracy. These different findings suggest multiple levels or a gradience of awareness in sociolinguistic perception, rather than just two separate processing routes.

D'Onofrio's (2016) model of gradience in automaticity fits a model of attitudes and evaluation developed by Cunningham et al (2007; see also Van Bavel, Xiao & Cunningham 2012) called the iterative reprocessing model, summarised in Figure 5 below (taken from Cunningham et al. (2007)). In this model an input stimulus is construed and then evaluated by feeding back into attitude representations. However, this evaluation can be processed repeatedly, each time adding more evaluative processes, and therefore consciousness and awareness to the process. Highly conscious and introspective processing then involves more evaluative processes and costs more time, whereas more automatic processes do not involve (as many of) these iterations. This can account for the differences in D'Onofrio's (2016) data by adding multiple levels of awareness due to the addition of more evaluative processes for more conscious processes.

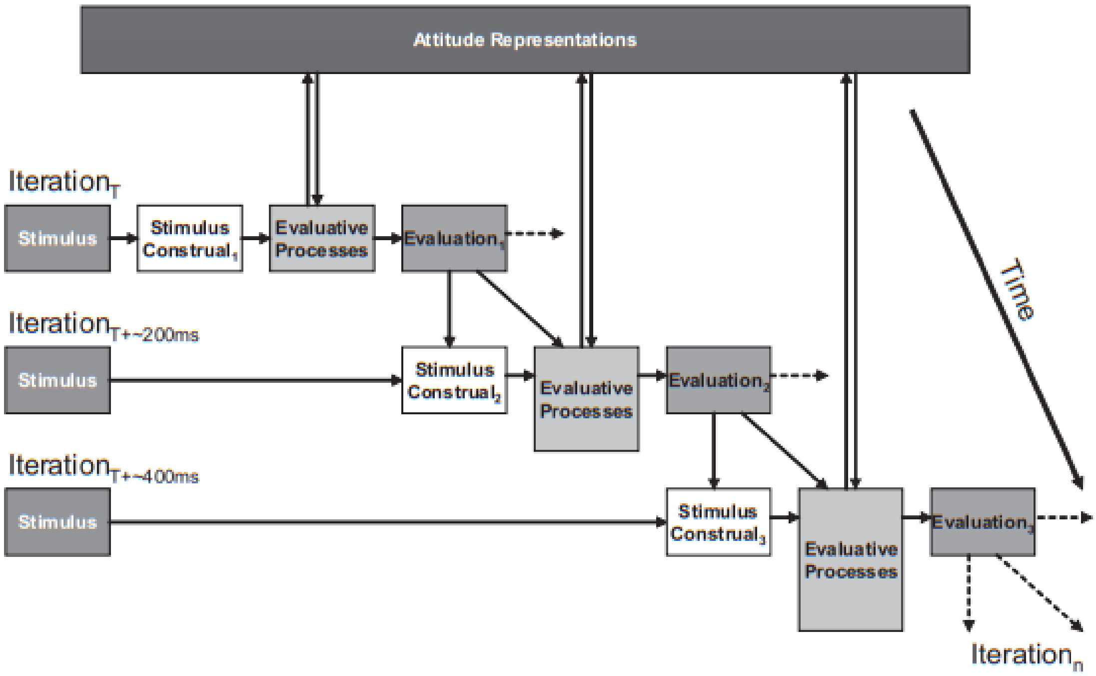

Figure 5. Cunningham et al.'s (2007) iterative reprocessing model. Each iteration adds new evaluative processes, creating a gradience of consciousness for some stimuli being processed with more iterations and some with less (taken from Cunningham et al. 2007).[4]

3.2.3 Automatic processing methodologies

In sociolinguistic perception research, especially in the last decade, researchers have aimed to measure more automatic perception processes (Drager & Kirtley 2016; Campbell-Kibler 2016), using a range of different methodologies along this automaticity parameter. Discussing these in the subsection below, I argue that it may be possible to use a new methodology to uncover even more automatic processing by means of a socio-contextual priming experiment. This enables further exploration of whether social meaning and linguistic variation are stored as a unit in memory, by testing if they are consistently co-activated even when as few evaluative processes as possible are involved in processing.

Matched guise technique

One of the methods traditionally used to access subconscious linguistic attitudes is the 'matched guise technique' (Lambert et al. 1960; Lambert, Anisfeld & Yeni-Komshian 1965; Lambert 1967). In this type of experimental design, listeners are asked to evaluate a speaker on a number of scales (e.g. friendly, confident, proper et cetera). Then, the researcher compares ratings of the same speaker in different 'guises', or sound fragments, in which the speaker uses one accent, dialect, or language, compared to sound fragments where the speaker uses another accent, dialect, or language. Because the only thing different between the guises is the language variety that is used, and not the speaker's voice, the difference in ratings indicates how listeners perceive the variety, and what they infer it means about its speakers. For example, Villarreal (2018) finds that men who exhibit the California Vowel Shift are perceived to be more confident than men without, by comparing ratings of audio recordings of the speakers with and without the California Vowel Shift present.

As the matched guise technique does not directly ask listeners to rate language varieties, but speakers instead, the technique minimises the risk of listeners giving socially desirable answers that listeners are consciously aware of. However, it still involves the explicit elicitation of social judgements. As

[4] Reprinted with permission from Guilford Publications.

D'Onofrio (2016) argues, this still leaves listeners with ample space and time for evaluative processes in order to create conscious social judgement. She finds differences between participants' responses in such a matched-guise task compared to more automatic implicit tasks: participants associated backed TRAP vowels with business personas in a matched guise task but not in an eye-tracking study. Campbell-Kibler (2012) also finds such differences in her study comparing matched-guise task responses and implicit association task responses (see below). This suggests that the matched guise technique is not a method which gets at the most automatic of perception processes.

Phoneme categorisation tasks

Another type of experiment that has often been used in work on the influence of social cues on perception is the phoneme categorisation task (Niedzielski 1999; Hay, Warren & Drager 2006; Hay, Nolan & Drager 2006; D'Onofrio 2016). In these studies listeners are asked to categorise phonemes whilst also being primed with social cues. For example Hay, Nolan & Drager (2006) use a phoneme categorisation task to argue that listeners' perceptual boundaries of those phonemes can shift based on whether the answer sheet in their experiment had the label "Australian" at the top or "New Zealander". Other studies (Koops, Gentry & Pantos 2008; D'Onofrio 2015) show these types of effects in eye-tracking experiments, which examine not the categorisation a participant lands on in the end, but their eye movements between the different options.

D'Onofrio (2016) finds that there are differences between the processing of backed TRAP in her eye-tracking and categorisation tasks compared to the more explicit matched guise techniques, with the 'valley girl' meaning of backed TRAP influencing categorisation tasks and eye-movement, but not the matched guise test results. She also found that the 'business persona' meaning came to the fore in the matched guise test and the categorisation task itself but not in the eye-tracking results of the categorisation task. This suggests that the categorisation task measures more automatic, more subconscious perception processes, and that eye-tracking measures these processes at an even more automatic level. Still, however, the categorisation tasks are relatively explicit in their presentation of the linguistic variation to experimental participants: they are instructed to listen to specific sounds and evaluate them.

The implicit association task

This explicit presentation of linguistic variation is also an aspect of the implicit association task (Greenwald, McGhee & Schwartz 1998) that Campbell-Kibler (2012) and Hay et al. (2019) use. In this task, participants categorise stimuli as one of two things. For example, in Campbell-Kibler's (2012) study, they are shown names of American states, and they have to classify these as either Northern states or Southern states. They do this by pressing one button for the Northern option (using one hand) and one button for the other (using the other hand). However, within the same task they also have to categorise another type of stimuli as one of two things too, with the same buttons. In Campbell-Kibler's (2012) study, these were sound stimuli ending in (-ing) which participants classified as either ending in [ɪn] or [ɪŋ]. One part of the task has the Southern classification on the same button as the [ɪn] classification, and in one part these are switched. If response times during the task are faster when Southern states and [ɪn] are on the same button than when Northern states and [ɪn] are on the same button, this implies that Northernness and [ɪn] are implicitly associated. This is because congruent pairs are more often activated together in the mind, resulting in faster activation and faster responses, whilst pairs that are incongruent for the listener are less often activated together, resulting in slower responses (see Greenwald, McGhee & Schwartz 1998).

Campbell-Kibler's (2012) implicit association task still incorporates a phoneme categorisation task like the ones described above, and therefore also implies some explicit evaluation of linguistic input and variation. Hay et al.'s (2019) methodology however circumvents this: their experiments combined the implicit association task with a lexical decision task. In one of their experiments, participants sorted real words and non-words on the same button as they sorted male faces and female faces. The real words then were either words that are more often used by men or by women. The fact that participants had faster response times when real 'male' words were on the same button as the male faces showed that 'male' words and male faces were implicitly associated.

One less automatic aspect here would be that again the social categories involved in the experiment are still engaged with in a very conscious way, as participants categorise input into two gender groups as part of their task. Hay et al. (2019) address this by asking their participants to categorise (gendered) objects not by gender, but by whether the object they were sorted was presented with a square frame around it. This coincided perfectly with the gendered category of the pictures, which the participants learned implicitly. This way, participants did not overtly categorise social categories (by sorting objects on the basis of their genderedness) and did so only indirectly by categorising on the basis of the presence or absence of a frame. This still showed the same effects of association between the objects and words that are used more often by one gender than another.

Lexical decision tasks

Walker & Hay (2011) as well as Kim & Drager (2018) use more simplified lexical decision task to investigate sociolinguistic processing. In this task, listeners categorise strings of sounds or letters as either real words or nonsense words. In their studies, they find that word recognition is faster when socio-phonetic detail and the social distribution of word stimuli are matched. In both their studies they found that listeners are quicker to recognise words which are more often used by older people than by younger people, when those words are pronounced in an older sounding voice. This methodology does not involve conscious evaluation of the linguistic variation ('old' and 'young' words, as well as the phonetic detail making a voice sound older or younger).

Priming

Finally, a number of studies have tried to measure the presence or absence of representations of sociolinguistic variants through priming studies (McLennan, Luce & Charles-Luce 2003; Sumner & Samuel 2009). As discussed above, in a priming experiment participants are exposed to one stimulus (the prime stimulus) and respond to a second stimulus (the target stimulus), in order to test if this exposure influences the response. When these effects are present they imply something about the absence, presence and relatedness between representations of the prime and target. In linguistics this is often done through lexical decision tasks where the target stimuli are words which listeners need to categorise as real words or nonsense words.

This way, Sumner and Samuel (2009) find that listeners with experience listening to New York City English (even if they do not speak the variety themselves) show faster response times to a word in a lexical decision task when they had been primed by that same word pronounced in a New York City English pronunciation, whilst listeners who do not have this experience did not. On the basis of this they are able to say there is a difference in the sociolinguistic representations these different listeners have.

One of the major advantages of priming experiments is that listeners are engaging with linguistic material in a way that does not direct them towards sociolinguistic evaluation of either the linguistic variant or the speaker. Instead the listener is preoccupied with deciding whether the items they are exposed to are real words or not, whilst still giving an insight into the mental representations they are using in processing.

At present, this technique has mostly been used to detect which linguistic representations are present in memory and used in linguistic processing (McLennan, Luce & Charles-Luce 2003; Sumner & Samuel 2005; Sumner & Samuel 2009; Pharao 2019), but not to detect the activation of the social meaning of linguistic variation. This is what I pursued in the experiments described in this chapter and in Chapter 4: I used variations on semantic priming experiments in order to uncover whether even in highly automatic processing, linguistic representations and their social meaning are always activated together or whether it is possible to activate one without the other. This can then provide evidence for or against the idea that social meaning and linguistic variation are stored separately in memory.

3.2.4 Response-time based priming

The experiment described in this chapter is a response-time based priming experiment which measures whether participants responded systematically faster or slower to a target stimulus when it was primed by a sociolinguistic variant and the target stimulus itself was a socio-contextually related concept. For example, it tested whether the regional lexical variant *y'all* primes the word GUNS for British participants. It also tested if this was the case when the prime was the name for an accent and the target was a sociolinguistic variant from the accent (for example, whether the word *American* primes the word CANDY).

Response-time based priming experiments are used quite commonly in psycholinguistic speech perception studies, in order to test whether activation in the brain spreads from one linguistic representation to another. More particularly, it has often been used to test aspects of exemplar theory, as it allows for insights into what linguistic representations are used and present in memory (Church & Schacter 1994; McLennan, Luce & Charles-Luce 2003; Norris et al. 2006). These response-time based priming experiments are distinct in their methodology from the larger body of work on sociolinguistic experiments where participants are exposed to social cues to test whether this influences linguistic processing (Niedzielski 1999; Hay, Warren & Drager 2006 and see Chapter 2 of this thesis), which is sometimes also called priming (as it also measures differences in responses to a stimulus because of previous exposure to another stimulus). I will take Sumner and Samuel's (2009) study (briefly discussed above) to illustrate how a priming experiment can provide new insight into sociolinguistic processing.

Sumner and Samuel (2009) conducted a number of priming experiments to see whether people with different levels of experience with New York City English all had the same representations for NYC English stimuli. They measured responses by asking participants to perform a lexical decision task. In one of their experiments, participants heard words being pronounced with either a General American accent or with a New York City accent (the prime) and then heard the same word again, either in the same or the other accent (the target). As can be seen in Figure 6 on the next page, in most pronunciation conditions in the experiment, this repetition of the same word caused a faster response time to the target word, compared to when prime and target were not the same word ('RT to Control Target – RT to Related Target', or 'Priming Effect' in the graph).

Where these priming effects become interesting from a sociolinguistic perspective is that not all participants show a priming effect in all conditions. The group of participants with little experience with NYC do not have priming effects between a General American and a NYC form or between two NYC forms – which suggests that they do not have the representations of NYC English pronunciations stored in memory that they would need to see activation spreading between GA [beɪkɚ] and NYC [beɪkə], or NYC [beɪkə] and NYC [beɪkə]. This prevents any priming effects from occurring, as can be seen in the rightmost bars in the table.

3.2.5 Semantic and socio-contextual priming

It has also been found that primes do not need to be the same stimulus as the target stimulus to result in priming effects. Prime-target pairs that have related meanings are responded to faster than prime-target pairs that do not have related meanings (Bordag et al. 2015; McNamara 2005; Radeau 1983; Meyer & Schvaneveldt 1971). This is called 'semantic priming'.

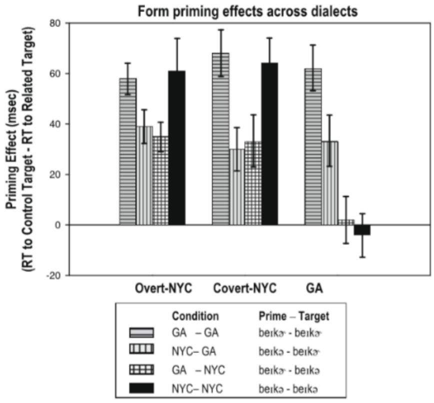

Figure 6. Priming effects for three participant groups, NYC English speakers ('Overt-NYC'), listeners with a lot of experience with NYC English ('Covert NYC'), and listeners with little experience with NYC English ('GA') across for priming conditions (taken from Sumner & Samuel 2009: 492)*.*[5]

As Sumner and Samuel (2009) show, priming experiments can yield new insights for sociolinguistic research by investigating what representations are activated in memory when sociolinguistic variants are processed. Whilst they investigate how the processing of linguistic units (responses to different pronunciations) is influenced by sociolinguistic variables (experience with an accent), the design of semantic priming experiments also seems to offer opportunities for measuring the activation of social meaning.

In the experiment described in this chapter (and those in Chapter 4), I extend the methodology of semantic priming to sociolinguistic information. Where traditional semantic priming measures the activation of one stimulus due to semantic similarity or overlap with another stimulus, these experiments measure whether this can also happen due to socio-contextually relatedness. That is to say, does a linguistic variant with a certain social meaning activate words describing that social meaning, and vice versa?

For example, when a British listener hears rhoticity in the word *car* or hears the word *y'all*, those stimuli may activate the concepts that a British listener associates with these forms, or that are part of the context or exemplars of these forms (see Chapter 2), like American-ness, loudness, and maybe even cowboys, guns, and capitalism. This type of activation could then be detected in a lexical decision task, by measuring how quickly participants in an experiment respond to target words relating to American-ness after they have been primed by the word *y'all*, and comparing this to response times to these target words when they were primed with an unrelated stimulus. Conversely, the prime could be the concept of the linguistic variety itself, like the word *American*, and the target could be a sociolinguistic variant from the variety, like SIDEWALK.

This methodology can therefore explore questions of the relationship between social meaning and linguistic information in memory storage. As is discussed in Chapter 2, it has been unclear whether social information and linguistic information are united in the same mental representations (such as exemplars, or higher level abstractions which unite the two) or have separate but connected mental representation. If there were evidence of consistent priming effects between accent names and stereotypical concepts and lexical variants in the experiment described below and the ones described in Chapter 4, this would make it more probable that social information and linguistic information are united in the same mental representations.

[5] Reprinted with permission from Elsevier.

3.2.6 Semantic priming mechanisms

There are different accounts of the mechanisms that underly semantic priming mechanisms. In some accounts semantic priming is thought to be caused by activation spreading from the mental representation of the prime stimuli to semantically related mental representations (Collins & Loftus 1975). In other accounts, semantic priming happens because because a prime word is part of the context and therefore the larger representation of a target word (Gillund & Shiffrin 1984). I focus on these below (but for other semantic priming models see McNamara 2005).

The first type of models are 'spreading activation' models (Collins & Loftus 1975; Anderson 1983): in these models retrieving an item from memory means activating its representation. This activation can then spread from one representation (e.g. *baker*) to other representations which are related to it (e.g. CAKE). The residual activation of these related representations then facilitates retrieval of those representations (e.g. allowing for a quicker response to CAKE).

If this approach is applied to socio-contextual priming, the activation of an accent variant would spread to concepts that are related to them, through links based on socio-contextual co-occurrence or ideology (e.g. the co-occurrence of *y'all* with American speakers and their stereotypes or the overt ideas people have about *y'all*). Activation may spread directly from one prime to target, but it may also be possible that the priming is 'mediated'. This is to say that the prime (e.g. *y'all*) may activate conceptualisations of its typical speakers (e.g. Americans), which then activate related targets (e.g. GUNS). Mediated priming effect have been attested, although generally with smaller effects than in direct priming (McNamara & Altarriba 1988; McNamara 2005).

Another account of semantic priming is given in compound cue models, such as the Search of Associative Memory model (Gillund & Shiffrin 1984). In this model long term memory contains 'images', which are unitised sets of features containing both information about an item and the context it appears in and associations with other images. The cue to retrieving such an image then does not just include the information about an items itself but also its context, as a *part* of the image. In a semantic priming task, then, the cue for retrieval of the target item is stronger if the context of the input includes not just the target item but also, in the context, the prime item.

This account resembles the exemplar theoretic approach to the storage of social meaning by sociolinguistic perception researchers (Hay, Nolan & Drager 2006; Walker & Hay 2011), i.e. that perceiving a social cue that matches linguistic information facilitates access to that linguistic information, as the linguistic representation of the exemplar includes contextual information (Foulkes & Docherty 2006; Johnson 2006). In this approach, socio-contextual priming would be strong, as it is specifically social *context* which would drive the effect.

One further note on semantic priming is important: with semantic priming it is relatively common to find inhibitory priming effects, where responses to related targets are slower than responses to unrelated targets, rather than as facilitatory effects, where responses to related prime-target pairs are faster than unrelated prime-target pairs (like in Sumner & Samuel's (2009) study discussed in Subsection 3.2.4). An example of inhibitory priming can be found in Bordag et al. (2015) who find that L2 learners of German were slower to respond to the newly learned German words BESEN ('broom') and GARTEN 'garden', if they had just been exposed to the word *Harke* ('rake'). Both main theories on this (the Distractor Inhibition Model by Tipper 1985; and the Episodic Retrieval Model by Neill & Valdes 1992) suggest that inhibitory priming happens because two representations are related but less strongly so than forms which cause facilitatory priming. Inhibition then happens to prevent these cognitively connected but distracting and incorrect representations from being selected in processing.

3.2.7 Design

The study described in this chapter explores whether socio-contextual priming effects can be found when regional lexical variants of accents are preceded by name for the regional variety (e.g. American

– CANDY), or whether concepts relating to that variety were primed by a regional lexical variant (e.g. y'all – GUNS). It did so using the highly automatic technique of a 'masked' priming experiment. In a masked priming experiment participants are exposed to primes for such a short period of time and in such a masked way that they do not realise they have been exposed to them (Kouider & Dehaene 2007; Bartlett 2012; Elgendi et al. 2018). Masked primes are presented for a very short period of time (for example 40ms – 60ms) and preceded by a 'mask', which generally consists of a row of hash marks (#####). This reduces participants' awareness of the presented primes (Balota 1983; Kouider & Dehaene 2007).

The experiment looked for priming effects between sociolinguistic variants like *y'all* and *sidewalk* and socio-contextually related (for British participants) words, like *guns* and *American*. Strong priming effects would support the idea that linguistic information and social meaning automatically activate each other and would make the notion more probable that they are stored within the same mental representation. The experiment tested this for seven different varieties: American, Tyneside, Yorkshire, Irish, Cornish, and Australian English.

I chose to present the primes and targets orthographically, rather than auditorily, as it was beyond the exploratory scope of the study to record auditory stimuli. This meant that the prime words that were chosen were either morpho-lexical variants (*y'all*) or accent names themselves (*American*, *Geordie*, et cetera). Overt and repeated exposure to these words would likely have made participants aware that the experiment was to do with accents (with every other word being an accent name or *y'all*) which would have interfered with the below-awareness processing the experiment aimed to test. To counter this the masked priming paradigm was chosen. This allowed these accent names and the word *y'all* to be repeated multiple times without participants noticing.

The masked semantic priming paradigm allowed for the following experimental research questions to be investigated:

1. Does the regional lexical variant *y'all* automatically cause activation of words that are socio-contextually related to its social meaning, like YANK and HAMBURGER?
2. Do terms describing accents like *American* and *Geordie* automatically cause the activation of sociolinguistic variants associated with the accents in question, like SIDEWALK and PET?

The automatic activation described in these research questions was intended to provide new evidence for or against a unified conceptualisation of social meaning and linguistic information in memory.

3.2.8 Procedure

The priming experiment consisted of a lexical decision task, presented through the priming software DmDX (Forster & Forster 2003). The experiment consisted of a practice phase with 19 items, followed by 288 critical items. Half of these were words and half of these were non-words. For each version half of the target words were primed with words related to the target word (e.g. *American* before the target MOVIE), and half of them were primed with an unrelated non-word. In all cases, the primes were masked with a row of hash marks, as is common in masked priming studies (Bartlett 2012: 13). The target words were also surrounded by three hash marks on both sides to also mask primes that were wider than their targets.[6] All targets were presented in uppercase and all primes were presented in lowercase, following Bartlett (2012), but started with a capital letter if that is how they would always be spelled (i.e. *Irish* and *Cornish*). This is summarised in Table 2 below.

[6] I am not aware of other priming studies which do this, but this made sure the primes which were wider than the targets stayed fully masked. Without the extra hash marks some of these became visible when I tested the experiment.

Table 2. Summary of stimulus presentation for both the related and unrelated prime conditions.

Phase	**Duration**	**Screen for example stimulus (related prime)**	**Screen for example stimulus (unrelated prime)**
Mask	612ms	########	########
Prime	40ms	American	flurp
Target	850ms	###MOVIE###	###MOVIE###

Table 2 also shows how long each of the three elements were shown in the experiment. The initial hash marks were shown for 612ms (36 of DmDX's standard measure for time, 17ms 'ticks'), followed by the masked prime which lasted for 40ms, and the target word which was shown for 850ms (50 ticks). Masked priming experiments tend to present their primes for between 10ms and 100ms (Elgendi et al. 2018), but effects can vary per type of priming. Bartlett (2012: 94–95) notes that orthographic priming effects can be found even for very brief exposures (e.g. 30ms), while phonological priming tends to require slightly longer exposure durations (e.g. 60ms). I tested the experiment with different durations and found that 40ms was the longest exposure duration that still prevented the possibility of being able to consciously read the prime words.

There were two versions of the experiment, each taken by one half of the participant group. Both versions had all the same targets, and the same order, but differed in whether each target was preceded by a related prime or not. This created a baseline of response times to targets which were *not* primed with a related word (e.g. *flurp* – MOVIE). Priming effects could then be calculated by comparing response times to targets which had been primed with a related word (e.g. *American* - MOVIE) to this baseline.

I followed McNamara (2005: 156) in taking non-words as the unrelated primes as a baseline measure. These were taken from the ARC Nonword Database (Rastle, Harrington & Coltheart 2002). An alternative to this would be to use 'neutral' primes such as a string of crosses (Neely 1977), or percent signs (Ferrand, Segui & Humphreys 1997; Schiller 1999), but these have been suggested to allow for quicker processing of the target word (Kinoshita & Hunt 2008) and therefore lead to overestimations of inhibitory priming and underestimations of facilitatory priming (Bartlett 2012). Non-words as primes do not seem to have this same effect, as Forster et al. (2003: 6) suggest, while they do prevent the accidental priming effects unrelated real words might cause.

Another alternative would have been to use unrelated real words (Forster 1998), but this has the risk that the unrelated word causes unexpected priming effects as well, clouding the view on whether the intended prime activated the target. Similarly primes could have been rotated between targets, so that for example the target word CROCODILE would be primed with the word *Australian* in the related condition and *Geordie* in the unrelated condition. I chose to avoid this because it would mean all primes would be repeated much more. Furthermore, a considerable number of target words are possibly related to multiple varieties (e.g. LAD or LASS being related to *Yorkshire* as well as *Geordie*), which would have muddied the results.

3.2.9 Stimuli

The experiment tested four different types of prime-target combinations, all in different items. The prime would either be the (American) regional lexical variant *y'all*, or an accent name (*American Geordie, Yorkshire, Irish, Cornish, Australian*), and the target could either be a concept that was socio-contextually related to an accent (e.g. GUNS or CAPITALIST for American English[7]) or a regional lexical variant (e.g. SWEATER or CANDY for American English). This is summarised in Table 3 below. The regional lexical variants were either exclusive (like the American English examples) or in some way typical for the accent in question (e.g. GRAND for *Irish* or PET for *Geordie*).

Table 3. Four types of prime-target relationships in the experiment.

Prime	Target	Example
Regional lexical variant	Regional lexical variant	*y'all – SWEATER*
	Accent-related concept	*y'all – GUNS*
Accent name	Regional lexical variant	*American – CANDY*
	Accent-related concept	*American – CAPITALIST*

The prime-targets where accent names as primes were combined with accent-related concepts (the bottom row of Table 3) were intended as a control group of sorts, as many other semantic priming experiments have shown this type of priming on the basis of semantically and associatively related words (see McNamara 2005).

The prime-target combinations used in this experiment were selected on the basis of brainstorm-style sessions between friends and colleagues.[8] This produced different numbers of regional lexical variants and accent related concepts, as not all accents had an equal number of prime-target combinations available, simply because some variants had fewer easily recognisable regional lexical variants and cultural associations available. There were 19 prime-target combinations with *American* as the prime word, 19 with *y'all*, 33 with *Geordie*, 21 with *Yorkshire*, 12 with *Irish*, 13 with *Cornish*, and 20 with *Australian* as the prime word. No primes were ever repeated within 7 consecutive items, in order to limit repetition effects. The full stimulus lists are given in Appendices Ia and Ib.

3.2.10 Participants

36 participants participated in the experiment at a computer either in a quiet postgraduate working space (participants 1-24 and 36), or a space specifically designed to conduct experiments in (participants 25-35). All participants were recruited through my own personal network and were all between the ages of

[7] I use this term American English here as a broad perceptual category for British listeners, to indicate what broad group of people sound like Americans, and by no means to suggest that American Englishes would be uniform.

[8] Thanks in particular to Matty Campbell and Holly Dann. The responsibility for the chosen stimuli is of course is my own.

19 and 29 (M=22.0). All of them were, or had at one point been, students at a Sheffield university. All participants had lived most of their childhood in the UK and were native speakers of English. As a proxy for gender, participants' pronouns were noted. 18 participants used the pronouns *he/him* (50%), 11 used *she/her* (31%), 5 used *they/them* (14%), and 2 used both *he/him* and *they/them* (6%). Participants were also asked to indicate the UK regions they had lived in for more than 10 years of their life. All but two people had lived in only one region for more than 10 years. The most common regions were the Midlands (26%), Yorkshire (24%), and the South-East (18%), followed by the North-West and the South-West (each 11%) and the North-East and Wales (each 5%). None of the participants had lived in Scotland for more than 10 years.

The participants were given participant information sheets about the ethics of the experiment and were asked to sign consent forms. The experiment went through the University of Sheffield's ethical review process and was approved. The participants were then asked to start the experiment, which gave them the instructions. The participants were told that they were going to be shown strings of letters on the screen and asked to press the right shift button if the string formed a real word, and the left shift button if the string did not form a real word. They were told that it was possible that the string of letters would disappear before they pressed a button, but to press the left or right shift button nonetheless. They were also reassured that they would most likely make a few mistakes, and to not worry about it. The participants were told that they could ask questions up until the end of the practice phase.

In some of the earlier runs someone would occasionally chat loudly in the background. This never lasted more than about a minute. Some of the participants would press the left shift button (for 'not a word') for the initially shown 7 hash marks as well as the actual targets. One participant used only one hand and was much slower.

3.3 Results

3.3.1 Overview

In order to convert all response times into reliable priming effects, all participants with error rates above 10% were discarded (N=4) (following Sumner & Samuel 2009). The participant who used only one hand during the experiment was also discarded. For each participant any response times more than 2.5 standard deviations away from the mean were discarded (following Sumner & Samuel 2009). When this left fewer than two thirds of responses per item, the item was discarded for all participants (N=2).

For each of the prime-target combination types (see Table 3), linear mixed-effects regression models (using the 'lme4' package (Bates et al. 2015) in R (R Core Team 2013)) were fitted to find the model best fitting the data. The reason separate models were used for each of the prime-target combinations was that the number of items for each of the combinations differed quite strongly ($N_{variant\text{-}concept} = 11$, $N_{variant\text{-}variant} = 7$, $N_{name\text{-}concept} = 78$[9], $N_{name\text{-}variant} = 25$).

These models were fitted with log-transformed response times as the dependent variable (following Hay et al. 2019) and were first fitted as maximal models and then reduced by taking out predictors which did not significantly improve the fit of the models, through ANOVA comparisons. If the relatedness of the prime word to the target word significantly improved the fit of the model this was taken as evidence for priming effects.

As is summarised in Table 4 below, inhibitory priming effects were found for prime-targets where the prime was a regional lexical variant, and the target was an accent related concept (e.g. *y'all* – GUNS),

[9] This, as a control group, is of course somewhat outsized compared to the other groups. This is because the experiment was originally set up to also measure differences within this group. This influenced the design of the experiment. However, this research question is not discussed in further detail, as the experiment turned out not to be fully equipped to answer that question, and it was not directly relevant to the larger aims of the PhD project.

but only if the prime had been repeated often in the experiment already. No priming effects were found for regional lexical variants when they were used as both the prime and the target (e.g. y'all-SWEATER). Facilitatory priming effects were found for pairs where the prime was an accent name and the target was a regional lexical variant (e.g. *American* - CANDY), but only if the prime had not been repeated much before. There was a priming effect between accent names and accent-related concepts (the control group), which was inhibitory, although it was necessary that the first character of the prime and target matched for this to happen.

Table 4. Summary of results by prime-target combination.

Prime	**Target**	**Example**	**Priming effect**
Regional lexical variant	Accent related concept	*y'all – GUNS*	Inhibitory with high number of repetitions of the prime
	Regional lexical variant	*y'all – SWEATER*	None
Accent name	Regional lexical variant	*American – CANDY*	Facilitatory with few prime repetitions
	Accent related concept	*American – CAPITALIST*	Inhibitory, if prime and target onsets matched

3.3.2 Priming effects for prime-target combination types

All models were fitted by starting with a maximal model with interactions between relatedness of the prime, the number of prime repetitions and whether the first character of the prime and target were the same. Target word and participant number were included as random effects. Relatedness of the prime was included as a slope on the random effect of target word. These maximal models can be summarised as follows, in the form of the code used in RStudio:

```
FullModel <- lmer(log(ReactionTime) ~ RelatednessOfPrime *
PrimeRepetitions * MatchingOnset + TargetLength +
log(LexicalFrequencyTargetWord) + ParticipantAge +
log(PreviousResponseTime) +

      (1 + RelatednessOfPrime | TargetWord) + (1 |
      ParticipantNumber), dat = [the relevant data frame])
```

These maximal models were reduced by first testing whether the slopes improved the fit of the model significantly, then the interactions, followed by the random effects and fixed effects. For none of the models adding the slope improved the fit of the model, but for all but one model including both target word and participant as random effects improved the fit of the model.

Priming between regional lexical variant and accent related concepts (*y'all* – GUNS)

The model with the best fit for the items which have a regional lexical variant as the prime and an accent related concept as the target, like *y'all* – GUNS, is shown in Table 5 on the next page (created by means of the tab_model function in the 'sjPlot' package in RStudio (Lüdecke 2019)). It shows there is an interaction between the relatedness of the prime and the number of repetitions in the prime.

In order to be able to see whether the interaction between prime repetitions and relatedness of the prime means there was a facilitatory priming effect with low repetitions and no priming effects with more repetitions, or the other way around (or a combination of both), the reaction times for these prime-target combinations were plotted against how often a prime had already been repeated in the experiment.[10] As can be seen in Figure 7, there are no priming effects for items which had not seen its prime repeated

[10] The primes for the unrelated priming conditions were not repeated, but assigned the same number as their counterparts with the related primes which were repeated, in order to be able to compare these to a baseline.

(much) yet, but there is an inhibitory priming effect which only shows up after the prime has been repeated more than six times.

Table 5. Best-fit model for items with a regional lexical variant as the prime and an accent related concept as the target (y'all – GUNS).

	Log(Reaction Time)		
Predictors	*Estimates*	*CI*	p
(Intercept)	6.05	5.82 – 6.28	**<0.001**
Relatedness of the prime	-0.05	-0.11 – 0.00	0.068
Prime repetitions	-0.00	-0.01 – 0.01	0.975
Lexical frequency (log)	-0.03	-0.05 – -0.00	**0.017**
Age	0.02	0.01 – 0.03	**0.002**
*Relatedness of the prime * prime repetitions*	0.01	0.00 – 0.02	**0.027**

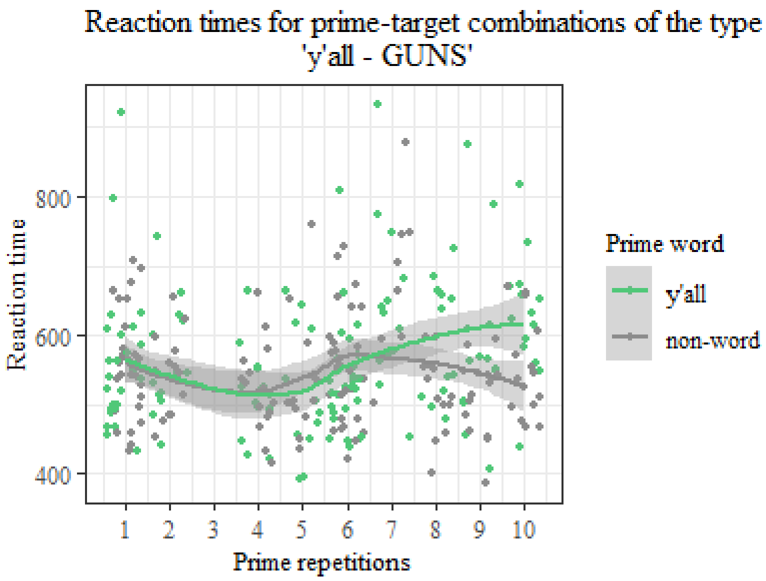

Figure 7. Reaction times for Lexical Variant – Related concept prime-target combinations, by number of prime repetitions. Individual data points are plotted as jitter for interpretability.

Priming between regional lexical variant and regional lexical variant (*y'all* – SWEATER)

The model with the best fit for the items which have a regional lexical variant for both the prime and the target, like *y'all* – SWEATER, was very small and seemed to struggle from the low number of items included (N=7). It is summarised in Table 6 on the next page. This was the only model where the inclusion of target word as a random effect did not improve the model.

Table 6. Best-fit model for items with regional lexical variants for both prime and target (y'all – SWEATER).

	Response time (log)		
Predictors	*Estimates*	*CI*	p
(Intercept)	4.85	4.11 – 5.59	**<0.001**
Previous response time (log)	0.23	0.12 – 0.35	**<0.001**

Priming between accent name and regional lexical variant (*American* – CANDY)

The model with the best fit for the items which have a regional lexical variant for both the prime and the target, like *American* – CANDY, shows a facilitatory priming effect, for item where the primes had not been repeated much yet, as is shown in the interaction between relatedness of the prime and the number of repetitions of the prime.

Table 7. Best-fit model for items with an accent name as the prime and a regional lexical variant as the target (American – CANDY).

	Response time (log)		
Predictors	*Estimates*	*Confidence Interval*	p
(Intercept)	5.22	4.86 – 5.58	**<0.001**
Relatedness of the prime	-0.05	-0.09 – -0.00	**0.032**
Prime repetitions	-0.01	-0.01 – 0.00	0.056
Lexical frequency (log)	-0.01	-0.02 – -0.00	**0.014**
Age	0.01	0.00 – 0.02	**0.005**
Previous response time (log)	0.14	0.09 – 0.19	**<0.001**
*Relatedness of the prime * prime repetitions*	0.01	0.00 – 0.01	**0.016**

As can be seen in Figure 8, the interaction between relatedness of the prime and the number of prime repetitions is caused by an early facilitatory priming effect with no or only one repetition of the prime, which disappears with a higher number of repetitions.

