Regional priming in Australian English KIT, DRESS and TRAP vowels

Michael Stuart Walker

Department of Linguistics

Faculty of Human Sciences

Macquarie University

A thesis submitted in partial fulfilment of the requirements for the degree of Master of Research (Linguistics)

10th October, 2016

Contents

List of Tables

List of figures

Abstract

The presence of culturally significant objects has been shown to induce perceptual biases in speech that are consistent with features of the implied dialect. This thesis details a pseudo-replication of Hay and Drager's (2010) stuffed toy study in an Australian context. 43 participants heard phrases of spoken Australian English (AusE) with a phrase-final word containing either a KIT, DRESS or TRAP target vowel. Each phrase was followed by audio presentation of a continuum consisting of six synthesised variants of the speaker's target vowel, from New Zealand English (NZE)-like to exaggerated AusE. Participants selected the synthesised variant they believed best matched the speaker's realisation of the target vowel. Participants were exposed to either one of two priming conditions, established by stuffed toy kiwis (New Zealand) and stuffed toy koalas (Australia), or a control. The priming effect observed in Hay and Drager (2010) was not found. Token selections did not differ significantly between the two priming conditions, even for participants who had indicated frequent exposure to NZE. It is suggested that the influence of regional priming on speech perception may be more limited than previously thought, however findings also raise questions inherent to the task design itself.

Declaration

I hereby declare that this thesis has not been submitted for a higher degree to any other university or institution. I have made every effort to clearly indicate the sources of information used and acknowledge the extent to which the work of others has been used in the text. The research presented in this thesis has been approved by the Macquarie University Faculty of Human Sciences Research Ethics Sub-Committee (ref: 5201600056).

Michael Stuart Walker

10.10.2016

Acknowledgements

First of all, my sincerest thanks and gratitude to my supervisors Felicity Cox and Anita Szakay. Your support, advice, enthusiasm and encouragement has been invaluable and this project is as much a reflection of your time and effort as it is mine. I couldn't ask for better mentors. I would like to thank everyone in the Macquarie Phonetics Lab for their ongoing encouragement and keeping me motivated with new ideas. I would also like to thank Titia Benders, Carman Kung and everyone in the writing group for advice, suggestions and feedback. You have all made me a better writer. Special thanks to all my fellow MRes students and our teaching staff from this year and last. Thanks to Katie Drager for sending us the stimuli sentences. Thanks to Mark Maiuolo for providing your voice to the stimuli and to all the participants for your time. I am grateful to all my family and friends for everything. Finally, to my partner Kate. Your patience, encouragement, ideas, understanding and willingness to listen to my endless rambling got me through this year.

1. Introduction

The role of indexical information has been largely overlooked by theories of speech perception in the generative tradition (Foulkes & Docherty, 2006; Docherty & Foulkes, 2014). However, there is widespread support for the idea that indexical information – information about a speaker as well as the social or physical speech context (Pierrehumbert, 2016) – does play a role in speech perception (Foulkes & Docherty, 2006; Hay, Nolan, & Drager, 2006). Recent proposals of representation have argued that linguistic experiences are retained as episodic memories that incorporate both indexical and acoustic information (Goldinger 1996; Pierrehumbert, 2001). This experience-based representation occurs concurrently with an abstract representation (Hay & Foulkes, 2016; Pierrehumbert 2016). Social interaction is seen as integral to the process of building linguistic knowledge, which includes rich phonetic detail and variation (Pierrehumbert, 2006). Familiarity with patterns between linguistic variation and social characteristics affords listeners a high level of sensitivity to the sociophonetic consequences of these characteristics and allows these listeners to anticipate how the speaker will sound. In experimental contexts, manipulation of indexical information has been shown to produce predictable shifts in phoneme identification (Niedzielski, 1999; Strand & Johnson, 1996).

Hay and Drager (2010) tested the extent to which the introduction of a culturally significant token could influence vowel identification. This built on the studies of Niedzielski (1999) and Hay et al. (2006), in which written regional labels were shown to influence selections in a vowel matching task. Hay and Drager (2010) found that the presence of stuffed toys had a similar regional priming effect. Toys related to Australia lead to more Australian English (AusE)-like vowels being selected. This result suggests that, for New Zealand English (NZE) speakers, reference to Australia through a cultural token could lead to the perception of more AusE-like vowels in recorded speech.

1.1 Motivation

The present study aims to investigate whether a regional priming effect could be observed in an Australian context. As New Zealanders have been shown to be influenced by Australian-themed priming conditions (Hay & Drager, 2010; Hay et al., 2006), we were curious to test whether Australians would show a similar influence when exposed to a New Zealand-themed priming condition. This would build upon our current understanding of the influence of indexical information on speech perception and the extent to which previously observed priming effects (i.e. Hay and Drager, 2010) could be generalised.

1.2 Thesis overview

This thesis will be divided into six chapters. Chapter 1 introduces the thesis topic and its motivation. Chapter 2 provides a review of previous research investigating sociophonetic influences on speech perception, including those studies investigating the regional priming effect. This chapter will also outline the central claims of exemplar theory and how it accounts for the regional priming effect. Background information on the dialects of English concerned in the present study – AusE and NZE – will also be provided. Chapter 3 provides a detailed outline of the stimuli, continua design and procedure used for the matching task. This section also provides a methodological rationale for differences between the present study and the previous studies. Chapter 4 details raw data from the matching task as well as results from statistical analyses. Chapter 5 discusses the results of the analyses and an interpretation of their relevance to the existing studies is offered. This chapter will also address the limitations of the study. Chapter 6 concludes the thesis and summarises the contribution to the field provided by this work, also outlining suggestions for further research.

2. Literature Review

2.1 Introduction

The following chapter surveys relevant literature to inform the methodological and theoretical paradigm in which the present study is situated. Section 2.2 provides an overview of sociophonetic studies investigating the influence of indexical information on the perception of speech. Section 2.3 will examine both the findings and methodological approach of a number of regional priming studies, including those of Niedzielski (1999), Hay et al. (2006) and Hay and Drager (2010). While these studies have, with one exception, reported a measurable influence of a regional prime in speech perception, there are a number of methodological and procedural concerns that this study seeks to address.

Sociophonetic priming has been widely argued to support an exemplar-based model of speech perception. As such, Section 2.4 summarises the main arguments made by exemplar theory and discusses how it can account for the observed regional priming effect. Section 2.5 provides a description of the two dialects concerned in this study, AusE and NZE. This section will include a review of literature investigating the extent to which speakers of AusE and NZE are sensitive to variation between the two dialects. Finally, the predictions and hypotheses of the present study will be outlined in Section 2.7.

2.2 Non-linguistic influences on speech perception

A large body of research has shown that listeners make use of linguistic cues to inform knowledge of a speaker's social or regional background suggesting a close relationship between linguistic and indexical information (see, e.g. Foulkes, Scobbie, & Watt, 2010; Thomas, 2002 for reviews). A number of studies have also shown this relationship to be bidirectional, suggesting that indexical information may also influence the way speech is perceived. As will be shown below, speaker gender, age, and ethnicity have each been shown to have a predictable effect on participants' performance in lexical access and phoneme identification tasks.

For example, the perceived gender of a speaker has been shown to align socially induced expectations with characteristics of their speech (Johnson, Strand, & D'Imperio, 1999; May, 1976; Strand & Johnson, 1996). May (1976) showed that the perceptual boundary between /s/ and /ʃ/ was higher in female speakers, reflecting differences in vocal tract size. Strand and Johnson (1996) were able to show that the perceived gender of a speaker could influence classification of /s/ and /ʃ/, even when it conflicted with acoustic information. Participants were able to accurately determine whether a speaker was male or female in an audio only context with their initial classification of tokens on an /s/-/ʃ/ continuum reflected the findings of May (1976). However, when a voice was presented with an image of a female, participants identified more tokens as /ʃ/ than when the same voice was presented with an image of a male. In a similar experiment, Johnson et al. (1999) found perceived speaker gender influenced the location of the perceptual boundary between /ʊ/ and /ʌ/ with the effect found to be greater when the voice was more stereotypically male or female. Therefore, categorisation of speech input has been shown to reflect preconceived expectations about gendered differences in speech.

Perceived age has also been shown to affect the categorisation of speech. Drager (2011) showed that the perceptual boundary between DRESS and TRAP vowels for NZE-speakers shifted according to the implied age of a speaker. When classifying tokens on a DRESS - TRAP continuum, listeners perceived more tokens as TRAP when the speech was accompanied by a photo of a younger person. Drager (2011) argues that this result suggests participants were accounting for a trend towards raised variants of both DRESS and TRAP in younger NZE speakers. Further, the effect was only found in older participants. This is attributed to older participants having experienced the ongoing short front vowel chain shift in NZE (Maclagan & Hay, 2007), resulting in a more salient difference between the production of these vowels in younger and older speakers. Although Drager (2011) concedes the effect is subtle, the result is supported by Hay, Warren, and Drager (2006), and Koops, Gentry, and Pantos (2008) who found that that perceived age influences anticipated vowel qualities in speakers.

Ethnicity of a speaker has been shown to influence perception, particularly with regard to the perceived intelligibility of speech. Rubin (1992) had participants listen to a recorded lecture, spoken by a single Standard American English (SAE) speaker, and included an image of either a female Caucasian or Asian face. The lecture was then rated on various intelligibility, accent and social homophily scores. Despite all participants hearing SAE speech, responses from those participants exposed to the image of an ethnically Asian woman were in a direction indicating nonstandard or 'accented' speech and suggested an undermining of listening comprehension. This result was supported in Babel and Russell (2015) who found that native English speaking Chinese Canadians were perceived as less intelligible when their voices were paired with static images of the speaker instead of participants only seeing a fixation cross. Together with the observed influence of gender and age, there is strong support for the claim that social information attributed to a speaker can influence the way their speech is perceived.

2.3 Speech perception and regional priming

Beginning with Niedzielski (1997), results from a series of phoneme and word identification tasks point to a regional priming effect (Hay & Drager, 2010; Hay et al., 2006; Jannedy, Weirich, & Brunne, 2011). Niedzielski (1997, 1999) demonstrated that Detroiters' recognition of Canadian Raising (CR) in recorded speech was dependent on where they believed a speaker was from. CR is the raising of the diphthongs /aɪ/ and /aʊ/ to become /ʌɪ/ and /ʌʊ/ before voiceless consonants and is a feature of multiple Canadian and American English dialects (Niedzielski (1997). It was found that while the dialect spoken by Detroiters included CR, they did not associate it with their speech. Detroiters believed their dialect to be equivalent to SAE, which does not feature CR. It was proposed that Detroiters would be more likely to perceive CR in the speech of another Detroiter if they believed they were actually listening to a Canadian. Conversely, if they believed the speaker was from Detroit they would perceive more SAE-like vowels. Similarly, Detroiters were found to be largely unaware of the influence of the Northern Cities Chain Shift (NCCS) on their vowel space (Niedzielski, 1997). The NCCS is a chain shift occurring in some vowels of speakers of American English centred in the Great Lakes Region (see Labov, 1991 for more details).

In order to test Detroiters' beliefs about their own dialect, a vowel matching task was designed to investigate whether belief about a speaker's regional origin would influence the way their speech was perceived (Niedzielski, 1999). Participants heard the recorded speech of a single Detroit speaker, who produced CR and NCCS vowels. Half of the participants were told the speaker was from Windsor, in Canada, and the other half were told Detroit. In addition, responses were recorded on a sheet with either *Canadian* or *Michigan* written clearly on it. Spoken phrases with identified target words, underlined on the response sheet, were played. After the sentence, a six-step continuum of re-synthesised tokens based on the vowel in the target word was played. For each target word, participants selected the closest matching token from the continuum.

For target words containing the diphthong /aʊ/, participants in the Canadian condition selected on average a more raised variant – consistent with CR – than those in the Michigan condition. For the NCCS affected vowel in *bed*, participants in the Michigan condition selected more 'standard' (SAE) tokens than those in the Canadian condition. The other NCCS affected vowels in the words *pop*, *last*, *close*, and *talking* also carried this trend but response variance between the conditions was not significant. Niedzielski (1999) argues that in the Michigan condition, participants anticipated an SAE vowel space, as they believe that is what the Michigan dialect sounds like. This expectation then biased perception towards variants that were congruent with their mental representations of SAE vowels. This suggests belief about a speaker's regional origin and their presumed dialect could not only influence expectation about their speech but the way their speech was perceived.

2.3.1 New Zealand replications

Two replications of the procedure outlined in Niedzielski (1999) were conducted in a New Zealand context: Hay, Nolan and Drager (2006) and Hay and Drager (2010). Both studies found support for the regional priming effect observed by Niedzielski (1999). These experiments targeted perception of KIT vowels, which represent the most salient difference between AusE and NZE, along with two additional short front vowels: DRESS and TRAP. Hay et al. (2006) used labelled response sheets to establish the two conditions, with either *Australia* or *New Zealand* written on top. Participants, all from New Zealand, heard recorded sentences spoken by a single NZE speaker. Each sentence contained a monosyllabic target word featuring a KIT, DRESS or TRAP vowel that was underlined on the answer sheet. Following each sentence, a six-step continuum was played. Each continuum represented a vowel trajectory of six synthesised tokens from AusE-like (token 1) to NZE-like (token 6) with the speaker's actual vowel represented by token 4. Hay et al.'s (2006) replication differed from Niedzielski's experiment in that participants were not explicitly told anything about the speaker's identity; the label on the answer sheet alone was used to infer nationality of the speaker. The task was also simplified, with only one target word per sentence and the position of the speaker's resynthesised token in each continuum remaining consistent.

In a follow-up study, and the primary focus of replication for the present study, Hay and Drager (2010) proposed that any culturally significant token might result in a priming effect. Participants completed the same task as reported in Hay et al. (2006) using identical auditory stimuli. Rather than a written label, stuffed toys were used to establish each condition, kiwis for New Zealand and kangaroos and koalas for Australia. Prior to the beginning of the perception task, the prime was introduced when the experimenter 'found' either the kangaroos and koalas or kiwis in the room and feigned surprise. The intention was to draw attention to the toys without indicating they were in any way related to the experiment, thus keeping participants naïve to the purpose of the toys. The toys then remained visible to the participant for the duration of the experiment.

In both Hay et al. (2006) and Hay and Drager (2010) token selections for KIT vowels showed significant variability congruence with the prime for female participants only, although the priming effect was more robust in the Hay et al. (2006) experiment. Female participants exposed to the Australian condition typically selected lower numbered tokens, representing more AusE-like variants of the target vowel, than those exposed to the New Zealand condition. The priming effect observed in Hay and Drager (2010) was limited to participants with higher social class index scores and in Hay et al. (2006) higher class participants were particularly influenced by the *Australia* label. Analysis in both instances used linear regression without taking subject into account therefore the results are potentially overstated, especially given the sample size in Hay and Drager (n = 26). In Hay et al. (2006) TRAP vowels also showed variability congruence however the opposite effect was observed in Hay and Drager (2010) (similar to how male participants performed overall, see below). Neither study reported a significant difference between conditions for DRESS vowels.

In both studies, male participants selected more AusE-like tokens when exposed to the New Zealand condition. No gender difference had been observed in Niedzielski's study and so Hay et al. (2006) and Hay and Drager (2010) attribute this to the competitive relationship, particularly in a sporting context, between Australians and New Zealanders. It is argued that men have a generally more negative view of Australia resulting in a divergence effect when exposed to the Australian prime (see also, Drager, Hay and Walker, 2010, discussed in further detail below).