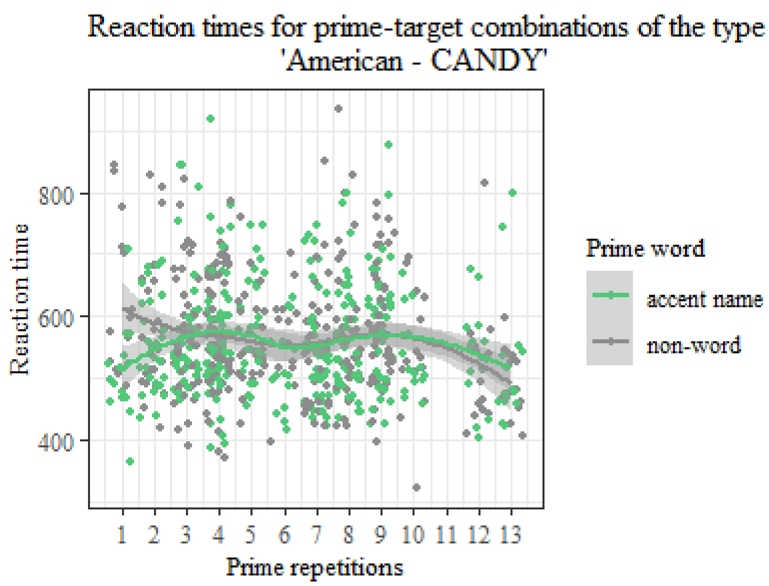

Figure 8. Reaction times for prime-target combinations with an accent name as the prime and a regional lexical variant as the target, by number of prime repetitions. Individual data points are plotted as jitter for interpretability.

Priming between accent name and related accent concept (*American* – CAPITALIST)

For the control group of items with an accent name as the prime word and a related accent concept as a target word, the best-fit model is shown in Table 8. It shows an inhibitory priming effect for items which had matching prime and target onsets.

Table 8. Best-fit model for items with an accent name as the prime and an accent related concept as the target (American – CAPITALIST).

	Response time (log)		
Predictors	*Estimates*	*Confidence Interval*	p
(Intercept)	5.44	5.19 – 5.70	**<0.001**
Relatedness of the prime	0.01	-0.00 – 0.02	0.077
Matching onset	-0.04	-0.08 – 0.01	0.087
Lexical frequency (log)	-0.01	-0.02 – -0.00	**0.003**
Age	0.01	0.00 – 0.02	**0.005**
Previous response time (log)	0.10	0.07 – 0.13	**<0.001**
*Relatedness of the prime * Matching onset*	0.09	0.03 – 0.15	**0.004**

The interaction reported in the model is caused more specifically by higher response times for the related prime word condition for items with matching onsets, rather than a difference for the items with non-matching onsets, as can be seen in Figure 9. Whilst the items with no matching onsets do not show much of a difference for the related and unrelated prime conditions, items with matching onsets and a related prime showed slowed response times, whilst those with matching onsets with an unrelated prime showed slightly speeded response times.

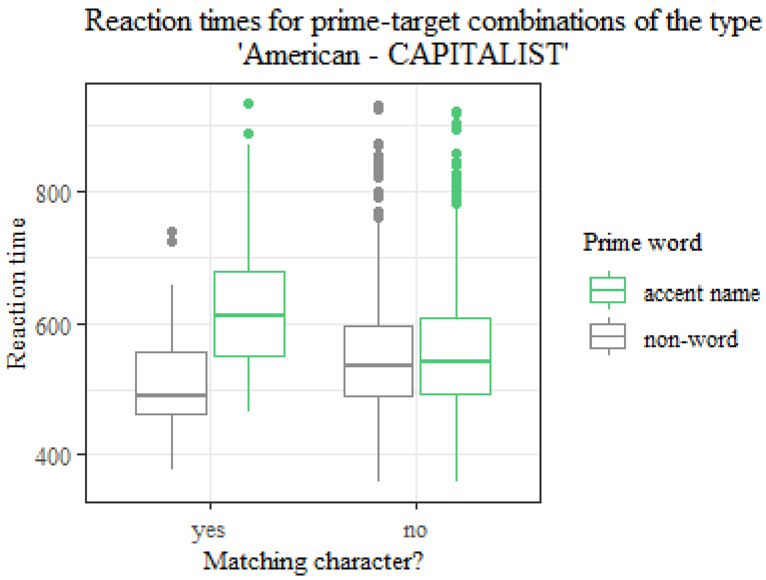

Figure 9. Reaction times to prime-target combinations with an accent name as the prime word and an accent-related concept as the target word.

3.4 Discussion

3.4.1 Evidence for separate storage of social and linguistic information

The results discussed above show a somewhat muddled picture with regards to whether the social context of a regional lexical variant and that lexical variant automatically activate each other in a highly automatic masked priming task. The most direct way in which this activation could be found, is if prime words which are accent variants – in this experiment the word *y'all* – would cause priming effects in the response times to target words related to the accent it was from (e.g. a quicker or slower response to words like GUNS or YANK after being primed by *y'all*). This was the case, but only when participants had been exposed to the prime word a number of times, in which case inhibitory priming occurred. In the other direction, accent names primed their regional lexical variants, facilitating access rather than inhibiting it, but only with very few exposures.

It is not entirely clear why the former priming effect would only happen after repeated exposure, as this experimental set-up is not very common. Most semantic priming studies only match up one prime with one target (see for example Perea, Duñabeitia & Carreiras 2008; Bueno & Frenck-Mestre 2008; Kusunose, Hino & Lupker 2016) or sometimes two (Bordag et al. 2015). It seems possible that the prime *y'all* has a relatively difficult orthography to parse (with a punctuation marker as the second character), and may not be very common in writing, which may have needed a number of repeated exposures to be able to fully be processed in the form of a 40ms masked prime.

Prime-target combinations with the lexical variant as the target word rather than the prime word showed a different picture. The items with regional lexical variants as both the prime word and the target word showed no effects, but the dataset likely had too few items to be able to analyse this well (N=7). When the prime word was an accent name however, there was a facilitatory priming effect when the prime had not been repeated many times before. This suggests that these accent names are a part of the context of its regional variants, or closely related to them (in the compound-cue model (Gillund & Shiffrin 1984) and spreading activation models (Collins & Loftus 1975; Anderson 1983) respectively). This is in line with the relatively large number of sociolinguistic studies which have found effects of social cues on linguistic processing (Niedzielski 1999; Hay, Warren & Drager 2006; Hay, Nolan & Drager 2006; D'Onofrio 2015), where it has been shown for example that simply reading the word 'Australian' can prime listeners to hear vowels as more Australian.

These two findings (inhibitory priming for *y'all* – GUNS type priming and facilitatory priming for *American* – CANDY priming) combined suggest that activation in accent processing may not always work in the same way in both directions. As discussed in Subsection 3.2.6, inhibitory priming is thought to happen when a prime and a target are less closely related than in facilitatory priming. Therefore the differences in those types of priming between linguistic information primed social meaning and when social meaning primed linguistic information, suggest that social meaning may be more strongly linked to linguistic variation than vice versa. These differences in the links imply separate storage requiring links.

This would be at odds with the notion that social meaning and linguistic information are a part of the same mental representations. If this was the case, activation of one would imply activation of the other, as a part of the same representation. The findings that the activation of one affects the activation of the other *in some cases*, whilst the activation of the latter affects the activation of the former differently *in other cases*, suggests more different links of different strength between them. This linked nature would be evidence for the idea that the two are stored separately and that the link between them is somewhat more flexible, as predicted by third-wave variation studies (Eckert 2008; Acton 2021).

3.4.2 Methodological issues

The study highlighted a number of issues to be addressed in further research in order to make stronger claims on the findings. Most of these are addressed in the follow-up experiments described in Chapter 4. A first possible issue is that the prime-target combinations were compiled through brainstorm-type sessions between me and friends and colleagues. As a non-native speaker who had only lived in Britain for 2 years at this point, I was central in this, but this may not have been an optimal way of finding regional lexical variants and concepts that participants themselves connected strongly to the accents in question. Furthermore, a rather wide net was cast with respect to the target words that were included. This means that it is possible that more prime-target combinations were included than participants actually have associations between. The items *Cornish* – FARMER and *Cornish* – TIN are probably not equal in the strength of their associations, and priming effects may have been weakened by the inclusion of items with an associative connection that was not strongly present for the participants. Therefore, a similar experiment which only uses prime-target combinations which have been found (through for example a survey like in Chapter 4) to be associatively connected to each other according to British participants, could strongly improve on the design.

Similarly, it would be an improvement to balance the four different types of prime-target combination types (sociolinguistic variant – sociolinguistic variant, sociolinguistic variant – socio-contextually related concept, accent name – sociolinguistic variant, accent name – socio-contextually related concept). This would allow for a better comparison between the groups, as the lack of balance prevented a real analysis of the ‘sociolinguistic variant – sociolinguistic variant’ type, and made an overall comparison in one model impossible.

The fact that the prime-target combination of the type ‘accent name – socio-contextually related concept’ (e.g. *American* – CAPITALIST) only showed priming effects when the onsets of the prime and target matched, means that onsets need to be controlled for in any further use of the current methods (as is done in the experiments in Chapter 4).

As discussed in Subsection 3.4.1, the prime-target combinations which had an accent-related word as the target and *y'all* as a prime (e.g. *y'all* – GUNS), only showed inhibitory priming effects when the prime had been repeated multiple times, which may have been caused by its relatively remarkable orthography and its relatively low frequency in written language. This issue can be addressed in two ways: firstly, it would be an improvement to expand the number of regional lexical variants included (e.g. *why-aye* or *howay* for Geordie English, or *ey up* or *mardy* for Yorkshire English) so that findings are based on a larger number of different primes.

Secondly, expanding a semantic priming study to measure activation in accent recognition to use auditory forms rather than written lexical forms can address the issue of the relative low frequency of forms like *y'all* in written language. This makes it possible to avoid repetition of the same prime word by priming participants with different words with the same phonological variant (e.g. a monophthongal fronted Yorkshire GOAT vowel in the words *stone, road, though*, et cetera.) before responding to the Yorkshire related words NORTH or STEEL.

Furthermore, such an auditory design means that repeating the same vowel does not stand out as much as repeating the same word. Therefore no masking would be necessary. Still, repetitions of a ‘prime vowel’ would happen. This may have effects similar to the repetition effects in the experiment in the current chapter. In order to isolate this effect, a future experiment can randomise how often a prime vowel has occurred already per target word. This was not done for the current experiment which means that if early or late items happen, by chance, to have stronger effects, this could lead to this being attributed to prime repetitions rather than the items themselves.

Finally, a simplification of the experiment would benefit the clarity of the outcomes. As White (2019) argues, semantic priming studies are highly sensitive and benefit from being kept as straightforward and simple as possible. This could include, for example, reducing the number of different accents tested to one, specifically one with a wide range of associations to provide enough possible target words. It could also mean making the category of ‘regional lexical variants’ more homogeneous, by separating lexical items typical for but not unique to a variety (*pet* for Geordie) and lexical items that are (relatively) unique to a variety (*candy* for American Englishes). The experiments described in Chapter 4 aim to address these issues.

3.5 Conclusion

The experiment described in this chapter explored the automatic activation of regional lexical variants through masked exposure to accent names as well as the automatic activation of accent stereotypes through masked exposure to regional lexical variants, in the form of a masked 'socio-contextual' priming study. This was a variation on semantic priming where social context formed the associative link between prime and target. The findings that there were some priming effects between primes and targets, but different ones and only at certain levels of repeated exposure to the prime, suggests that the two types of information are stored separately in memory and sometimes closely connected and sometimes less closely connected. More specifically, the difference between facilitatory priming effects for prime target-pairs of the type '*American* - CANDY' and inhibitory priming effects for pairs of the type '*y'all* – GUNS' suggests that there is a link from accent information to its variants which is stronger than the link from variants to information about their variety. The following chapter explores this separation and linking in storage further and improves on some of the methodological issues found in this chapter's experiment.

4 Socio-contextual priming between regional lexical variants and sublexical accent features

4.1 Introduction

The socio-contextual priming experiment of Chapter 3 showed some priming effects which suggested different relationships between regional lexical variants and accent names and stereotypes depending on the direction of activation (whether the variant primed stereotypes or the accent name primed the variants), and only at different levels of exposure to the prime. Its findings that accent information primed linguistic variants in a facilitatory way, but that a linguistic variant primed accent information in an inhibitory way suggests that these two types of information are linked in different ways depending on the directionality of activation. This in itself implies that those types of information are stored separately, so that there can be (differences in) links between the two. This is in contrast with the idea that they would be part of the same mental representation. This is further explored in two twin experiments described in the current chapter. These were different from the experiment in Chapter 3 because the primes contained auditory sociolinguistic information in the form of General American[11] or Standard Southern British English (SSBE) variants, such as rhoticity and the LOT vowel.

The aim of the current experiments was to test whether processing linguistic accent information would automatically activate socio-contextually related lexical items or stereotype based social meanings. The outcomes from the experiment reported in Chapter 3 suggest that words relating to social information prime associated lexical variants in a facilitatory way, in some conditions, whilst in other conditions lexical variants prime social information in an inhibitory way, suggesting separate storage of social and linguistic information. This finding was further explored in the current study by using auditory primes that included socio-phonetic information, such as General American rhoticity, or an open LOT vowel. The current chapter describes a pair of highly similar auditory semantic priming experiments that tested whether hearing an American accent before responding to American lexical items, such as SIDEWALK and EGGPLANT, and words for American stereotypes, such as GUNS and RELIGIOUS, resulted in priming effects in in a lexical decision task. They were also different in that the design of the experiments were more structured, and simplified on the basis of the issues encountered with the experiment in Chapter 3.

The experiments did not find evidence of priming effects for prime stimuli that included sublexical accent features with target stimuli formed from accent stereotypes (e.g., GUNS) or regional lexical variants (e.g., SIDEWALK). The lack of effects (whilst statistically unprovable in a frequentist approach) supports the assumption that social and linguistic information are stored separately and do not always automatically activate each other.

---

[11] I follow Wells (1982: 470) in using the term ‘General American’ to mean a wider group of American accents which do not show marked Southern or Eastern Characteristics. He argues that whilst this is not a unified accent, it has it uses as it corresponds to non-linguists perceptions of American English that is not regionally marked. This percept is what I aim to test in the experiments in this thesis. I use ‘Standard Southern British English’ to talk about Standard English in the South of English, as a less evaluative term than ‘Received Pronunciation’, following Hughes, Trudgill & Watt (2012: 3).

In both the experiments in this chapter, participants were asked to recognise strings of letters as either real words or non-words in a lexical decision task. Before seeing these strings of letters, however, they would hear an audio clip of a singular word or non-word, which was either pronounced in a General American accent or an SSBE accent. The experiments aimed to uncover whether hearing an American accent in the audio stimulus (the prime) influenced response times to the following visual, orthographic word (the target) when they related to American accents. If such a priming effect existed this would suggest that when an General American accent is processed, its social meaning is also automatically activated, either as a part of the same mental representation, or because the two are very closely connected. The fact that this was not the case suggests that the two are separately stored and that one can be activated without necessarily activating the other.

Furthermore, the experiments aimed to test whether there was a difference in priming effects depending on whether the prime stimulus was a real word or a non-word. If priming effects had been found when caused by real word primes but not by non-word primes this could have indicated that the mental representations used to process social meaning are word-sized, rather than sub-lexical representations (for a further investigation of the use of non-words to investigate the use of sub-lexical representations see Chapter 6). Similarly it was explored whether there was a difference between priming effects on response times to targets which were a part of the General American variety (lexical variants such as SIDEWALK) and targets words which were associatively connected to its speakers (lexical items such as GUNS). This would shed more light on what information would be a part of the context of socio-phonetic information or closely connected to it. As priming effects were absent across the board, the experiments were not able to provide new insight into this.

The experiments also cast a relatively wide net methodologically as 'twin' experiments: they were identical experiments, with only one methodological difference. This was the inter-stimulus interval (ISI) – the time period between being exposed to a prime and a target. Different ISIs have been found to influence the presence and size of priming effects, with experiments with short ISI conditions (<400ms) showing priming effects attributed to spreading activations, and experiments with longer (>400ms) showing priming effects more likely to be caused by (more strategic) expectancy-based mechanisms (De Groot, Thomassen & Hudson 1986; Del Toro 2000; Carter et al. 2011). As the use of semantic priming experiments to measure accent activation is novel, it was uncertain which of these mechanisms was more likely to underlie a potential effect.

4.2 Methodology

4.2.1 Experimental design

The methodology used in the experiments described in this chapter, like in Chapter 3, is what I am calling socio-contextual priming, a variation on semantic priming where the connection between prime and target is the social context of linguistic variation. This functions similarly to semantic priming proper, which tests whether the response to a target stimulus (e.g. DOG) is influenced by a semantically related prime stimulus preceding it (e.g. *cat*). For example, in many semantic priming experiments it has been found that participants are either quicker or slower to respond to a stimulus word after being exposed to another stimulus which is semantically related (McNamara 2005; Mayr & Buchner 2007). As discussed in Section 3.2, in spreading activation models of semantic priming such effects imply that activation of the prime stimulus in the brain has spread to the target stimulus to such a degree that it influences responses to the target stimulus (Collins & Loftus 1975; Anderson 1983), whilst compound-cue models would account for this as the prime being a part of the larger (context-included) image of the target in memory (Gillund & Shiffrin 1984). In the case of socio-contextual priming, these priming effects would suggest a close connection between the two types of information in the spreading activation models, or that the two types of information are a part of each other's context, and stored in the same mental representation.

In both the experiments in this chapter, responses were elicited through a lexical decision task which included words as well as non-words. Native speakers of British English were asked to classify these as real words or nonsense words. Differently from the experiment in Chapter 3, the current experiments are cross-modal (like Holcomb & Anderson 1993; Carter et al. 2011), using auditory primes rather than orthographic ones, preceding the still orthographic targets. The auditory primes were sound clips of speakers with either a General American or SSBE accent pronouncing words and non-words. They preceded target words which were words closely associated with Americanness (e.g. RELIGIOUS) or words that are only typically used in American Englishes (e.g. DIAPER).

The design of the experiments was reduced in complexity (following from the issues found in Chapter 3, and following White (2019)): the current experiments used only two accents, they randomised the order of the items to remove order as a potential effect, they prevented any prime target combinations with overlapping onsets, and only targets associated with one variety were used. Furthermore, the strength of the associations between the targets and Americans and American accents was more homogeneous across targets because these were selected on the basis of the pre-test, which asked British participants about their associations with Americans and the American accent.

The main research question of both experiments was to test whether processing sublexical accent information implied the automatic activation of accent stereotypes and lexical variants which are part of the same variety (General American). This is summarised in research question 1 below:

1. Do listeners respond differently (quicker, slower, or more accurately) to target words related to American accents if they have been primed by an auditory prime, pronounced in an American accent?

The secondary research questions explored whether there were differences in any found priming effects depending on what relationship a target word had to the prime accent (2), and whether the prime stimulus was a word or a non-word (3). The former would be able to shed light on what information is stored as a part of the context of socio-phonetic information or closely connected to it (accent stereotypes and other aspects of the variety), whilst the latter would be able to shed light on the mental representations involved in the processing of social meaning (discussed in more detail in Chapters 5 and 6).

2. Does it make a difference whether the target word is a concept that is associated with Americans and the United States (such as 'GUNS' or 'LOUD') or a lexical variant of General American (that is not present in British English e.g. 'SIDEWALK')?
3. Does it make a difference whether the prime stimulus is a word or a non-word?

The final research question compared the two separate twin experiments, which differed only in the ISI between prime and target:

4. Does it make a difference whether primes and targets were presented without an ISI, or with an ISI of 800ms?

Comparing the two experiments would allow for more insight into what a potential measuring of the activation of social meaning relied on: joint or spreading activation or expectancy-based mechanisms (Del Toro 2000). If priming effects only occurred with an ISI of 800ms, this would suggest that expectancy based strategies play a role in facilitating a link between the two types of information. In that case the two types of information would be stored separately and need to be linked.

In summary, the experiments tested whether General American accented audio stimuli preceding responses to American accent related target words influenced response times to those target words. It tested this across target conditions (accent associations or variety specific words) and prime conditions (words or non-words), as well as procedural parameters (0ms ISI or 800ms ISI).

4.2.2 Procedure

The priming experiments were hosted through the online experimental software Gorilla (Anwyl-Irvine et al. 2019). 113 suitable participants (described below) took part in the 0ms ISI experiment, and 35 in the 800ms ISI experiment. Participants of the former experiment were recruited through Twitter and Facebook, whilst participants of the latter experiment were recruited through a University of Sheffield mailing list for students, as well as a smaller effort to distribute the experiment through Twitter. The participants were presented with an online participant information sheet about the ethics of the experiments and were asked to confirm that they consented.[12]

After filling out a short demographic survey, the participants were told to classify written words as either real words or nonsense words by selecting either the F-button on their keyboard (for real words) or the J-button on their keyboard (for nonsense words). They were told they would hear audio before each word but to not respond to this. They were given a practice round to learn how the experiment worked, and were told at the end of the practice round how many items they got right. After this, they proceeded to the main experiment, which contained 72 items, with a break in the middle. They were told how many items they got right both in the middle and at the end of the experiment.

Each item consisted of either three or four parts, depending on the inter-stimulus interval. In all cases, first a fixation cross was shown for 250ms, which was followed by the auditory prime. In the 0ms ISI experiment, this was then immediately followed by the orthographic target. In the 800ms ISI, a fixation cross was shown for 800ms, before the target was shown. If a participant did not respond to the target within 4000ms, the screen would move on to the next item.

4.2.3 Stimuli

The experimental items that were used included 24 different auditory prime stimuli and 24 orthographic target stimuli. The auditory primes were either pronounced in a General American accent (related to the American targets) or an SSBE accent (the control group, unrelated to the American targets). Half of each were words, and half of each were non-words, in order to be able to compare between priming effects caused by words and non-words. The target words were all words relating to Americans. 12 of the target words were words that were commonly associated with Americanness by participants in a pre-test which asked about these associations. The other 12 of the target words were words typical for American Englishes which were most often mentioned as American English words by participants in the same pre-test. This is summarised Table 9 on the next page.

The primes and targets were combined randomly, whilst avoiding matching initial letters and sounds (due to the confounding onset effects found in Chapter 3), in 24 different versions of the experiments (the combinations were kept the same across both the 0ms ISI and the 800ms ISI experiment). These randomised lists were automatically created by means of a Python script. The script ensured that each of the target words had a relatively even balance of General American primes and the SSBE control primes. This ensured that there would be enough datapoints for targets both in the related prime conditions (General American accent) and the unrelated control conditions (SSBE accent). Each of the

[12] The experiments went through the University of Sheffield's ethical review process and was approved.

targets was combined with a General American prime at least 9 times and at most 15 times (and vice versa).[13]

Table 9. Number of primes and targets by type.

<table>
<tr><th colspan="4">Auditory primes</th><th>randomly assigned to</th><th colspan="2">Orthographic targets</th></tr>
<tr><td rowspan="2">Gen. Am. (related)</td><td rowspan="2">12</td><td>Words</td><td>6</td><td rowspan="4"></td><td rowspan="2">Associations</td><td>6</td></tr>
<tr><td rowspan="2">Non-words</td><td>6</td><td>6</td></tr>
<tr><td rowspan="2">SSBE (unrelated)</td><td rowspan="2">12</td><td>6</td><td rowspan="2">Variety-specific words</td><td>6</td></tr>
<tr><td>Words</td><td>6</td><td>6</td></tr>
<tr><td colspan="7">Examples</td></tr>
<tr><td rowspan="2">Gen. Am. (related)</td><td rowspan="2">12</td><td>work</td><td>6</td><td rowspan="4"></td><td>ARROGANT</td><td>6</td></tr>
<tr><td>zor</td><td>6</td><td>BIG</td><td>6</td></tr>
<tr><td rowspan="2">SSBE (unrelated)</td><td rowspan="2">12</td><td>twore</td><td>6</td><td>DUMPSTER</td><td>6</td></tr>
<tr><td>door</td><td>6</td><td>DIAPER</td><td>6</td></tr>
</table>

In addition to the critical stimuli, 24 items were included which had non-word targets (sourced from Rastle, Harrington & Coltheart (2002)), combined with different SSBE and General American prime stimuli, in order to create a balanced mixture of word and non-word items, to facilitate the lexical decision task. As at this point all target words were in some way related to Americanness, another 12 items were included that were real word targets that were unrelated to Americanness, as well as another 12 items which were non-word targets, bringing the total number of items to 72. This was to prevent participants from becoming aware that the experiments tested processing of words relating to American accents.

Auditory prime stimuli

Speakers

The prime stimuli were all taken from a larger set of audio recordings also used for the experiments described in Chapters 5 and 6. This set of recordings consisted of reading list recordings of 92 speakers including SSBE, General American speakers (included in the current experiment and those of Chapters 5 and 6) as well as Yorkshire English speakers and speakers of a range of other accents (only used in Chapters 5 and 6).

From these 92 speakers I selected those who were as homogeneous as possible in their realisations of the recognisable sublexical variants used in the experiments, as well as their regional origin, for as far as this was possible. The speaker characteristics are summarised in Table 10 below. For the American group only General American voices were used, and for the Standard English group SSBE voices were used. As two of the SSBE speakers were not themselves from the South of England, their accents were perhaps technically ‘regionless Received Pronunciation’ (cf. Halfacre 2019), but they did not differ from their SSBE peers in the pronunciation of their stimuli (by my own auditory judgement). Similarly, one of the General American speakers was from the South of the US but spoke with a General American accent. The group of speakers was made to be as mixed as possible for gender, to create as much variety in voices as possible.

[13] The automatic randomised creation of 24 stimulus lists where numbers were perfectly balanced was beyond the the capabilities of the computer I used, due to the very small chance that this balance would be created in all 24 lists in the same randomisation.

Table 10. Summary of speaker demographics for all speakers with critical stimuli, by accent and by potentially recognisable sublexical variant.

Accent	*Sublexical variant*	*Speaker location during youth*	*Age*	*Gender*	*Education*
General American	LOT vowel	South Carolina, US	20s	Female	University
		Rhode Island, US	20s	Female	University
		Arizona, US; England	20s	Male	University
	Rhoticity	Outside UK/US; Washington D.C., US; California, US	40s	Female	University
		New York, US; Massachusetts, US; Outside UK/US	20s	Female	University
		Michigan, US	20s	Female	University
SSBE	BATH vowel	Berkshire	20s	Male	University
		Kent	20s	Male	University
		Cumbria	20s	Female	University
	Lack of rhoticity	Hampshire	20s	Male	University
		Lincolnshire	20s	Male	University
		Surrey	40s	Female	University

Recognisable sublexical variants

The phonetic variants in the single-word stimuli used in the experiments were selected to be well-recognisable as part of their variety. In this thesis, I use 'recognisable sublexical variant' to describe variants of a sublexical – sound – variable which are systematically (even if not always ubiquitously) realised in a way that is different in one variety of English from at least some other varieties of English, and therefore could be recognised. Below I discuss the recognisable sublexical variants used in the auditory prime stimuli in the experiments in this chapter.

General American: Rhoticity

General American is uncontroversially a rhotic variety of English (Wells 1982: 125–126; Svartvik & Leech 2006: 163–164; Trudgill & Hannah 2002: 38,40), despite rhoticity not being present in all varieties of American English. This contrasts with most varieties of English in England, including SSBE and most (although not all) Yorkshire varieties of English (Wells 1982). Whilst there is not a large body of research on non-linguist British listeners' perceptions of General American, rhoticity seemed like a very recognisable feature of General American for British listeners, as it involves the presence of an additional sound. This was confirmed in the accent recognition tasks described in Chapters 5 and 6, where rhoticity was well recognised.

General American: LOT vowel

The vowel in the LOT lexical set in General American is unrounded (Trudgill & Hannah 2002: 39; Wells 1982: 123–124), and is described by Svartvik & Leech (2006: 164) as a 'longish /ɑ/', which is in contrast with SSBE [ɒ]. As with General American rhoticity, there is no body of research on non-linguist British listeners' perceptions of General American on which to base expectations of ease of recognition. Nonetheless, findings from the accent recognition tasks in Chapters 5 and 6 confirmed that it is a strong cue.

SSBE: BATH vowel

The first selected recognisable feature for SSBE was the BATH vowel. The presence of a split between the vowels in the BATH lexical set and the TRAP lexical set is one of the main things that sets Southern British Englishes, including SSBE, apart from other British accents (Wells 1982: 353–356; Gupta 2005; Leemann, Kolly & Britain 2018). As discussed further in Chapter 5, the SSBE realisation of BATH as [ɑː] has very strong associations with Southern Britishness. This makes it highly recognisable, especially in contrast to American Englishes where it is absent.

SSBE: Lack of rhoticity

Finally, the last recognisable feature used for the critical stimuli in SSBE was the lack of rhoticity, as a counterpart to General American presence of rhoticity. As with the above sublexical variants, this was well recognised in the accent recognition tasks in the following chapters.

Selected stimuli

Half the General American critical stimuli included rhoticity, and the other half included LOT vowels. For the SSBE critical stimuli, half the stimuli included a lack of rhoticity and the other half a BATH vowel. The first 24 non-word target items included more General American and SSBE prime stimuli. The other 24 distractor items (12 word targets, 12 non-word targets) were preceded by prime stimuli from the ‘Other accent’ audio clips, collected using the same methodology as described for the General American and SSBE stimuli. The distractor stimuli included Tyneside, Scottish and Liverpool English clips. Half of these were words and half of these were non-words. One speaker’s voice would be used a maximum of 2 times in the critical stimuli, and a maximum of another 2 times in the distractor stimuli.

Target stimuli

In order to improve on the semantic priming experiment described in Chapter 3, the experiments selected their target words on the basis of a pre-test, which asked a total of 45 native speakers of British English (mean age = 32, SD = 13)[14] to list their associations with Americanness and American accents. The participants were asked the following questions:

1. What are the first **7 words** that come to mind when you think of words that **Americans would use**, but **British people would not** generally use? If you can think of more, please feel free to add them as well. (Please separate your answers with a comma.)
2. What are the first **5 things** that come to mind when you think about **America**? If you can think of more, please feel free to add them. (Please separate your answers with a comma.)
3. What are the first **5 things** that come to mind when you think about **American people**? If you can think of more, please feel free to add them. (Please separate your answers with a comma.)

The answers to the bottom two questions were combined in processing, as there was considerable overlap (e.g. ‘racism’ for associations with America as well as American people). The answers were automatically processed in Python by separating all answers at commas and indentations, all converted to lower case, and all spaces were removed. The most frequent entries were listed in descending order and the 12 most frequent items were selected for the associations with America and American people combined, and the 12 most frequently listed American English lexical items. Any names were removed, as well as words or phrases with spaces in them. Furthermore, American English items which also occur in British English, but with a different meaning, were removed. This resulted in the following list of target words, as shown in Table 11:

---

[14] Other than native speaker status, nationality and age these participants were not asked further demographic questions.

Table 11. Selected target words and how frequently they were listed in the pre-test.

Accent-related concepts		**American lexical variants**	
Target	*N*	*Target*	*N*
Loud	21	Sidewalk	36
Guns	18	Diaper	15
Friendly	15	Eggplant	15
Racism	10	Trash	11
Confident	9	Faucet	10
Patriotic	8	Garbage	9
Religious	7	Soccer	9
Enthusiastic	6	Highway	8
Arrogant	5	Candy	7
Fat	5	Restroom	7
Flag	5	Dumpster	5
Big	4	Mall	5

These were used as the target stimuli in the experiments. A full list of all stimuli used in the experiments described in this chapter can be found in Appendix II.

4.2.4 Participants

Experiment 1: ISI = 0ms

138 participants took part in the first experiment where the inter-stimulus interval was 0ms. 113 of these were native speakers of British English who had not lived in the United States. These participants were almost exclusively university students and alumni, as can be seen in Table 12 below. Roughly a third of the participants had a background in linguistics, as can be seen in Table 13. They were still included in the experiment, as conscious knowledge of linguistics was not assumed to influence the subconscious processes involved in priming experiments, although they may have been slightly more likely to understand the purpose of the experiment if they happened to be familiar with priming techniques. The age distribution of the participants is shown in the histogram in Figure 10.

Table 12. Participant educational background for experiment 1.

Education	**Count**	**Percentage**
University	107	95%
College (UK)	4	4%
Secondary school	1	1%
Other	1	1%
Total	113	100%

Table 13. Participant linguistic background for experiment 1.

Linguistic background	**Count**	**Percentage**
No	70	62%
Yes	42	37%
Prefer not to say	1	1%
Total	113	100%

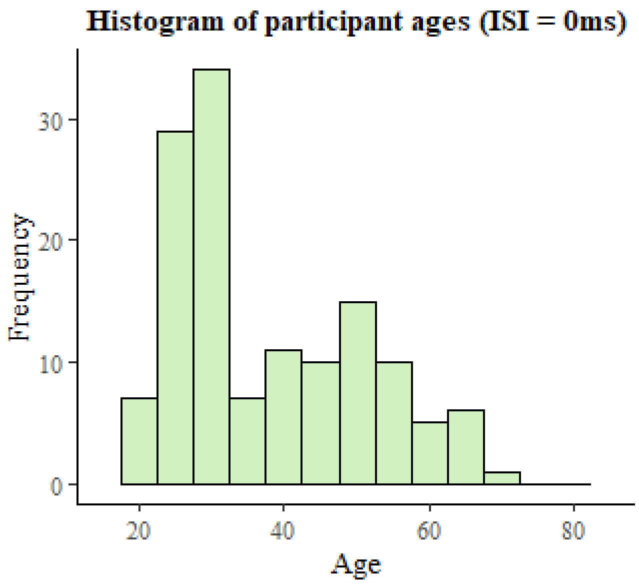

Figure 10. Histogram of participant ages for experiment 1.

Experiment 2: ISI = 800ms

The sample for experiment 2 was relatively similar, albeit considerably smaller, with most participants having a university background (as can be seen in Table 14) and in this case roughly a quarter of participants having a background in linguistics (Table 15). This sample, however, was much younger and showed less of a spread in ages, likely due to the nature of the distribution, which was mostly through a university student volunteer list (Figure 11). This means that in the statistical analysis it could have been slightly more difficult to disentangle effects caused by the different ISIs from effects caused by age (although this was not the case, as can be seen in Section 4.3 below.)

Table 14 Participant educational background for experiment 2.

Education	**Count**	**Percentage**
University	34	97
College (UK)	1	3
Total	35	100%

Table 15. Participant linguistic background for experiment 2.

Linguistic background	**Count**	**Percentage**
No	26	74
Yes	9	26
Total	35	100%

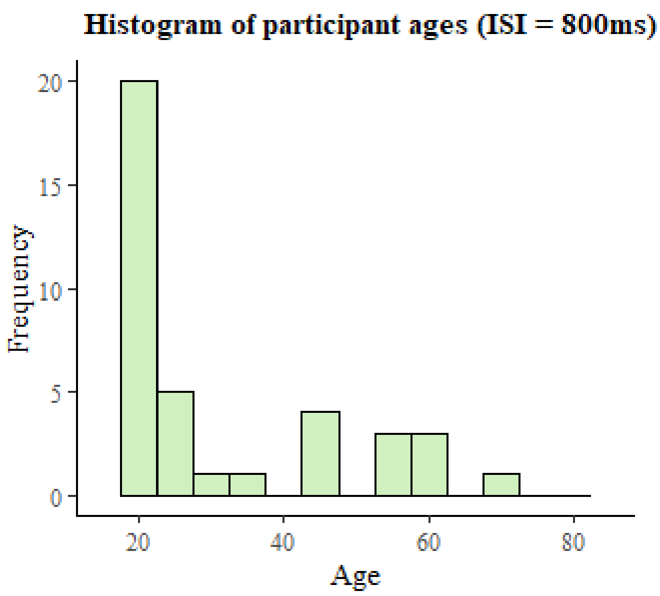

Figure 11. Histogram of participant ages for experiment 2

4.3 Results

4.3.1 Overview

Neither of the experiments found a priming effect that indicated that American accented audio stimuli preceding responses to American-accent related target words influenced responses to those target words. This was true across target conditions (accent associations or variety specific words) and prime conditions (words or non-words), as well as experimental parameters (0ms ISI or 800ms ISI). This means there is a lack of support in the findings of the current experiments for social and linguistic information being stored in the same mental representations.

4.3.2 Data processing

In order to process the data in the experiment, the mean accuracy for each of the prime-target items was checked to see if any were responded to poorly. None of the items were responded to with an accuracy of below 90%. Next, all participants who were native speakers of British English were selected, unless they had lived in the United States for a year or longer. Then any participants were removed whose accuracy was 2.5 standard deviations lower than the mean for either the practice round or the full task (N = 8 for the first experiment, N = 3 for the second). This is the data set that was used to measure response accuracy.

In order to look at the response time accuracy, however, incorrect responses were disregarded. Then the distribution of response times for both experiments was explored, which showed that there was a group of very quick responses, somewhat separate from the main bulk of responses, as can be seen in Figure 12.

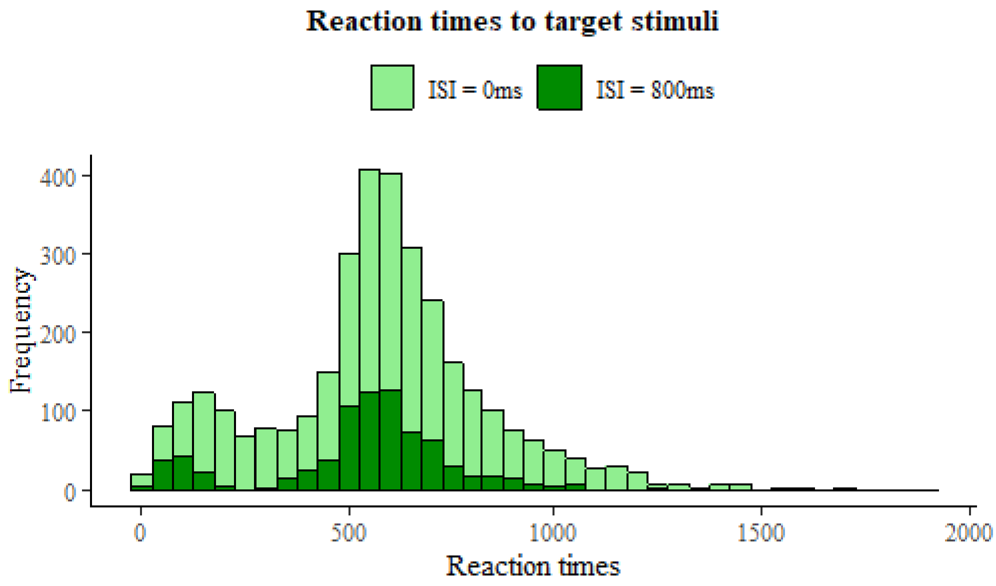

Figure 12. Reaction times to target stimuli.