It is important to recognise that the results do not suggest participants in the Australian condition selected AusE tokens while those in the New Zealand condition select NZE tokens, the effect is simply not that strong. The key finding was that female participants in the Australian condition selected *more AusE-like* KIT vowels. The significance of these results lies not in the actual tokens selected, rather, it lies in the differences between responses across the two conditions. These differences are directionally consistent with differences between AusE and NZE when the vowels are compared in an F1/F2 plane. In addition, the results from the two New Zealand experiments reflect those differences that are the most socially salient, hence the strongest effect being in KIT vowels. Further, in Hay et al. (2006), a free-choice question about the speaker's nationality revealed that all participants (except one) knew they were listening to a New Zealander. This meant that the regional label was not influencing belief about the speaker's origin, rather, Hay et al. (2006) argue, exposure to the *concept* of Australia was enough to initiate a perceptual shift. This was confirmed when the stuffed toy experiment (Hay & Drager, 2010) also resulted in a similar shift in the same demographic. It seems unlikely that the presence of stuffed toy kangaroos and koalas would cause participants to think the speaker was Australian, however the toy would certainly raise the concept of Australia.

For this reason, the priming effect observed in the two New Zealand experiments is actually subtly different to that observed in Niedzielski's (1999). Niedzielski's (1999) results confirmed the prediction that Detroiters do not believe CR is a feature of their speech, but do associate it strongly with the speech of a Canadian. This suggests that participants in the Canadian condition actually believed the speaker was Canadian. In Niedzielski's (1999) experiment, participants were explicitly told the origin of the speaker and would arguably have no reason to doubt this information. However, Hay et al. (2006) showed that the label activated the concept of Australia but participants did not actually believe the speaker was Australian.

Hay et al. (2006) and Hay and Drager's (2010) finding that only females carried the trend suggests that the priming effect may not be consistent nor generalizable, and therefore should not be overstated. Especially considering males showed the opposite effect. What is unclear from these results, is whether it should be expected that males and females perform differently in the perceptual task (they did not in Niedzielski's experiment), or if the reason males are speculated to have performed differently, competitiveness, are unique to an Australia/New Zealand context.

2.3.2 Additional replication

Jannedy et al. (2011) also found support for Niedzielski's (1999) findings in a similar experiment testing perception of sociophonetic variation in German. In the multi-ethnolect spoken in large urban areas of Germany, /ç/ is realised as palatalised [ʃ], a variant said to be well-known, "quite pervasive and noticeable" as well as "mocked, and stigmatized" (Jannedy et al., 2011: p. 1). Written prompts were used to imply the speaker was from one of two Berlin neighbourhoods – Kreuzberg where the [ʃ] variant is common and Zehlendorf, where it is not. Participants were presented tokens from a continuum of words from Fichte /fɪçtə/ to fischte /fɪʃtə/ and asked to identify them. Participants in the Kreuzberg condition perceived more tokens as fischte than those in the Zehlendorf condition. Colouring of the preceding vowel from /ɪ/ to /ʏ/ was seen to cue listeners to the palatalised variant, increasing the effect. These results indicate participants responded to the inferred social characteristics of the two neighbourhoods and the assumed dialect of their speakers. This result broadens the scope of the priming effect observed in Niedzielski (1999), Hay et al. (2006), and Hay and Drager (2010) and supports both the notion that listeners are highly sensitive to fine sociophonetic variability and that expectations about when variability will occur are intrinsically tied to regional labels.

Contrasting these findings however, Lawrence (2015) found no evidence for regional labels influencing responses in BATH and STRUT vowels in speakers of Standard Southern British English. The BATH and STRUT vowels are said to be highly salient regional markers of

Northern and Southern British English (Lawrence, 2015). In Southern speakers BATH is realised as [ɑː] and in Northern speakers as [a]. The Southern STRUT is realised as [ʌ] and Northern, either [ʊ] or [ə]. In a replication of Niedzielski (1999) and Hay et al. (2006), 40 monolingual participants heard sentences produced by a speaker from Sheffield, Northern England. Target words contained either a BATH or STRUT vowel in sentence final position of sentences that contained no other instances of the focus vowels. Consistent with Hay et al. (2006), token 4 of each continuum was a resynthesised version of a vowel produced by the speaker in an isolated hVd context. Continua represented a six-token trajectory from a Southern (token 1) to exaggerated Northern (token 6) variant. Priming conditions were established with an on-screen label (*Sheffield, Northern England* or *London, Southern England*). Participants were also explicitly informed of the speaker's 'origin'. Results from chi-squared tests comparing token selection proportions between the two conditions revealed no significant difference in selection.

This result raises further doubt regarding the generalisability and strength of a regional priming effect. While it could be concluded that the lack of variability was due to a less socially salient variation than was the focus of other experiments, such as Jannedy et al. (2011) and Hay et al. (2006), the BATH and STRUT vowels are "widely acknowledged as highly salient markers of regional identity in British English" (Lawrence, 2015: p. 1). It appears as though participants in Lawrence's (2015) experiment did exhibit variability consistent with the priming condition, however this was not significant. It is certainly plausible that another trial would produce a significant result. It is also possible that additional trials in the Detroit or New Zealand contexts would not produce significant results. For this reason, further research is certainly warranted to investigate this effect, and the extent to which it can be generalised.

2.4 Exemplar Theory

Results from the sociophonetic research discussed in Sections 2.2 and 2.3 present strong evidence that listeners make use of indexical information to anticipate how a speaker will, or should, sound. Listeners have displayed an understanding of how speech varies according to age, gender, ethnicity and other social features, particularly when this variation is socially salient. As seen in Drager (2011), Rubin (1992), and Strand and Johnson (1996) for example, manipulation of available social information may result in perceptual biases in phoneme and word recognition (Hay & Foulkes, 2016). This effect of indexical information on perception has been largely overlooked in traditional generative models of phonology (Foulkes & Docherty, 2006; Pierrehumbert, 2006). In response, exemplar models of perception have emerged to provide a usage-based account of statistical learning and situational indexing (Pierrehumbert, 2006) that have been supported by the increasing body of sociophonetic experimentation.

Exemplar theory proposes that phonological knowledge is represented in memory as a continually updated aggregation of phonetically rich perceptual memories (Bybee, 2006; Goldinger, 1996; Johnson, 1997; Lacerda, 1997; Pierrehumbert, 2001; Wedel, 2006). Importantly, exemplar theory also accounts for the indexical context in which speech input was encountered. An exemplar representation therefore encompasses simultaneous indexing of the propositional, allophonic and indexical properties of speech. In this way, the auditory properties that distinguish speakers, as well as the causes of these properties such as age, dialectal background or gender are retained (Johnson, 1997).

Within an exemplar system, categories emerge from clusters of similar input formed within a cloud of remembered exemplars (Pierrehumbert, 2001; Wedel, 2006). These categories represent a class of equivalent perceptual experiences and are shaped by frequency information and density distributions (Pierrehumbert, 2001; 2006). Accordingly, more frequently encountered categories become more substantially and richly represented and a lack of invariance in input builds a more explicit representation of variation (Pierrehumbert, 2006). Social categories are similarly represented and are acquired "through cluster analysis over perceived properties of people and social interactions" (Pierrehumbert, 2006: p. 527). Over time, associations between linguistic information and relevant indexical information begin to develop (Foulkes & Docherty, 2006). These associations are strengthened when the link between the phonological variant and social category is more transparent (Foulkes & Docherty, 2006). Although the precise mechanisms for category assignment vary between exemplar models, it is generally agreed that all novel input automatically updates the entire category system in some way (Wedel, 2006). A number of exemplar models also allow for remembered tokens to contribute simultaneously to multiple categories (Pierrehumbert, 2001; Wedel, 2006). Category assignment of novel input is thought to be made by analogically comparing new stimuli with previously categorised exemplars (Bybee, 2006; Johnson, 1997; Walsh, Möbius, Wade, & Schütze, 2010) however, comparison to an abstract prototype has also been proposed (Goldinger 1996; Hintzman 1986).

More recent applications of exemplar-based perception support the proposed role of abstraction, arguing for multi-level representation and fully incorporating elements from traditional abstractionist theory (McLennan, 2007; Pierrehumbert, 2006; 2016). This hybrid model can be traced back to Goldinger (1998) however it is most clearly outlined in Pierrehumbert (2006; 2016). Functionally, the hybrid model imports the central claims of exemplar theory but borrows, from generative models, the concept of multiple levels of representation. Pierrehumbert (2016) argues that an abstract level of representation is necessary because phonological representation is inherently abstract. Further, abstract representation is required for the processing of novel word forms. At this level, phonetic detail and indexical information is likely to be disregarded. However, as we have seen, there is enough evidence to suggest a second level of representation, where fine phonetic detail and indexical features are retained. This undercuts theories of minimal representation and demands an approach that can account for both levels. As Pierrehumbert (2016) argues, “the ability to perceive indexical features, produce them more or less in different contexts, and generalize them to new words and new interlocutors means that they must be cognitively represented” (p. 42).

Hay et al. (2006) and Hay and Drager (2010) contend that the observed regional priming effect can be explained by an exemplar model of speech perception. A prime, such as the stuffed toy koalas, would activate social categories associated with ‘Australia’ prior to the listener receiving any speech input. This would then raise their resting activation level, in a sense anticipating AusE input. When speech input is received, provided there is enough acoustic similarity, the phonological categories associated with the activated social category (i.e. those labelled as AusE) would then be the quickest to reach full activation, resulting in a system bias (Drager, 2011; Hay et al. 2006; Hay & Drager, 2010).

In the experimental context of the matching task this bias results in misidentification, however, misidentification is generally seen to be beneficial to the listener as it reduces recognition and retrieval time as well as improving identification of variant acoustic signals (Hay et al. 2006). Precisely how this miscategorisation would result in the selection of, for example, a slightly more AusE-like token in the matching task is unclear. A possible explanation is that the weight of other remembered tokens in the assigned category influence memory of the vowel to a more AusE-like quality than it actually was.

This account is certainly consistent with the implications of exemplar theory. Exposure to indexical variability improves a listener’s ability to identify the ways in which social characteristics influence speech (Foulkes & Docherty, 2006). Further, dialect is considered to be a highly important indexical feature (Pierrehumbert, 2016). However, the explanation put forth in Hay et al. (2006) and Hay and Drager (2010) does rely on a listener having the sufficient amount of exposure required to develop the requisite associations between social and phonological categories. It would seem apparent that the participants in Niedzielski

(1999), Hay et al. (2006) and Hay and Drager (2010) did have sufficient representations of the social and phonological categories relevant to the experiment. However, the extent to which this assumption can be extended and generalised remains to be seen. As seen in the results reported in Lawrence (2015), as well as in a morphosyntactic context by Squires (2013), experimental manipulation of speaker information does not necessarily result in significant perceptual shifts.

The present study aims to investigate whether the regional priming effect can be generalised. By replicating Hay and Drager (2010) in an Australian context, we well be able to establish whether the AusE/NZE priming effect is bidirectional. That is, whether Australians show the same sensitivity to a New Zealand prime and NZE as New Zealanders did with the Australian prime and AusE. Further, exemplar theory predicts that any influence on perception should be restricted to listeners who have the relevant indexical and linguistic category associations. Thus, performance in the matching task described in the present study may be able to predict the level of exposure Australian listeners have to NZE.

2.5 Australian and New Zealand English

As discussed, the present experiment relies on speakers of AusE being sensitive to variation between AusE and NZE. This section will outline differences between the two dialects as well as the salience and recognisability of these differences. We will focus on the short front vowels, justifying their selection as the target vowels for the present study. For consistency and clarity, Wells' (1982) lexical set labels will be used when referring to vowels.

As dialects of English spoken by peoples with close geographical and socio-cultural ties, AusE and NZE share many features (Watson, Harrington, & Evans, 1998). In general, both dialects use a consonant inventory and vowel system consistent with other non-rhotic dialects of English, such as Received Pronunciation (RP) (Bauer, Warren, Bardsley, Kennedy, & Major, 2007). There are, however, differences in the realisation of the short front vowel monophthongs discussed below (Hay et al. 2006; Watson et al., 1998). Other notable differences between AusE and NZE include a more centralised FOOT, and a NEAR-SQUARE merger, both features of NZE (see, e.g. Holmes, 1997; Maclagan & Gordon, 1996; Watson, Maclagan, & Harrington, 2000).

2.5.1 Comparison of the KIT, DRESS and TRAP vowels in AusE and NZE

Consistent with Hay et al. (2006) and Hay and Drager (2010), the target vowels for this study are the short front vowels KIT, DRESS and TRAP. Table 2.1 outlines the IPA representation of the AusE and NZE vowels, as recommended by Harrington, Cox, and Evans (1997) and (Bauer et al., 2007) respectively. Figure 2.1 outlines the approximate relative positioning of the KIT, DRESS and TRAP vowels in AusE and NZE.

Table 2.1. Target vowels comparing Wells (1982) lexical set, AusE transcriptions (Harrington et al., 1997) and NZE transcriptions (Bauer et al., 2007).

Wells	AusE	NZE
KIT	/ɪ/	/ɘ/
DRESS	/e/	/e/
TRAP	/æ/	/ɛ/

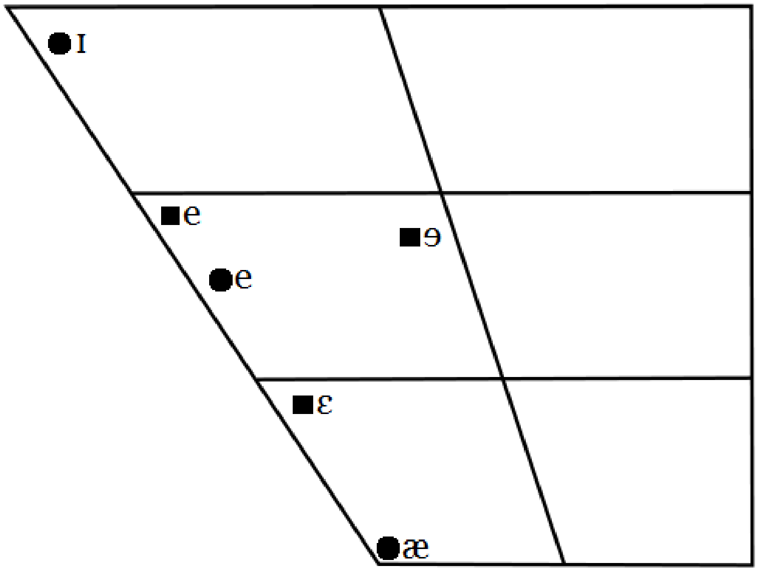


Figure 2.1. Approximate positions for target vowels in AusE (circles) (Cox & Palethorpe, 2007) and NZE (squares) (Bauer et al. 2007).

The KIT vowel is relatively centralised in NZE (Bauer et al., 2007). Easton and Bauer (2000) suggest that the ongoing centralisation of KIT in NZE has mostly manifested as diachronic lowering, however, Watson et al. (2000) argue that the distinctiveness of the NZE KIT is due to ongoing retraction. In AusE, KIT is a high front vowel. AusE KIT is believed to have raised throughout the 20th century however it is now showing signs of lowering (Cox & Palethorpe, 2008).