This suggested that some participants were not always responding to the target stimulus but instead responding to the preceding prime stimulus (even when after an 800ms ISI). To be able to distil which participants were doing this and which ones were not, I did not just look at participants' mean response times, but also at their standard deviations. If this was very high, it could be due to participants sometimes responding to the prime stimulus (very low response times), and sometimes to the target stimulus (normal response times). This would indicate that their data is unreliable, and not measuring the intended response time. Figure 13 plots participant standard deviations against their mean response times. It confirms that there seem to be such participants in the dataset who behave differently from the group always responding to the target stimuli.

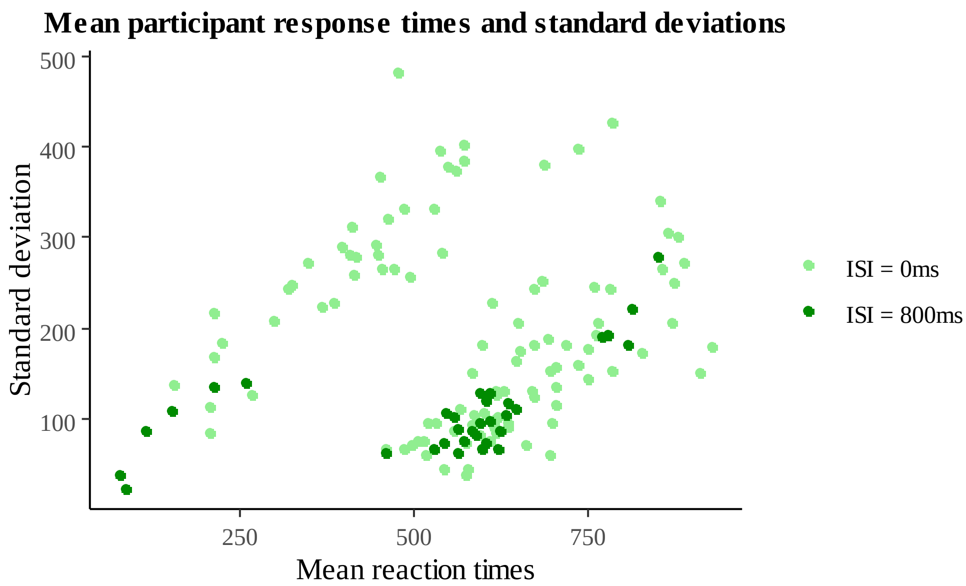

Figure 13. Mean participant response times by standard deviation.

There are two distinct diagonally distributed groups in this plot: the left diagonal cloud in the plot starts in the left bottom corner with very low mean reaction times and similarly low standard deviations. These are the participants (almost) exclusively responding to the prime stimulus.

As the mean reaction times go up within this diagonal cloud, the standard deviations go up rapidly as well, implying that even as some of the participants in this cloud have more and more normal response times (of around 600ms) they still have enough very low response times (likely to the prime rather than the target) to create large standard deviations. The right diagonal cloud shows participants with more expected reaction times of around 600ms and standard deviations of mostly 200ms.

This suggests that there are two separate groups of participants in the sample: some who responded very early (presumably to the previous stimulus) a considerable amount of the time, and some who just responded to the target stimuli. This pattern is confirmed when we look at a distribution of values created by subtracting one standard deviation from a participant's mean reaction time, as an indication of what the lower (but frequent) range of response times was for a participant. This is shown in the histogram in Figure 14, showing two separate groups.

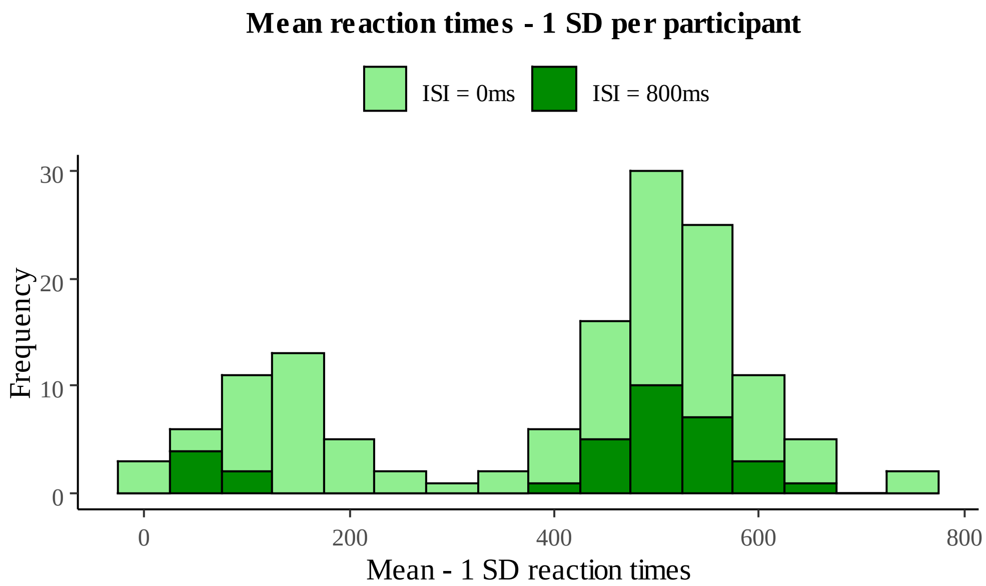

Figure 14. Distribution of participant mean reaction times minus one standard deviation, as an indication of the lower ranges of participant response times.

In order to take out these participants responding to the wrong stimulus, any participants were removed who had a lower mean value than 300ms for this measure. This left 71 participants in the 0ms ISI condition, and 27 in the 800ms ISI condition. It yielded the distribution of reaction times shown in Figure 15 on the next page.

After this, all reaction times more than 2 standard deviations away from the mean were removed to exclude outliers.

As in the previous chapter, logistic and linear mixed effects regression models were run for the accuracy and response time data respectively. The models were fitted by starting out with maximal models and reducing these by removing fixed effects which did not improve these models significantly (tested by means of ANOVAs). This was carried out using the 'lme4' package (Bates et al. 2015). The data points from both experiments were combined in a single model, as the only difference between the two experiments was the ISI.

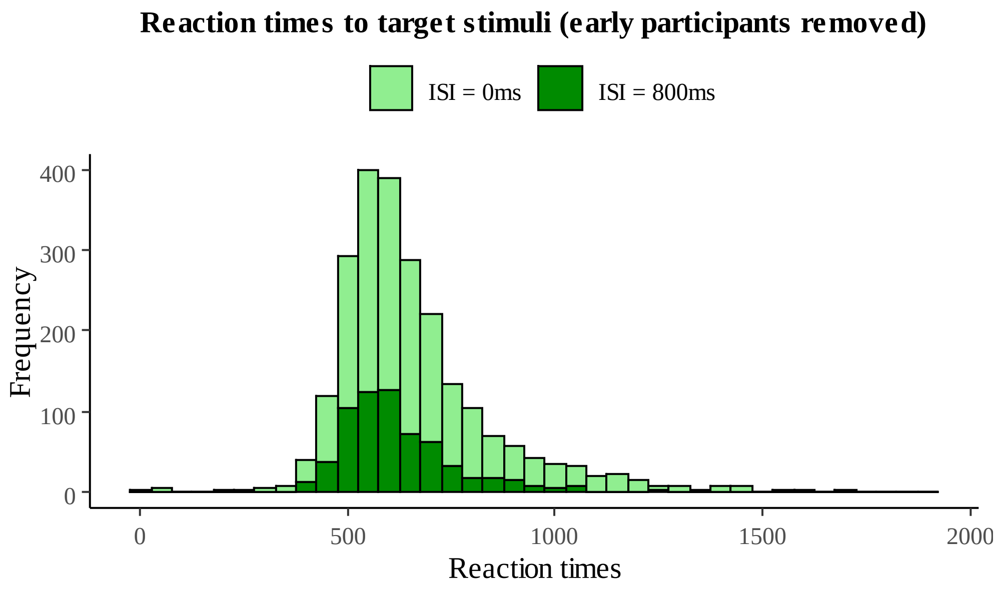

Figure 15. Distribution of reaction times to target stimuli, with the participants with unrealistically early responses removed.

4.3.3 Accuracy

Participants were highly accurate in the lexical decision task, with a mean accuracy of 98% and a standard deviation of 3 percentage points. A logistic mixed-effects regression model was used to measure priming effects on the accuracy of participants' responses to the target words. All numerical values were log-transformed. The starting maximal model included the accent of the prime, the type of target, the ISI, and whether the prime stimulus was a word or not. These were all fixed effects relating to the research questions of this chapter. The other fixed effects that were included as control variables were the strength of the association (measured by the number of participants mentioning the word in the pre-test), the age of the participant, the length of the target word, the trial number, the frequency of the target word (in the BNC), and its comparative frequency between the BNC and the Corpus of Contemporary American English. The code for this in R was the following:

```
ModelAccuracy <- glmer(CorrectResponse ~ PrimeAccent * TargetType *
ISI * TargetRelationshipStrength + WordStatusPrime + log(Age) +
log(TargetLength) + TrialNumber + log(FrequencyBNC) +
log(FrequencyRatioBNCvsCOCA) +
      (1 | Participant) + (1 | TargetWord) + (1| PrimeWord) ,
      family="binomial", data = dat, control =
      glmerControl(optimizer = "bobyqa",
      optCtrl=list(maxfun=100000)))
```

This model was a singular fit. This means that there were not enough datapoints for all of the variables to be included. Because of this it was reduced to exclude prime word as a random effect, as well as target relationship strength (the number of times a target was mentioned in the pre-test) as well as the ratio between the target word's frequency in the British National Corpus and the Corpus of

Contemporary American English, until the fit of the model was not singular. These effects were removed first as they were they least immediate to the chapter's research questions.

Then, fixed effects were removed if they did not improve the fit of the model. This was the case for all predictors, including the accent in which the prime was pronounced (Chisq. = 1.78, *p* = 0.18). This means there was no priming effect on participants' accuracy in the lexical decision task.

4.3.4 Response times

Linear mixed-effects regression models were fitted to measure differences in response times depending on the accent in which prime words were pronounced in. The code for the maximal model was as follows:

```
ModelRT <- lmer(log(ReactionTime) ~ PrimeAccent * TargetType * ISI *
TargetRelationshipStrength * WordStatusPrime + log(Age) +
log(TargetLength) + log(TrialNumber) + log(FrequencyBNC) +
log(FrequencyRatioBNCvsCOCA) +
        (1 | Participant) + (1 | TargetWord) + (1 | PrimeWord), data =
        ReactionTimes)
```

In order to prevent a singular fit, the random effect of prime word was removed, as well as target relationship strength. None of the interactions improved the model, and all of them were removed: firstly, the interaction of reaction time and the word status of the prime did not improve the fit of the model (Chisq. = 2.33, *p* = 0.93), which means that it did not matter whether the prime was a word or a non-word. Furthermore, the interaction between the ISI and the accent of the prime did not improve the model, which means that there was no difference in how much priming there was between the two twin experiments (Chisq. = 3.43, *p* = 0.32). Furthermore, the interaction between the prime accent and the type of target (association or variety specific word) did not improve the model's fit and was removed (Chisq. = 1.18, *p* = 0.27). Then, fixed effects were removed one by one if they did not improve the fit of the model. Crucially, this included the accent in which a prime was pronounced (Chisq. = 1.95, *p* = 0.16). This means that participants did not respond differently to American English related targets after hearing an American English accent than after hearing SSBE and that there was no statistically significant priming effect. The mean response times and priming effects for each of the conditions are summarised in Table 16 on the next page.

The distribution of response times is further illustrated in Figures 16 and 17 below. Figure 16 shows boxplots for response times to the target words by whether they were primed or not (the prime accent) and whether this prime was a word or a non-word (word status) across ISI conditions. Comparing the boxplots with unbroken lines and dotted lines (related primes versus unrelated primes) within their conditions, no real differences can be seen. The same is true for Figure 17 on the next page, which shows differences between response times in related and unrelated prime conditions, by the target type for both the 0ms and 800ms ISI experiments.

Table 16. Mean response times to target stimulus by target type, ISI, prime word status and prime accent.

Target Type	ISI (ms)	Prime	Prime Accent	Mean RT (ms)	SD (ms)	Priming effect (ms)
Association	0	Non-word	SSBE	648	150	
Association	0	Non-word	American	617	147	31
Association	0	Word	SSBE	633	161	
Association	0	Word	American	619	141	13
Association	800	Non-word	SSBE	585	93	
Association	800	Non-word	American	587	119	-2
Association	800	Word	SSBE	594	113	
Association	800	Word	American	597	118	-3
Variety-Specific Word	0	Non-word	SSBE	675	163	
Variety-Specific Word	0	Non-word	American	678	158	-3
Variety-Specific Word	0	Word	SSBE	670	154	
Variety-Specific Word	0	Word	American	654	152	16
Variety-Specific Word	800	Non-word	SSBE	638	125	
Variety-Specific Word	800	Non-word	American	649	135	-11
Variety-Specific Word	800	Word	SSBE	638	136	
Variety-Specific Word	800	Word	American	634	130	4

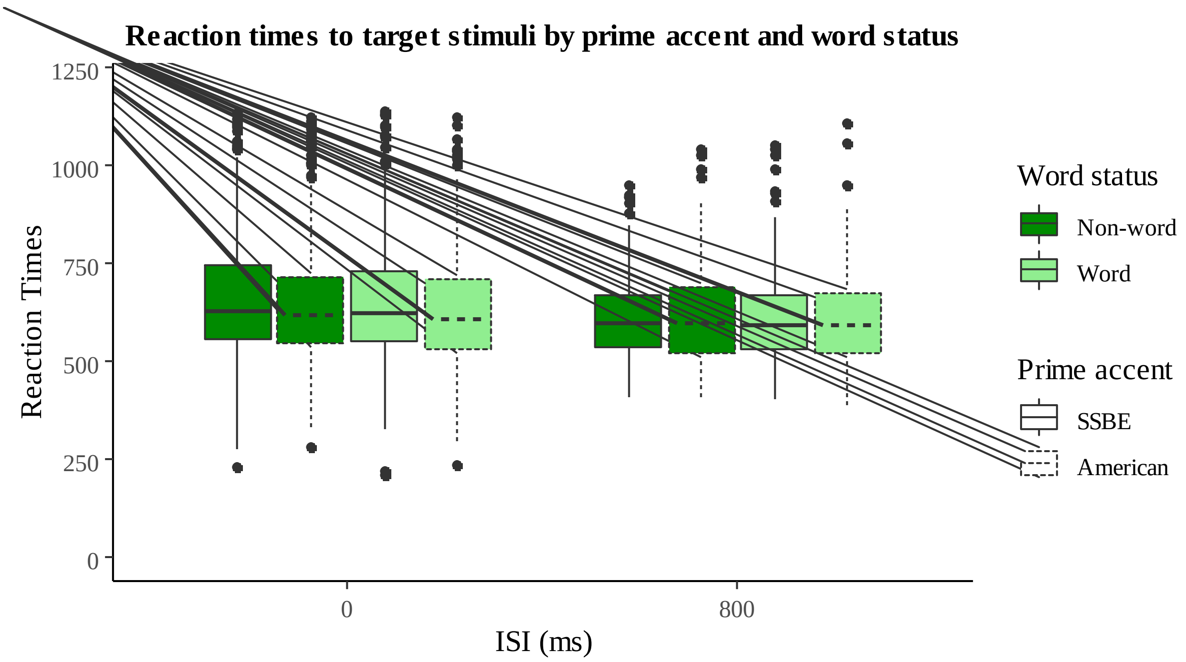

Figure 16. Reaction times to target stimuli by prime accent and prime word status in both the 0ms and the 800ms experiment.

The fixed effects in the linear mixed-effects model that did influence response times are summarised in Table 17 below, created by means of the sjPlot package in RStudio (Lüdecke 2019). It shows that older participants are slower. It also showed that responses were slightly faster if they occurred later in the experiment, and that response times were somewhat faster if the target word was more frequent. This is illustrated in Figures 18, 19, and 20 which plot response times against age, lexical frequency, and trial number.

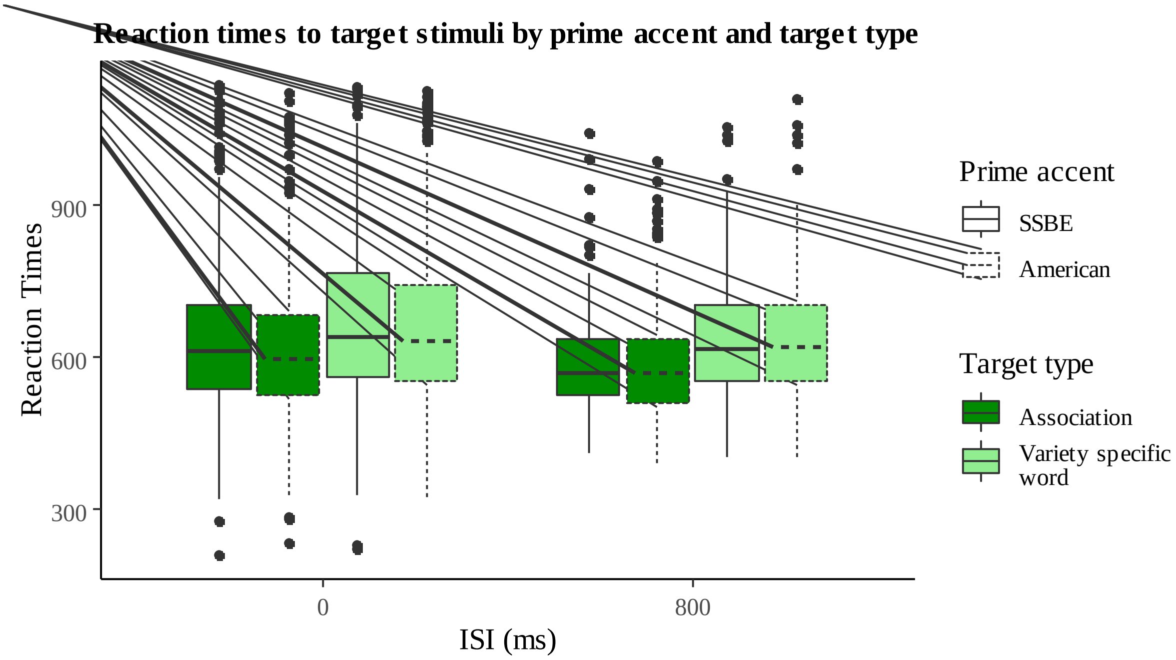

Figure 17. Reaction times to target stimuli by prime accent and target type in both the 0ms and the 800ms experiment.

Table 17. Overview of fixed effects in best-fit linear mixed-effects regression model for log-transformed response times to target words.

Predictors	**Estimated effect on log-transformed RT**	**Confidence interval**	***p*-value**
(Intercept)	6.12	5.38 – 6.41	**<0.001**
Age (log)	0.13	0.05 – 0.21	**0.002**
Trial number	-0.001	-0.001 – -0.001	**<0.001**
Frequency BNC (log)	-0.02	-0.03 – -0.01	**<0.001**

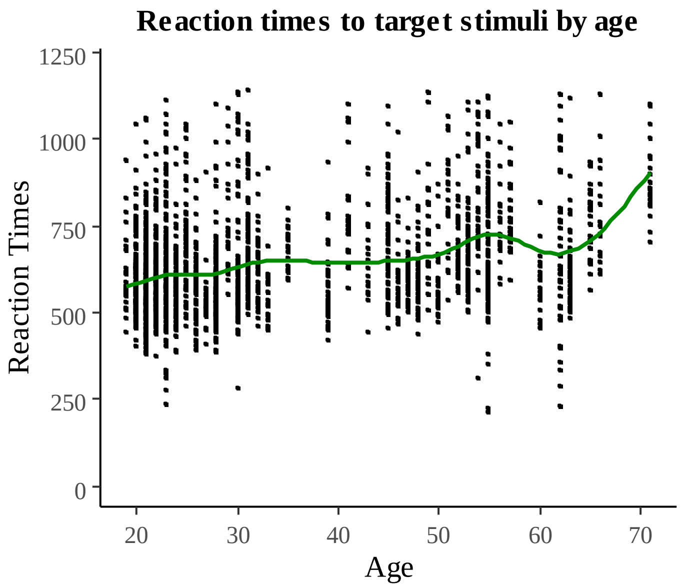

Figure 18. Reaction times to target stimuli by age.

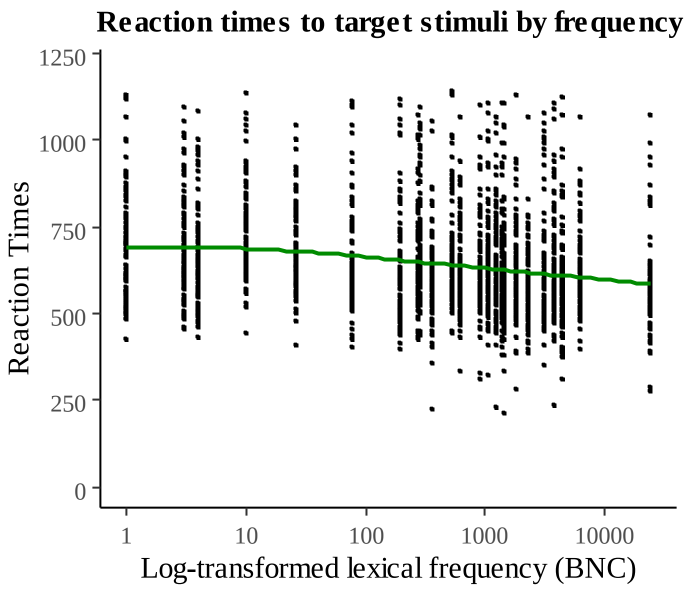

Figure 19. Reaction times to target stimuli by lexical frequency.

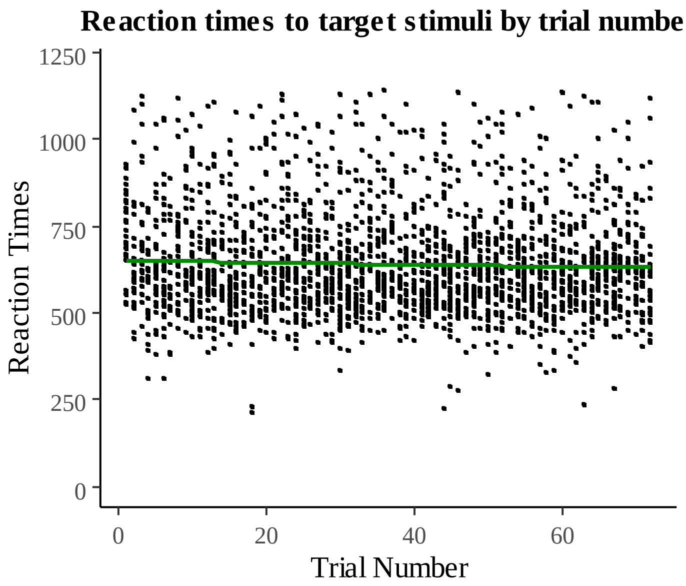

Figure 20. Reaction times to target stimuli by trial number.

The reduction of the models above to achieve the best-fit model excluded the interaction between prime accent and the strength of the associations between Americanness and the target word as found in the pre-test, because this caused a singular fit. In order to still examine whether the closeness of association (a potential methodological issue, as discussed in Chapter 3) influenced priming effects, a model was fitted which included an interaction between the target association strength and whether the prime was in a related accent. This interaction was added to the best-fit model found above. The modified model is summarised below:

```
ModelRT <- lmer(log(ReactionTime) ~ PrimeAccent *
TargetRelationshipStrength + log(Age) + log(TrialNumber) +
log(FrequencyBNC) + log(FrequencyRatioBNCvsCOCA) +
    (1 | Participant) + (1 | TargetWord), data = ReactionTimes)
```

This interaction did not improve the fit of the model either (Chisq. = 0.30, p = 0.86).

As the priming experiments above did not show any clear priming effects, I attempted to uncover whether this may have been caused by priming effects happening in opposite directions, with some participants showing facilitatory (faster response times) and some showing inhibitory (slower response times) priming effects. In order to do so a mixed-effects linear regression model was fitted which put the prime accent as a slope on the random effect of participants. This means the model created coefficients (indications of how much slower or quicker data points are deduced to be on the basis of a variable) for how the prime accent affected different participants' response times. This was done on the basis of a slightly larger dataset where incorrect items and outliers were not removed, as the cleaned up dataset yielded a singular fit even with the smallest possible model with a slope; for the same reason, the model itself was kept minimal, as is summarised below:

```
ModelRTWithSlope <- lmer(log(ReactionTime) ~ PrimeAccent +
    (1 + PrimeAccent | Participant) + (1 | TargetWord), data =
    ReactionTimesIncludingOutliers)
```

The distribution of the random slopes for prime accent for each participant is plotted below, in Figure 21. As there is no clear binomial distribution here, there is no evidence that there were different participant groups exhibiting two different forms of priming, cancelling out each other's effects.

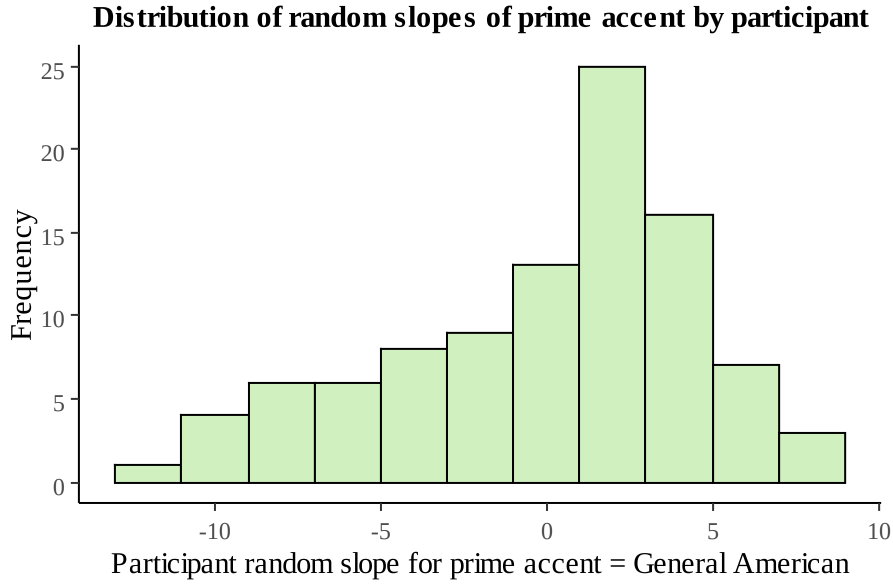

Figure 21. Distribution of the random slopes for prime accent for each participant for the prime accent being General American.

4.4 Discussion

The twin priming experiments described in this chapter did not find any priming effects as the result of participants hearing a General American accent before responding to a target word relating to an American accent. This means that the experiments provide no support for the idea that social and linguistic information are stored in the same mental representations. As there were no priming effects overall, the results are not able to answer the further, more detailed, questions about any differences in priming effects between non-word primes and word primes, or the type of target that followed the primes.

Taken together with the findings from Chapter 3, where some priming effects were found between social and linguistic information, but in different directions, the findings in this chapter suggest a looser link between social information and linguistic information than the concept of exemplars in exemplar theoretical speech processing (Johnson 2006; Docherty & Foulkes 2014) implies. Instead the findings in these two chapters fits third-wave approaches to social meaning better, which posit a more fluid link between the two (Eckert 2008; Acton 2021), and suggests that listeners in a hybrid exemplar model have separate abstract representations of social meaning and linguistic information.

This comes with a number of caveats. Firstly, it is statistically not possible (with the traditional frequentist statistical methods used in this chapter) to find evidence for a lack of effects. This means that the strongest claim that can be made about the findings in this chapter is that there is no support for the storage of social and linguistic information within the same representations, rather than that there is evidence against it.

Another caveat is the potential distance between primes and targets used in the experiment. Even if the target words used in the current experiments were carefully selected and carried the strongest associative links between accent and accent-related concepts and lexical variants (as was done through the pre-test survey), it may be that the contextual information linked to linguistic utterances that listeners store in memory is different from these specific words. For instance, maybe it is not the word LOUD that is part of the contextual information of American accent features, but a non-verbal concept of loudness.

Still, for the regional lexical variants which were used as targets (such as SUBWAY and DIAPER), it would be those exact verbal forms which may be a part of the direct context of such accent features. Here, however, the question may be whether those lexical variants are frequent and important enough to be a part of the context of stored memories of American utterances.

A final caveat is that the experiments used a total of 24 targets which related to American accents and Americans. This may be too wide of a spread. In most semantic priming experiments, a prime is only combined with one or two highly related targets, as discussed in Section 3.4. This may mean that the 24 used in the experiments are either not closely related enough, or fall outside of the stored context of exemplars (or more abstract representations uniting linguistic information and social context). In this case there may be some social information that is stored within linguistic representations, which these experiments were not able to uncover. There is no suggestion of this from the current experiments however, as there were no effects of the strength of association between prime and target on the accuracy or response times to the targets.

4.5 Conclusion

This chapter described two cross-modal socio-contextual priming studies which aimed to investigate whether the processing and activation of two General American sublexical features – rhoticity and an open LOT vowel – automatically primed American English lexical variants and accent associations for British English listeners. The experiments found no evidence of this, neither in an experiment which had no inter-stimulus interval nor in an experiment which had a 800ms inter-stimulus interval. Similarly, there were no effects for either items with lexical variants as targets or items with the accent associations as targets. Whether the prime was a non-word or a word also had no effect. On the basis of these findings, as well as those of Chapter 3, I draw a tentative conclusion that social and linguistic information is stored in separate representations in memory. The next question of this thesis, then, is which linguistic representations are used to access social meaning (presumably through associations rather than by being united with it in the same representation). This is explored further in Chapters 5 and 6, which investigates the size and detail of such representations.

5 Testing the role of word-sized representations in accent recognition

5.1 Introduction

In this chapter I describe an accent recognition experiment which aimed to find out whether listeners are better at recognising phonetic features of accents in high-frequency words than in low-frequency words. The aim of this experiment was to gain further insights into what linguistic representations are decisive in social meaning processing: if lexical frequency influences accent recognition this would be an indication that lexical representations stored in memory play a decisive role in phonetic accent processing. It would suggest that these representations are the ones that are used to access social meaning. This was not the case, suggesting that word-sized representations may not play a central role in accent processing.

As discussed in Section 2.3, discussions of exemplar theory in sociolinguistics often centre around the word as the central linguistic unit in language processing (Hay 2018), especially very detailed exemplar representations. Most research on this topic, however, has been done on the basis of how social cues influence the processing of the linguistic form (cf. Hay, Nolan & Drager 2006; Walker & Hay 2011), or only whether listeners have mental representations of sociolinguistic variants (McLennan, Luce & Charles-Luce 2003; Sumner & Samuel 2009). For example, Walker & Hay (2011) find that older sounding voices facilitate lexical access to words which are more often used by older people, and Sumner & Samuel (2009) find evidence that 'fluent' listeners of New York City English have different representations for NYC English forms than listeners with less experience perceiving the variety.

There is a gap in knowledge about which linguistic forms and representations influence the processing of social meaning. This is an important gap, as these processes may differ from how social cues influence linguistic processing: the semantic priming experiment in Chapter 3, for example, showed that there was a difference in priming effects on the basis of the direction of the social or linguistic cue: a sociolinguistic variant priming sociolinguistically related concepts (*y'all* – GUNS) caused inhibitory priming, occurring only after a large number of repetitions of the prime, and a social concept priming sociolinguistic variants (*American* - CANDY) caused facilitatory priming after few or no repetitions of the prime. Therefore, finding effects of social cues on linguistic processing at the lexical level (e.g. Walker & Hay 2011) does not automatically mean that it is at the lexical level that linguistic processing provides social cues. This means that it is still unclear what exact linguistic representations are used to process social meaning.

The experiment described in this chapter aims to fill this gap by testing whether participants were able to recognise accents more easily in highly frequent words than in infrequent words. By testing whether lexical frequency had an effect on accent recognition, the experiment explored whether lexical representations play a role in recognition. It was expected that highly frequent words would be easier to recognise, as listeners would have more and stronger memories or representations of those words pronounced in the relevant varieties (see Section 2.3). This would then provide evidence for the notion that the representations that are used in accent recognition, and in the processing of social meaning, are lexical in nature.

The experiment, however, did not find an effect of lexical frequency on accent recognition accuracy, raising questions around the exact role of the word in sociolinguistic processing. The results show that there were large differences in how likely listeners were to recognise different sublexical variants (e.g. a Yorkshire GOAT vowel, or an SSBE BATH vowel), which suggests that it is in fact sub-lexical representations (representations smaller than words, such as phonetic variants) which dominate the processing of social meaning. This is explored further in Chapter 6.

5.2 Methods

5.2.1 Accent recognition experiments

The method used in this experiment is that of the accent recognition task. There is a relatively extensive tradition of accent and dialect recognition research, in which listeners are asked to recognise regional varieties of languages on the basis of audio recordings (e.g. Preston 1993; Van Bezooijen & Gooskens 1999; Williams, Garrett & Coupland 1999; Clopper 2004; Clopper & Pisoni 2004; Pinget, Rotteveel & Van de Velde 2014; Jeffries 2016). Most of this type of research has focussed on relatively abstract questions, particularly the questions of how good listeners are at recognising accents, and whether this is influenced by listeners' experience with and proximity to the accents.

In response to the first question, most research finds that listeners are sometimes relatively good at recognising broader categories of accents, and sometimes find it more difficult. They are best with broader categories and struggle more with accents that are geographically and linguistically closer to each other (Preston 1993; Van Bezooijen & Gooskens 1999; Clopper & Pisoni 2004; Pinget, Rotteveel & Van de Velde 2014). In response to the second question, most research finds that listeners are better at recognising accents from their own region (Williams, Garrett & Coupland 1999; Pinget, Rotteveel & Van de Velde 2014; Jeffries 2016; Mepham & Evans 2019; Watt et al. 2019). Clopper (2004), however, finds no difference in overall accuracy in her accent recognition task based on where participants had lived.

Whilst most of the studies described above look at the recognition of accents as a whole, rather than specific aspects of them, other studies have focussed on questions to do with socio-phonetic detail: what exact allophonic realisations of a phoneme are recognised as what accents? Plichta & Preston (2005) synthesised a continuum from more to less monophthongal realisations of the PRICE vowel in US accents and asked participants to classify them on a continuum of locations from the North to the South of the US. Their classifications matched the vowel's distribution between more northern and more southern cities. Lawrence's (2014) study focussed on four sociolinguistic variables: BATH, STRUT, FACE, and GOAT in British English, and asked participants to place different realisations of these vowels on a map. He found that participants placed Sheffield English realisations of BATH and STRUT in the North more generally, whilst they placed Sheffield realisations of FACE and GOAT more specifically in the North-East, showing that not all allophonic variants of the same variety are always processed in the same way.

In most cases, when accent recognition studies focus on more specific socio-phonetic realisations, these studies use recordings of smaller linguistic units. Whilst many of the studies looking into larger accuracy patterns (and experience and proximity) use sentences, narratives, and longer audio clips (Van Bezooijen & Gooskens 1999; Williams, Garrett & Coupland 1999; Pinget, Rotteveel & Van de Velde 2014), other studies use single-word stimuli: for example the allophone focussed studies described above – as well as ongoing research by Watt et al. (2019) and Walker, Van Hell & Bowers (2019). I follow this in the experiment described in this chapter.

5.2.2 Experimental design

The experiment described in this chapter was an online multiple choice task designed and presented in the online experimental software Gorilla (www.gorilla.sc; Anwyl-Irvine et al. 2019), and distributed through social media. It was able to attract a large sample because the experiment was gamified and gave participants a score at the end of the experiment (see below). Participants were encouraged to share this score online which promoted its distribution.

In this task participants were asked to classify single-word stimuli as one of 4 accent categories: 'Yorkshire', 'American', 'Standard English from the South of England' (see the Subsection 'Accents' in Subsection 5.2.5 below for a discussion of the accent labels), and a distractor group of other accents, named 'any other accent'. This last category included Tyneside, Scottish and Liverpool English accents, amongst others. Participants heard 42 different speakers and responded to 88 items, as well as a practice round with 12 different speakers and 12 items. These were taken from www.forvo.com and www.soundcomparisons.com. They were used for non-commercial purposes and without modifying the material.

The experiment aimed to measure the effect of lexical frequency on accent recognition. This meant that the stimuli could not be sentences or fragments with multiple words: this would mean that the different lexical frequencies of multiple words would be a part of a singular stimulus, making it much more complicated to establish which of these did or did not influence participants' responses. Instead, isolated word stimuli were used as the stimuli, in order to be able to focus on one word's frequency per stimuli, and to allow for the highest level of control over other factors in the stimuli (e.g. larger numbers of phonemes in full sentences or phrases).

All critical items fell into one of two categories: high-frequency words (more than 250 occurrences in the spoken section of the British National Corpus) or low-frequency words (less than 40 occurrences). This prevented any overlap between frequency categories, with the widest possible interval between the categories that still left enough possible stimuli for each of the categories. The separated categories made it possible to test, for example, whether listeners are better at recognising an accent as being from Yorkshire in the highly frequent word *run*, than they are at recognising a Yorkshire accent in the much less frequent word *shun*. This is despite the fact that, at a more abstract level, they contain the same Northern STRUT vowel (see Subsection 5.2.5). This tested the main research question of the experiment:

1. Do listeners recognise accents more accurately in high-frequency words than in low-frequency words?

Finding a difference in accent recognition between high and low-frequency words would imply that lexical exemplars play an important role in accent recognition: for example, such an effect might arise if listeners process a word in a certain accent by matching it with exemplars of that word in that accent. For high-frequency words, these would likely be available – most British listeners will have quite a large number of memory traces (i.e. exemplars) of Americans using the word 'or' for example. For low-frequency words, they would not be as likely to be available, or not as strongly present, as the chance that a listener has stored exemplars of Americans using the word 'orc' is probably lower.