Stressed KIT is arguably the most identifiable and socially salient difference between the two dialects (Maclagan, Gordon, & Lewis, 1999; Watson et al., 1998). Indeed, the difference between the respective KIT vowels is the source of much humour and mockery between Australians and New Zealanders, perhaps best exemplified by the exaggerated imitation of the others' pronunciation of 'fish and chips'. Australians are said to pronounce "feesh and cheeps", while New Zealanders, "fush and chups" (Bauer & Warren, 2004). Australians have been reported as using their STRUT vowel to imitate or parody the NZE KIT (Bauer & Warren, 2004) indicating the variation is well-known, if not exaggerated.

The DRESS vowel is raised in both AusE and NZE relative to other dialects of English (Watson et al., 1998) however the NZE variant is typically more raised. NZE DRESS is similar in F1 and F2 to AusE KIT (Watson et al., 1998). There is ongoing raising in the NZE DRESS, particularly in younger speakers for whom DRESS overlaps with, and may even be higher than, FLEECE (Easton & Bauer, 2000; Maclagan & Hay, 2007; Watson et al., 2000). Similar to KIT, the AusE DRESS is lowering (Cox & Palethorpe, 2008).

Both varieties of TRAP are also raised relative to RP (Watson et al., 1998). TRAP is rising in NZE (Easton & Bauer, 2000) although not as much as DRESS (Watson et al. 2000). AusE TRAP has extensively lowered and retracted over a period of at least 25 years (Cox, 1999; Cox & Palethorpe, 2008).

2.5.2 Recognition and attitudes between AusE and NZE speakers

It has been shown that dialect recognition is facilitated by familiarity and listeners have been shown to make use of reliable acoustic-phonetic properties to identify a dialect (Clopper & Pisoni, 2004). Even naïve listeners, despite showing low recognition rates when identifying an unfamiliar speaker by dialect, are able to distinguish "between dialect groups using broad perceptual categories" (Clopper & Pisoni, 2004: p. 137; see also: Preston, 1993; Williams, Garrett, & Coupland, 1999). With regard to AusE and NZE, speakers of both dialects are said to consider the other to be not only easily identified, but unpleasant (Weatherall, Gallois, & Pittam, 1998). Australians and New Zealanders have been shown to be generally accurate at identifying both dialects (e.g. Bayard, Weatherall, Gallois, & Pittam, 1998; Ludwig, 2007; Weatherall et al., 1998) although speakers were misattributed in some instances.

Bayard et al. (2001) showed that both Australians and New Zealanders were accurately able to identify both male and female AusE and NZE speakers. However, New Zealanders were more likely to label a male Australian speaker as a New Zealander in an accent evaluation task. Results also indicated that New Zealanders tended to downgrade a NZE-speaker in terms of power and solidarity when compared with Australian, American and English speakers, yet Australians did not downgrade AusE. Similarly, New Zealanders rated NZE-speakers lower in competency than Australians did AusE-speakers (Bayard et al., 2001). A similar experiment was reported in Weatherall et al. (1998), with both Australians and New Zealanders accurate at identifying speakers. However, the female cultivated speaker and male general speakers of both NZE and AusE were frequently misattributed as being speakers of the other dialect.

Ludwig (2007) found that both Australians and New Zealanders were able to accurately differentiate between AusE and NZE productions of isolated words, although New Zealanders generally performed better. Participants rated synthesised words from 'definitely New

Zealander' to 'definitely Australian' on a six-point scale. Regarding Australians' perception of the three target vowels in the present study, KIT, DRESS and TRAP, Australians were found to be more accurate at identifying DRESS and TRAP vowels than New Zealanders. KIT was less accurately identified by Australians however it was the Australian KIT that was perceived as more NZE-like while the NZE KIT was correctly identified with a high level of accuracy. Consistent with expectations, Australians found KIT and SQUARE vowels to be the most salient identifiers of NZE (Ludwig, 2007).

2.5.3 Priming effects in speakers of AusE and NZE

It has been widely observed that speakers in conversation make spontaneous adjustments to their speech as a way to increase, reduce or maintain social distance with an interlocutor (Bell, 1984; Bourhis & Giles, 1977; Giles & Ogay, 2006; Giles & Smith, 1979; Labov, 1966; 1972). Babel (2010) showed that speakers of NZE shifted to more AusE-like vowel productions in the presence of an Australian, however, the degree of accommodation shown was related to a participant's score on an Implicit Association Task (Greenwald, McGhee, & Schwartz, 1998). Participants with a positive attitude towards Australia demonstrated a greater willingness to accommodate. Sanchez, Hay, and Nilson (2015) showed that a shift in production by NZE speakers towards more AusE-like vowels could be initiated even without the presence of an AusE speaker. Temporal proximity to Australia-related topics and lexical items caused a shift in the production of some NZE speakers' KIT and TRAP vowels to more AusE-like qualities. Production of DRESS vowels preceded by an Australia-related word was found to shift towards more AusE-like qualities only for participants with a higher level of exposure to AusE. This suggests that familiarity to the primed dialect may determine the extent to which an individual shows accommodation.

Drager, Hay, and Walker (2010) showed that exposing NZE speakers to either 'good' or 'bad' facts about Australia could influence vowel realisations. Those participants labelled as 'sports fans' (as determined from questionnaire responses) produced a more AusE-like KIT vowel after exposure to 'bad' facts about Australia relative to the control group and the group exposed to 'good' facts. Non-fans showed the opposite effect, converging to AusE after exposure to 'good' facts. Drager et al. (2010) argue that sports fans have an inherently more competitive and negative view of Australia, thus divergence is not entirely unexpected when the concept of Australia is introduced positively. When 'Australia' is introduced negatively, it conforms to a negative baseline attitude and thus becomes a cultural prime, resulting in normal convergence. Non-fans may be more neutral toward Australia therefore negative priming results in divergence and positive priming in convergence.

When considered alongside the results from Hay et al. (2006) and Hay and Drager (2010), there certainly seems to be evidence that the differences between short front vowels in NZE and AusE are salient and phonetically identifiable, at least for New Zealanders. The extent to which these results can be generalised to speakers of AusE remains to be seen. These experiments are useful in that they reveal the level of prime required to successfully elicit NZE, as well as the extent to which attitude and exposure may moderate the amount of linguistic sensitivity one displays. At present, more research is required in Australia to investigate how speakers of AusE might accommodate to NZE and whether accommodation interacts with familiarity and attitude towards NZE. If we assume sufficient exposure to NZE, given results discussed above, it would not be unexpected that speakers of AusE would be sensitive enough to the differences between the two dialects to show a priming effect.

2.6 Summary

It has been widely shown that listeners are sensitive to the sociophonetic consequences of a speaker's age, gender, ethnicity and dialect area. Results from the perceptual tasks described above suggest that indexical information may influence the perception of speech. This may result in biases that are, in some instances, consistent with anticipated features of speech, rather than actual features. This suggests that social information about a speaker plays a role in perception and has been argued as evidence for exemplar theory (Hay and Drager, 2010).

NZE speakers and listeners have displayed a high level of sensitivity to the phonetic variation between the short front vowels of AusE and NZE. A small amount of research has indicated that speakers of AusE may be similarly sensitive to these differences (Bayard et al., 1998; Ludwig, 2007; Weatherall et al., 1998), and as such may perform in a similar manner in the matching task reported in Hay and Drager (2010).

2.7 Predictions

Following the results reported in Niedzielski (1999), Hay et al. (2006) and Hay and Drager (2010), and assuming an exemplar model of speech perception, it is predicted that stuffed toys will influence participants' performance in a vowel matching task. AusE-speaking participants exposed to a New Zealand prime should select more NZE-like continuum tokens than those exposed to an Australian prime. Consistent with Hay et al. (2006) and Hay and Drager (2010) this shift should be present for target words containing a KIT vowel and may also be observed for target words containing DRESS or TRAP vowels but to a lesser extent.