At the same time, the experiment allowed the more explorative investigation of some other potentially influencing factors. These explorative research questions were the following:

2. Do participants with different levels of experience with the accents used in the experiment behave differently?
3. Does comparative lexical frequency of words in an American compared to a British spoken corpus help recognition in General American stimuli?

If different levels of experience influenced lexical effects in accent processing in this experiment, this could shed light on the development and importance of lexical exemplars in accent processing. The comparative frequency measure could shed light on another way that lexical exemplars could affect accent processing. The different ways in which lexical representations could affect accent processing are discussed further below.

5.2.3 Measuring activation of word-sized units

A key question for the design of this experiment was how to measure lexical effects, as there are multiple conceivable ways in which the different levels of activation of word-sized mental representations could lead to measurable effects on accent recognition: absolute frequency, recency, and comparative frequency. I argue that the best measure for this is absolute frequency.

Pierrehumbert (2001: 141) and Goldinger (1998) argue that the strength of an exemplar is a combination of its frequency and its recency. The former could influence accent recognition as follows: if a listener has frequently processed input consisting of American speakers using the word *car*, they have strong exemplars of this, making them more easily activated when hearing new input similar to this (e.g. another American voice pronouncing the word). As exemplars include the social context of the auditory input (the knowledge that the word was pronounced by an American speaker), this aspect of the exemplars would also be more easily activated. This way, the frequency of an exemplar can facilitate the access of the social meaning of the input. This would make it easier to recognise an 'American accent' in the word *car* (frequent word which most British listeners will have heard pronounced by Americans and therefore have strong exemplars of) than in the word *mar* (infrequent word, which listeners may not have any (strong) exemplars of). If the representations used in processing are more abstract but still word sized, the same mechanism would apply: more frequently processed and used higher-up representations would be quicker to be activated as well.

The argument made by Pierrehumbert (2001: 141) and Goldinger (1998) includes another factor: recency. This could influence recognition as follows: because a listener heard an SSBE speaker say the word *stance* recently (e.g. just before starting the experiment), its SSBE exemplars are more active still and more likely to be recognised – even if it is less frequent than the word *dance*. In practice, this separate influence would most likely coincide with absolute frequency measures in an experiment with a large sample, as on average participants are more likely to have recently heard and processed frequent words than infrequent words before the experiment. In any case, establishing what exact linguistic input has been recent for participants is difficult to control for without the experiment exposing participants to input before carrying out the main experiment. This makes it a less suitable variable to use in this chapter's experiment.

Another possibility is that it does not matter so much how frequent an exemplar is, as much as how many exemplars a listener has for each of the varieties; in Walker & Hay's (2011) study, words which were comparatively more frequent in older speaker's speech were recognised quicker in older sounding voices. Similarly, we might expect someone to hear the word *president* and skew towards hearing that as American, if they have more exemplars of Americans saying *president* than they have exemplars of British speakers saying the word. In this case it is not the frequency of the word *president* itself that would cause an effect on accent recognition but the ratio of American exemplars to British exemplars.

A downside of the comparative frequency measure is that finding an effect may not be direct evidence for word-sized units playing an active role in accent recognition, but may instead be caused by semantic associations. For example, the word *president* might simply have a semantic association with Americanness, based on world knowledge and extra-linguistic information. This could prime listeners to recognise a stimulus as American. American accents however are not just recognised in utterances with American-associated words. Another downside to using comparative frequency is that there is no corpus of spoken English which incorporates General American, SSBE and Yorkshire English, which can effectively compare frequency ratios of words between the varieties – instead the only possibility is to compare different corpora which are compiled differently. This makes this variable a less suitable one for the purpose of this experiment as well.

For these reasons, absolute lexical frequency was chosen as the primary measure of the experiment, as the most feasible and precise measurement to investigate the role of lexical exemplars in accent recognition. All critical items were selected to fall into either a high or low absolute lexical frequency condition.

Still this does not mean that the other two influencing factors, recency and comparative frequency would not affect the experiment. Recency, however, should mostly coincide with absolute frequency, as argued above, and comparative frequency was still also tested, albeit with a less precise measure, and without the experiment being designed around it. The General American stimuli were given a score for their relative frequency between the spoken section of the Corpus of Contemporary American English and the British National Corpus as a rough measure of of whether a word is more frequent in British or American Englishes and therefore whether listeners comparatively have more or stronger exemplars of a word in either variety. This measure was tested as well, allowing for an exploration of whether it influenced recognition of those stimuli.

In addition to the question of the type of frequency, there is the matter of the degree of ‘low’ frequency: this can either be low-frequency words, or non-words. Non-words have the advantage that they create a more fundamental difference between the two conditions: a difference between a high number of representations and no possible representations at all, whilst the difference between high-frequency and low-frequency words is gradual.

However, for this experiment the set-up with low-frequency real-word stimuli was chosen. Non-words could conceivably cause more general linguistic processing costs compared to words (Newman & Twieg 2001; Raettig & Kotz 2008). This may mean listeners have fewer resources and less time to process the social meaning of the stimuli – even if they do recognise accents on the basis of the sublexical variants in the non-words. Finding an effect with low-frequency words would have been more unambiguous, which is why this research design was chosen for the experiment in this chapter.

Furthermore, using non-words in the task might have forced listeners to listen at the level of abstract sounds, being unable to rely on lexical representations (if they are in fact used). Including these in a task which also includes real words might have forced listeners to apply a more analytic strategy that could spread across conditions, and even out lexical effects if they are there. Therefore, a non-words based accent recognition task needed to be done separately; this is described in Chapter 6.

5.2.4 Procedure

Participants were recruited online through social media and mailing lists. First, the participants were presented with an online participant information sheet about the ethics of the experiment and were asked to confirm that they consented.[15] Then filled out a questionnaire on their demographics and whether they were doing the experiment for the first time. The participants were specifically asked about whether they had experience living in the South East of England, Yorkshire, and the United States, whether they interacted with people with different accents often (on a 5-point scale), and how good they themselves thought they were at recognising accents (on a 5-point scale). Participants were then told they were going to hear different words pronounced by different voices, which they were to categorise in four broad accent categories. Figure 22 shows a screenshot of how one of the stimuli was presented in the browser, as participants heard the practice stimulus ‘holy’.

[15] The experiment went through the University of Sheffield’s ethical review process and was approved.

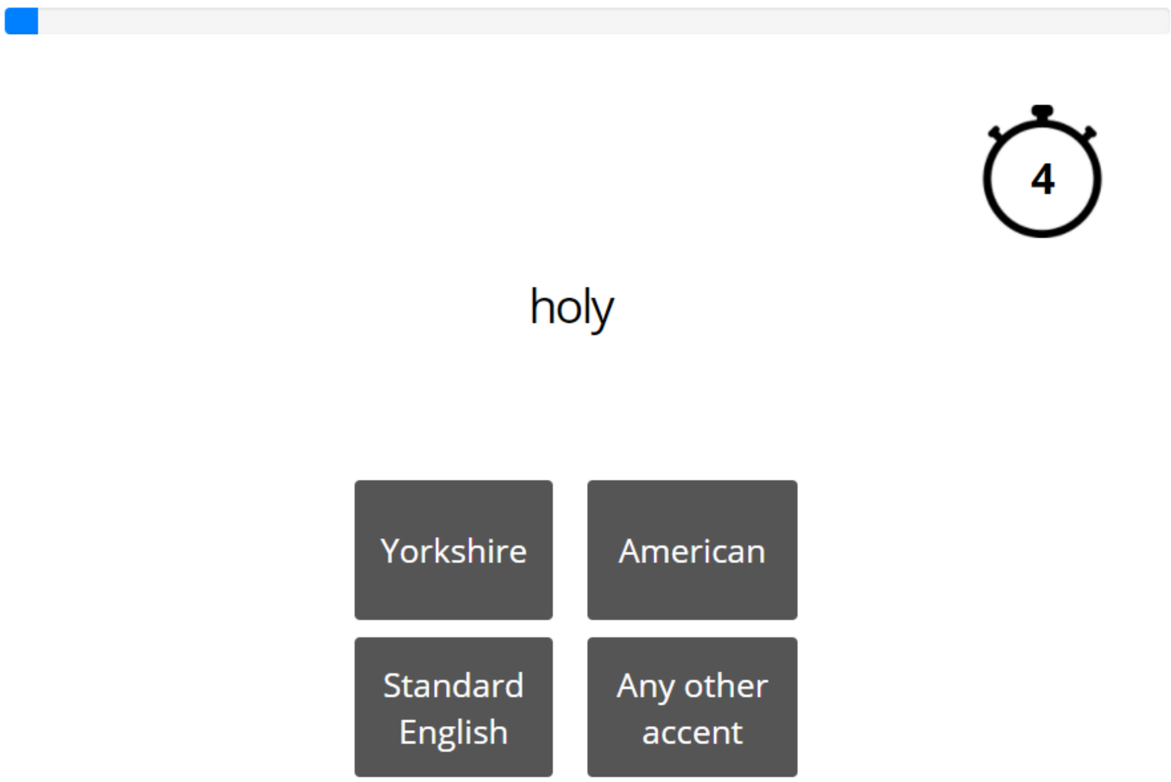

Figure 22. Screenshot of the practice stimulus 'holy' at the start of the experiment.

The participants were told to pay close attention to the experiment, as they would only hear each stimulus once and had only 4 seconds to respond to each item. They were told to not worry too much about getting some wrong, and that the quiz was quite difficult. After playing a practice round, they were given their score for that round. Participants were not given any other feedback during the experiment, with the exception of the final score given at the end. To encourage completion a progress bar was shown at the top of the screen from the instructions onwards. Participants were then asked to test their sound with an example sound file that could be replayed as often as was desired.

After the instructions and the practice round, participants were played a block of 44 items in a randomised order, with a set of three distractor items at the start. After the first block they were given the opportunity to take a break, before being played the second block of 44 items in randomised order, again starting with a set of distractor items. The order of the blocks was also randomised.

In order to speed up listener responses (in line with Walker, Van Hell & Bowers 2019; D'Onofrio 2016) participants were given 4 seconds to respond to each stimulus, and were shown a fixation cross between stimuli for 600ms to prepare for the next stimulus. In piloting the experiment, I found that this was a pace that gave participants enough time to not feel stressed, but also to not give them too much time to analyse their own choices, and find patterns in the stimuli (for example, the fact that every monophthongal GOAT vowel was from Yorkshire). For the same reason, listeners were not able to replay the sound they heard. Whilst the speed of the task may have prevented elaborate conscious thinking processes about the task, the task was not intended to be fully subconscious, and it was not expected that participants could not improve during the task.

Whilst many accent recognition studies (Lawrence 2014; Pinget 2019) use map-based tasks, the current one used four answer buttons ('American', 'Yorkshire', 'Standard English', and 'Other', following Pinget, Rotteveel & Van de Velde (2014). This allowed for analysis of accuracy in a relatively straightforward way, i.e. as a percentage of correct responses. This circumvented the need to transform geographical distance to linguistic distance, which Pinget (2019) points out can be extremely complicated, with coasts and country border influencing error rates in a non-linear way.

Participants were also shown the orthographical representation of the stimulus they heard, in order to avoid confusion. This did mean participants were not exclusively processing the input auditorily, but also visually, which might have affected their responses, as these are different perception processes (see for example Holcomb & Neville 1990). However, it would have done so across all stimuli. Without orthographical help, some of the participants of the experiment in the piloting stage heard a number of low-frequency words as non-words. Including the orthography of the word ensured that as many of the low-frequency words as possible were recognised, and ensured that words with a realisation of one accent were not interpreted as different words with the realisation of a different accent (e.g. hearing *foe* with a Yorkshire fronted monophthong could possibly be heard as a SSBE realisation of the word *fur*). This could influence responses in a skewed way: low-frequency words with ambiguous sublexical variants may be more likely to be interpreted as a different word in a different accent than high-frequency words, and therefore create lower accuracy, not on the basis of lexical influence on accent processing, but on the basis of confusion stemming from lexical processing itself.

The final feedback given to the participants was calculated based on items which were coded to be relatively doable (e.g. stimuli with Yorkshire FACE and GOAT vowels), rather than all items (including some of the distractor items with no strongly recognisable sublexical variants, like General American *deep* or *fade*). They were given a score which corresponded to four categories (from excellent to chance level). After this, participants were asked in a short post-hoc question whether there were any stimuli they heard which they did not recognise as real words. This was to avoid blurred lines between low-frequency words and non-words (which are looked at in the next experiment).

5.2.5 Stimuli

All critical stimuli in the study fell into two lexical frequency conditions: high frequency or low frequency. All other potential influencing factors were controlled for as much as possible, either by keeping them the same (or as similar as possible) or by balancing them out across conditions. All critical stimuli were monosyllabic, for example, and all of them were recorded in the same sound booth. For other variables, the high-frequency and low-frequency stimuli were matched as closely as possible within each run of the experiment. This is illustrated in Table 18.

Table 18. Example set stimuli, showing matched stimuli for a given run of the experiment.

			Matched stimuli		
Accent	**Potentially Recognisable Sublexical Variant**	**Speaker**	**High frequency**		**Low frequency**
Yorkshire	STRUT vowel	John	*run*	⇔	*shun*
		Mary	*much*	⇔	*Dutch*
		Adam	*bus*	⇔	*suss*
	FACE vowel	Sarah	*day*	⇔	*fray*
		etc.			
	GOAT vowel	Liam	*road*	⇔	*ode*
				etc.	
American	Rhoticity	Betty	*poor*	⇔	*gore*
		Danny	*work*	⇔	*curse*
		Sandra	*or*	⇔	*orc*

etc.

As the table illustrates, if the high-frequency stimulus *run* was pronounced by a speaker, John[16], who has a Yorkshire accent, this would be matched with a stimulus with the same recognisable sublexical variant in it (a Northern STRUT vowel) also pronounced by John: the low-frequency stimulus *shun*. This means that for these stimuli the speaker was the same, which meant that accent, voice quality, speech rate, and reading style were matched, so that these were not able to affect how easily listeners recognised the accents between the high and low-frequency conditions. This is illustrated in Table 18 below. A full list of stimuli is given in Appendix III.

Lexical frequency

The main measure of absolute lexical frequency was operationalised in the experiment as two conditions: high-frequency words and low-frequency words. Stimuli qualified as high-frequency if there were more than 250 occurrences of the whole word in the spoken section of the British National Corpus. Stimuli were qualified as low-frequency words if there were fewer than 40 occurrences of them in the spoken section of the British National Corpus. The thresholds ensured that it was possible to select enough stimuli whilst keeping the frequency categories as distinct as possible. The British National Corpus is a large corpus of British English which includes a spoken section. This made it particularly suitable to the experiment with its British participants and its spoken stimuli.

Whilst the categorisation as two frequency groups allowed for maximal control over speaker voice and potentially distracting influences of voice quality, it is of course also possible to measure lexical frequency more precisely, i.e. by using it as a continuous variable. In order to establish this measure the raw number of whole-word occurrences in the British National Corpus was taken and log-transformed (following Hay et al. 2019). These are summarised in the boxplots in Figure 23 below. As can be seen in this figure, the spread of stimuli on a log-transformed scale is slightly wider for the high-frequency stimuli. Most importantly, however, the two categories do not overlap.

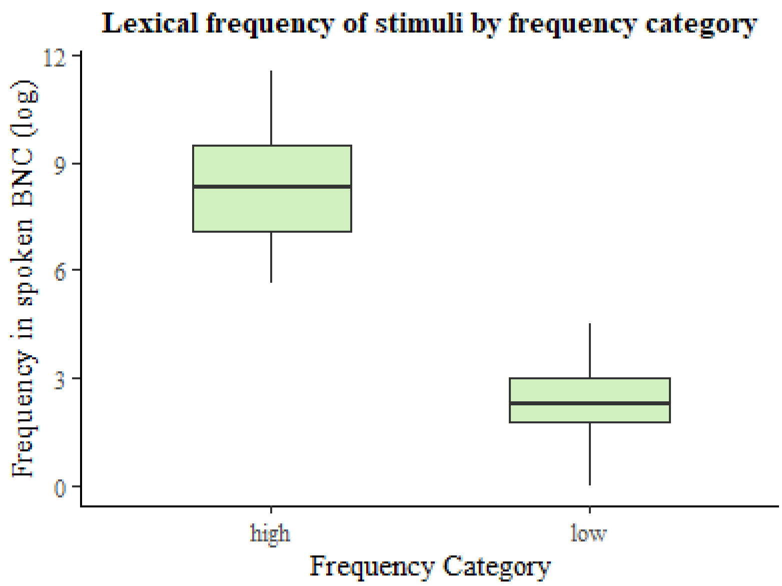

Figure 23. Lexical frequency of stimuli by frequency category.

Speakers

Controlling for speaker voice was a relatively complicated issue in this experiment. On the one hand, it was important to have enough different speakers included in the experiment to make sure participants recognised the accents on the basis of the linguistic cues rather than recognising the same voice. On the other hand, the potentially small differences in recognition required a tightly controlled within-subject

[16] All names used here are pseudonyms.

design, in which listeners heard stimuli in both critical conditions – with as much as possible being kept equal between the conditions.

Repeating speakers' voices was one of the main ways to keep other influencing factors equal. Two stimuli spoken by the same person, in the same sound booth, on the same day, ensured that pitch, voice quality, and base articulatory settings were as similar as possible between the stimuli in the different conditions. These have been found to vary between accents (Wieling & Tiede 2017) and might very well vary within accents, and could otherwise strongly interfere with accent recognition. Repeating the voices between these conditions solved this issue.

In order to still create enough variation in the speaker voices to prevent participants from being helped or hindered by recognising the same voice, a large number of different speakers were recorded for the experiment (92 speakers in total, 42 of which were selected for the final experiment). Voices were repeated a maximum number of two times per run of the experiment, once for each lexical frequency condition. Different versions of the experiment counterbalanced which lexical frequency condition would be heard first for each speaker, so that if listeners were able to recognise a stimulus on the basis of having heard a stimulus by that speaker before, this would happen as often with a high-frequency as a low-frequency as the second stimulus.

For each accent (category) seven speakers were selected from a larger number of recruited speakers, who had been recorded reading word lists in the same sound attenuated booth at the University of Sheffield. The wordlists for this were presented to speakers through the experimental software PsychoPy (Peirce et al. 2019), which allowed for self-paced reading, as words appeared on the screen separately. This prevented the varying intonation typical of enumerations in lists, and instead kept them the same within speakers. The monotony created by this design also minimised differences in speed and voice quality within speakers. Differences in volume were normalised to have the same peak amplitude (-3.0 dB) by means of the audio software Audacity[17], although some participants still reported noticing differences in volume across stimuli, possibly due to some stimuli including segments with a higher peak amplitude than others (e.g. a loud plosive).

Accents

The experiment contained three critical answer options and one distractor option: 'Yorkshire accents', 'Standard English', 'American', and 'any other accent'. The three main accent groups were chosen as similarly broad and well-known accents for British listeners. How broadly or narrowly an accent is defined can make a large difference in an accent recognition task (Van Bezooijen & Gooskens 1999), which meant that using accent categories that were similarly broad kept conditions relatively similar.

In the experiment the broad accent categories were defined as follows in the instructions:

- 'Yorkshire accents' (shortened to 'Yorkshire' in the task)
- 'American accents' ('American' in the task)
- 'Standard English from the South of England' ('Standard English' in the task)
- 'A broad category of other accents' ('Any other accent' in the task)

The labels 'Yorkshire accents' and 'American accents' were deemed transparent enough for non-linguists, as they referred to geographic regions. The geographic labels 'Yorkshire' and 'American' are relatively clear cut: there are no varieties of Yorkshire or American English that would not fall into the category of 'Yorkshire accents' or 'American accents'.

The label 'Standard English (from the South of England)' was chosen over the geographic label 'Southern British English' however, as the latter would also cover a wide range of non-standard accents from the South of England, such as Cockney or Cornish English, some of which were included in the

[17] https://www.audacityteam.org/

'Any other accent' category. 'Standard English (from the South of England)' was chosen as the best option, as it has been used in perceptual dialectology before as a label for Standard Southern British English (Coupland & Bishop 2007). It could be somewhat of a specialist term, but not as much as 'SSBE' or 'Received Pronunciation', and it has the advantage of being more accurate than terms like 'the Queen's English' or 'BBC English'.

As in Chapter 4, speakers were selected mainly based on the sublexical variants they produced (see Subsection 4.2.3 above and the Subsection 'Recognisable Sublexical Variants' below), and I selected speakers who were as homogeneous as possible in their regional origin and accent beyond the exact sublexical variants used in the experiment. This is summarised in Table 19 below. For the Yorkshire group mostly speakers from South-Yorkshire were used, for the American group only General American voices were used, and for the Standard English group SSBE voices were used. As in Chapter 4, the group of speakers was made to be as mixed as possible for gender, to ensure variety in the voices.

Table 19. Summary of speaker demographics for all speakers with critical stimuli, by accent and by potentially recognisable sublexical variant.

Accent	*Sublexical variant*	*Speaker location during youth*	*Age*	*Gender*	*Education*
American	KIT vowel	Virginia, US	20s	Female	University
		Virginia, US	20s	Female	University
		New York, US	30s	Male	University
	LOT vowel	South Carolina, US	20s	Female	University
		Rhode Island, US	20s	Female	University
		Arizona, US; England	20s	Male	University
	Rhoticity	Outside UK/US; Washington D.C., US; California, US	40s	Female	University
		New York, US; Massachusetts, US; Outside UK/US	20s	Female	University
		Michigan, US	20s	Female	University
SSBE	BATH vowel	Berkshire	20s	Male	University
		Kent	20s	Male	University
		Cumbria	20s	Female	University
	GOAT vowel	Berkshire	20s	Female	University
		Isle of Wight	20s	Female	University
		East Sussex	20s	Non-binary	University
	Lack of rhoticity	Hampshire	20s	Male	University
		Lincolnshire	20s	Male	University
		Surrey	40s	Female	University
Yorkshire	FACE vowel	South Yorkshire	30s	Female	University
		South Yorkshire	40s	Male	High-School
		South Yorkshire	18-20	Female	College
	GOAT vowel	West Yorkshire	20s	Female	University
		South Yorkshire	20s	Male	University
		South Yorkshire	20s	Female	University
	STRUT vowel	South Yorkshire, Derbyshire, Scotland	20s	Female	University
		West Yorkshire	20s	Female	University
		South Yorkshire	20s	Female	University

Recognisable sublexical variants

One of the most important variables that needed to be controlled for in the accent recognition experiment was which sublexical variants were potential cues for accent recognition. By comparing recognition of high-frequency and low-frequency stimuli between stimuli with the same recognisable sublexical variants (e.g. Yorkshire English *run* and *shun* and General American *poor* and *gore*) it was possible to rule out that frequency effects were actually caused by the recognisability of the sublexical variants included in the stimuli.

Furthermore, these stimuli always rhymed with an equivalent stimulus in the other lexical frequency condition, to keep phonetic environment as similar as possible. It was not possible to use splicing to control for phonetic environment (like Lawrence 2014 did), as the used stimuli were different words, which would have different transitions between consonants and vowels. There was a risk that the changing surrounding sounds were a potential cue to accent recognition themselves, or the combination of those sounds with the relevant recognisable sublexical variant in the stimulus. In order to prevent this, the surrounding sounds in the paired stimuli were as similar as possible and as 'neutral' as possible, i.e. they are not known to differ between accents. For example, the General American stimulus pair *poor* and *gore* only differ in their onsets, which are both plosives, not known to differ in pronunciation between General American and British Englishes.

As there is no exhaustive body of research on the precise features of which accents are easily recognisable to naïve listeners, it was necessary to cast a relatively wide net of potentially more or less easy to recognise features, ranging from the highly salient SSBE BATH vowel to the rarely commented on slightly lengthened KIT vowels before voiced consonants in General American.[18] The chosen features are listed in Table 20 below.

Table 20. Selected recognisable features in the accents used in critical conditions.

Accent	Recognisable sublexical variants
Yorkshire	Monophthongal FACE vowel
	STRUT vowel realised as [ʊ]
	Monophthongal GOAT vowel
General American	Rhoticity
	Lengthened KIT vowel before voiced consonant
	Unrounded LOT vowel
Standard Southern British English	Lack of rhoticity
	Diphthongal GOAT vowel with a central onset
	BATH vowel realised as [ɑː]

These recognisable sublexical variants were not chosen to directly compare how easy or difficult it was to recognise the variants themselves, or to see which of them were the most recognisable. This would not be possible as it would depend strongly on the accents of the speakers pronouncing the stimuli, and was not part of the research aim of the current experiment (finding out if lexical frequency improves accent recognition). Rather, the recognisable variants were chosen to form a varied range of more and less easily recognised variants, so that participant responses would not all be at chance level or at ceiling level. Performance at chance level would mean no accent recognition is occurring and there would be no effects on it anyway. Performance at ceiling level may level out any effects of lexical frequency on accent recognition which may be there when the task is more difficult. For example, if a stimulus is very easy to recognise (e.g. an SSBE BATH vowel), it might be possible for a weak exemplar to still

[18] These decisions were supported in part on my own intuitions and conversations with other linguists (in particular Holly Dann and Johanna Grace at the University of Sheffield). The responsibility for these choices is my own.

reach the threshold for correct recognition for all participants, when it would not do so for stimuli with less recognisable variants. The features and their different levels of recognisability are discussed below. Some of these have been discussed in Subsection 4.2.3 when describing the recognisable sublexical variants used in the experiments in Chapter 4.

Yorkshire

GOAT vowel

The first selected recognisable sublexical variable for the Yorkshire accents was the vowel in the GOAT lexical set (Wells 1982). There is a lot of variation in Sheffield GOAT variants: Stoddart, Upton & Widdowson (1999) list 9 different variants: [ɔː], [oʊ], [əʊ], [ɔʊ], [ɒʊ], [ʊə], [ɔɪ], [ɒɪ], and [ʊ]. Many of these are however restricted to phonological contexts or specific lexical items. The most common variants reported are monophthongal [ɔː], and diphthongal [oʊ]. These are also the main variables reported by Wells (1982: 358) in his description of the North more generally.

In addition to this, the fronted monophthong [ɵː] has been found to be on the increase across Yorkshire, even beyond Hull, where it is traditionally found: Watt and Smith (2005) find it in West-Yorkshire, Watt and Tillotson (2001) find it in Bradford, and Khattab (2007) in Leeds and York. Whilst some of those findings are more impressionistic observations, Finnegan (2015) finds more systematic evidence for some GOAT fronting in middle-class younger women in Sheffield.

The importance of the increasing frequency of [ɵː] in perceptions of Yorkshire English can be seen in the enregisterment of GOAT fronting as a typical Yorkshire feature (Cooper 2017: 361). Cooper (2017) finds a considerable amount of commodification of GOAT fronting in relation to Hull English (t-shirts, mugs, and fridge magnets being sold which say 'err nerr' for 'oh no' for example). He points out that this is not limited to Hull, and is found in Yorkshire more broadly as well, with some of his respondents commenting on GOAT fronting in his online survey asking participants to list language features they thought of as Yorkshire (Cooper 2017: 357, 362). He makes a case that GOAT fronting is becoming enregistered as Yorkshire English. Another anecdotal but not entirely trivial argument is that the two sisters from Leeds in the TV programme Gogglebox (3-5 million viewers in the latest season at the time of conducting the experiment)[19] have strong GOAT fronting.

Therefore, there were three potential realisations of the GOAT vowel that could be picked to be used in the experiment: monophthongal [ɔː], monophthongal fronted [ɵː], and diphthongal [oʊ]. The 16 South and West Yorkshire speakers recorded for the experiment exhibited all of these three options, which meant that it was up to me as the researcher to decide which of them to use. Speakers with the diphthongal realisation were not included, because it was too similar to SSBE [əʊ], and General American [oʊ]. Lawrence (2014) finds, for example, that a diphthongal realisation of the GOAT vowel is placed relatively consistently in the South of England.

This left only monophthongal realisations. The speakers that were recorded for this experiment showed large variation in the degree of frontedness in their monophthongal realisations: some speakers had no fronting whatsoever, some had strong fronting, and others had something between the two. Whilst the unfronted realisation [oː] might be the more traditional and more common form for many Yorkshire varieties of English (Wells 1982; Stoddart, Upton & Widdowson 1999), the commodification of the fronted variant and its enregisterment (Cooper 2017) might be more crucial in a perception experiment like the one described here. Lawrence (2014) for example, finds that [oː] is most often placed in the North-East of England specifically, rather than Yorkshire. Furthermore, Jeffries (2016) uses fronted realisations for Yorkshire speakers in her accent recognition experiment with young children. Whilst her stimuli were sentence length and included more features than just the GOAT vowel, her participants were able to correctly classify these stimuli. I made the decision to use the speakers with the slightly

---

[19] https://en.wikipedia.org/wiki/List_of_Gogglebox_episodes (retrieved 21st of November 2019)

fronted realisations as a middle ground between committing to a new, but enregistered form that is on the increase, and more traditional forms.

The stimuli that were chosen for this recognisable feature excluded words which often exhibit other Yorkshire variants. For example, a word like *go*, which is often realised as [gʊ] in Yorkshire Englishes (Stoddart, Upton & Widdowson 1999) was not included.

FACE vowel

The FACE vowel in Yorkshire Englishes is less variable than the GOAT vowel. In Sheffield, Stoddart, Upton & Widdowson (1999) find mostly monophthongal [eː], sometimes more open, and sometimes slightly diphthongal, as is also reported by Wells (1982: 357) about the North in general and by Hughes, Trudgill & Watt (2012: 104) in Bradford. There are some lexical exceptions to this: words like *make* and *take* can have short [ɛ] and words like *eight* and *straight* can have the diphthong [ɛɪ] (Stoddart, Upton & Widdowson 1999). The variant chosen for the experiment was the one furthest away from SSBE: monophthongal [eː]. To avoid confusion, no stimuli were used where Yorkshire English traditionally has the full diphthong [ɛɪ] in words like *eight*.

A complicating factor in this case is that Lawrence (2014) finds that listeners place monophthongal [eː] in the North-East of England, more than in Yorkshire, whilst the stimuli were produced by speakers from Sheffield. If participants in my experiment heard these as quite North-Eastern, this could push them towards 'any other accent' more than would be desirable. However, the fact that Yorkshire is an explicitly labelled option that listeners are listening out for, as opposed to Tyneside English which is just one of the options mentioned in the instructions as being in the 'any other accent' category, was anticipated to prevent or alleviate this problem. This was the same for other recognisable sublexical variants which had overlap between one of the three main accents and the other accents included (e.g. [ʊ] in the STRUT lexical set which is used in Tyneside and Liverpool English as well as Yorkshire English).

STRUT vowel

The vowel realisations for the STRUT lexical set are relatively uncontroversial for Yorkshire varieties of English. The STRUT and FOOT lexical sets have the same vowel in most of the North, including Yorkshire (Wells 1982: 351; Stoddart, Upton & Widdowson 1999; Leemann, Kolly & Britain 2018; Williams & Escudero 2014). At the same time, a vowel sounding close to [ə] may be heard among middle class speakers, rather than [ʊ] (Wells 1982: 353; Stoddart, Upton & Widdowson 1999). Still, Lawrence (2014) finds that listeners reliably place [ʊ] realisations of STRUT in the North of England, as opposed to [ʌ], which is reliably placed in the South of England. Furthermore, all Yorkshire speakers who were recorded for the current experiment had [ʊ] for the STRUT lexical set.

General American

Rhoticity

As described in Subsection 4.2.3, General American is a rhotic variety of English, in contrast to most varieties of English in England, including SSBE and most (although not all) Yorkshire varieties of English (Wells 1982).

LOT vowel

As described in Subsection 4.2.3, the vowel in the LOT lexical set in General American is unrounded, which is in contrast with Yorkshire English and SSBE [ɒ].

KIT vowel before voiced consonants

The General American realisation of the vowel in the KIT lexical set stands in some contrast to that of rhoticity and the LOT vowel. Its deviation from the varieties in England does not involve the presence of an additional sound, or a stark difference in vowel quality, but instead a slightly longer duration, and in some cases diphthongisation of the vowel to [ɪə] (Wells 1982: 485). The minor, exclusively durational differences between these realisations and those in SSBE and Yorkshire English likely make this one of the least recognisable out of the recognisable sublexical variants. It was still selected, as it was not certain before conducting the experiment how accurate participants would be, so less recognisable sublexical variants were also included, to avoid the possibility that all participants would score at ceiling level for all stimuli.

Standard Southern British English

BATH vowel

The first selected recognisable feature for SSBE was the BATH vowel, which was also discussed in contrast to General American in Subsection 4.2.3. Whilst Southern [ɑː] (as opposed to Northern [a]) is not exclusive to SSBE, there is some evidence that it is enregistered as specifically indicative of both Southernness and 'poshness'. Gupta (2005: 25) finds, for example, that many Northerners are hostile to the Southern variant, calling it, amongst others, 'snobbish' and 'pompous' and Mugglestone (2003: 78) points at it as a marker of sounding 'standard' and 'proper'.

Some linguists point at the strength of its associations with Southernness: Trudgill (1990: 18) argues that the stereotype of it being a Southern form is too strong for many Northerners to use it themselves, and Wells (1982: 354) suggests it would be a denial of Northern identity to use [ɑː]. The strength of this seems to be reflected in Gupta's (2005) findings that even middle-class speakers in the North and South of England are extremely consistent in the use of their respective variants. Finally, Lawrence (2014) finds that listeners reliably place [ɑː] in the South of England, and [a] in the North of England. Because of this, the SSBE BATH vowel can be expected to be very easily recognised.

GOAT vowel

The SSBE GOAT vowel is quite clearly distinct from other English variants in England: it is a clear diphthong starting in an unrounded mid-central position [əʊ] (Wells 1982: 146; Bjelaković 2017: 505; Hughes, Trudgill & Watt 2012: 53). Diphthongal realisations of the vowel do also occur in middle-class environments in the North (Finnegan 2015; Watt 2000). However, SSBE monophthongs are traditionally start more centrally, and Lawrence (2014) still finds that listeners locate a diphthongal realisation of the GOAT vowel in the South of England more than anywhere else.

There was some degree of variation in the pronunciation of GOAT vowels in the speakers recorded for the experiment, as the onset was more central for some than for others. I chose to use the speakers with more central onsets, to make sure the stimuli would not overlap with General American realisations which are realised with an onset that is further back, as [oʊ] (Trudgill & Hannah 2002: 39). This also meant that the SSBE GOAT vowels were somewhat closer to – but still distinct from – the Yorkshire GOAT vowels included in the experiment, which were fronted monophthongs, relatively close to the SSBE onset in GOAT.

Lack of rhoticity

Finally, the last recognisable feature used for the critical stimuli in SSBE was the lack of rhoticity, as also discussed in Subsection 4.2.3. There are not a lot of stand-out features of SSBE for the target audience of British participants, as this is the standard variety in Britain. This meant it was necessary to include variants which were shared with at least one other variety. In order to use a variant which is in clear contrast with one of the other varieties, there were two main options: either the open SSBE STRUT vowel, contrasting with the Yorkshire variant, but the same as the General American variant, or the lack of rhoticity in SSBE, contrasting with General American, but not with Yorkshire English. The latter was chosen as the option with more stimuli available to it.

This lack of rhoticity is also found almost without exception in Yorkshire accents, but also a large range of the accents in the 'Other' category used in the accent quiz (including Scouse, Geordie and Australian English), although Wells (1982) points out that in some cases there might be differences between Southern English and Northern English vowels which were historically followed by /r/. In some cases [iːə] can be found for NEAR, for example, as opposed to the standard variant [ɪə]. However, the Standard English forms are also very common in Yorkshire varieties of English, which needs to be taken into account as a methodological confound.

Surrounding sounds

In order to make all stimuli with the same sublexical variant phonetically as recognisable as each other, the surrounding variants in the stimulus words were selected to be as neutral as possible in order for these surrounding variants to not be an extra clue for recognition. The stimulus list for the speaker recordings aimed to avoid any surrounding sounds that had accent specific realisations, on the basis of the work cited throughout the sections above (most notably Wells 1982; Trudgill & Hannah 2002; Hughes, Trudgill & Watt 2012). This meant that the Yorkshire word list excluded /h/ dropping, for example. The recorded stimuli were then reviewed for whether there were still any surrounding variants which I perceived to be too helpful. This meant that none of the final SSBE stimuli had word-final /d/, as most SSBE speakers had higher levels of aspiration in this sound than non-SSBE speakers, which seemed fairly noticeable and risked improving the chances of recognition based on features other than the intended recognisable sublexical variant.

Distractor items

In addition to all of the critical items for each of the three main broad accents, 4 distractor items were included per accent, in order to distract from the repeated recognisable accent features (e.g. 6 items with FACE vowels for Yorkshire English). Two of these were particularly easy and contained multiple cues for recognition. For example, one of the Yorkshire distractors was the word *spicy* which was pronounced with a monophthongal PRICE vowel, and happY laxing. The two others were particularly difficult to recognise and contained no obvious cues for recognition. For example, one of the Yorkshire distractors was the word *skill*. The speakers of the distractor items were always different from the speakers of the critical items, and only used once, with the exception of the General American items. These speakers were used for two distractor items, because of the sparseness of General American speakers near the recording booth in Sheffield, UK.

The fourth broad accent category 'Any other accent' consisted exclusively of distractor items, pronounced by speakers with a range of different accents: Scottish, Liverpool, Welsh, South-West English, Northern Irish, and Tyneside English. All of these contained relatively clear cues for recognition, for example glottal reinforcement for Tyneside voices (e.g. in *chicken*), and geminated consonants for Welsh English (e.g. in *rotten*).

5.2.6 Participants

A total of 2238 people were recruited in the experiment through online distribution on Facebook, Twitter, and Reddit[20], and distribution in a local newspaper.[21] 1592 of these completed the full task and scored better than 40% both in the practice round and in the critical items of the main experiment. Although the experiment was geared towards non-linguist native speakers of English from Great Britain, this was not a requirement for participation. This meant a wide range of experience levels with different accents was available in the data allowing for an exploration of the effects of participant experience, where non-native speakers from outside the Anglosphere functioned as a low-experience participant group.