However, we also predict that the priming effect will be limited to those participants who indicate some level of familiarity to New Zealand and NZE. This level of familiarity will be assessed from responses to a post-task questionnaire. Previous research has indicated that speakers of AusE are able to identify NZE (Bayard et al., 1998; Ludwig, 2007; Weatherall et al., 1998). In addition, the short front vowels, particularly KIT, are well-known signifiers of NZE for Australians (Maclagan et al., 1999; Watson et al., 1998). As indicated in Lawrence (2015), as well as the limitations of the results reported in the New Zealand experiments, there is reason to believe the effect may not be consistent, or even present for all participants.

Thus, our hypotheses are as follows:

H1. Stuffed toy kiwis will influence Australian participants' selections in the vowel matching task.

H2. Token selections for KIT vowels will exhibit the most variability between the Australian and New Zealand conditions. However,

H3. The priming effect will be limited to those participants who indicate some level of familiarity to New Zealand and NZE.

3. Method

3.1 Introduction

The primary aim of the present study was to conduct a replication of the experiment reported in Hay and Drager (2010) in an Australian, rather than a New Zealand, context. Hay and Drager's (2010) experiment was itself a continuation of the paradigm established by Niedzielski (1997, 1999) and replicated in Hay et al. (2006). Results from these studies suggest that responses in a perceptual vowel matching task could be influenced by priming participants towards a particular regional dialect. Consistent with Hay and Drager (2010) and Hay et al. (2006), the present study focuses on KIT, DRESS and TRAP vowels in two dialects of English: AusE and NZE. Participants completed a matching task where they selected, from six-step continuum of isolated synthesised vowels, what they believed was the closest match to a target vowel in a sentence of recorded speech.

To introduce the concept of AusE and NZE, stuffed toys, similar to those used in Hay and Drager (2010) were used as regional primes. In addition to replicating the procedure in the opposite dialectal context to previous experiments, we also aimed to address a number of concerns related to the design and methods used in Hay and Drager (2010), as well as Hay et al. (2006) and Niedzielski (1997, 1999). These concerns relate to control of the stimuli, design of the continua and the lack of a control condition. The intention was to improve the strength of any claims made by producing a more tightly controlled study. Methodological and procedural differences will be detailed in the following sections. The present study should be regarded as a pseudo-replication.

3.2 Participants

Participants in the present study were 43 female undergraduate students from Macquarie University in Sydney, Australia. Participants were judged to be native speakers of AusE on the basis that they were born in Australia and had received all of their primary and high school education in Australia. Age range was 18-27 years with a mean age of 19.5 years (*sd* = 1.49). At the time of the experiment, all participants were enrolled in either Linguistics or Speech and Hearing Sciences units at Macquarie University and received course credit in an eligible unit for their time. The task took approximately 45 minutes in total.

The distribution of participants across conditions was as follows:

- Australian condition – 14
- New Zealand condition – 15
- Control condition – 14

Although 52 participants originally completed the matching task, nine were excluded. One participant was excluded as they were not born in Australia. One participant in the New Zealand condition stated after the experiment that she recognised the task and the purpose of the toy kiwis. Two participants claimed to recognise the speaker's voice. Two female participants whose ages were greater than two standard deviations above the mean were also excluded. As only three males volunteered for the study, male participants were also excluded as they were too small a sample for any meaningful comparison to be made between genders.

3.3 Stimuli and Materials

3.3.1 Target vowels and target words

Consistent with Hay et al. (2006) and Hay and Drager (2010), the present study focuses on the KIT, DRESS and TRAP vowels (Wells, 1982). Each of the three target vowels were presented in 10 different monosyllabic, monomorphemic words (see Table 3.1) represented by 30 target words in total. Target words were either a noun or an adjective. The preferred form for target words was /CVt/ however, due to lexical restrictions, five target words had a complex onset, /CCVt/ (*grit*, *skit*, *slat*, *Brett*, *threat*). This set of target words represents a more controlled set of stimuli than those used in the Hay et al. (2006) and Hay and Drager (2010) experiments, which used monosyllabic target words with multiple coda consonants including one complex coda. As discussed in Section 2.5.1, the word *fish* is a particularly canonical distinguisher of AusE and NZE speakers. Hay et al. (2006) found that when *fish* appeared as a target word, the priming effect was stronger. However, this effect was not observed in Hay and Drager (2010) and as *fish* violates the constraints of our target words, it was excluded. The target words used in Niedzielski (1999) also varied in number of syllables and coda identity. The use of a constant coda for all target words ensured target vowels were presented in the most consistent phonetic environment possible.

Table 3.1. Target words in KIT, DRESS and TRAP sets.

KIT	DRESS	TRAP
bit	bet	bat
fit	Brett	cat
grit	debt	chat
hit	jet	fat
kit	net	gnat
knit	pet	hat
mitt	set	Matt
pit	threat	rat
skit	vet	slat
wit	wet	vat

3.3.2 Sentences

Target words were each embedded in a unique carrier sentence. To facilitate target identification by participants, all carrier sentences contained the target word in phrase-final position. This represents a divergence from the stimuli used by Niedzielski (1999) which contained up to four target words per sentence. Phrase-final target word positioning also ensured participants would not be exposed to any additional vowels between the target vowel and continuum vowels. Unlike the stimuli used in Hay et al. (2006) and Hay and Drager (2010), sentences contained no other instances of the target vowels or NEAR/SQUARE vowels in stressed position. This minimised any additional priming effect by reducing overt identifiers of AusE when contrasted with NZE. Example sentences, with identified target words, are shown below (for a complete list, see Appendix A):

1. The new movie was a huge summer **hit**
2. She is studying to become a **vet**
3. She called her mum for a short **chat**

A 19-year-old male monolingual speaker of Standard AusE from Sydney was recorded reading the carrier sentences. The speaker and his parents were Australian-born and the speaker had completed all of his primary and high school education in Australia. At the time of recording, the speaker was an undergraduate student at Macquarie University. Sentences were recorded in a soundproofed room with an AKG C535 condenser microphone and a PreSonus StudioLive 16.4.2 digital mixer using Pro Tools 11.3.1 at a 48kHz sampling rate. The room contained no potential regional primes that might otherwise influence the speaker's vowel production beyond typical variation. The speaker's vowels were compared acoustically with mean vowel formant values provided by Cox, Palethorpe, Miles, and Davies (2014) and were determined to be a representative of Standard AusE by a highly experienced AusE phonetician.

3.3.3 Continua design

In general, continua design was modelled on that reported in Hay et al. (2006) and Hay and Drager (2010) in which continua represented six vowel tokens ranging from AusE-like (token 1) to exaggerated NZE-like (token 6). As the present experiment was conducted in the opposite context to Hay et al. (2006), and Hay and Drager (2010) (i.e. in Australia rather than New Zealand), the tokens in each continuum ranged from a NZE-like token 1 to an exaggerated AusE token 6. In all experiments, including the present, the synthesised version of the speaker's vowel was token 4.

While the structure of each continuum was consistent with those used in Hay et al. (2006) and Hay and Drager (2010), some important departures from their design were made. In the previous studies, continua tokens for all target words were based on single KIT, DRESS and TRAP vowel produced by the speaker in an isolated /hVd/ frame. These vowels were synthesised to create token 4, then first and second formant values (F1 and F2) were manipulated to create the additional five tokens. This meant that continua tokens were consistent across all target words but (as acknowledged in Hay et al. (2006)) token 4 would not align exactly with the speaker's vowel in each target word – "While the tokens are generally positioned near token 4 in the continua, some individual tokens are closer to tokens 3 or 5" (p. 9). We elected to create a unique continuum for each target word whereby the F1 and F2 values of token 4 matched those of the vowel produced by the speaker in each target sentence. This represents a slight variation in what the task requires of participants. Rather than matching a target vowel to a consistent set of continuum tokens, participants in the present study match the target vowel to continuum tokens that are based on the speaker's actual production of that vowel.

3.3.4 Vowel synthesis

The speaker's production of the target vowel from each sentence was analysed in Praat (Boersma & Weenink, 2016). First and second formant values (F1 and F2) were hand measured with values extracted at the vowel target using criteria from Cox (2006). These F1 and F2 values were used to define the parameters for a synthesised version of the target vowel, to be used as token 4 in each continuum, and the baseline from which each of the other five continuum tokens was derived. A total of six synthesised vowel tokens made up each continuum.

Step intervals between tokens were calculated to ensure token 1 would be plausibly NZE-like. We determined that the F1 and F2 values for token 1 should be within 2 standard deviations of mean formant values for the relevant NZE vowel published in Easton and Bauer (2000), and comparable with values taken from vowel plots in Watson et al. (2000). It should be noted, however, that some F1 and F2 values did violate these constraints. For example, the coarticulatory effects on the post-rhotic KIT vowel in *grit* predictably resulted in a much lower F2. This then affected the F2 value for all tokens in the *grit* continuum resulting in an F2 for token 1 that violated the two standard deviation range. In addition to F2 for *grit*, this violation also occurred with F1 values for *bit*, *grit*, *knit* and *debt*. However, these violations were deemed acceptable on the basis that the original vowels represented appropriate values given any coarticulatory effects, and the vowel steps maintained a consistent perceptual distance between tokens. Further, any adjustments to the step intervals resulted in additional violations elsewhere. Mean AusE values provided by Cox et al. (2014) were used as a reference for tokens 4, 5 and 6, however, values for these tokens were ultimately determined by the values for token 4 and the subsequent two step intervals.

Step intervals between tokens were calculated using the bark scale, a psychoacoustical scale that divides the frequency range perceptible to humans into critical bands representing perceptually equal distances (Zwicker, 1961). The extracted F1 and F2 values were converted into bark values in order to define the parameters of each continuum. Step intervals between tokens in each continuum were equal within each vowel context but differed between continua for KIT, DRESS and TRAP vowels.

Following the criteria discussed above (i.e. that the F1/F2 values for token 1, would remain within 2 standard deviations from the mean of published NZE data), for KIT vowels, F1 intervals were 0.35 bark steps and F2 intervals 0.6 bark steps. For DRESS the intervals were 0.8 (F1) and 0.3 (F2) and TRAP, 0.5 (F1) and 0.5 (F2). Step intervals are different for F1 and F2 and vary across the vowels KIT, DRESS and TRAP because dialectal differences between these vowels are not uniform. For example, AusE and NZE DRESS vowels are much closer together in an F1/F2 plane than AusE and NZE KIT. Bark values for each token were converted back to Hz to give formant values for each continuum. Table 3.2 shows example formant values of the continuum steps for *hit*, *vet* and *bat* and Figures 3.1 – 3.3 show spectrograms of the continuum tokens for the same vowels.

Table 3.2. Example continuum formant values. Token 1 represents NZE values, token 4 represents the synthesised speaker's token and token 6 represents exaggerated AusE values.

Continuum Tokens - hit		
Token	**F1**	**F2**
1	484	1627
2	445	1781
3	407	1948
4	370	2130
5	334	2330
6	300	2551

Continuum Tokens - vet		
Token	**F1**	**F2**
1	380	2078
2	466	1987
3	560	1900
4	660	1817
5	769	1737
6	887	1661

Continuum Tokens - bat		
Token	**F1**	**F2**
1	640	1997
2	706	1853
3	775	1719
4	848	1595
5	925	1479
6	1007	1370

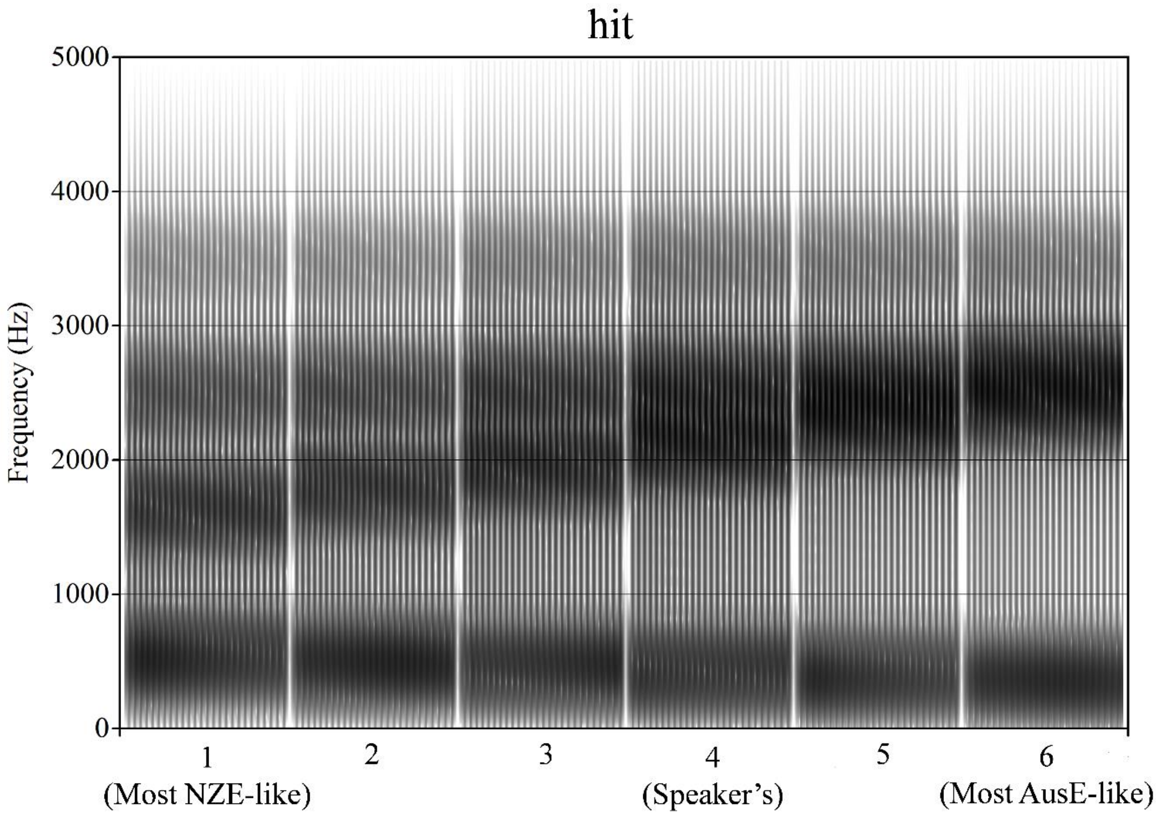


Figure 3.2. Spectrograms of continuum tokens – hit.