[20] https://www.reddit.com/r/SampleSize/comments/dzhxmp/academic_how_good_are_you_at_recognising_accents/

[21] https://www.thestar.co.uk/retro/sheffielders-test-how-well-you-recognise-accents-student-hielke-vriesendorps-online-quiz-1350492

Excluding these non-target groups, 545 participants were in the target group of non-linguist native speakers of English. This can be seen in Table 21 which summarises the numbers of linguists and native speakers in the larger sample.

Table 21. Cross-tabulation of number of participants for whether they were native speakers of English, and whether they had a background in linguistics. The selected target audience sample has been marked by being boxed and in bold.

	Native speakers	**Non-native speakers**	**Prefer not to say**	**Total**
Linguist	204	268	4	476
Non-linguist	**545**	558	7	1110
Prefer not to say	5	1	0	6
Total	754	827	11	1592

The majority of my analysis was performed on the original, more homogeneous, target audience sample of 545 native speakers of English who had no background in linguistics, in order to avoid convergence issues (which the wider data set encountered in regression modelling). This was caused by the fact that non-native speakers were often sparsely populated across demographic variables that native speakers were asked about (e.g. very few non-native speakers selected 'College (UK) as their education level'). The discussion of the sample below therefore describes this target audience sample.

The target audience sample had a large spread with regard to participant location, but with most participants living in the South East of England, and in Yorkshire (as can be seen in Table 22 below), and it was quite balanced for gender (Table 23). It did include mostly people in university or with a university degree (Table 24), and was relatively young, albeit with a relatively wide range of ages overall (Figure 24).

Table 22. Number of participants in the target audience sample by location. Answering options that were smaller (e.g. different parts of Yorkshire) are added up within the table, to show the frequency of their larger regions.

Location	**Count**	**Percentage**
Scotland	31	6%
North of England:	156	29%
North-West	*40*	*7%*
North-East	*21*	*4%*
Yorkshire	*95*	*17%*
Midlands:	54	10%
East Midlands	*31*	*6%*
West Midlands	*23*	*4%*
Wales	9	2%
South of England:	162	30%
South East	*119*	*22%*
South West	*43*	*8%*
Northern-Ireland	13	2%
Republic of Ireland	6	1%
USA	26	5%
Other	85	16%
Total	545	100%

Table 23. Number of participants in the target audience sample by gender.

Gender	Count	Percentage
Male	288	53%
Female	228	42%
Non-binary	26	5%
Prefer not to say	3	1%
Total	545	100%

Table 24. Number of participants in the target audience sample by type of education.

Education	Count	Percentage
University	449	82%
College (UK)	42	8%
Secondary school	33	6%
Other	18	3%
Prefer not to say	3	1%
Total	545	100%

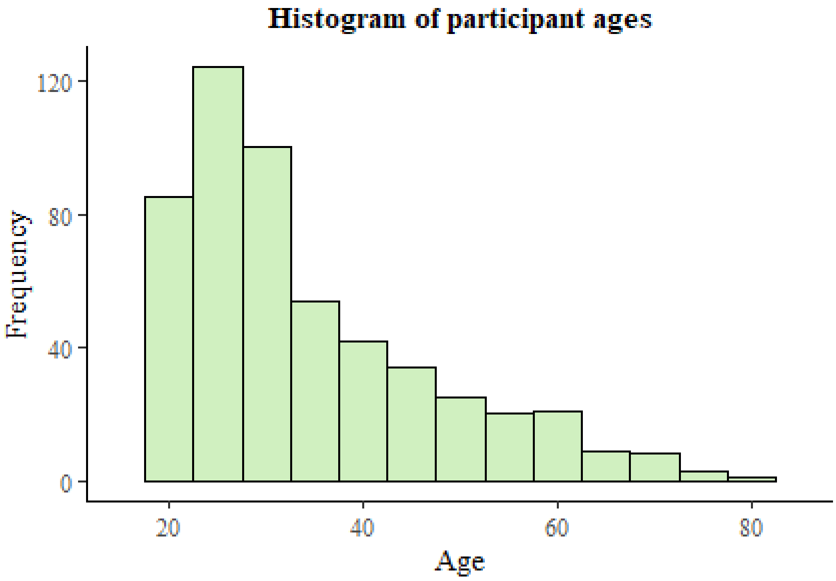

Figure 24. Histogram of participants in the target audience sample by age.

The participants were asked some further information about their experience with different accents, especially those used in the experiment. 36% of the participants had lived in Yorkshire for at least one year of their lives, 49% had lived in the South-East of England, and 12% had lived in the USA (Table 25). There was a relatively even spread in how often participants were exposed to very different accents on a weekly basis, with few participants at the extremes, as can be seen in Table 26. As can be seen in Figure 25, participants were relatively positive in the assessment of their own accent recognition abilities, with most participants thinking they were neutral or good at recognising accents.

Table 25. Number of participants in the target audience sample who had experience of living in one of the three regions where the critical accents of the experiment were spoken, for at least one year.

Participants with experience living in	**N**	**Perc.**
Yorkshire	194	36%
South-East of England	267	49%
USA	66	12%

Table 26. Share of weekly contacts with a very different accent, as indicated by participants.

Indicated share of weekly contacts with very different accent	**N**	**Perc.**
None of them	17	3%
A couple of them	194	36%
About half of them	128	23%
Most of them	152	28%
All of them	51	9%
Prefer not to say	3	1%
Total	545	100%

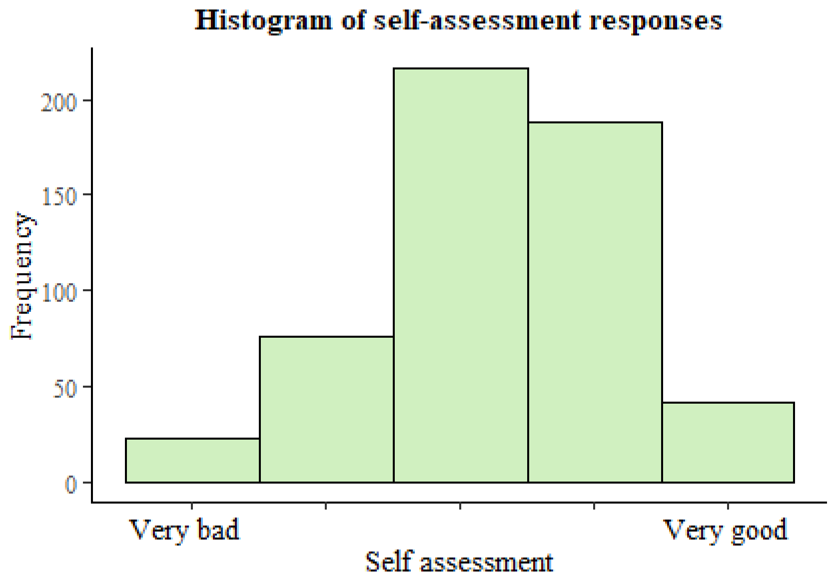

Figure 25. Histogram of participant self-assessment of their accent recognition abilities.

5.3 Results

5.3.1 Overview

The main aim of the accent recognition experiment was to test whether there was an effect of lexical frequency on the recognition of accents in isolated words. It did not find such an effect. This was true across listener groups: neither participants with a lot of experience with the accents in the quiz, nor those without much experience were better at recognising accents in high frequency words.

The results also find that listeners were not more likely to correctly recognise an American accent in stimuli that are more frequent in American English (on the basis of corpus frequencies). Furthermore, the experiment found large differences between how easily different recognisable sublexical variants were recognised correctly. Anecdotal evidence from listener comments also suggests that listeners perceived there to be fewer voices involved in the task than was the case.

After discussing the way the data were processed, I discuss the main analysis of the effects of lexical frequency and recognisable sublexical variants on the basis of the data from the target audience sample. After discussing this, I explore if different demographics show different effects for lexical frequency and recognisable sublexical variants and explore effects of relative frequency and voice specificity.

5.3.2 Data processing

The data from the experiment were processed as follows: all participants who finished the full experiment were selected and were included in the data set if they performed above 40% in the practice round (i.e. comfortably above chance; this removed 95 participants), and above 40% in the critical items (this removed another 21 participants). All low-frequency items that more than 5% of respondents reported to not recognise as real words in the questions of the experiment were removed, as well as the high-frequency they rhymed and were matched with in order to limit the influence of the recognisable sublexical variant's phonetic environment on recognition.

The data were analysed using mixed-effects logistic regression models (using the 'lme4' package (Bates et al. 2015), and the 'bobyqa' optimiser in R (R Core Team 2013)), with accuracy as the binary dependent variable, and lexical frequency and recognisable sublexical variant as the main independent variables. Separate models were fitted to test the effect of lexical frequency as a categorical variable ('high' or 'low') and lexical frequency as a continuous variable (log-transformed N of occurrences in the spoken section of the British National Corpus). As this never changed the outcome only the models with the continuous variable are reported below. As in previous chapters, the best-fit models were reduced from maximal models by removing effects if they did not significantly improve the fit of the model.

Demographic variables which had values with very few data points were reshaped to create broader categories with more data points, in order to prevent convergence issues. This meant that any 'prefer not to say' or 'other' responses to demographic questions were removed, that more fine-grained participant locations were grouped together, and that experience living in the three accent regions was transformed into a yes-or-no question rather than an exact number of years of experience (with one year counting as 'yes').

5.3.3 Lack of lexical frequency effects

To answer the main research question of the experiment on the effect of lexical frequency on accent recognition I fitted mixed-effects logistic regression models with lexical frequency of the stimulus as the dependent. The most complex models included an interaction between lexical frequency and recognisable sublexical variant. This interaction was added in order to see if there were effects of lexical frequency on the recognition of stimuli which were more difficult to recognise, as in previous research some exemplar effects have only been found when processing was more difficult or delayed (McLennan, Luce & Charles-Luce 2003; McLennan & Luce 2005).

The maximal model was reduced to include only predictors which significantly improved the fit, by taking away one predictor step by step. This meant that a participant's share of regular contacts with people with other accents was taken out, as well as experience with any of the accent regions. Speaker was removed as a random effect as it did not improve fit of the model. Furthermore, the interaction between recognisable sublexical variant and lexical frequency did not improve fit of the model (Chisq. = 10.167, p = 0.25). The model that was the best fit is summarised in Table 27. The 'Odds Ratios' column shows how much more likely it is that someone answered a stimulus correctly (e.g. participants who went to university in this model were found to be 1.22 times more likely to correctly respond to a stimulus than those who did not).

Table 27. Best fit model for correct accent recognition in stimuli (target audience sample). The baseline for the fixed effect of recognisable sublexical variant is responses to the well-recognised SSBE BATH *vowel.*

Predictors	**Odds ratios**	**Confidence interval**	***p*-value**
(Intercept)	18.84	10.91 – 32.52	<0.001
Recognisable sublexical variant = FACE	0.11	0.06 – 0.20	<0.001
Recognisable sublexical variant = GOAT (SSBE)	0.05	0.03 – 0.10	<0.001
Recognisable sublexical variant = GOAT (Yorkshire)	0.19	0.10 – 0.36	<0.001
Recognisable sublexical variant = KIT	0.01	0.00 – 0.01	<0.001
Recognisable sublexical variant = LOT	0.28	0.14 – 0.52	<0.001
Recognisable sublexical variant = rhoticity (SSBE)	0.05	0.02 – 0.09	<0.001
Recognisable sublexical variant = rhoticity (Gen. Am.)	0.14	0.07 – 0.26	<0.001
Recognisable sublexical variant = STRUT	0.16	0.08 – 0.32	<0.001
Self-assessed accent recognition skill	1.23	1.14 – 1.32	<0.001
Went to university = yes	1.22	1.02 – 1.46	0.027
First speaker occurrence = no	1.22	1.14 – 1.31	<0.001

The most important finding is that lexical frequency did not significantly improve the fit of the model. Because of this it is not included in Table 27. ANOVA comparisons between the best fit model and the models that had lexical frequency as a predictor showed that they did not improve the fit of the model significantly (Chisq. = 0.2062, $p = 0.64$). Figure 26 illustrates the lack of difference between the high and low frequency conditions in the model which included categorical lexical frequency in an estimated marginal means plot (calculated with the 'emmeans' package (Lenth 2020) in RStudio). Estimated marginal means are the mean responses adjusted for the other variables in the model. As can be seen here, the error bars overlap considerably, indicating there is no significant effect between the groups.

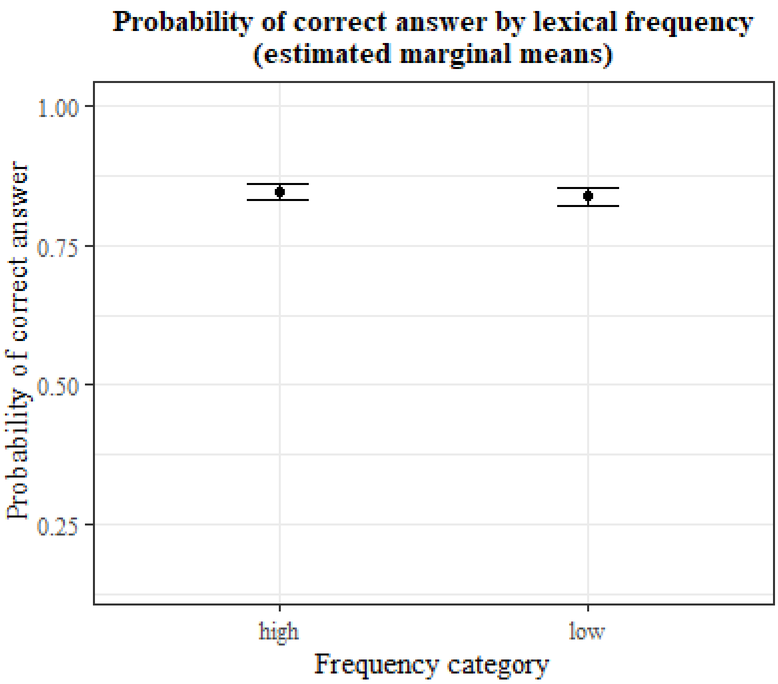

Figure 26. Estimated marginal means for lexical frequency in the most reduced model which still included lexical frequency as a predictor.

The best fit shown in Table 27 above included the random effects of participant and stimulus word, which both improved the fit of the model (for participant as a random effect, Chisq. = 687.44, $p < 0.001$;

for stimulus word as a random effect, Chisq. = 357.3, p <0.001). Furthermore, the fixed effects self-assessment, university background and whether listeners heard a voice they had heard once before were found to significantly improve the fit of the model.

In the model in Table 27, the recognisable sublexical variant included in a stimulus is significant for all different sublexical variants when the baseline is the highly recognisable SSBE BATH vowel. Estimated marginal means were calculated in order to compare between the sublexical variants. These are plotted in Figure 27 below.

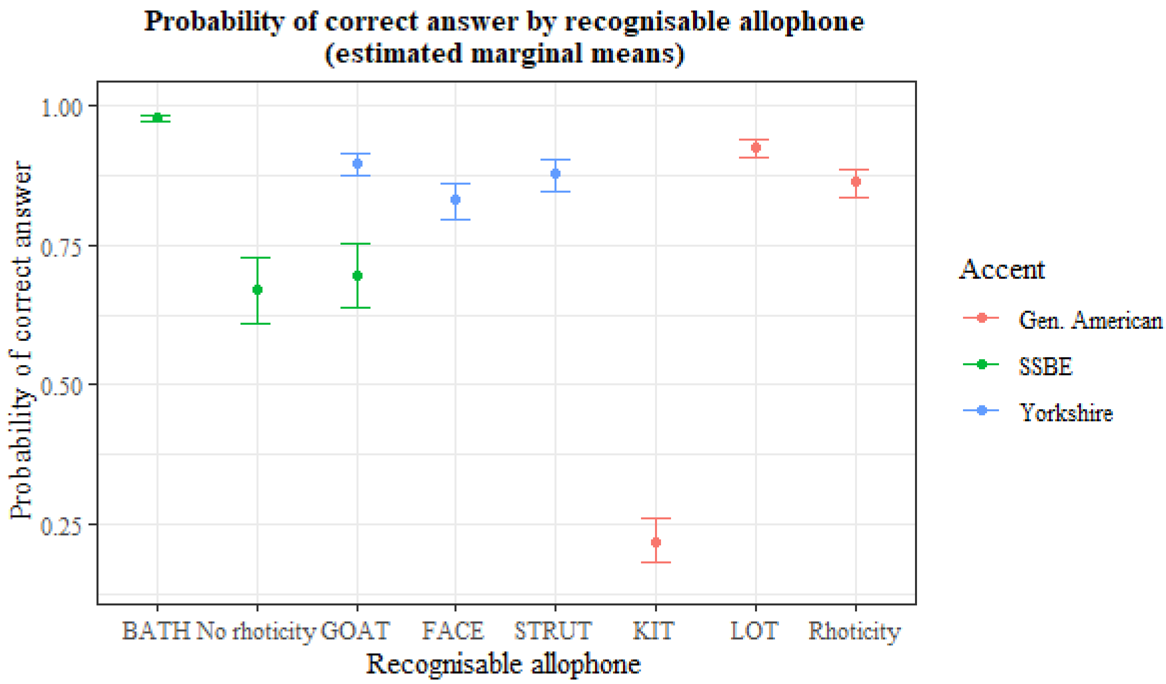

Figure 27. Estimated marginal means of the probability of a correct answer per recognisable sublexical variant.

As can be seen in the figure, the SSBE BATH vowel is recognised at ceiling level and the General American KIT vowel seems to be recognised at chance level (25% for the four answering options). The other sublexical variants float somewhere in the middle, with the General American LOT vowel seeming to be slightly easier than most, and the SSBE GOAT vowel somewhat more difficult, as well as the SSBE lack of rhoticity.

The sublexical variants that were recognised at ceiling and at chance level were removed in a subsequent logistical mixed-effects model, to check whether these stimuli lessened any lexical frequency effects in stimuli with recognisable sublexical variants with more variation. This was not the case. The inclusion of a lexical frequency predictor did not improve fit of the model in which these sublexical variants were excluded (Chisq. = 1.4259, p = 0.2324). This means that the lack of a frequency effect on accent recognition is not due to this potential methodological issue.

5.3.4 Lexical frequency across experience levels

In order to test whether there are differences in how listeners with different experience levels process accents and recognise them, a smaller dataset was compiled which included only responses to the Yorkshire stimuli, with demographic groups which were more or less likely to be familiar with the Yorkshire accent: a low-experience group (non-native speakers, from places other than the British Isles or the US, with no experience living in Yorkshire, the US, or the South East of England), a medium-experience group (native-speakers from the British Isles, not living in Yorkshire, and with no experience living there), and a high-experience group (native-speakers living in Yorkshire, who have lived there for at least a year). Linguists were excluded across experience groups. The maximal models that were run included all significant fixed effects from the target audience sample model and an interaction between experience group and lexical frequency (as a continuous variable):

```
FullYorkshireExperienceDataSet <- glmer(CorrectResponse ~
LogFrequencyBNC * ExperienceGroup + RecognisableSublexicalVariant
+ SelfAssessment + FirstSpeakerOccurrence

   (1 | Word) + (1 | ParticipantNumber), family = “binomial”, data
   = YorkshireExperienceDataSet)
```

The interaction between the lexical frequency predictors and the groups with experience levels with Yorkshire English did not improve the fit of the model (Chisq. = 0.3157, p = 0.854). The model with the best fit is summarised in Table 28 below and included self-assessment and first speaker occurrence as variables that improved the fit of this model, but excluded recognisable sublexical variant as a predictor.

Table 28. Best fit model for responses to the Yorkshire critical stimuli, with experience level included as a predictor. The baseline for experience is the high-experience group.

Predictors	**Odds ratios**	**Confidence interval**	***p*-value**
(Intercept)	3.53	2.07 – 6.01	<0.001
Experience = medium	0.91	0.67 – 1.23	0.535
Experience = low	0.37	0.27 – 0.51	<0.001
Self-assessed accent recognition skill	1.29	1.16 – 1.43	<0.001
First speaker occurrence = no	1.26	1.13 – 1.41	<0.001

As can be seen here, there was no difference in recognition between the high-experience group (the baseline) and the medium-experience group, whilst the high-experience group did perform better than the low-experience group (the low-experience group is 0.37 times as likely to correctly classify a stimulus). This is also shown in the estimated marginal means plot below:

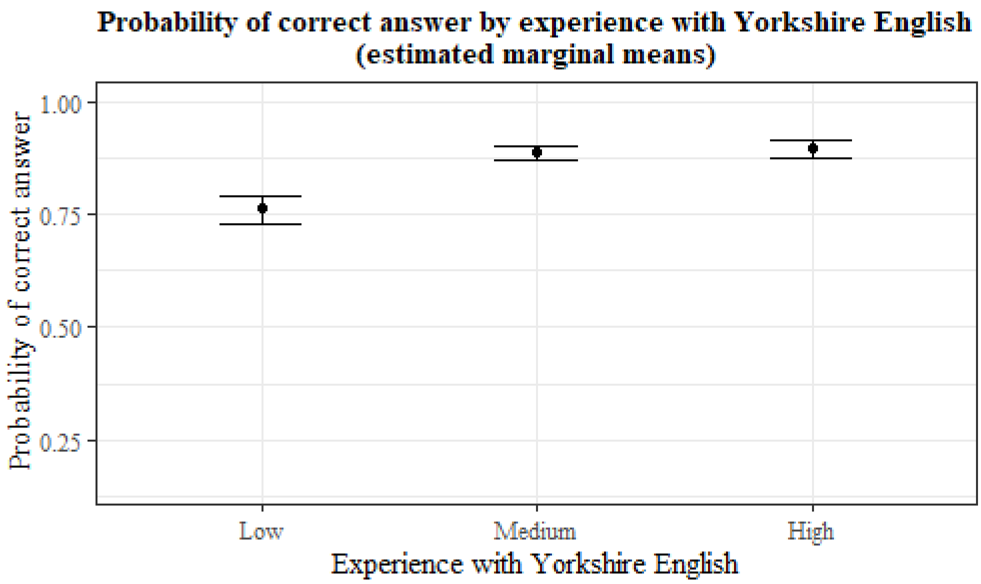

Figure 28. Estimated marginal means of the probability of a correct answer by level of experience with Yorkshire English.

This same procedure was repeated for the General American stimuli. This time, two experience groups were used: a low-experience group (non-native-speakers from outside the US and the British Isles) and a high-experience group (native speakers who were living in the US and had at least a year of experience living in the US). The following full model was run:

```
FullAmericanExperienceDataSet <- glmer(CorrectResponse ~
LogFrequencyBNC * ExperienceGroup + RecognisableSublexicalVariant
+ SelfAssessment + FirstSpeakerOccurrence +

   (1 | Word) + (1 | ParticipantNumber), family = “binomial”,
   data = AmericanExperienceDataSet)
```

Here, again, the interaction did not improve fit of the model either (Chisq. = 0.0555, p = 0.8137). The best fit model is summarised in Table 29 below.

Table 29. Best fit model for responses to the General American critical stimuli. The baseline for the recognisable sublexical variants is the KIT *vowel.*

Predictors	Odds ratios	Confidence interval	*p*-value
(Intercept)	0.27	0.15 – 0.51	**<0.001**
Recognisable sublexical variant = LOT[22]	5.54	3.07 – 10.01	**<0.001**
Recognisable sublexical variant = rhoticity	8.92	4.92 – 16.17	**<0.001**
Self-assessed accent recognition skill	1.35	1.15 – 1.58	**<0.001**

Interestingly, in the case of the General American stimuli, experience itself did not improve the fit of the model (Chisq. = 1.6576, p =0.1979). High-experience listeners did recognise the General American stimuli better, but not significantly, as can be seen in Figure 29 below. This suggests that the amount of experience most participants have with American Englishes (e.g. through TV and films if they were not American themselves) is enough to reliably recognise it.

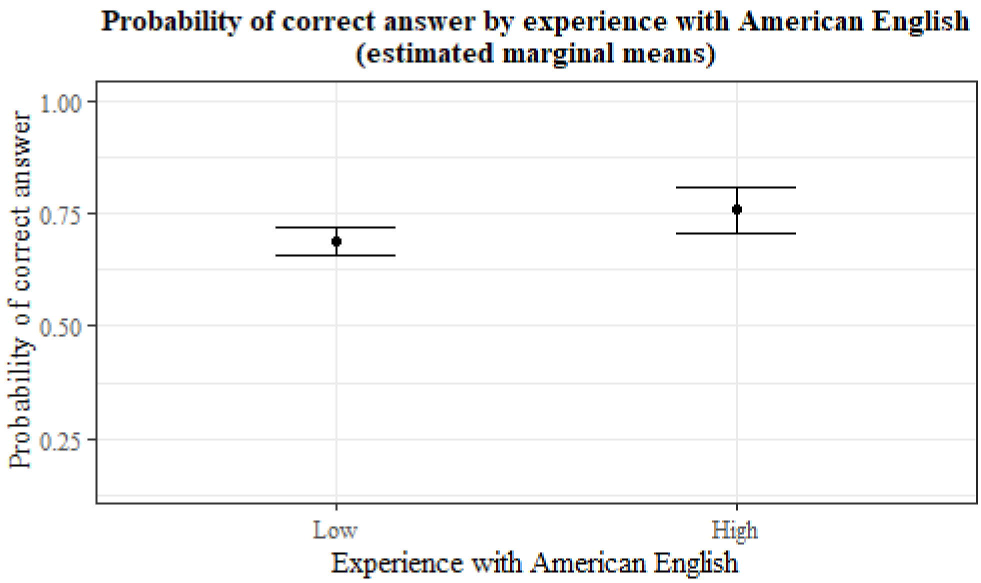

Figure 29. Estimated marginal means of the probability of a correct answer by level of experience with American English.

[22] The odds ratios for the recognisable allophones of the unrounded LOT vowel and rhoticity are very high, but this can be explained by the fact that the KIT vowel allophone was recognised at chance levels.

5.3.5 Other measures

In order to explore whether there might be an effect of the relative lexical frequency of different stimuli in different accents, I ran a model exclusively on the American stimuli where I introduced the ratio of occurrences in the spoken sections of the British National Corpus to occurrences in the spoken section of the Corpus of Contemporary English as a predictor. It included all predictors that significantly improved the fit of the above model based on American stimuli, but with the target audience sample as the sample again. This again did not improve the fit of the model (Chisq. = 0.3892, p = 0.5327).

In addition to the measures reported by the experimental software itself, the high number of participants and its distribution online also meant that a relatively large number of people gave explicit commentary on the quiz, which provided some further insight into how participants consciously experienced the tasks. Despite not systematically collecting feedback, I received numerous comments on social media platforms where participants reported that they started to recognise the voices after a while, making the experiment easier. When asked how many voices they perceived to be included in total, most believed there to be fewer than 10 voices included in the experiment - much lower than the actual 44 voices that were used. This suggests that listeners may generalise (more than is accurate even) over speaker voices and that the sociolinguistic representations used to process the social meaning of speakers may not be specific enough for listeners to distinguish between individual speakers.

5.4 Discussion and conclusions

5.4.1 Lexical frequency

The accent recognition experiment found no effect for the lexical frequency of a stimulus word on the likelihood of it being recognised correctly. This was the case for the target audience sample of non-linguist native speakers of English, as well as for participants with different levels of experience: there was no interaction between lexical frequency and participant experience with the accents used in the experiments. This suggests that the explicit recognition of accents is not helped by the use of word-sized representations in the processing of socio-indexical meaning, as is sometimes hypothesised in exemplar models of sociolinguistic processing (Walker & Hay 2011; Kim & Drager 2018). What does it mean that no such effect was found in the current experiment?

Most importantly, it suggests that word-sized linguistic representations may not play a very central role in accent recognition. It seems that when we look at how linguistic information influences social processing, rather than how social information influences linguistic processing (like in Walker & Hay 2011; Kim & Drager 2018), the role of lexical representations is different. This matches up with findings from Chapter 3 which found inhibitory priming effects when linguistic information was the prime and the social information the target and facilitatory priming when this was reversed.

Beyond the directionality of the effect there is another reason for this difference with Walker & Hay's (2011) and Kim & Drager's (2018) findings. The current experiment tested the relationship between linguistic information and social meaning with a stricter focus on social meaning itself. The participants in Walker & Hay's (2011) and Kim & Drager's (2018) experiments were presented with social meaning in a relatively linguistic way: the phonetic, and therefore *linguistic*, detail which indicated that someone's voice was old. In my experiment, participants were asked to indicate the accent they thought they heard, which circumvents this linguistic aspect. Maybe the phonetic and lexical aspects of the stimuli in Walker & Hay's (2011) and Kim & Drager's (2018) experiments facilitated each other at the level of socially meaningful variation on the basis of distributional patterns of what words are more often used by older people and what words are more often used by younger people, rather than the social meaning of the patterns. Such an effect would then not show up in an experiment about the processing of social meaning itself (the current one).

The idea that lexical units do not play a central role in accent processing may be a problem for a pure form of exemplar theory where only word-sized exemplars are behind linguistic and sociolinguistic processing. One of the ways this pure exemplar theory elegantly fits sociolinguistic processing is that these exemplars are thought to include the socio-contextual detail of the word-memories, and therefore the sociolinguistic information (see Foulkes & Docherty 2006). In such a model it is these word-memories (lexical exemplars) which carry socio-phonetic detail and sociolinguistic meaning. It would then be expected that the activation of this sociolinguistic meaning is influenced by those lexical exemplars. The current experiment did not find evidence of lexical frequency affecting the recognition of social meaning, in order to support this notion.

Hybrid models of exemplars, however, open up the possibility that it is not exclusively these exemplars that are used in linguistic processing but also representations of different levels of abstraction and different sizes (see McLennan, Luce & Charles-Luce 2003) (even if in these models size and abstraction are not always disentangled, as I argued in Section 2.3). The lack of lexical effects in the current experiment seems in support of the size aspect of this: accent recognition does not seem to exclusively rely on lexical representations.

Still, the lack of an effect found in this experiment cannot be taken as direct evidence that there is no role whatsoever for word-sized units in accent recognition. As I argued in Section 5.2.3, there are at least three main ways in which word-sized representations could influence accent processing, including absolute lexical frequency, recency, and comparative frequency between varieties.

The experiment was designed around the first measure: absolute frequency, and did not find a difference in recognition based on this as a variable. The other three measures were also explored: recency most likely on average overlaps with absolute lexical frequency (see Section 5.2.3), for which there was no effect. Comparative frequency was explored within the set of stimuli that was used to measure absolute frequency, by using its comparative frequency in a British and an American spoken corpus (BNC and COCA), and was not found to have a significant effect on recognition. However, an experimental design which takes recency or comparative frequency as its main variable could provide more certainty on the matter. In this case, it would be possible to create more clear-cut conditions for these variables, and filter out more of the noise that was present in the current experiment, where other conditions such as recognisable sublexical variants were not counterbalanced between stimuli which are more frequent in British Englishes than American Englishes.

Finally, the lack of an effect does not necessarily imply there is no effect. It is of course possible that there were methodological issues which prevented such an effect from showing up in the results. Therefore, to further explore the role of lexical and sub-lexical representations in accent processing, it is important to now test the alternative account: do sub-lexical representations play the most central role in accent recognition? This is tested in the next chapter.

5.4.2 Detail of mental representation

Whilst the experiment was not specifically set up to answer questions around the detail of the mental representations used in the processing of sociolinguistic variation, some of the results point towards a central role of speaker-level social meaning abstractions. Although no speakers were repeated more than once in the experiment, participants (anecdotally) perceived the speakers to be repeated much more than this. This means that, in a task which had a wide variation of 42 speaker voices, explicitly asking listeners to categorise stimuli as one of four broad sociolinguistic categories did not prevent participants from categorising the input into speaker-level generalisations. These speaker-level generalisations were not supported by the large variation in the data.

Furthermore, when a speaker was repeated, listeners were significantly better at recognising the correct accent. This may have to do with participants getting better at, and becoming more familiar with, the experiment design. However, this may not be the sole influencing factor. It is also possible that listeners are not simply matching a second instance of the same voice to the exemplar of the first instance of that voice (then performance would be the same), but potentially averaging over a larger group of similar voices with similar accents as one voice, and using this larger pool to improve recognition.

Such a process would be in line with Cai et al.'s (2017) speaker-model account of spoken word recognition and the importance of the speaker in sociolinguistic perception. Within an account of hybrid exemplar theory in which the detail and size of the linguistic units used for sociolinguistic processing are separate, it could provide support that the representations used in accent recognition are relatively abstract, and would be in the range of speaker representations or even slightly less specific than that (causing the conflation of multiple speakers into one). This is further explored in Chapter 6 which explicitly asked participants how many speakers they perceived to hear in a similar experiment.

5.4.3 Demographic factors

The demographic factors which affected the correct perception of the different accented stimuli in the experiment contained no major surprises: it was helpful for listeners to be a native speaker, to go to or have gone to university, and to have a background in linguistics. However, it was not necessarily helpful to have experience living in any of the three accent regions (Yorkshire, the US, or the South-East of England) or to (self-report to) regularly be in contact with people with different accents. This may have been because of the broad accent categories which were used in the experiment: SSBE, General American and Yorkshire English may be so widely known and different from each other that no intimate knowledge of the varieties was required to recognise them. This is in line with findings from other accent recognition research which finds that broader accent categories are recognised much more accurately (Van Bezooijen & Gooskens 1999).

The only effect of experience with varieties was found in the analysis of correct responses to Yorkshire stimuli only, where participants with very low experience with the variety (non-native speakers, from places other than the British Isles or the US, with no experience living in Yorkshire, the US, or the South East of England) performed worse than participants with medium experience (native-speakers from the British Isles, not living in Yorkshire, and with no experience living there) or high experience (native-speakers living in Yorkshire, who had lived there for at least a year). There was no difference between the high-experience group and the medium-experience group however, and even the difference between the low-experience group and the others may be based more on nativeness than experience itself (if those can be separated at all). For American stimuli, not even a difference between similar low and high-experience groups was found, indicating that non-native speakers who had never lived in the US or Yorkshire or the South-East of England were able to recognise American accents as well as native speakers living in the US.

Taken together this suggests that listeners build up notions of these broad accent categories relatively quickly and easily, without improving on them very much with more experience. This may again be evidence against the idea that lexical units are central to accent recognition: many of the low-frequency items in the experiment are words which low-experience and medium-experience listeners of a variety are unlikely to have ever heard in that variety, like *orc* and *ode*, whilst they would be familiar with the sublexical variants of those stimuli in that variety, like rhoticity and the GOAT vowel. The fact that this did not negatively influence their performance suggests that it is not necessary to have lexical memory representations to perform as well as listeners who are more likely to have those representations.

5.5 Conclusion

A rapid-answer online accent recognition quiz was used to test whether listeners recognise accents better in high-frequency words than in low-frequency words. The experiment found no evidence for this: participants were not better at recognising accents in high-frequency words. This, together with the findings of Chapter 3 and 4 which suggest that social information and linguistic information may not be stored within exemplars, casts some doubt on the centrality of the lexical exemplar as the main vehicle for social meaning in sociolinguistic adaptations of exemplar theory. Effects were absent across all experience groups. There was also no evidence that listeners were better at recognising accents in stimuli which were comparatively more frequent in the variety they were pronounced in.

Furthermore, commentary from the experiment's participants suggest that listeners made inaccurate speaker-based generalisations, despite the experiment's focus on broad accent categories, suggesting that listeners automatically categorise sociolinguistic information through the lens of speaker information, or even slightly less specific representations than speakers. These anecdotal findings are further explored in Chapter 6.

As the experiment which focussed on the role of words on accent recognition did not find word effects, the question is now whether an experiment which focusses on sub-lexical representations can find sub-lexical effects on recognition. This is explored in the next chapter.

6 Testing the role of sub-lexical representations in accent processing

6.1 Introduction

In this chapter, I describe an accent recognition experiment which used non-words as its stimuli to investigate the possibility that accents are recognised on the basis of sub-lexical mental representations rather than lexical ones: in the absence of lexical representations for these non-word stimuli, listeners need to rely on sub-lexical information (Vitevitch & Luce 1998; Vitevitch et al. 1999; see below). This ties into the previous experiment described in Chapter 5, which found that the lexical frequency of stimuli had no effect on whether they were accurately recognised. This lack of a lexical effect suggests that sub-lexical units may be used to recognise accents. The experiment described in the current chapter tests that hypothesis explicitly by using non-word stimuli.

The experiment finds that participants perform much better than chance would predict when they are asked to recognise accents in the non-word stimuli, suggesting that the sub-lexical units the stimuli were comprised of facilitated accent recognition, without the need for word-sized representations. However, it was also true that the participants in the experiment did not perform as well with the non-word stimuli as they did with the word stimuli in the experiment described in the previous chapter. This may indicate that the involvement of lexical representations and word-sized input does boost accent processing, alongside the use of sub-lexical representations, or that lexical representations form an anchor point for processing sub-lexical information.

6.2. Methodology

6.2.1 Using non-word stimuli

The non-word accent recognition experiment is in as many ways as possible a duplicate of the word-based accent recognition experiment described in Chapter 5. In this experiment however, participants were asked to classify non-word stimuli, rather than word stimuli, as one of 4 broad accent categories: 'American', 'Standard English from the South of England', 'Yorkshire' and 'Any other accent'. The non-word stimuli used in the experiment were used in order to measure accent recognition in stimuli without lexical units, so that it was possible to investigate to what degree sub-lexical units are used to recognise accents. As is argued by Vitevitch and colleagues (Vitevitch & Luce 1998; Vitevitch et al. 1999), sub-lexical units play a more important role in word processing in non-words than in words, whilst lexicality, they find, does not play a role in non-word recognition.

Vitevitch & Luce (1998) find in their study that whilst listeners are slower to process real words when they have high phonotactic probability (common sound combinations), they are quicker to process non-words when they have high phonotactic probability. In processing real words, listeners need to select the right word out of a range of 'neighbours' (see the Neighborhood Activation Model by Luce & Pisoni 1998). These are word representations which are similar in their acoustic-phonetic characteristic. The selection from the range of neighbours makes it harder to process a word if it has a high number of competitor words due to the common sound combinations in the words. The fact that this does not hinder non-word processing is evidence that non-words are not processed by selecting a representation from a group of competing lexical units. Instead, as the helpfulness of having more probable sub-lexical sound combinations shows, non-words are processed on the basis of their sub-lexical information.

Staum Casasanto (2009)[23] brings the use of non-words for this purpose into practice in a reaction time experiment which investigated whether sociolinguistic expectations skew phonetic perception in non-words in the same way that they do with words. She found that participants responded differently to heard t/d-deletion on the basis of whether they were shown a White or a Black face, and she found that this effect was present both in word stimuli and in non-word stimuli. She argues that this is evidence that linguistic processing can be influenced by social factors at the level of the sub-lexical chunk.

Whilst the distinction between word and non-word processing seems relatively distinct and clear-cut, there are also theories of non-word processing which imply that there is a role for lexicality in non-word processing. For example, Gathercole (1995; Archibald & Gathercole 2006) finds that children are better at non-word repetition if non-words are more wordlike (not a very dissimilar finding from Vitevitch & Luce's (1998)), but hypothesise that this is because it is helpful to have many lexical neighbours activated which then facilitates the processing of the similar input. Gathercole & Adams (1993) also find that children with larger vocabularies are better at non-word repetition. This would be an alternative explanation to that of Vitevitch et al.'s (1999) that these effects are caused by phonotactic probability.

Furthermore, Newman and Twieg (2001) find that, compared to real-word processing, non-word processing elicits significantly more activation in the posterior cortical regions of the brain, which is where previous studies (Petrides 1996) have suggested lexical access takes place. These higher levels of activation would be caused by an increased demand on the system because it is struggling to find a matching lexical item. A counter-argument to this might be that, if the lexical system is activated because it struggles to find matching representations, this does not necessarily imply that one is actually found and used in processing. Still, the nuance that there might be some degree of lexical processing involved in non-word processing needs to be taken seriously.

One of the consequences of using non-word stimuli is that they are generally processed more slowly than word stimuli (Stanners, Jastrzembski & Westbrook 1975; Whaley 1978), in line with findings that non-words require higher demands on lexical processing systems (Newman & Twieg 2001; Raettig & Kotz 2008). This may mean that participants in a non-word based task are slower and less accurate, simply because it costs more effort to process non-words. In the context of the current experiment, this could create a gap between performance in the current non-word based experiment and in the word based one described in Chapter 5. However, it is not clear how large this gap would be, as the observed behavioural differences between word and non-word processing are not necessarily very large: the differences between response times for non-word and word stimuli in the experiment in Chapter 3 for example was only 49ms (similarly small differences are found in Stanners, Jastrzembski & Westbrook 1975; Whaley 1978).

6.2.2 Experimental design

Like the previous experiment, the current experiment was an online multiple choice task presented through the online software Gorilla (www.gorilla.sc; Anwyl-Irvine et al. 2019) which was distributed through social media. Participants were asked to classify the accent they heard in 44 non-word stimuli as one of 4 broad accent categories: Yorkshire, American, Standard English and a distractor group of "other" accents. The 44 items used in the experiment were pronounced by the same 42 speakers used

[23] Staum Casasanto (2009) used the term 'nonce words' rather than 'non words' and some of the papers cited below use 'pseudowords'. Some studies (e.g. Stark & McClelland 2000) use the terms 'non-words' and 'pseudowords' to mean different things: 'pseudowords' is then used to mean non-existing words which are phonotactically legal, whereas 'non-words' is then used to mean non-existing words which are phonotactically impossible. In this PhD thesis I do not follow this distinction and use 'non-words' to describe phonotactically possible non-existing words, as is most common in linguistic papers.

in the previous experiment,[24] and a new a practice round was constructed with 12 items (all different new speakers).

All stimuli in the current experiment were non-words created to match the high and low-frequency words from the previous experiment as closely as possible in their phonetic form. If a speaker's recordings of the stimuli *run* and *shun* were included in the previous experiment, that same speaker's recording of the non-word *zun* (from the same recording sessions) was included in this experiment. This was done in order to keep as many circumstantial influences – recognisable sublexical variant, phonetic environment, speaker voice – as constant as possible. This means that the non-words were not selected from a non-word database like the ARC Non-Word Database (Rastle, Harrington & Coltheart 2002) which was used for the priming experiments in Chapters 3 and 4. However, the created non-words did follow the restrictions used in the creation of this database: all stimuli were phonotactically legal in English, following Rastle, Harrington & Coltheart's (2002) methodology. As the stimuli were created to rhyme with the high-frequency stimuli from Chapter 6, they all had at least one lexical neighbour, and were relatively word-like. As Raettig & Kotz (2008) argue, this may mean they were processed more similarly to how real words are processed. The full list of stimuli is shown in Appendix IV.

The design aimed to test two experimental research questions:

1. Are listeners able to recognise accents in non-words better than chance would predict?
2. How does the recognition of accents in non-words in this experiment compare to the recognition of accents in real words in the previous experiment?

These two questions will be able to give new insight into the role of sub-lexical representations in accent recognition, especially when compared to the role of lexical representations. Listeners being able to recognise these accents above chance levels implies that they can rely on non-lexical representations. However, recognition being much lower in the non-word experiment than it is in the word-based experiment implies that lexical representations do play an important role. One caveat here is that some of this could be caused by higher processing costs involved in processing non-words.

In order to further investigate how listeners evaluate their own perception processes, this experiment included a post-hoc questionnaire which asked them about their own experience processing the non-word stimuli. As self-perception of the automatic processes involved in accent recognition is not a reliable source for insights in these processes (see for example Campbell-Kibler 2016), the responses to these questions cannot be taken as direct evidence for the use of lexical or sub-lexical units in accent processing. However, if these self-perceptions were to influence participant accuracy, this could be taken as stronger evidence.

Finally, the post-hoc questionnaire asked listeners to indicate how many different speakers they perceived to be included in the experiment, in order to further explore the anecdotal findings from Chapter 5 that listeners perceived hearing the same speakers much more often than was the case. This yielded two further experimental research questions:

3. What linguistic units did participants perceive their accent recognition to be based on in non-words? Did this influence their performance?
4. How many different speakers did participants think they heard? Did this influence their performance?

6.2.3 Stimuli and procedure

The stimuli and procedure of the experiment were mostly identical to those of the previous experiment. Again the participants were first presented with an online participant information sheet about the ethics

[24] As in the previous experiment, two speakers were reused once for a distractor item, due to the slightly lower number of available General American speakers.

of the experiment and were asked to confirm that they consented.[25] The same accents, recognisable sublexical variants, speakers, and variants were used for the stimuli, and the procedure of the task was identical: listeners had 4 seconds to respond to each item and were given 4 answering buttons. They were presented with the orthography of the stimuli during the experiment. Figure 30 shows a screenshot of how a practice stimulus was presented in participants' browsers whilst they heard the stimulus 'leabring'.

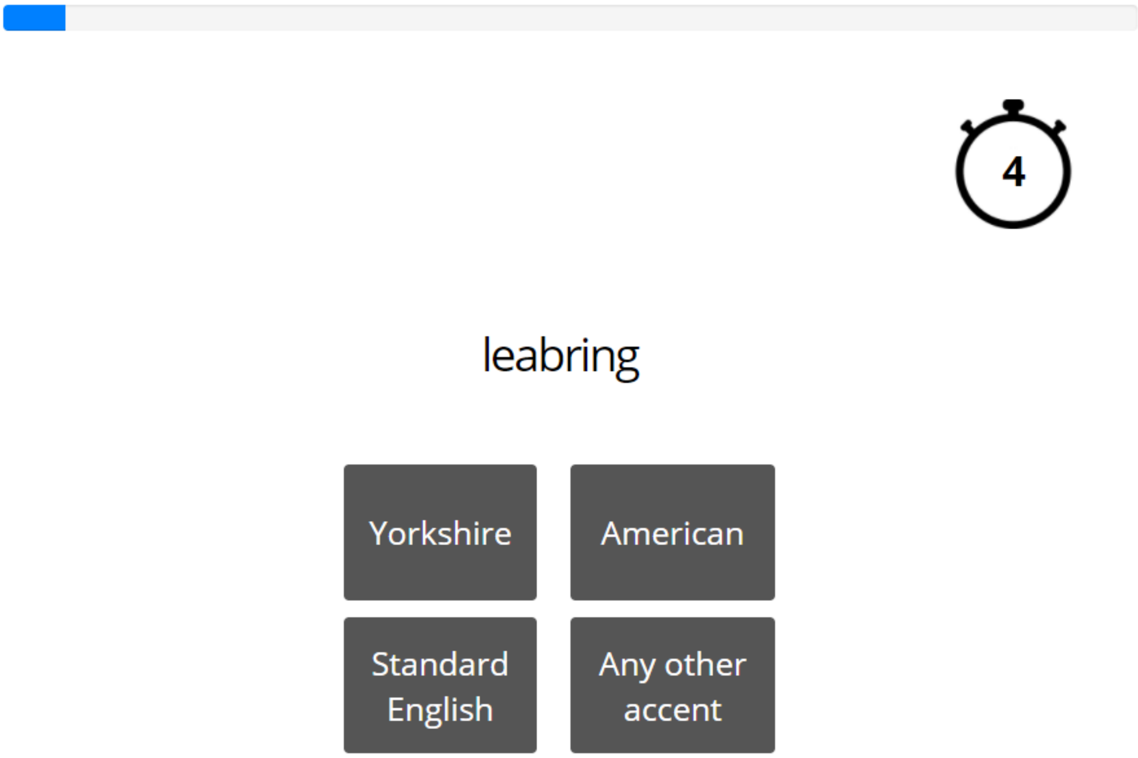

Figure 30. Screenshot of the practice stimulus 'leabring' at the start of the experiment.

Participants were given a score at the end of the quiz and were encouraged to share this online for further promotion of the experiment. As the experiment did not compare between a high and a low frequency conditions, the number of items was halved to only be 44 items, with none but two distractor items being pronounced by the same voice more than once.

After the main task, a new series of post-hoc questions was introduced to investigate participants' own perception of their recognition processes. Firstly, the participants were asked how many different voices they perceived hearing in the stimulus items, on a scale from 0 to 44 items (the total number of items). This was to be able to quantify participant observations about recognising the same speaker multiple times. The next question concerned how listeners felt they were able to recognise the different accents in the non-words in the experiment. There were four answering options for this:

- It was intuition.
- I focussed on the sounds of the fake word.
- I thought about what real words the fake words sounded like
- Other (please specify)

[25] The experiment went through the University of Sheffield's ethical review process and was approved.

Participants were able to select multiple options. Finally, this section included a comment box for any further comments.

6.2.4 Participants

The participant sample for the current experiment was smaller than the previous experiment described in Chapter 5: 155 participants in total. As the experiment did not gain as much online traction as the previous experiment did, participants were more actively recruited and often individually asked if they would be interested in participating. The number of linguists and native speakers is cross tabulated in Table 30 below.

Table 30. Cross table of linguist and native-speaker status for the participant sample. The target audience sample is boxed and in bold.

	Native speakers	Non-native speakers	Prefer not to say	Total
Linguist	43	17	0	60
Non-linguist	$\boxed{\mathbf{61}}$	33	1	95
Total	104	50	11	155

As in Chapter 5, the non-linguist native speakers were taken as a 'target audience' sample, on which the sample summaries below are based. The main part of the target audience sample consisted of participants from South-Yorkshire and the South-East of England (as can be seen in Table 31 below). The sample was quite balanced for gender (Table 32), although the sample included almost exclusively people with a university degree and university students (Table 33) and was quite young (Figure 31).

Table 31. Number of participants in the target audience sample by location. Answering options that were smaller (e.g. different parts of Yorkshire) are added up within the table, to show the frequency of their larger regions.

Location	Count	Percentage
Scotland	3	5%
North of England:	34	56%
North-West	*5*	*8%*
North-East	*5*	*8%*
Yorkshire:	*24*	*39%*
North Yorkshire	*2*	*3%*
East Yorkshire	*0*	*0%*
West Yorkshire	*1*	*2%*
South Yorkshire	*21*	*34%*
Midlands:	3	5%
East Midlands	*3*	*5%*
West Midlands	*0*	*0%*
Wales	1	2%
South of England:	18	30%
South East	*15*	*25%*
South West	*3*	*5%*
Northern-Ireland	0	0%
Republic of Ireland	0	0%
USA	0	0%
Other	2	3%
Total	61	100%

Table 32. Number of participants in the target audience sample by gender.

Gender	Count	Percentage
Male	32	52%
Female	27	44%
Non-binary	2	3%
Total	61	100%

Table 33. Number of participants in the target audience sample by type of education.

Education	Count	Percentage
University	56	92%
College (UK)	3	5%
Secondary school	2	3%
Total	61	100%

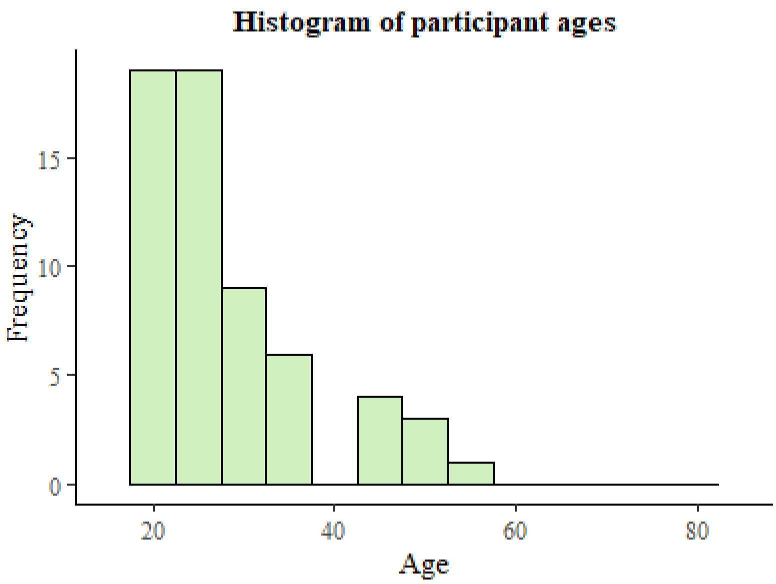

Figure 31. Histogram of participants in the target audience sample by age.

A large share of participants had lived in one of the British accent regions that were used in the experiment, as can be seen in Table 34: 59% of participants reported to have had at least a year of experience living in Yorkshire, and 46% in the South-East of England. Only three reported having lived in the United States. Few participants reported being in contact with no speakers with different accents or exclusively speakers with different accents, with the most respondents indicating that a couple of their weekly contacts had different accents (see Table 35).

Table 34. Number of participants in the target audience sample who had experience living in one of the three regions where the critical accents of the experiment were spoken, for at least one year.

Participants with experience living in	N	Perc.
Yorkshire	36	59%
South-East of England	28	46%
USA	3	5%

Table 35. Self-reported share of weekly contacts with a very different accent.

Indicated share of weekly contacts with very different accent	**N**	**Perc.**
None of them	1	2%
A couple of them	26	43%
About half of them	17	28%
Most of them	14	23%
All of them	3	5%
Total	61	100%

Participants' self-assessments of their ability to recognise accents skewed neither very positively nor very negatively, as can be seen in Figure 32. This means the sample is somewhat less positive about itself than the sample for the previous experiment. This may be because the current experiment's participants had to be recruited more actively, where participating was likely seen as more of a favour to the researcher, whilst the previous experiment managed to reach and attract a high number of curious participants with no relation to the researcher, based on enthusiasm, and perhaps confidence in their own accent recognition skills. As self-assessment was a significant predictor of performance in the previous experiment, the current sample may simply be somewhat worse at accent recognition, because of the different recruitment styles.

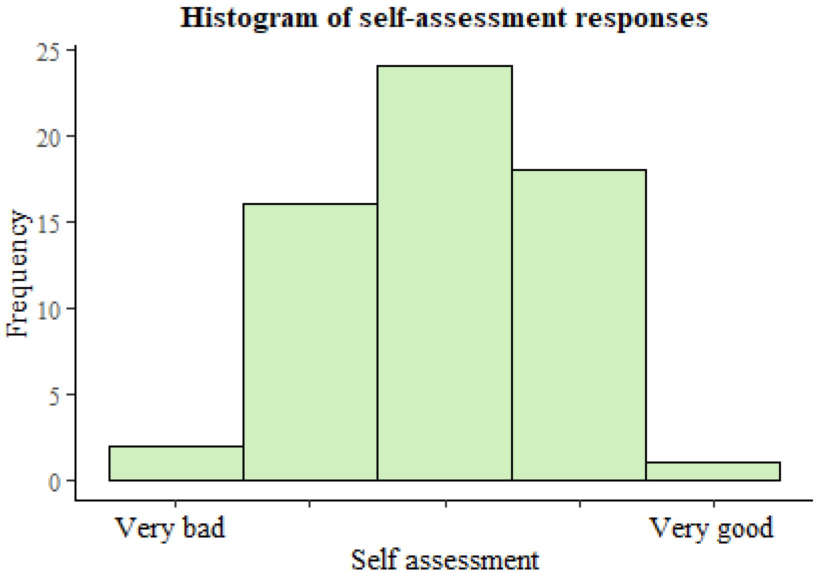

Figure 32. Histogram of participant self-assessment of their accent recognition abilities.

Finally, participants self-reported whether they had taken part in the previous experiment, described in Chapter 5. Of the participants 84% indicated not to have participated in that experiment, and 16% indicated they had, as is shown in Table 36 below. This means that, for the majority of participants, there was no possibility of them improving their accuracy on the basis of having previous experience with the task.

Table 36. Number of participants in the target audience sample by participation in previous accent recognition experiment.

Participation in previous experiment	**Count**	**Percentage**
Yes	51	84%
No	10	16%
Total	61	100%

6.3 Results

6.3.1 Overview

The experiment found that participants performed better than chance would predict in the non-word accent recognition task, although they performed significantly worse than in the previous experiment with word stimuli. Participants themselves mostly expressed that they recognised the accents on the basis of sub-lexical information, and confirmed the anecdotal evidence from Chapter 5, that they perceived fewer voices than were actually used in the experiment.

After discussing the processing of the data, I discuss the finding that participants in the current experiment performed much better than chance. After this I compare performance in this task to that in the experiment of Chapter 5. Finally, I discuss the post-hoc questions that participants were asked regarding their experience of their recognition strategies and how many speakers they perceived to have heard.

6.3.2 Data processing

Unlike the data from the previous experiment, the current experiment's data were not filtered by whether participants performed above chance in the practice round and across the board, as the main research question was whether participants performed above chance. In order to compare between the current experiment and the previous experiment's data, Chapter 5's data were also re-wrangled in the same way, including participants who scored below 40% in the practice round or overall, to create an appropriately comparative dataset.

When applicable, the data were analysed statistically using mixed-effects logistic regression models (using the 'lme4' package (Bates et al. 2015), and the 'bobyqa' optimiser in R (R Core Team 2013)), with accuracy as the binary dependent variable. Models were reduced from maximally large models by removing effects if they did not significantly improve the fit of the model.

As in Chapter 5, demographic variables which had values with very few data points were reshaped to create broader categories with more data points, in order to prevent convergence issues. This meant that any 'prefer not to say' or 'other' responses to demographic questions were removed, that more fine-grained participant locations were grouped together, and that experience living in the three accent regions was transformed into a yes-or-no question rather than an exact number of years of experience.

This chapter used two separate regression models on the same outcome data: one to compare the current experiment to the experiment in Chapter 5, using both experiments' data, and one to investigate the influence of participants' responses to the post-hoc questions, using only the current chapter's data (as the post-hoc questions were only asked in this experiment. To correct for these multiple comparisons a Bonferroni correction was applied to create a new alpha (*p*-value cut-off point). This was 0.025 (0.05 divided by the 2 comparisons). Whilst the Bonferroni correction is sometimes criticised to be too conservative (Narum 2006), it only eliminated one otherwise significant effect, which was not central to the main research questions of this chapter's experiment: American experience would have otherwise been significant in the model using only the current experiment's data ($p = 0.045$).

6.3.3 Non-word performance

The participants in the experiment almost exclusively performed better than chance would predict, as can be seen in the histogram in Figure 33, which summarises the distribution of mean participant accuracy in the critical items. As can be seen in the figure, only 4 out of the 61 participants performed below, or close to chance level, and the bulk of participants were correct close to or more than three quarters of the time.

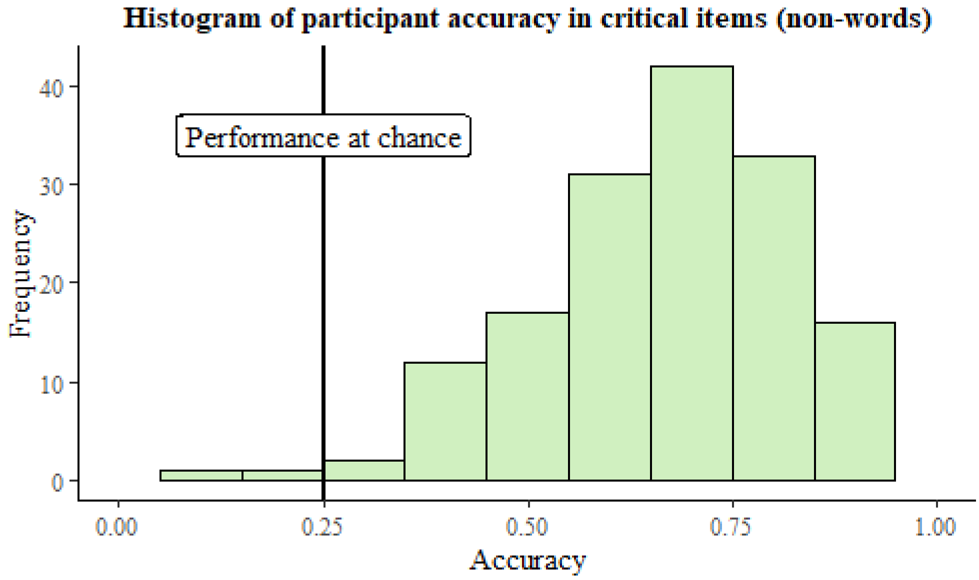

Figure 33. Histogram of participant accuracy in critical items. The frequency axis denotes the number of participants within a range of accuracy scores.

Whilst performance is clearly above chance, it is lower than it was for the word-based experiment described in Chapter 5, as can be seen in Figure 34 which summarises participant performance in that experiment. Here, the distribution is clustered further towards the higher end of the accuracy scale.

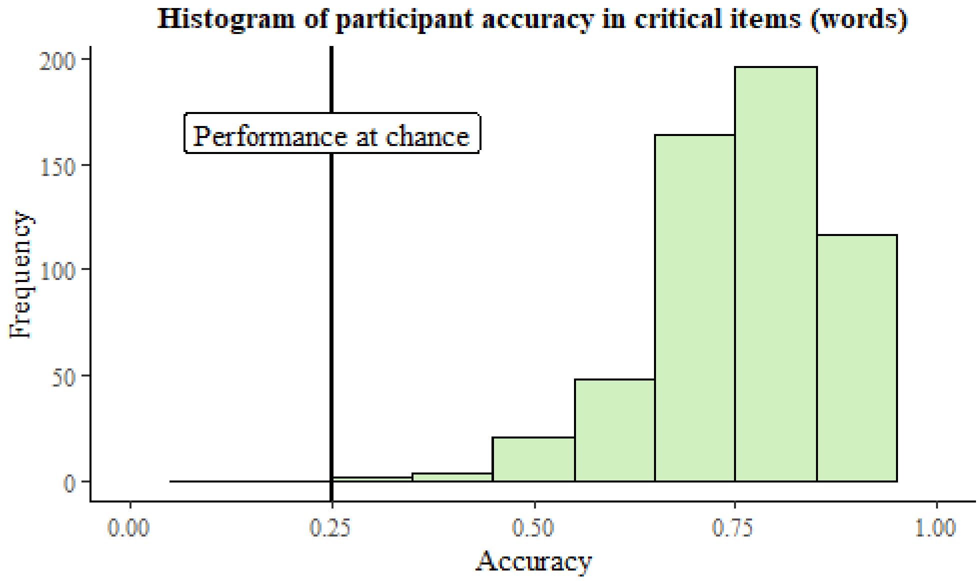

Figure 34. Histogram of participant accuracy in the critical items in the word-based accent recognition experiment described in Chapter 5. The frequency axis denotes the number of participants within a range of accuracy scores.

6.3.4 Comparing non-word and word performance

In order to compare performance between the two experiments and between the different (non-word versus word stimuli) the target audience sample datasets from both experiments (i.e. non-linguist native speakers) were combined into one. A logistic mixed-effects model was then fitted with correctness of the answers as the dependent variable and non-word status as a fixed effect, as well as the recognisable sublexical variant in the stimulus, education, experience with other accents in friends, and experience living the US, the South-East, and Yorkshire. Self-assessment is included as a random variable in order to control for the difference in mean self-assessment between the two participant groups. As this measure predicted recognition accuracy in Chapter 5, it may compensate for a potential difference in the general accent recognition skills in the different samples. The code for the maximal model is summarised below:

```
FullComparativeModel <- glmer(CorrectResponse ~ NonWordStatus +
RecognisableSublexicalVariant + Education + AccentFriends +
AmericanExperience + SouthEastExperience + YorkshireExperience +

    (1 | Speaker) + (1 | Word) + (1 | ParticipantNumber) + (1 |
    SelfAssessment), data = CombinedDatasets, family ="binomial",
    control = glmerControl(optimizer = "bobyqa",
    optCtrl=list(maxfun=100000)))
```

The model was then reduced by taking out predictors which did not significantly improve the fit of the model, which meant that experience with the different accents and the accent of participants' regular contacts were removed. The best-fit model is summarised in Table 37.

Table 37. Best-fit model for responses in the combined target audience samples of the word-based and the non-word based accent recognition tasks. The baseline for the fixed effect of the recognisable sublexical variant is responses to the well-recognised SSBE BATH vowel.

Predictors	**Odds Ratios**	**Confidence interval**	***p*-value**
(Intercept)	14.27	7.71 – 26.42	<0.001
Non-word = yes	0.53	0.38 – 0.74	<0.001
Recognisable sublexical variant = FACE	0.12	0.06 – 0.26	<0.001
Recognisable sublexical variant = GOAT (SSBE)	0.11	0.05 – 0.26	<0.001
Recognisable sublexical variant = GOAT (Yorkshire)	0.22	0.10 – 0.48	<0.001
Recognisable sublexical variant = KIT	0.01	0.00 – 0.02	<0.001
Recognisable sublexical variant = LOT	0.34	0.16 – 0.72	0.005
Recognisable sublexical variant = rhoticity (SSBE)	0.06	0.03 – 0.13	<0.001
Recognisable sublexical variant = rhoticity (Gen. Am.)	0.20	0.09 – 0.43	<0.001
Recognisable sublexical variant = STRUT	0.13	0.06 – 0.29	<0.001
Went to university = yes	1.27	1.07 – 1.51	0.006

As can be seen in the table, non-word status was a significant predictor, with non-words being recognised much less well than words. The estimated marginal means are plotted in Figure 35 to illustrate this. This is despite the fact that the regression model controls for the fact that in the non-word experiment participants thought they were worse at recognising accents (which in Chapter 5 was predictive of actual performance as well). This means that the difference in accuracy is not likely to be based on a difference in accent recognition skill between the two samples.

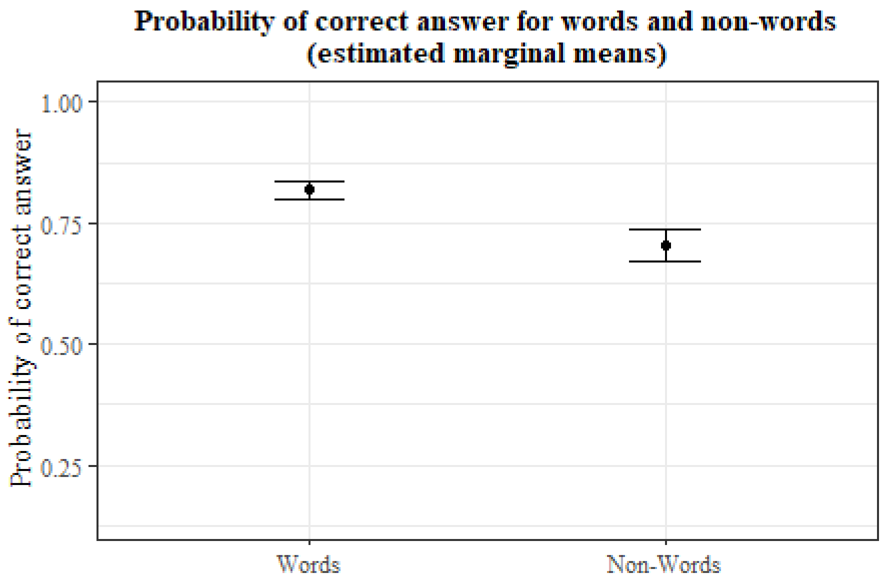

Figure 35. Estimated marginal means for words and non-words in the best-fit combined model of accent recognition in the word and non-word based accent recognition experiments.

When the estimated marginal means between the two experiments are compared at the level of the recognisable sublexical variant we see that the worse performance for non-word stimuli happens almost across the board. As can be seen in Figure 36, the estimated marginal means for each of the recognisable sublexical variants is lower in the non-word based experiment, with the exception of the General American KIT vowel, which both participants groups performed at chance for, and the SSBE stimuli with a lack rhoticity, which was one of the most ambiguous recognisable sublexical variants, as it could be a part of Yorkshire and other British varieties.

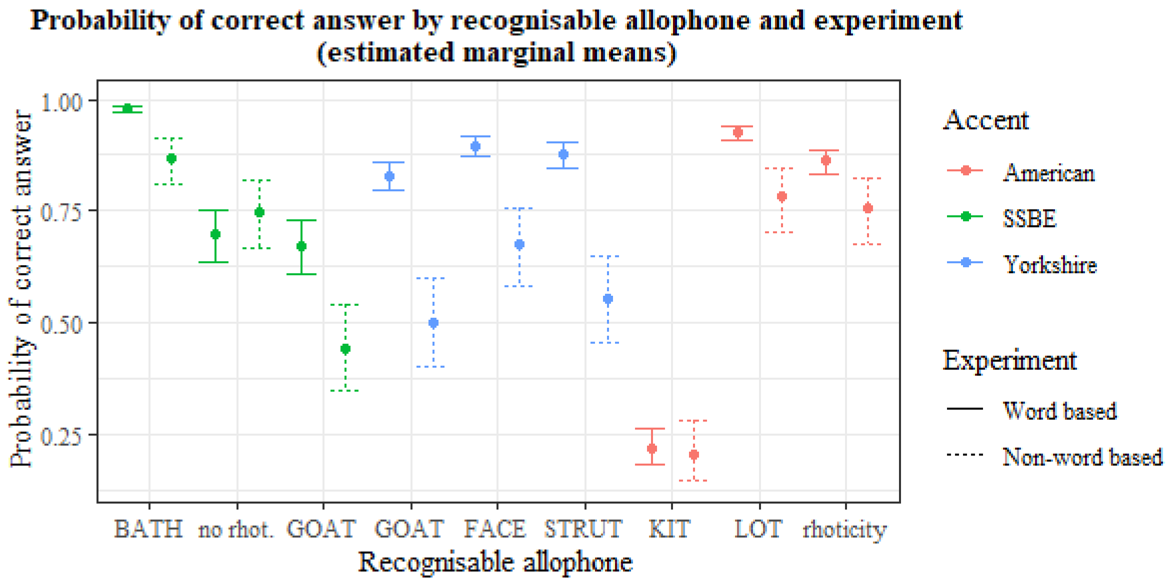

Figure 36. Estimated marginal means of the probability of a correct answer per recognisable sublexical variant, by experiment.

As the drop in recognition seems larger for the Yorkshire variants in Figure 36, a new logistic mixed-effects model was fitted to test whether there was an interaction between accent and non-word status. It was necessary to fit a new model for this, as the fixed effect of accent would overlap too much with the fixed effect of recognisable sublexical variant in the previous model. As a significant predictor in the previous model education was retained as a fixed effect. The KIT vowel items were removed from this analysis because performance for those items were at the level of chance. The new model was the following:

```
InteractionModel <- glmer(CorrectResponse ~ NonWordStatus *
Accent + Education +
   (1 | Speaker) + (1 | Word) + (1 | ParticipantNumber) + (1 |
   SelfAssessment), data = CombinedDatasetsNoKIT, family
   ="binomial", control = glmerControl(optimizer = "bobyqa",
   optCtrl=list(maxfun=100000)))
```

The interaction here did not significantly improve the fit of the model (Chisq. = 3.807, p = 0.149). This means the drop in recognition was not significantly larger for one of the different accents.

The reaction times to the items differed somewhat between the word and non-word experiment. They were somewhat slower in the non-word experiment (1944ms) than for the word experiment (1687ms), as can be seen in the reaction time histogram shown in Figure 37. This suggests that there were, as expected, higher processing costs involved in the non-word experiment than in the word experiment. However, the difference in reaction times was not so large that many participants will have timed out because of it (which would be at 4000ms).

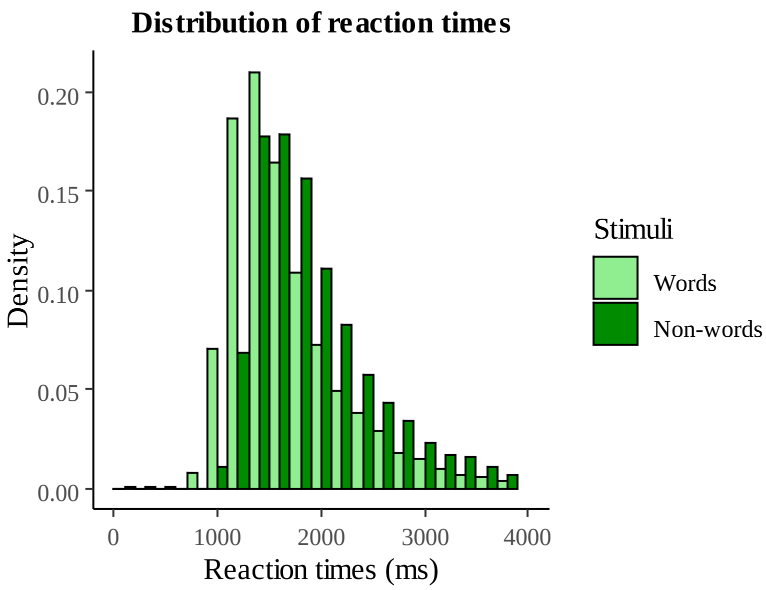

Figure 37. Density plot of reaction times to the stimuli presented in the word-based and non-word based experiments.

6.3.5 Participant strategies

In the post-hoc questions of the experiment, a majority of the participant sample indicated that the strategy they used to recognise the accents in the non-words was to use the sounds of the non-words (72%), whilst only around a fifth of the sample indicated they compared the non-word to similar words (21%), as is summarised in Table 38 below. The option of basing answers on intuition was selected by close to half the participants (48%).

Table 38. Self-reported participant strategy to accent recognition in non-words. Multiple responses were possible.

Strategy	N	Percentage
Intuition	29	48%
Using the sounds of the non-words	44	72%
Comparing to similar words	13	21%

It was tested whether these differing strategies had an effect on participant performance by fitting a logistic mixed-effects model to the non-word experiment responses with participant strategies as predictors, as well as the perceived number of speaker voices discussed in Section 6.3.6 below. The maximal model used to process the non-word only data (this time by itself) was as follows:

```
FullNonWordModel <- glmer(CorrectResponse ~ SoundsStrategy +
SimilarWordStrategy + RecognisableSublexicalVariant +
AmericanExperience + SouthEastExperience + YorkshireExperience +
SelfAssessment26 + AccentFriends + PerceivedNOfVoices +

    (1 | Speaker) + (1 | word) + (1 | ParticipantNumber), data =
    NonWordExperiment, family ="binomial", control =
    glmerControl(optimizer = "bobyqa",
    optCtrl=list(maxfun=100000)))
```

Whilst removing the predictors which did not significantly improve the fit of the model both strategy predictors were removed (for the sound-based strategy Chisq. = 1.467, p = 0.226; for the similar-word strategy Chisq. = 0.604, p = 0.437). This means that the strategy listeners reported using did not significantly affect their performance in the experiment. The best-fit model is summarised in Table 39 below.

Table 39. Best fit model for recognition accuracy in the non-word. The baseline for the fixed effect of the recognisable sublexical variant is responses to the well-recognised SSBE BATH *vowel.*

Predictors	Odds Ratios	Confidence interval	*p*-value
(Intercept)	3.30	1.26 – 8.66	0.015
Recognisable sublexical variant = FACE	0.15	0.05 – 0.42	<0.001
Recognisable sublexical variant = GOAT *(SSBE)*	0.45	0.16 – 1.27	0.131
Recognisable sublexical variant = GOAT *(Yorkshire)*	0.31	0.11 – 0.87	0.027
Recognisable sublexical variant = KIT	0.04	0.01 – 0.11	<0.001
Recognisable sublexical variant = LOT	0.54	0.19 – 1.53	0.247
Recognisable sublexical variant = rhoticity (SSBE)	0.12	0.04 – 0.33	<0.001
Recognisable sublexical variant = rhoticity (Gen. Am.)	0.47	0.17 – 1.33	0.154
Recognisable sublexical variant = STRUT	0.19	0.07 – 0.52	0.001
Self-assessed accent recognition skill	1.44	1.18 – 1.76	<0.001

[26] Self-assessment is reintroduced here as a fixed effect rather than a random effect, as it did not function as a control between the two participant groups anymore.

As can be seen in the table, self-assessment and some of the recognisable sublexical variants were found to be significant predictors in the non-word based experiment, like in the word-based recognition experiment.

6.3.6 Speaker recognition

The experiment confirmed the anecdotal evidence from the word-based experiment that listeners perceived the wide variety in speaker voices to be much smaller than the real number of voices used in the experiment. Whilst 44 different voices were used, none of the participants perceived there to be 44 different voices. Most participants guessed there were 20 or fewer voices in the experiment, as can be seen in the histogram in Figure 38.

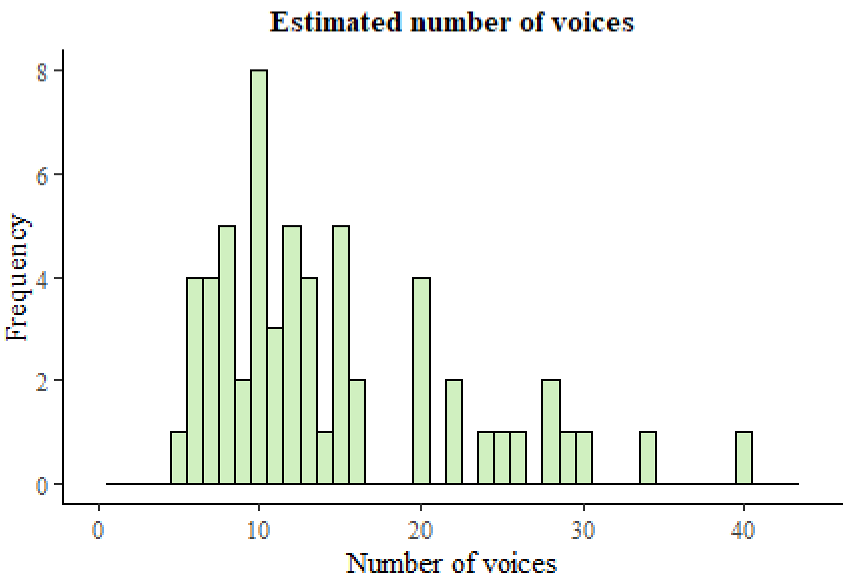

Figure 38. Histogram of participant estimations of the number of voices used for the experiment.

Whilst the spread of this estimated number of voices was relatively large, this had no effect on participant performance either: adding this as a fixed effect to the best-fit logistic regression model in the model described above in Section 6.3.5 did not improve its fit (Chisq. = 0.275, p = 0.6).

6.4 Discussion and conclusions

6.4.1 The role of lexical and sub-lexical units in accent recognition

The experiment described in the current chapter found support for the hypothesis that accent recognition can be based on sub-lexical mental representations: participants recognised accents in isolated non-word stimuli at levels much higher than chance would predict. Furthermore, at an explicit, conscious level participants mostly indicated that they based their recognition on the sounds of the words rather than searching for similar words, although this did not affect accuracy. This ties into hybrid models of exemplar theory which argue that, in linguistic processing, not just lexical exemplars are used, but also different, more abstract representations (Pierrehumbert 2016).

Whilst the current experiment did not directly test the level of abstraction of the mental representations used to recognise accent information, it did test the size parameter of the linguistic representations used for the processing of social meaning: it is not necessary to have lexical input to process the social meaning of that input, suggesting that sub-lexical units are used to access social meaning (in line with Staum Casasanto’s (2009) finding that the processing of non-words can be influenced through social cues).

In addition to this, it seems that these representations are sub-lexical units, rather than general articulatory setting differences between accents (Wieling & Tiede 2017). If general articulatory setting difference were to help the perception of social meaning, we would expect this to occur for any stimulus, regardless of the sublexical variants included. However, in both experiments recognition of the American KIT vowel was at chance levels. This suggests that with the lack of a reliably recognisable sublexical variant, listeners are not able to categorise stimuli as the correct accent, and are not relying on different information like a speaker's articulatory setting (or at least for the KIT vowel in General American English).

Whilst listeners are able to recognise accents on the basis of non-word stimuli, there is still a significant gap between non-word performance and word performance. This suggests that lexical representations do boost recognition. This is in some contrast with the findings from the experiment described in Chapter 5, where recognition was not worse in low-frequency words than it was in high-frequency words. As it seems unlikely that non-Yorkshire listeners have solid representations of Yorkshire speakers saying words like *ode*, *foe*, and *sway*, it could be seen as strange that they still recognise a Yorkshire accent better in these stimuli than they did in equivalent non-word stimuli. How can this be explained?

A first possible explanation is that the data used in the comparison are from two different experiments which were conducted at different points in time by different groups of participants. Whilst as much as possible was kept the same across the experiments, and the regression models controlled for self-assessment as a stand-in for accent recognition skills, it is possible that there are still differences in the data because of this.

Another possible explanation for the difference is that processing non-words is a more difficult task which requires more time and resources, independently from whether recognising the accent it is pronounced in is more difficult. As discussed above, non-words are slower to be recognised in lexical decision tasks than real words are. For example, the mean response times to non-words in the semantic priming experiment described in Chapter 3 was 49ms slower than to words. It is possible that this slowed participants down in a recognition task where they only had limited time to respond. At the same time, the distribution of response times in both the experiments (see Figure 37 above) did not suggest that the higher processing cost led to participants timing out on the stimuli, other than in very small number of cases at the tail of the reaction time distribution.

A final, more theoretical possibility is that, for the recognition of the recognisable sublexical variants, it is necessary that listeners know which variable these sublexical variants are a sublexical variant of, as a necessary reference point. This would mean that having some sort of lexical representation of words like *ode*, *foe*, and *sway* to retrieve sublexical information from to compare the input to, would boost recognition, as it speeds up having a reference point to compare the input to. This means that even listeners who do not have strong memory representations of *ode*, *foe*, and *sway* being pronounced in a Yorkshire accent would be helped by having non-accent specific lexical representations: the phonological forms of *ode*, *foe*, and *sway*, for example, to 'slot' a more detailed sub-lexical input into. This would imply that the actual phonetic accent detail used for recognising accents is not stored in highly specific lexical representations (lexical exemplars), but in smaller, more detailed sub-lexical exemplar representations which then work in tandem with much more abstract phonological forms of words functioning as an anchor point.

In such an account, the dichotomy between sub-lexical abstractions and highly detailed lexical exemplars is turned on its head: the sub-lexical representations used in sociolinguistic processing would be relatively specific, including phonetic detail and its social meaning, whereas the word-sized representations would be highly abstract phonological forms without much detail. I expand on this new conceptualisation in the next chapter.

6.4.2 Speaker recognition

Participants perceived that they heard fewer different voices in the experiment than they actually did. Whilst it is possible that the isolated non-words used in the experiment did not contain enough information to be able to distinguish between all the different voices, the reduction found here seems extreme, especially considering the variation in speaker voices along the lines of gender, age, and region.

This suggests there is a reduction of information and detail at the level of speaker-level sociolinguistic representations: listeners did not distinguish much between the differences in socio-phonetic detail of the different speakers. This is in contrast to pure exemplar models of sociolinguistics in which only highly specific and highly detailed exemplars would be used for processing (Foulkes & Docherty 2006; Johnson 2006). In such models no detail would be filtered out, whilst there does seem to be a clear loss of detail implied by the findings in this chapter.

For the recognition of unfamiliar voices, listeners seem to therefore use sociolinguistic representations slightly more abstract than those at the level of speaker specificity. This level may be that of the character type. Social cues of this type have been found to be able to assist linguistic processing (D'Onofrio 2015) and would still be more specific than just accent level representations – after all, most listeners still perceived more voices than the number of accents involved in the quiz.

New speaker input would be matched with such slightly more abstract character type representations, and similar sounding voices (with similar sounding accents) would be grouped as one, without retaining the more specific detail of the individual voices. This way hearing a new stimulus which is processed through the same character type representation may be interpreted as the same voice.

Still, these are relatively tentative findings, as the data reported here were all self-assessed and based on conscious introspection, elicited through explicit questions. The next step for future research would be to investigate how specific the linguistic units used in sociolinguistic processing are in experiments which have this as their main focus and are able to test this through less explicit methods, in order not to rely on self-evaluation.

6.4.3 Conclusions

This chapter described an accent recognition experiment on the basis of isolated non-word stimuli, which was in as much as possible a duplicate of the word-based accent recognition experiment reported in Chapter 5. The experiment found that listeners were able to recognise accents in these stimuli at a rate much higher than chance, providing evidence that sub-lexical units can facilitate accent recognition. In comparison to the word-based experiment, however, performance was worse, which suggests that being able to use lexical representations in accent processing does boost performance. The fact that this difference exists, whilst no such difference was found in between high and low-frequency word stimuli, suggests that it helps accent recognition to have an abstract phonological forms of words, to 'slot' smaller more detailed sub-lexical exemplars into, or to contextualise these sub-lexical exemplars with.

The experiment's post-hoc questions showed that listeners were most likely to perceive their own accent recognition to be based on the individual sounds of the non-words in the experiment, rather than by finding similarly sounding real words. However, no evidence was found that participants who thought they used lexical strategies performed better than participants who thought they used sub-lexical strategies. The post-hoc questions also showed that listeners perceived the stimuli in the experiment to be pronounced by the same voices when they were not. This suggests that accent recognition, even in a task which does not ask about or require speaker recognition, may use representations which are slightly less specific than speaker-based information, for example representations that connect linguistic detail with social representations of character types.

7 Discussion and conclusion

7.1 Introduction

This PhD thesis used five experiments to investigate the way social meaning is stored in cognition. The two main research questions it focussed on were:

1. How is social meaning connected to linguistic information?
2. What linguistic representations are used to retrieve social meaning?

These questions were explored by using accent-based socio-contextual priming experiments (Chapters 3 and 4) and explicit, rapid-answer accent recognition tasks (Chapters 5 and 6). The socio-contextual priming experiments provide new evidence that the social meaning of linguistic variation may not be stored as a part of the linguistic memory representations, as is implied in exemplar theoretic accounts, as I discuss below in Section 7.2. The accent recognition tasks provide new information about the size and detail of the representations that are used to access social meaning. On the basis of those findings it seems most likely that social meaning is carried by sub-lexical representations which are contextualised by the phonological forms of the words they occur in, as I discuss below in Section 7.3. Taken together, this thesis provides evidence against the idea that lexical exemplars are the main carriers of social meaning in cognition.

7.2 Separate storage of social and linguistic information

7.2.1 Evidence for separate storage

As I discussed in Section 2.2, exemplar theoretic approaches and third-wave variationist approaches can have somewhat conflicting ideas about the storage of social meaning, and how social meaning is connected to representations of linguistic variation. Firstly, third-wave of variation studies argue that the social meaning of linguistic variation is linked to representations of linguistic variation in a flexible way, using inferential indexical links to connect the two (Eckert 2008; Acton 2019). As Campbell-Kibler (2021) argues, these links may even be different depending on whether it is social information that is influencing linguistic perception or vice versa. This conceptualisation, summarised in Chapter 2 in Figure 1, and repeated as Figure 39 below, would take different aspects of American Englishes and conceptualise them as each linked to a social meaning ‘American’ through context-specific indexical links. In this approach, representations of social information are stored separately from representations of linguistic variation.

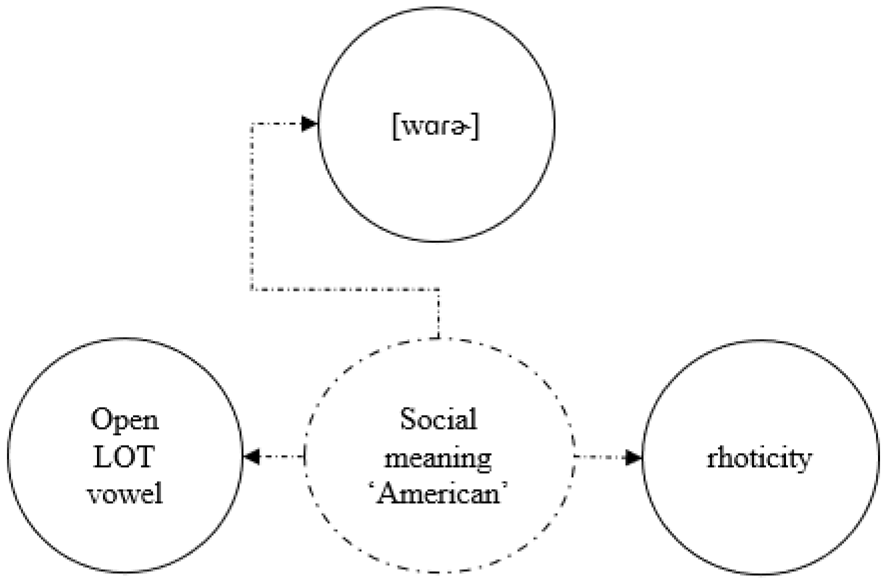

Figure 39. Schematic representation of social meaning and linguistic variation as separate cognitive units connected through indexical links.

In contrast to this early versions of exemplar theory account for the storage of social meaning by positing that speech is processed on the basis of episodic memories of utterances (mostly assumed to be words (e.g. Johnson 2006)), or 'exemplars'. These memories include detail about the speaker and social context of the utterance. Consequently, when processing new linguistic variation, the input is compared to those exemplars including social information, meaning that retrieval of information about the input includes not just denotational meaning but also the social meaning of the exemplars most closely resembling the new input. In this conceptualisation, social information and linguistic information are united within the same representation: the exemplar. This is schematically shown in Figure 40 (Figure 2 in Chapter 2) on the next page (taken from Docherty & Foulkes 2014: 44).

However, more recent hybrid exemplar models include the possibility that, on the basis of exemplars, listeners make generalisations, which form more abstract, less detailed representations, which are also used in speech processing and production (Goldinger 2007; German, Carlson & Pierrehumbert 2013; Ernestus 2014; Pierrehumbert 2016). How these 'hybrid' exemplar models work sociolinguistically – that is which representations are used to access social meaning, and in what detail – has not been explored in detail. This thesis proposed two ways that these more abstract generalisations could be used to access social meaning:

1. The generalisations over exemplars are made separately for social information and linguistic variation, but those are closely connected through indexical links.

The generalisations over exemplars are made following the same structure as exemplars themselves, incorporating both social and linguistic information within the same mental representations.

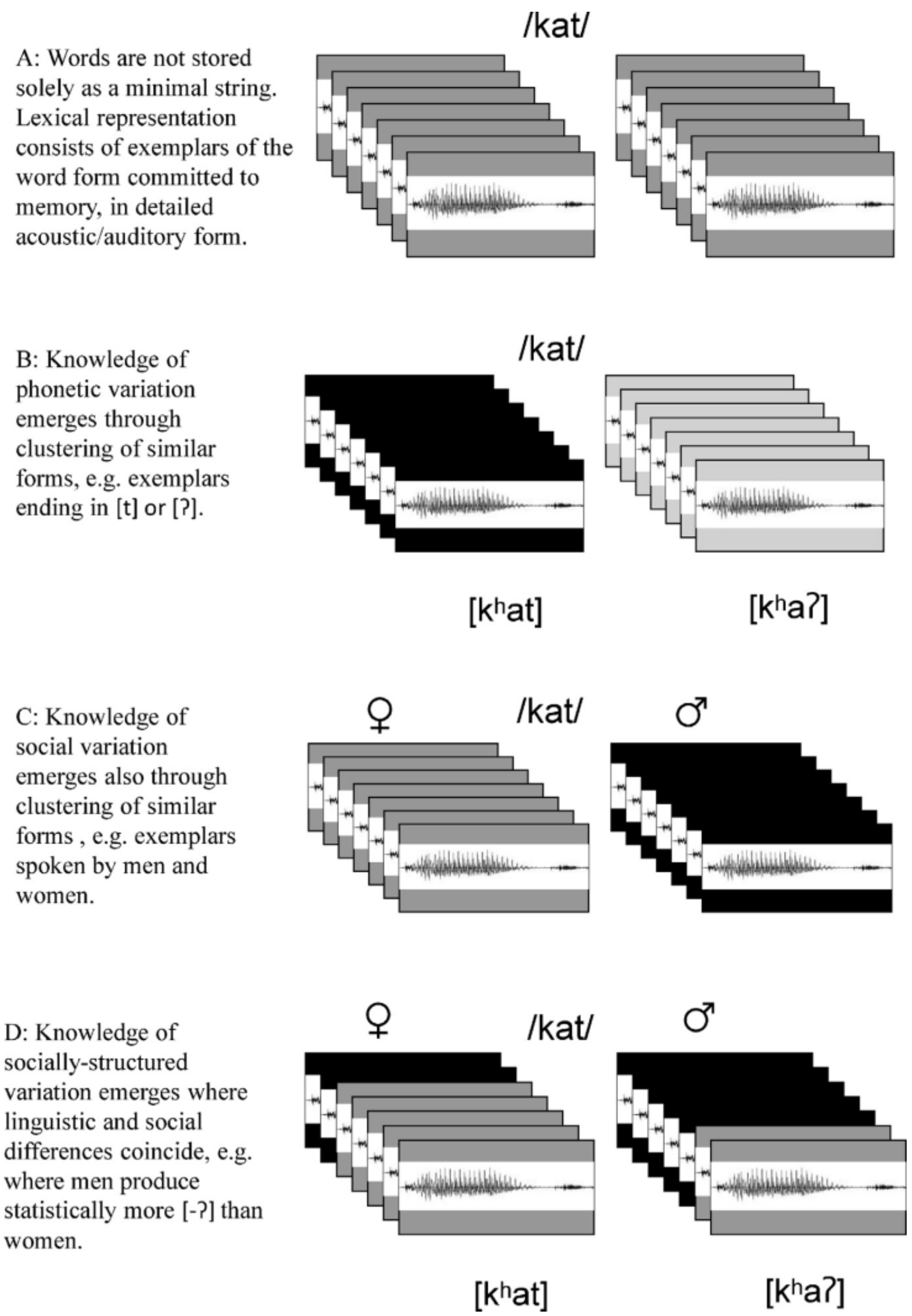

Figure 40. Schematic illustration of socio-phonetic variability within an exemplar-model (Docherty & Foulkes 2014: 44)[27]

The latter option is summarised in Figure 3 from Section 2.2 below, repeated as Figure 41. In this conceptualisation the representations would be connected through indexical and iconic links (where these representations look like each other, because the social information between the representations look similar).

---

[27] Reprinted with permission from Elsevier.

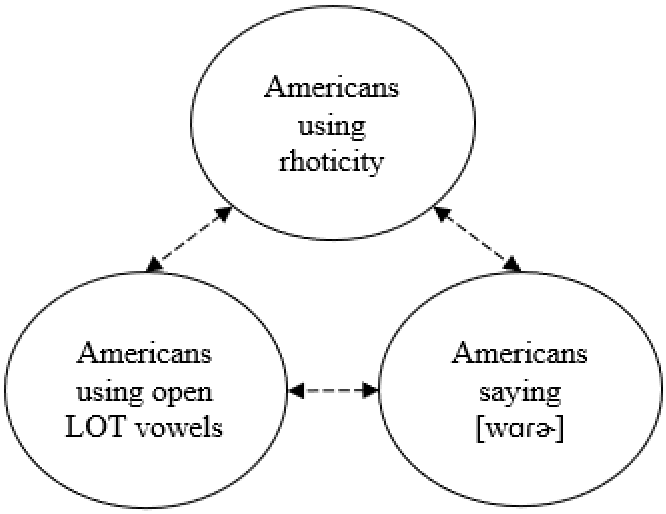

Figure 41. Abstract representations of social meaning and linguistic information where social and linguistic information are united in the same representations, whilst being less specific than exemplars. The representations are linked through iconicity and indexicality.

The findings from the experiments in Chapters 3 and 4 point towards a conceptualisation of social meaning where it is stored in representations separate from representations of linguistic variation, in line with third-wave variation studies, and in line with a hybrid model of exemplar theory which supposes that generalisations over exemplars create more abstract representations of social meaning and linguistic information separately from each other.

Chapter 3's experiment it was found that, in most cases, it was possible to activate regional lexical variants or accent names without also activating corresponding accent stereotypes or regional lexical variants. This was done in an exploratory masked socio-contextual priming experiment. The only times priming effects were found was under certain conditions:

1. The regional lexical variant *y'all* only primed accent related concepts (such as GUNS) if the prime had already been repeated many times. This priming was inhibitory, suggesting relatively weak connections between prime and target.
2. Accent names such as *American* and *Geordie* only primed regional lexical variants such as CANDY and LAD if the primes had not been repeated often yet. This priming however was facilitatory, suggesting a closer connection between prime and target.

These findings not only show that it is possible to process linguistic variation without activating its social meaning, and vice versa (implying that these are stored in separate representations), but they also show that the way the former connects to the latter (inhibitory priming) can differ from the way the latter connects to the former (facilitatory priming). This again suggests a separation between the representations of social and linguistic meaning; it would be expected that if the two are part of the same representation, they would connect to each other in the same way.

Chapter 4's experiments were less exploratory simplified auditory versions of this socio-contextual priming experiment. These found no priming effects whatsoever between socio-phonetic accent primes (General American rhoticity and open LOT vowels) and accent related words (e.g. LOUD and GUNS) or matching regional lexical variants (e.g. SUBWAY and DIAPER). This suggests that it is possible to process phonetic variation without activating related social meaning.

7.2.2 Implications for exemplar theory

Whilst there is tentative evidence in this thesis that social meaning and linguistic variation are not intrinsically a part of the same memory representations, this breaks with some other sociolinguistic perception studies arguing that social meaning is stored within the same exemplar representations as linguistic variation. In these studies, the effects of social cues on the perception of linguistic stimuli is attributed to the activation of the social information part of exemplar memories by those social cues, making retrieval of the matching linguistic information more likely. As discussed in Section 2.2, Walker & Hay (2011) and Kim (2016) find that listeners recognise words which are more typically used by old people if they are pronounced with older sounding voices, and Niedzielski (1999) finds that listeners are more likely to hear Canadian raising when they have been told they will be listening to a Canadian speaker.

However, the tentative evidence from this thesis suggests that this may not be the way perceptions are influenced by social cues, if, like Chapter 3 and 4 suggest, social meaning and linguistic information can be activated separately from each other and are stored separately. Then, the reason for the effects on social cues on speech perception may be caused through spreading activation between the closely linked social and linguistic information. Being told you will be hearing a Canadian speaker might activate the more general social meaning of 'Canadian-ness' or 'Canadian speakers', which then causes activation to spread to linguistic forms indexically connected to Canadianness (e.g. Canadian raising). The same would happen in the case of the studies by Walker & Hay (2011) and Kim (2016): old sounding voices activate the social meaning 'old' which then activates words typically used by older speakers, facilitating access to those words.

A final possibility is that exemplars themselves form not one unit of representation but rather a combination of memory representations which encompass different aspects of an episodic memory in different parts of cognition, which in access are activated mostly together. In this case it is possible that in highly automatic processing only some aspects of the episodic memories are activated, which would account for the lack of automatic activation of social meaning by processing linguistic variation and vice versa in the two experiments.

7.2.3 Methodological caveats

There are three methodological caveats to these findings. Firstly, the experiment described in Chapter 3 faced some methodological issues. Its main findings only occurred at certain levels of exposure to the prime. This meant that these findings are based on relatively few datapoints (those with low prime exposure for one condition, and those with high prime expose for the other). As the experiment did not randomise trial order (and therefore prime exposure), the effects reported in that chapter are at risk of simply have occurred because of a small number of idiosyncratic items showing priming effects, rather than showing more general patterns.

Secondly, the findings from Chapter 4 statistically only show a lack of support for the idea that social meaning and linguistic information are stored within the same representations in memory, rather than evidence for the claim that they are stored separately. Thirdly, the lack of priming effects in the experiments in this chapter may be that the target words that were used were not a part of the social meaning of the variants of the accent primes. Part of this may be that the large number of target words were not all that strongly associated with Americanness, and part of this may be the form in which this social meaning was presented – i.e. in orthographic form, rather than, for example, through icons, pictures, or even spoken words.

Future research on the potential separate storage of social and linguistic information in memory would address the caveats discussed above: the evidence for this would be stronger if socio-contextual priming experiments which use different target stimuli also do not find effects. These different target stimuli would either be more immediately part of the social context of a variety, or presented in a different modality, such as auditory verbal targets (e.g. a voice saying 'subway') or visual non-verbal targets (e.g. an American flag or a picture of American stereotypes).

7.3 The representations used to access social meaning

7.3.1 Representation size

The second main research question of this thesis was to find out which mental representations are used in social meaning perception processes, more particularly what size they are, how specific, and how social meaning and linguistic information are connected in storage. The experiments in Chapters 4, 5, and 6 looked at this. The experimental designs for these experiments were created to support most directly the first element of this question: what size are the linguistic representations which are used for social meaning perception?

The experiments in Chapter 4 did not find priming effects across the experiment and were therefore not able to provide new information on the matter of whether non-words or words are more dominant or helpful in below-awareness accent recognition. However, the more explicit elicitation of the accent recognition experiments in Chapters 5 and 6 did provide new insights: participants were equally accurate at recognising accents in high-frequency and low-frequency words but performed worse with non-word stimuli. However, they still were much more accurate in non-words than chance would predict.

On the basis of these results, I argue below that the main mental representations that carry the social meaning of phonetic variants are sublexical variants, which are contextualised within the abstract phonological forms of words. This accounts for participants' ability to recognise accents in non-words. It also accounts for the difference between accuracy in non-words and real words, as well as the highly similar accuracy between high and low lexical frequency stimuli, by positing that the presence of an abstract phonological form of low-frequency words, allows for the contextualisation of its sublexical variants in a way that does not occur for non-words. Below I discuss the different possible linguistic representations which may be able to be used to access social meaning, based on the experimental findings in this thesis.

Word-sized representations

As discussed in Section 2.3, one conceptualisation would be that word-sized units are most dominant in social meaning processing. This would be in line with exemplar theoretical accounts of social meaning processing (Johnson 2006; Foulkes & Docherty 2006; Walker & Hay 2011; Hay 2018) in which the word-sized exemplar is posited to play the main role in speech processing and, more specifically, is where socio-contextual detail is assumed to be located and united with linguistic detail. Whilst the findings from Chapters 3 and 4 cast some doubt on the idea that social and linguistic information would be stored within the same exemplar representations in memory, it would still be possible that somewhat more abstract word-sized representations, or highly detailed linguistic representations strongly linked to highly detailed social information (see Section 7.2.2 above) play a role in processing.

The evidence for a dominant role of these word-sized representations in social meaning processing would be that, in the experiments in Chapters 5 and 6, listeners were better at recognising accents in real words, for which they would have word-sized exemplars, than they were in non-words, for which they would not. However, there are also some strong counterpoints to this position in the results: firstly, there was no evidence that participants were better at recognising accents in high-frequency stimuli than in low-frequency stimuli, despite many of the low-frequency stimuli being unlikely to have been heard at all in listeners' experience: for example words like *orc* (in a General American accent) or *ode* (in a Yorkshire accent) will have been exceedingly rare for most British English listeners without a background in either of these accents.

Secondly, the fact that participants were able to recognise accents well above chance in the accent recognition task which exclusively used non-word stimuli, suggests that listeners do not need to have lexical exemplars or other lexical representations to be able to recognise accents, and the social meaning of linguistic variation.

However, it is also possible that the lexical representations used for accent processing do not need to match the input phoneme by phoneme. As Johnson (2006) posits, the retrieval of social meaning in (word-based) exemplar theory happens through 'clouds' of exemplars which are the most similar to the input being processed. If it is not a requirement that these exemplars are necessarily the same lexical unit, with the same phonemes, this would allow for a degree of social meaning processing for non-word and low-frequency word input on the basis of neighbours: a Yorkshire voice pronouncing the word *ode* may cause resonance with exemplars of Yorkshire voices pronouncing the word *road*. An American voice pronouncing the word *zor* may cause resonance with exemplars of Americans pronouncing the word *or*. Still, this would be more effortful, and more difficult to process, as there would be less resonance caused by the more dissimilar input. This would then explain the gap in accuracy between the word based experiment and the non-word based experiment.

However, it is not entirely clear why then there is no gap in accuracy between the low-frequency stimuli and the high-frequency stimuli. The difference between the low-frequency stimuli and the non-word stimuli in this conceptualisation is only gradual: for the low-frequency items, the resonance cloud of exemplars would mostly be lexical neighbours as exemplars (of the word *or* for example), as well as the occasional exemplar with the matching lexeme (*orc* exemplars), whereas for the non-word items there would not be the occasional lexeme matching exemplar (just *or* exemplars, no *zor* exemplars). It is unclear therefore why low-frequency items would behave like high-frequency items and differ so starkly from non-word items.

Sub-lexical representations

The second option is that it is sub-lexical representations which are used to access social meaning in accent processing, which would be in line with more hybrid approaches to exemplar theory (McLennan & Luce 2005; German, Carlson & Pierrehumbert 2013; Ernestus 2014; Pierrehumbert 2016). The evidence for this is that listeners performed above chance in the non-word based accent recognition task. It would also be supported by the lack of differences found in the word based accent recognition task with high-frequency and low-frequency stimuli. Whilst there were no differences on the basis of this lexical information, there was very large variation between recognition in the different recognisable sublexical variants included in the experiment. Finally, it may be worth mentioning that participants themselves indicated they thought they recognised accents on the basis of separate sounds rather than words in the non-word experiment, although this self-reporting is, of course, not necessarily reliable.

The main problem for this approach is the fact that there was a large and significant difference in accuracy between the word based and the non-word based experiments. This suggests that lexical information does facilitate accent recognition in some sort of a way, despite sub-lexical units being able to facilitate accent recognition as well. The fact that there was a difference between recognition in non-words and words, but not between high-frequency and low-frequency words, suggests that there is a fundamental difference between whether listeners have access to general lexical information or not, rather than the number of lexical exemplars or representations available to a listener. One way in which this may be explained is if accent recognition does occur on the basis of sub-lexical information such as sublexical variants, but that these need to be contextualised by information about the phonological form of a word. For example, hearing the sublexical variant [ɵː] may not carry any social information, as long as it is not clear whether it occurs in the /əʊ/ position in /təʊtəl/ for a Yorkshire speaker saying the word 'total' or in the /ɜː/ position in /tɜːtəl/ for an SSBE speaker saying the word 'turtle'.

In this conceptualisation, the gap between accuracy in the non-word experiment and accuracy in the word experiment would be caused by the lack of lexeme information to 'anchor' the sublexical variation to, resulting in a poorer performance. In processing, listeners would potentially rely simply on the sublexical variants themselves – and have less reliable information – or create a stand-in lexeme on the basis of the non-word – which would be more effortful, and would allow for an extra step in the process to produce errors. In this conceptualisation, social meaning would still (at least predominantly) be carried by sub-lexical units rather than lexical units, and only need to interact with highly abstract phonological information about a word. This account best explains the findings from this thesis.

Whilst this interactiveness between lexeme information and sub-lexical representations has not yet – to my knowledge – been applied to sociolinguistic processing, this does tie into a debate in more general speech processing research. This debate is one where some models argue that speech processing always flows in one direction: first phonetic features are processed, then on the basis of that phonemes, and then on the basis of that words (e.g. Massaro 1989; Norris, McQueen & Cutler 2000). Here, feedback from a higher level (that is to say, further down this linear process) is seen as never necessary or used. The opposing view is that in speech processing different levels of processing can influence each other (e.g. McClelland & Elman 1986; Grossberg 2013). Sub-lexical information in these models can activate lexical information which in turn can facilitate sub-lexical information. As Samuel (2011) argues, it can be difficult to find evidence for one or the other theory. However, if in sociolinguistic processing it is necessary to have lexeme information as an anchor point, that would in fact provide some evidence for the interactive approach.

7.3.2 Representation specificity

The experiments described in Chapters 5 and 6 also found some exploratory evidence for the dominance of representations with approximately enough social and linguistic information to recognise speaker character types, although the evidence for this was more incidental than the findings for representation size which the experiments were directly designed for. It consists of the anecdotal findings that listeners heard fewer different speakers than there were in the experiment in Chapter 5, and the more systematic but still self-reported findings from the post-hoc questions at the end of the experiment in Chapter 6.

There are three different ways to explain these findings:

1. In accent recognition, listeners do use the most detailed exemplars they have available to them.
2. In accent recognition, listeners use representations which are approximately as detailed as representations of a certain speaker using a linguistic form
3. In accent recognition, listeners use representations which are approximately as detailed as representations of certain sociolinguistic character types using a linguistic form.

The first possibility is that listeners do use 'exemplars' for accent recognition, i.e. the maximally detailed memory traces that input could be compared to in accent processing – despite Chapters 3 and 4's tentative evidence against storage of social and linguistic information within the same representation. In this case, the reason that listeners heard fewer speakers in the accent recognition experiments than were actually used, could either be that these exemplar memories are much less detailed than they are often conceptualised as (Johnson 2006; Goldinger 2007; Docherty & Foulkes 2014). In this case there would be so little detail in these exemplars that it is not possible to pinpoint the exact speaker who used the utterance on the basis of the information included in the exemplar. This seems relatively unlikely as most listeners without impairments are generally quite accurate at voice recognition (Maguinness, Roswandowitz & von Kriegstein 2018).

The other option could be that exemplars are detailed enough to include voice information, but that they are not clustered in very clear or distinct speaker exemplar clouds, or in speaker exemplar clouds which are less than completely accurate (especially if the input is only a singular word). In this approach the fact that listeners heard fewer different voices than were used is not due to the abstraction of the representations used in accent recognition, but rather due to how the most specific representations are clustered.

Another approach to this would be to assume that in accent recognition higher-level abstractions are used, at the level of the speaker, or slightly above: if listeners have representations stored in memory which are approximately as specific as a memory of how one speaker would pronounce a word, or a sound or a part of a word, this would explain why it is possible that those can be combined and generalised into fewer speakers than there really were in the experiments. It is possible that these representations are not always kept completely distinct or are not accurate enough, causing confusion.

However, if this confusion is due to a lack of specificity in these representations, it may be better to hypothesise a level of specificity that is slightly more abstract than speaker-level representations. In this case, it may be better to hypothesise character type level representations. In such representations, there would not be enough detail to distinguish each speaker individually (especially not on the basis of singular words), but instead have enough detail to distinguish between groups of relatively similar speakers. The participants in the study were able to distinguish a higher number of voices than just the number of different accents.

On the basis of this I hypothesise that the ‘character type’ (Starr 2021) plays a central role in listener’s perception of linguistic variation: a socially recognised figure that is less specific than a speaker but more specific than just a regional variety: for example ‘valley girls’, or ‘jocks’ or ‘gay divas’. Following Sumner et al. (2014), character type level information in new input (or new exemplars) would be what most attention would be paid to in processing, resulting in weighted storage, and ready availability in the retrieval of social meaning. In Johnstone’s (2009) and Agha’s (2003, 2005) approaches, these may be some of the most important varieties or registers that are enregistered.

The idea that character type level representations play an important role in the processing of social meaning aligns well with D’Onofrio’s (2016) work on the role of character type based information on perception.[28] As discussed in Section 3.2, she finds that listeners perceive backed TRAP vowels differently on the basis of different social cues that are given to them. For example, in an eye-tracking study, listeners who were exposed to an icon of stereotypical ‘valley girl’ items were quicker to hear a backed TRAP vowel as a TRAP vowel rather than a LOT vowel. This shows that character type information (e.g. the ‘valley girl’) can influence speech perception and that it can play an important role in sociolinguistic perception. She also finds effects of character type cues on perception in phoneme categorisation tasks, and explicit memory tasks.

At a more explicitly evaluative level, Vriesendorp & Rutten (2019: 338) find that listeners, when answering open-ended questions about their impressions of speakers, often described them in terms of character types (e.g. ‘typical white blonde hockey girl’ or even ‘the type of person who would enthusiastically bring home-made tiramisu to Christmas dinner’). This also suggests that ‘character types’ are central in more conscious processing.

D’Onofrio (2020) argues that personas (perhaps as individualised versions of character types) are interactional constructs which function as mediators of macro-social information (such as class, race, gender, place of origin et cetera) by making them specific in interaction, and providing structure to the heterogeneity of the macro-social information about speakers. If this is the case, it follows that having mental representations of social meaning and the linguistic representations they are connected to that are approximately as specific as is necessary to recognise those personas or character types (or at least that those representations are dominant in processing).

7.3.3 Recombining size and specificity

As set out in the theoretical framework detailed in Chapter 2, this PhD thesis aimed to disentangle the parameters of size and specificity of the potential mental representations used in social meaning

[28] As mentioned in Chapter 2, D’Onofrio uses the term ‘persona’ for what in Starr’s (2021) categorisation would be called a character type.

processing by looking at these parameters separately. However, with the new information gleaned from the experiments, it is now possible to recombine the parameters to settle on the representations that, on the basis of my findings, seem the most likely to play a role in accent processing and the retrieval of social meaning. In the experiments with unfamiliar voices in Chapter 5 and 6, representations with specificity at the level of approximately the character type seemed to be dominant, whilst for the size parameter it seemed most likely that sub-lexical segments are dominant, whilst lexeme information is also used for contextualisation.

To summarise the model most likely on the basis of the findings of the thesis, the most likely representations to be dominant in sociolinguistic processing are a combination of character type level specificity sublexical variants, used in tandem with lexemes as anchor points. These representations are marked in green in the Table 1 from Chapter 2 (repeated as Table 40 on the next page).

The main representations used in the retrieval of social meaning are most likely stored separately from the representations of linguistic variation they are connected to (Chapters 3 and 4), and attached to sublexical linguistic representations that are as detailed as phonetic accent variants (e.g. the Standard Southern British BATH vowel [ɑː]), whilst being contextualised by more abstract lexical representations. The most important social meanings attached to this seem to be approximately as specific as those of character types.

The exact role of the exemplar itself in this model is not something the findings of the study can clearly speak on: within a hybrid exemplar theoretical model, it would be assumed that exemplar units are the starting point for processing (e.g. German et al. 2013), on the basis of which more abstract representations form. This process was not investigated in this thesis, as it investigated which representations play the most central role in social meaning processing in listeners who were already familiar with the varieties in question.

Table 40. Schematic representation of intersections of the detail parameter and size parameter of specificity and abstraction. The most likely representations, on the basis of the current findings, to be used in sociolinguistic processing are marked in green.

Size → **Detail↓**	**Utterance**	**Word**	**Syllable**	**Segment**
Very low (underlying form)	/teɪk ðə məʊtəweɪ/	/məʊtəweɪ/	/weɪ/	/eɪ/
Potential social meanings	*British English - quaint - posh - direct*	*British English – neutral – quaint – posh*		
Low (widest sociolinguistic category)	[te:k ðə mo:ʔəwe:] (and similar realisations)	[mo:ʔəwe:] (and similar realisations)	[we:] (and similar realisations)	[e:] (and similar realisations)
Potential social meanings	*North of England – friendly – authentic – direct*	*North of England – friendly – authentic*	*North of England – friendly – authentic*	
Some (medium-width sociolinguistic category)	[tɪ:k ðə mo:ʔtəwɪ:] (and very similar realisations)	[mo:ʔtəwɪ:] (and very similar realisations)	[wɪ:] (and very similar realisations)	[ɪ:] (and very similar realisations)
Potential social meanings	*North-East of England – funny – friendly – authentic – direct*	*North-East of England – funny – friendly – authentic*	*North-East of England – funny – friendly – authentic*	
Character type (narrow sociolinguistic category)	[tɪək ðə mʊəʔtəwɪə] + accent wide articulatory setting	[mʊəʔtəwɪə] + accent wide articulatory setting	[wɪə] + accent wide articulatory setting	[ɪə] + accent wide articulatory setting
Potential social meanings	*Very strong Mackem accent – authentic – friendly – working class – masculine – direct*	*Very strong Mackem accent – authentic – friendly – working class – masculine*	*Mackem lad – authentic – friendly – working class – masculine*	
High (speaker)	[tɪəʔk ə mʊəʔtəwɪə] + voice quality (abstracted)	[mʊəʔtəwɪə] + voice quality (abstracted)	[wɪə] + voice quality (abstracted)	[ɪə] + voice quality (abstracted)
Potential social meanings	*George – Mackem – neighbour – old – grumpy – working class – masculine – direct*	*George – Mackem – neighbour – old – grumpy – working class – masculine*		
Very high (speaker- specific persona)	[tɪəʔk ə mʊəʔtəwɪə] + voice quality (somewhat abstracted)	[mʊəʔtəwɪə] + voice quality (somewhat abstracted)	[wɪə] + voice quality (somewhat abstracted)	[ɪə] + voice quality (somewhat abstracted)
Potential social meanings	*#George – Mackem – neighbour – old – working class – masculine – when he's relaxed and friendly*			
All that is processed (exemplars)	[tɪəʔk ə moəʔtəwɪə] + voice quality + speech rate + pitch + volume maximally (but not fully) specific	[mʊəʔtəwɪə] + voice quality + speech rate + pitch + volume maximally (but not fully) specific	[wɪə] + voice quality + speech rate + pitch + volume maximally (but not fully) specific	[ɪə] + voice quality + speech rate + pitch + volume maximally (but not fully) specific
Potential social meanings	*George, my old, grumpy, masculine, working-class neighbour from Sunderland, when he told me to take the motorway as we're having a relaxed friendly cup of tea in his nice and quiet Sunderland home, with a low pitch and relatively slow speech rate and average volume.*			
All (acoustic signal)				
Potential social meanings				

7.3.4 Future research

To gain further insight in the exact representations used in the processing of social meaning on the basis of linguistic information, there are three main lines of further research to pursue:

1. To further investigate the possibility of lexical effects through variables other than absolute frequency.
2. To further investigate the possibility that for social meaning processing sub lexical information is used in tandem with phonological forms of words.
3. To look at the specificity of representations with experimental methods designed to look at specificity.

The first line of research would serve to provide more certainty about the influence of lexical representations on accent recognition, and the current findings that high-frequency and low-frequency lexical stimuli are equally easily recognised, but much better than non-word stimuli. As was argued in Chapter 5, there are also different possible methods to investigate lexical effects, such as comparative frequency between accents and recency. If those measurements show the same results this would strengthen the arguments made in this chapter to support the idea of segmental units being used to access social meaning whilst using lexeme information as an anchor point.

The second line of research would be to further investigate whether there is a mechanism where sub-lexical information is contextualised by means of the phonological forms of words in order to establish how this sub-lexical information is to be interpreted. One way this could be done is to compare the recognition of accents in stimuli across non-word and word stimuli, by whether the recognisable accent information is sublexical (where knowing which variable a sublexical variant is a variant of is helpful) or a different type of sub-lexical information, such as prosody. If there is a difference in accuracy for the stimuli with recognisable sublexical information but not for those with recognisable prosody, this would suggest there is a contextualisation process for sublexical cues.

The third line of investigation would be to more directly investigate how specific or abstract the representations are which are used to recognise social meaning. One important question in particular would be to investigate whether the character type level effects found in the current experiments (listeners perceiving there to be more speakers in the experiment than the number of accents, but fewer than the real number of speakers) can also be found in experiments directly designed to ascertain the level of specificity of the representations used in sociolinguistic processing. At a more conscious level of processing, it may be possible to investigate the ways speakers describe voices in perception experiment when they are asked open-ended questions: do they talk about macro-social categories, or character types. How specific are these descriptions? Whilst open-ended perception questions have been asked in previous research (e.g. Campbell-Kibler 2005, Pharao et al. 2014, Vriesendorp & Rutten 2019), this has often been to investigate more content-based (is a variant heard as masculine or feminine?) questions than how specific they are. A more systematic investigation may be able to shed light on this. At a more subconscious level it is not immediately evident how this would be implemented.

7.4 Conclusion

This chapter addressed the main two research questions of this thesis on the basis of the data from the previous four experimental chapters:

1. How are linguistic variation and its social meaning connected in memory?
    a. Are they stored in separate representations, but closely connected?
    b. Or are they stored in the very same representations, as a unit?
2. Which linguistic representations are used in the retrieval or interpretation of the social meaning of linguistic variation?
    a. What size are these units?
    b. How detailed are these units?

I argued that given the data from the experimental chapters it is most likely that social meaning is stored separately from linguistic information in memory, as Chapter 3 and 4 found that in most cases it was possible to activate social information without activating corresponding lexical variants, and vice versa. In the cases in Chapter 3 where they did activate each other this happened in different ways: some only with limited exposure to the first stimulus, facilitating access to the second, and others only with repeated exposure to the first, inhibiting access. This suggests that the third-wave variationist approach to social meaning holds up cognitively: the theorised fluid links between social meaning and variation are possible if social and linguistic information are stored separately. Thus, in hybrid exemplar theory the abstract representations that listeners establish on the basis of exemplars are likely separate for linguistic and for social information.

Furthermore, I argued that, given the data from Chapters 5 and 6, the linguistic representations used in the retrieval of social meaning are most likely to be sub-lexical units, which are contextualised by means of phonological information about the words they occur in. This explains why listeners were able to recognise accents in non-word stimuli (without lexical information) in Chapter 6, and did not perform worse in low-frequency word stimuli than in high-frequency word stimuli in the experiment in Chapter 5. Finally, the fact that the experiments in these chapters were able to recognise fewer voices than were actually present in the experiments suggests that the main representations being used to access social meaning are approximately specific enough to distinguish between character types but not always between speakers.

All in all, on the basis of the findings in this thesis, social meaning is most likely carried by sublexical variants, which are contextualised in abstract phonological forms of words. These are then closely connected to their social meaning, but do not incorporate them within their own representations.

Appendices

There are five appendices to this PhD thesis, stored in separate pdf files, due to their size. These are the following.

For Chapter 3:

- Appendix Ia – Stimulus list priming experiment Chapter 3 – Version A
- Appendix Ib – Stimulus list priming experiment Chapter 3 – Version B

For Chapter 4:

- Appendix II – Stimulus list priming experiments Chapter 4

For Chapter 5:

- Appendix III - Stimulus list accent recognition task Chapter 5

For Chapter 6:

- Appendix IV - Stimulus list acent recognition task Chapter 6

References

Acton, Eric. 2021. Pragmatics and the Third Wave: The social meaning of definites. In Lauren Hall-Lew, Emma Moore & Robert Podesva (eds.), *Social Meaning and Linguistic Variation*. Cambridge: Cambridge University Press.

Acton, Eric K. 2019. Pragmatics and the social life of the English definite article. *Language* 95(1). 37–65.

Agha, Asif. 2003. The social life of cultural value. *Language & Communication* 23(3–4). 231–273.

Agha, Asif. 2005. Voice, Footing, Enregisterment. *Journal of Linguistic Anthropology* 15. 38–59.

Anderson, John R. 1983. A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior* 22(3). 261–295.

Anwyl-Irvine, Alexander L., Jessica Massonnié, Adam Flitton, Natasha Kirkham & Jo K. Evershed. 2019. Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*. (30 October, 2019).

Archibald, Lisa MD & Susan E. Gathercole. 2006. Nonword repetition: A comparison of tests. *Journal of Speech, Language, and Hearing Research*. ASHA.

Balota, David A. 1983. Automatic semantic activation and episodic memory encoding. *Journal of Verbal Learning and Verbal Behavior* 22(1). 88–104.

Bartlett, Kirsten E. 2012. *Processing SMS Shortcuts : The Contribution of Phonology, Orthography and Semantics*. Sheffield: Sheffield Hallam University, PhD Thesis.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *arXiv.org*.

Benor, Sarah Bunin. 2001. The learned /t/: phonological variation in Orthodox Jewish English. *University of Pennsylvania Working Papers in Linguistics* 7(3). 2–16.

Bjelaković, Andrej. 2017. The vowels of contemporary RP: vowel formant measurements for BBC newsreaders. *English Language & Linguistics* 21(3). 501–532.

Bordag, Denisa, Amit Kirschenbaum, Erwin Tschirner & Andreas Opitz. 2015. Incidental acquisition of new words during reading in L2: Inference of meaning and its integration in the L2 mental lexicon. *Bilingualism: Language and Cognition* 18(3). 372–390.

Bradlow, Ann R., Lynne C. Nygaard & David B. Pisoni. 1999. Effects of talker, rate, and amplitude variation on recognition memory for spoken words. *Perception & Psychophysics* 61(2). 206–219.

Bucholtz, Mary. 1999. “Why be normal?”: Language and identity practices in a community of nerd girls. *Language in Society* 28(2). 203–223.

Bueno, Steve & Cheryl Frenck-Mestre. 2008. The activation of semantic memory: Effects of prime exposure, prime-target relationship, and task demands. *Memory & Cognition* 36(4). 882–898.

Bybee, Joan. 2002. Phonological evidence for exemplar storage of multiword sequences. *Studies in second language acquisition* 24(2). 215–221.

Cai, Zhenguang G., Rebecca A. Gilbert, Matthew H. Davis, M. Gareth Gaskell, Lauren Farrar, Sarah Adler & Jennifer M. Rodd. 2017. Accent modulates access to word meaning: Evidence for a speaker-model account of spoken word recognition. *Cognitive Psychology* 98. 73–101.

Campbell-Kibler, Kathryn. 2005. *Listener Perceptions of Sociolinguistic Variables: The Case of (ING)*. Stanford University PhD Thesis.

Campbell-Kibler, Kathryn. 2012. The implicit association test and sociolinguistic meaning. *Lingua* 122(7). 753–763.

Campbell-Kibler, Kathryn. 2016. Towards a cognitively realistic model of meaningful sociolinguistic variation. In Anna Babel (ed.), *Awareness and Control in Sociolinguistic Research*, 123–151. Cambridge: Cambridge University Press.

Campbell-Kibler, Kathryn. 2021. The Cognitive Structure Behind Indexicality: Correlations in Tasks Linking /s/ Variation and Masculinity. In Lauren Hall-Lew, Emma Moore & Robert Podesva (eds.), *Social Meaning and Linguistic Variation*. Cambridge: Cambridge University Press.

Carter, Matthew D., Monica S. Hough, Andrew Stuart & Michael P. Rastatter. 2011. The effects of inter-stimulus interval and prime modality in a semantic priming task. *Aphasiology*. Taylor & Francis 25(6–7). 761–773.

Casasanto, Laura Staum. 2009. How Do Listeners Represent Sociolinguistic Knowledge? In *Proceedings of the Annual Meeting of the Cognitive Science Society*, vol. 31.
Church, Barbara A. & Daniel L. Schacter. 1994. Perceptual specificity of auditory priming: implicit memory for voice intonation and fundamental frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 20(3). 521–533.
Clopper, Cynthia G. 2004. *Linguistic Experience and the Perceptual Classification of Dialect Variation*. Indiana University PhD Thesis.
Clopper, Cynthia G. & David B. Pisoni. 2004. Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of Phonetics* 32(1). 111–140.
Collins, Allan M. & Elizabeth F. Loftus. 1975. A spreading-activation theory of semantic processing. *Psychological Review* 82(6). 407–428.
Cooper, Paul. 2017. “Turtely Amazing”. The enregistrement of ‘Yorkshire’ dialect and the possibility of GOAT fronting as a newly enregistred feature. In Chris Montgomery & Emma Moore (eds.), *Language and a Sense of Place: Studies in Language and Region*, 348–367. Cambridge: Cambridge University Press.
Coupland, Nikolas & Hywel Bishop. 2007. Ideologised values for British accents. *Journal of Sociolinguistics* 11(1). 74–93.
Cunningham, William A., Philip David Zelazo, Dominic J. Packer & Jay J. Van Bavel. 2007. The iterative reprocessing model: A multilevel framework for attitudes and evaluation. *Social Cognition*. Guilford Press 25(5). 736–760.
Cutler, Anne. 2008. The abstract representations in speech processing. *The Quarterly Journal of Experimental Psychology* 61(11). 1601–1619.
De Groot, Annette MB, Arnold JWM Thomassen & Patrick TW Hudson. 1986. Primed-lexical decision: The effect of varying the stimulus-onset asynchrony of prime and target. *Acta Psychologica*. Elsevier 61(1). 17–36.
Del Toro, Janice Feagin. 2000. An examination of automatic versus strategic semantic priming effects in Broca’s aphasia. *Aphasiology*. Taylor & Francis 14(9). 925–947.
Dobs, Katharina, Leyla Isik, Dimitrios Pantazis & Nancy Kanwisher. 2019. How face perception unfolds over time. *Nature Communications*. Nature Publishing Group 10(1). 1–10.
Docherty, Gerard J. & Paul Foulkes. 2014. An evaluation of usage-based approaches to the modelling of sociophonetic variability. *Lingua* 142. 42–56.
D’Onofrio, Annette. 2015. Persona-based information shapes linguistic perception: Valley Girls and California vowels. *Journal of Sociolinguistics* 19(2). 241–256.
D’Onofrio, Annette. 2016. *Social Meaning in Linguistic Perception*. Stanford University PhD Thesis.
D’Onofrio, Annette. 2020. Personae in sociolinguistic variation. *WIREs Cognitive Science* 11(6).
D’Onofrio, Annette. 2021. Sociolinguistic signs as cognitive representations. In Lauren Hall-Lew, Emma Moore & Robert Podesva (eds.), *Social Meaning and Linguistic Variation*. Cambridge: Cambridge University Press.
Drager, Katie. 2010. Sociophonetic variation in speech perception. *Language and Linguistics Compass* 4(7). 473–480.
Drager, Katie & Joelle Kirtley. 2016. Awareness, salience, and stereotypes in exemplar-based models of speech production and perception. In Anna Babel (ed.), *Awareness and Control in Sociolinguistic Research*, 1–24. Cambridge: Cambridge University Press.
Eckert, Penelope. 2008. Variation and the indexical field. *Journal of Sociolinguistics* 12(4). 453–476.
Eckert, Penelope. 2012. Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology* 41. 87–100.
Eisner, Frank & James M. McQueen. 2006. Perceptual learning in speech: Stability over time. *The Journal of the Acoustical Society of America* 119(4). 1950–1953.
Elgendi, Mohamed, Parmod Kumar, Skye Barbic, Newton Howard, Derek Abbott & Andrzej Cichocki. 2018. Subliminal priming—state of the art and future perspectives. *Behavioral Sciences* 8(6).
Ernestus, Mirjam. 2014. Acoustic reduction and the roles of abstractions and exemplars in speech processing. *Lingua* 142. 27–41.
Ernestus, Mirjam, Harald Baayen & Rob Schreuder. 2002. The recognition of reduced word forms. *Brain and Language* 81(1–3). 162–173.

Escudero, Paola, Rachel Hayes-Harb & Holger Mitterer. 2008. Novel second-language words and asymmetric lexical access. *Journal of Phonetics* 36(2). 345–360.
Ferrand, Ludovic, Juan Segui & Glyn W. Humphreys. 1997. The syllable's role in word naming. *Memory & Cognition* 25(4). 458–470.
Finnegan, Katie. 2015. Sheffield. In Raymond Hickey (ed.), *Researching Northern English*, 227–50. Amsterdam: John Benjamins Publishing Company.
Forster, Kenneth I. 1998. The pros and cons of masked priming. *Journal of Psycholinguistic Research* 27(2). 203–233.
Forster, Kenneth I. & Jonathan C. Forster. 2003. DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers* 35(1). 116–124.
Forster, Kenneth I., Kathleen Mohan, Jo Hector, S. Kinoshita & S. J. Lupker. 2003. The mechanics of masked priming. *Masked Priming: The State of the Art* 3–37.
Foulkes, Paul & Gerard Docherty. 2006. The social life of phonetics and phonology. *Journal of Phonetics* 34(4). 409–438.
Gathercole, Susan E. 1995. Is nonword repetition a test of phonological memory or long-term knowledge? It all depends on the nonwords. *Memory & Cognition* 23(1). 83–94.
Gathercole, Susan E. & Anne-Marie Adams. 1993. Phonological working memory in very young children. *Developmental Psychology*. American Psychological Association 29(4). 770.
German, James S., Katy Carlson & Janet B. Pierrehumbert. 2013. Reassignment of consonant allophones in rapid dialect acquisition. *Journal of Phonetics* 41(3). 228–248.
Gillund, Gary & Richard M. Shiffrin. 1984. A retrieval model for both recognition and recall. *Psychological Review*. American Psychological Association 91(1). 1–67.
Goldinger, Stephen D. 1996. Words and voices: episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22(5). 1166–1183.
Goldinger, Stephen D. 1998. Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105(2). 251–279.
Goldinger, Stephen D. 2007. A complementary-systems approach to abstract and episodic speech perception. In *Proceedings of the 16th International Congress of Phonetic Sciences*, 49–54.
Goldinger, Stephen D., David B. Pisoni & John S. Logan. 1991. On the nature of talker variability effects on recall of spoken word lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 17(1). 152–162.
Greenwald, Anthony G., Debbie E. McGhee & Jordan LK Schwartz. 1998. Measuring individual differences in implicit cognition: the implicit association test. *Journal of Personality and Social Psychology*. American Psychological Association 74(6). 1464.
Grossberg, Stephen. 2013. Adaptive Resonance Theory: How a brain learns to consciously attend, learn, and recognize a changing world. *Neural Networks* (Twenty-Fifth Anniversay Commemorative Issue) 37. 1–47.
Grossberg, Stephen & Gregory Stone. 1986. Neural dynamics of attention switching and temporal-order information in short-term memory. *Memory & Cognition* 14(6). 451–468.
Gupta, Anthea Fraser. 2005. Baths and becks. *English Today* 21(1). 21–27.
Halfacre, Caitlin. 2019. Regional or regionless? Investigation RP with privately educated speakers in the North East and the South East. In. London.
Hall-Lew, Lauren, Emma Moore & Robert Podesva. 2021. *Social Meaning and Linguistic Variation*. Cambridge: Cambridge University Press.
Hawkins, Sarah. 2003. Roles and representations of systematic fine phonetic detail in speech understanding. *Journal of Phonetics* 31(2). 373–405.
Hay, Jennifer. 2018. Sociophonetics: The role of words, the role of context, and the role of words in context. *Topics in Cognitive Science* 10(4). 696–706.
Hay, Jennifer, Aaron Nolan & Katie Drager. 2006. From fush to feesh: Exemplar priming in speech perception. *The Linguistic Review* 23(3). 351–379.
Hay, Jennifer, Janet Pierrehumbert & Mary Beckman. 2004. Speech perception, well-formedness, and the statistics of the lexicon. *Papers in Laboratory Phonology* VI. 58–74.

Hay, Jennifer, Abby Walker, Kauyumari Sanchez & Kirsty Thompson. 2019. Abstract social categories facilitate access to socially skewed words. *PLOS ONE* 14(2). e0210793.
Hay, Jennifer, Paul Warren & Katie Drager. 2006. Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics* 34(4). 458–484.
Holcomb, Phillip J. & Jane E. Anderson. 1993. Cross-modal semantic priming: A time-course analysis using event-related brain potentials. *Language and Cognitive Processes*. Taylor & Francis 8(4). 379–411.
Holcomb, Phillip J. & Helen J. Neville. 1990. Auditory and visual semantic priming in lexical decision: A comparison using event-related brain potentials. *Language and Cognitive Processes*. Taylor & Francis 5(4). 281–312.
Hughes, Arthur, Peter Trudgill & Dominic Watt. 2012. *English Accents and Dialects: An Introduction to Social and Regional Varieties of English in the British Isles*. 5th edn. New York: Routledge.
Jeffries, Ella. 2016. *Children's developing awareness of regional accents: A socioperceptual investigation of pre-school and primary school children in York*. University of York PhD Thesis.
Johnson, Keith. 2006. Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics* 34(4). 485–499.
Johnson, Keith. 2007. Decisions and mechanisms in exemplar-based phonology. *Experimental Approaches to Phonology* 25–40.
Johnstone, Barbara. 2009. Pittsburghese shirts: Commodification and the enregisterment of an urban dialect. *American Speech*. Duke University Press 84(2). 157–175.
Joos, Martin. 1948. Acoustic phonetics. *Language* 24(2). 5–136.
Kemps, Rachèl, Mirjam Ernestus, Robert Schreuder & Harald Baayen. 2004. Processing reduced word forms: The suffix restoration effect. *Brain and Language* 90(1–3). 117–127.
Khattab, Ghada. 2007. Variation in vowel production by English-Arabic bilinguals. *Laboratory Phonology* 9. 383–410.
Kim, Jonny. 2016. Perceptual associations between words and speaker age. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 7(1).
Kim, Jonny & Katie Drager. 2018. Rapid Influence of word-talker associations on lexical access. *Topics in Cognitive Science* 10(4). 775–786.
Kinoshita, Sachiko & Louise Hunt. 2008. RT distribution analysis of category congruence effects with masked primes. *Memory & Cognition* 36(7). 1324–1334.
Kleinschmidt, Dave F., Kodi Weatherholtz & T. Florian Jaeger. 2018. Sociolinguistic perception as inference under uncertainty. *Topics in cognitive science*. Wiley Online Library 10(4). 818–834.
Koops, Christian, Elizabeth Gentry & Andrew Pantos. 2008. The effect of perceived speaker age on the perception of PIN and PEN vowels in Houston, Texas. *University of Pennsylvania Working Papers in Linguistics* 14(2). 12.
Kouider, Sid & Stanislas Dehaene. 2007. Levels of processing during non-conscious perception: a critical review of visual masking. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 362(1481). 857–875.
Kraljic, Tanya, Susan E. Brennan & Arthur G. Samuel. 2008. Accommodating variation: Dialects, idiolects, and speech processing. *Cognition*. Elsevier 107(1). 54–81.
Kraljic, Tanya & Arthur G. Samuel. 2006. Generalization in perceptual learning for speech. *Psychonomic Bulletin & Review* 13(2). 262–268.
Kusunose, Yuu, Yasushi Hino & Stephen J. Lupker. 2016. Masked semantic priming effects from the prime's orthographic neighbours. *Journal of Cognitive Psychology* 28(3). 275–296.
Lambert, W. E., R. C. Hodgson, R. C. Gardner & S. Fillenbaum. 1960. Evaluational reactions to spoken languages. *The Journal of Abnormal and Social Psychology* 60(1). 44–51.
Lambert, Wallace E. 1967. A social psychology of bilingualism. *Journal of Social Issues* 23(2). 91–109.

Lambert, Wallace E., Moshe Anisfeld & Grace Yeni-Komshian. 1965. Evaluation reactions of Jewish and Arab adolescents to dialect and language variations. *Journal of Personality and Social Psychology* 2(1). 84.

Lavan, Nadine, Sarah Knight & Carolyn McGettigan. 2019. Listeners form average-based representations of individual voice identities. *Nature Communications* 10(1). 1–9.

Lawrence, D. 2014. Phonetic variation as a cue to regional identity: an experimental approach. In *Proceedings of the second Postgraduate and Academic Researchers in Linguistics at York (PARLAY 2014) Conference*, 42–63.

Leemann, Adrian, Marie-José Kolly & David Britain. 2018. The English Dialects App: The creation of a crowdsourced dialect corpus. *Ampersand* 5. 1–17.

Lenth, Russell. 2020. *emmeans: Estimated marginal means, aka least-squares means. R package*.

Levon, Erez. 2007. Sexuality in context: Variation and the sociolinguistic perception of identity. *Language in Society* 36(4). 533–554.

Levon, Erez. 2017. Sociolinguistics: Theoretical debates. *Journal of Sociolinguistics*.

Luce, Paul A. & David B. Pisoni. 1998. Recognizing spoken words: The Neighborhood Activation Model. *Ear and Hearing* 19(1). 1–36.

Lüdecke, Daniel. 2019. *sjPlot: Data visualization for statistics in social sience*.

Maguinness, Corrina, Claudia Roswandowitz & Katharina von Kriegstein. 2018. Understanding the mechanisms of familiar voice-identity recognition in the human brain. *Neuropsychologia*. Elsevier 116. 179–193.

Massaro, Dominic W. 1989. Testing between the TRACE model and the fuzzy logical model of speech perception. *Cognitive Psychology* 21(3). 398–421.

Mayr, Susanne & Axel Buchner. 2007. Negative priming as a memory phenomenon. *Zeitschrift für Psychologie / Journal of Psychology* 215(1). 35–51.

McClelland, James L & Jeffrey L Elman. 1986. The TRACE model of speech perception. *Cognitive Psychology* 18(1). 1–86.

McLennan, Conor T. & Paul A. Luce. 2005. Examining the time course of indexical specificity effects in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 31(2). 306–321.

McLennan, Conor T., Paul A. Luce & Jan Charles-Luce. 2003. Representation of lexical form. *Journal of Experimental Psychology. Learning, Memory & Cognition* 29(4). 539–553.

McNamara, Timothy P. 2005. *Semantic Priming: Perspectives from Memory and Word Recognition*. London: Psychology Press.

McNamara, Timothy P. & Jeanette Altarriba. 1988. Depth of spreading activation revisited: Semantic mediated priming occurs in lexical decisions. *Journal of Memory and Language*. Elsevier 27(5). 545–559.

Mepham, Alex & Bronwen Evans. 2019. Hæ? Exploring factors influencing the identification and judgement of two Norwegian dialects. In.

Meyer, David E. & Roger W. Schvaneveldt. 1971. Facilitation in recognizing pairs of words: evidence of a dependence between retrieval operations. *Journal of experimental psychology* 90(2). 227–234.

Moore, Emma & Robert Podesva. 2009. Style, indexicality, and the social meaning of tag questions. *Language in Society* 38(4). 447–485.

Mugglestone, Lynda. 2003. *Talking Proper: The Rise of Accent as Social Symbol*. Oxford: Oxford University Press on Demand.

Narum, Shawn R. 2006. Beyond Bonferroni: Less conservative analyses for conservation genetics. *Conservation Genetics*. Springer 7(5). 783–787.

Neely, James H. 1977. Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General* 106(3). 226–254.

Neill, W. Trammell & Leslie A. Valdes. 1992. Persistence of negative priming: Steady state or decay? *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18(3). 565–576.

Newman, Sharlene D. & Donald Twieg. 2001. Differences in auditory processing of words and pseudowords: An fMRI study. *Human brain mapping*. Wiley Online Library 14(1). 39–47.

Niedzielski, Nancy. 1999. The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology* 18(1). 62–85.

Norris, Dennis, Anne Cutler, James M. McQueen & Sally Butterfield. 2006. Phonological and conceptual activation in speech comprehension. *Cognitive Psychology* 53(2). 146–193.

Norris, Dennis, James M. McQueen & Anne Cutler. 2000. Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences* 23(3). 299–325.

Norris, Dennis, James M. McQueen & Anne Cutler. 2003. Perceptual learning in speech. *Cognitive Psychology* 47(2). 204–238.

Nygaard, Lynne C., Mitchell S. Sommers & David B. Pisoni. 1994. Speech perception as a talker-contignent process. *Psychological Science* 5(1). 42–46.

Ochs, Elinor. 1992. Indexing gender. In *Rethinking context: Language as an interactive phenomenon*, 335. Cambridge: Cambridge University Press.

Peirce, Charles Sanders. 1991. *Peirce on signs: Writings on semiotic*. Chapel Hill, NC: UNC Press Books.

Peirce, Jonathan, Jeremy R. Gray, Sol Simpson, Michael MacAskill, Richard Höchenberger, Hiroyuki Sogo, Erik Kastman & Jonas Kristoffer Lindeløv. 2019. PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods* 51(1). 195–203.

Perea, Manuel, Jon Andoni Duñabeitia & Manuel Carreiras. 2008. Masked associative/semantic priming effects across languages with highly proficient bilinguals. *Journal of Memory and Language* 58(4). 916–930.

Petrides, Michael. 1996. Specialized systems for the processing of mnemonic information within the primate frontal cortex. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*. The Royal Society London 351(1346). 1455–1462.

Pharao, Nicolai. 2019. Processing words in a different accent: The role of tonal contours in lexical access. In. Münster.

Pharao, Nicolai, Marie Maegaard, Janus Spindler Møller & Tore Kristiansen. 2014. Indexical meanings of [s+] among Copenhagen youth: Social perception of a phonetic variant in different prosodic contexts. *Language in Society* 43(1). 1–31.

Pierrehumbert, Janet B. 2001. Exemplar dynamics: Word frequency, lenition and contrast. *Typological Studies in Language* 45. 137–158.

Pierrehumbert, Janet B. 2002. Word-specific phonetics. In Carlos Gussenhoven & Natasha Warner (eds.), *Laboratory Phonology 7*, 101-139. Berlin: Mouton de Gruyter.

Pierrehumbert, Janet B. 2016. Phonological representation: Beyond abstract versus episodic. *Annual Review of Linguistics* 2. 33–52.

Pinget, Anne-France. 2019. Accent recognition in the Netherlands. In.

Pinget, Anne-France, Marjolein Rotteveel & Hans Van de Velde. 2014. Standaardnederlands met een accent. Herkenning en evaluatie van regionaal gekleurd Standaardnederlands in Nederland. *Nederlandse taalkunde* 19(1). 3–45.

Plichta, Bartlomiej & Dennis R. Preston. 2005. The/ay/s have It the perception of/ay/as a north-south stereotype in United States English. *Acta Linguistica Hafniensia* 37(1). 107–130.

Podesva, Robert. 2004. On constructing social meaning with stop release bursts. In. Newcastle upon Tyne.

Podesva, Robert. 2007. Phonation type as a stylistic variable: The use of falsetto in constructing a persona. *Journal of Sociolinguistics* 11(4). 478–504.

Preston, Dennis R. 1993. Folk dialectology. *American Dialect Research* 333–378.

R Core Team. 2013. R: A language and environment for statistical computing.

Radeau, Monique. 1983. Semantic priming between spoken words in adults and children. *Canadian Journal of Psychology/Revue canadienne de psychologie* 37(4). 547–556.

Raettig, Tim & Sonja A. Kotz. 2008. Auditory processing of different types of pseudo-words: An event-related fMRI study. *NeuroImage* 39(3). 1420–1428.

Ranbom, Larissa J. & Cynthia M. Connine. 2007. Lexical representation of phonological variation in spoken word recognition. *Journal of Memory and Language* 57(2). 273–298.

Rastle, Kathleen, Jonathan Harrington & Max Coltheart. 2002. 358,534 nonwords: The ARC nonword database. *The Quarterly Journal of Experimental Psychology Section A* 55(4). 1339–1362.

Samuel, Arthur G. 2011. Speech Perception. *Annual Review of Psychology* 62(1). 49–72.

Schacter, Daniel L. & Barbara A. Church. 1992. Auditory priming: Implicit and explicit memory for words and voices. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18(5). 915–930.
Schiller, Niels O. 1999. Masked syllable priming of English nouns. *Brain and Language* 68. 300–305.
Silverstein, Michael. 2003. Indexical order and the dialectics of sociolinguistic life. *Language & Communication* 23(3–4). 193–229.
Snell, Julia. 2010. From sociolinguistic variation to socially strategic stylisation. *Journal of Sociolinguistics*. Wiley Online Library 14(5). 630–656.
Stanners, Robert F., James E. Jastrzembski & Allen Westbrook. 1975. Frequency and visual quality in a word-nonword classification task. *Journal of Verbal Learning and Verbal Behavior* 14(3). 259–264.
Stark, Craig EL & James L. McClelland. 2000. Repetition priming of words, pseudowords, and nonwords. *Journal of Experimental Psychology. Learning, Memory & Cognition*. American Psychological Association 26(4). 945.
Starr, Rebecca Lurie. 2021. Changing language, changing character types. In Lauren Hall-Lew, Emma Moore & Robert Podesva (eds.), *Social Meaning and Linguistic Variation*. Cambridge: Cambridge University Press.
Stoddart, Jana, Clive Upton & John DA Widdowson. 1999. Sheffield dialect in the 1990s: revisiting the concept of NORMs. In Paul Foulkes & Gerard Docherty (eds.), *Urban Voices: Accent Studies in the British Isles*, 72–89. London and New York: Routledge.
Strand, Elizabeth A. 1999. Uncovering the role of gender stereotypes in speech perception. *Journal of Language and Social Psychology*. Sage Publications Sage CA: Thousand Oaks, CA 18(1). 86–100.
Sumner, Meghan & Arthur G. Samuel. 2005. Perception and representation of regular variation: The case of final /t/. *Journal of Memory and Language* 52(3). 322–338.
Sumner, Meghan & Arthur G. Samuel. 2009. The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language* 60(4). 487–501.
Sumner, Meghan, Seung Kyung Kim, Ed King & Kevin B. McGowan. 2014. The socially weighted encoding of spoken words: a dual-route approach to speech perception. *Frontiers in Psychology*. Frontiers 4.
Svartvik, Jan & Geoffrey Leech. 2006. Linguistic change in progress: Back to the inner circle. In *English*, 206–221. Springer.
Tipper, Steven P. 1985. The negative priming effect: Inhibitory priming by ignored objects. *The Quarterly Journal of Experimental Psychology* 37(4). 571–590.
Trudgill, Peter. 1990. *The Dialects of England*. Oxford: Basil Blackwell.
Trudgill, Peter & Jean Hannah. 2002. *International English. A Guide to the Varieties of Standard English*. 4th Ed. London: Arnold.
Van Bavel, Jay J., Yi Xiao & William A. Cunningham. 2012. Evaluation is a dynamic process: Moving beyond dual system models. *Social and Personality Psychology Compass*. Wiley Online Library 6(6). 438–454.
Van Bezooijen, Renée & Charlotte Gooskens. 1999. Identification of language varieties: The contribution of different linguistic levels. *Journal of Language and Social Psychology* 18(1). 31–48.
Villarreal, Dan. 2018. The Construction of Social Meaning: A Matched-Guise Investigation of the California Vowel Shift. *Journal of English Linguistics* 46(1). 52–78.
Vitevitch, Michael S. & Paul A. Luce. 1998. When words compete: Levels of processing in perception of spoken words. *Psychological Science*. SAGE Publications Inc 9(4). 325–329.
Vitevitch, Michael S., Paul A. Luce, David B. Pisoni & Edward T. Auer. 1999. Phonotactics, neighborhood activation, and lexical access for spoken words. *Brain and Language*. Elsevier 68(1–2). 306–311.
Vriesendorp, Hielke & Gijsbert Rutten. 2017. Omg zo fashionably english. Codeswitching to English as an identity practice in the chatspeak of Dutch young gay men. *Taal en Tongval* 69(1). 47–70.

Vriesendorp, Hielke & Gijsbert Rutten. 2019. Het einde van het Poldernederlands?*: De perceptie van/ɛi/en/ai/in standaardtalig Nederlands in Nederland. *Nederlandse Taalkunde*. Amsterdam University Press 24(3). 323–356.

Walker, Abby & Jen Hay. 2011. Congruence between ‘word age’ and ‘voice age’ facilitates lexical access. *Laboratory Phonology* 2(1). 219–237.

Walker, Abby, Janet Van Hell & Mike Bowers. 2019. Listener ratings of accentedness (somewhat) predict ease of lexical access. In.

Walsh, Michael, Bernd Möbius, Travis Wade & Hinrich Schütze. 2010. Multilevel exemplar theory. *Cognitive Science* 34(4). 537–582.

Watt, Dominic JL. 2000. Phonetic parallels between the close-mid vowels of Tyneside English: Are they internally or externally motivated? *Language Variation and Change* 12(1). 69–101.

Watt, Dominic, Carmen Llamas, Peter French, Almut Braun & Duncan Robertson. 2019. Listener sensitivity to localised accent features using the Geographical Association Test (GAT). In.

Watt, Dominic & Jennifer Smith. 2005. Language change. In M.J. Ball (ed.), *Clinical Sociolinguistics*, vol. 36, 101–119. Oxford: Blackwell Publishing Ltd.

Watt, Dominic & Jennifer Tillotson. 2001. A spectrographic analysis of vowel fronting in Bradford English. *English World-Wide* 22(2). 269–303.

Weber, Andrea & Anne Cutler. 2004. Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language* 50(1). 1–25.

Wedel, Andrew. 2012. Lexical contrast maintenance and the organization of sublexical contrast systems. *Language and Cognition* 4(4). 319–355.

Wells, John C. 1982. *Accents of English*. Cambridge: Cambridge University Press.

Whaley, C. P. 1978. Word—nonword classification time. *Journal of Verbal Learning and Verbal Behavior* 17(2). 143–154.

White, Yosiane. 2019. Auditory semantic priming: How can we use it to study linguistic variation? In. University of Münster.

Wieling, Martijn & Mark Tiede. 2017. Quantitative identification of dialect-specific articulatory settings. *The Journal of the Acoustical Society of America* 142(1). 389–394.

Williams, Angie, Peter Garrett & Nikolas Coupland. 1999. Dialect recognition. In Dennis R. Preston (ed.), *Handbook of Perceptual Dialectology*, 345–358. Philadelphia: John Benjamins.

Williams, Daniel & Paola Escudero. 2014. A cross-dialectal acoustic comparison of vowels in Northern and Southern British English. *The Journal of the Acoustical Society of America* 136(5). 2751–2761.

Zhang, Qing. 2008. Rhotacization and the ‘Beijing Smooth Operator’: The social meaning of a linguistic variable. *Journal of Sociolinguistics* 12(2). 201–222.