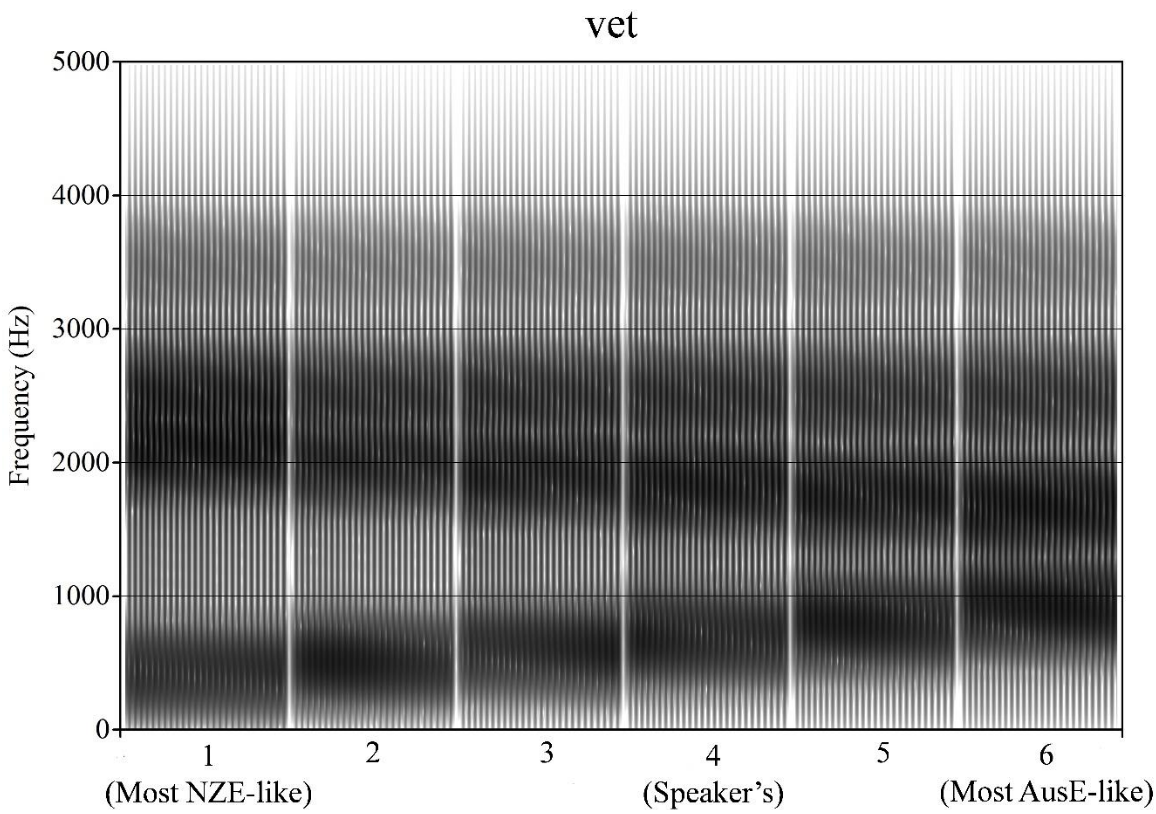


Figure 3.3. Spectrograms of continuum tokens – vet.

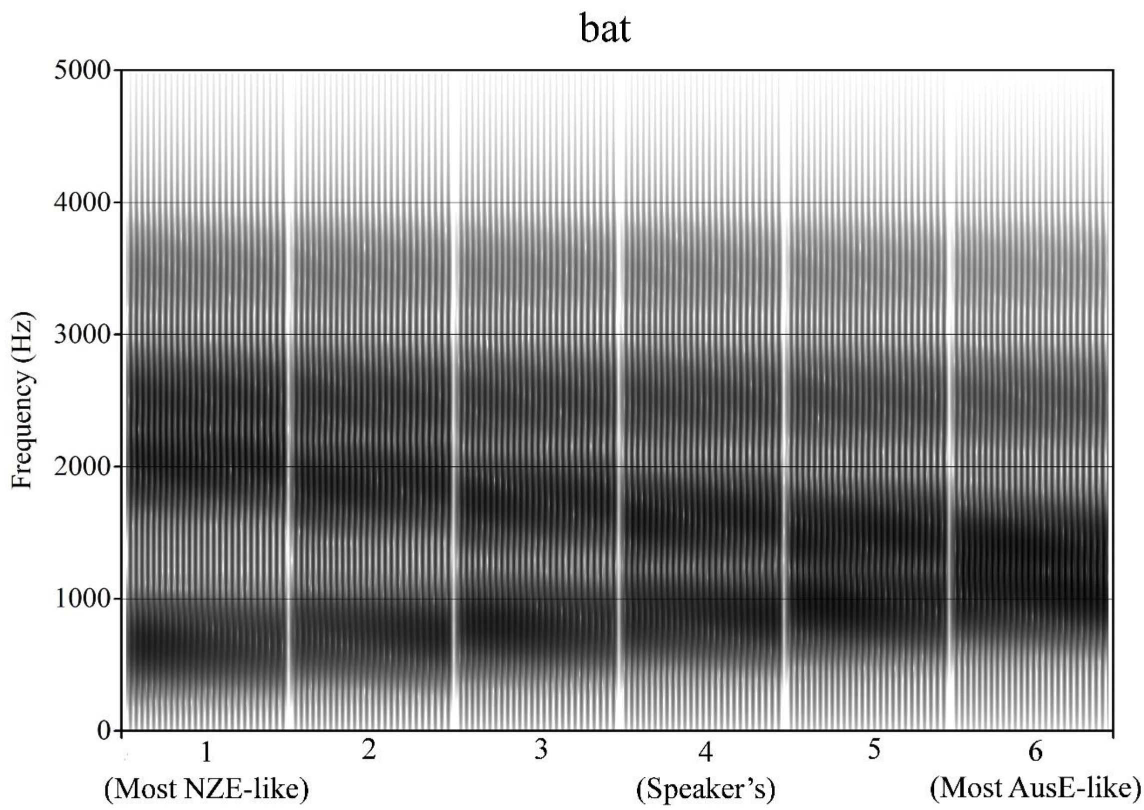


Figure 4.3. Spectrograms of continuum tokens – bat.

All synthesised tokens were 180ms. The task was piloted with token lengths consistent with that of the speaker's original vowel, as well as 180ms and 250ms. Pilot participants indicated that 180ms was the most suitable length for the task. The vowels that were synthesised with the same duration as the vowel in the target word were judged to be too short, making the task extremely difficult. The 250ms tokens were said to be unnecessarily long. F0 at the vowel onset was 160Hz. This was the mean initial F0 for the speaker's 30 tokens. An F0 slope of -1 octaves per second was also applied to all synthesised vowels. The F0 slope was applied following feedback from pilot participants that suggested the change in pitch resulted in a more 'natural' sounding vowel, making comparison to the recorded speech easier. Vowels were synthesised using the vowel editor function in Praat (Boersma & Weenink, 2016) which generates a vowel from inputted formant values as well as F0 and duration parameters. Consistent with Niedzielski (1999), Hay et al. (2006), Hay and Drager (2010), and Lawrence (2015), F3 was not manipulated in any way.

3.4 Procedure

3.4.1 Priming

Participants completed the perception task in one of three conditions: Australian, New Zealand and a control. Priming for the conditions was cued by the presence of stuffed toy koalas (Australian) or kiwis (New Zealand) with the control group exposed to no toys. Primes consisted of two koalas and two kiwis of approximately equivalent combined size.

Figure 3.4. Stuffed toy koalas and kiwis used in the experiment.

To introduce the prime, the experimenter went to a drawer in the room where the experiment was taking place to find the headphones required for the task. The headphones were 'found' under the stuffed toys. The experimenter mentioned the toys were being used in another experiment and placed them on the table, in the participant's line of sight, where they remained for the duration of the experiment. The toys were placed at approximately arm's length from the participant, alongside the laptop being used for the task. The intention was to draw attention to the prime, ensuring it had been seen, without making it obvious that the toy was related in any way to the experiment. This was similar to the procedure used in Hay and Drager (2010) where the experimenter 'found' the toys in a cupboard containing response sheets.

None of the previous iterations of the experiment had used a control, however, it was deemed important for a group to complete the perception task in an un-primed context. This would allow us to determine if the New Zealand and Australian primes were both producing an effect and, if so, which effect was the strongest. For example, there is no way of knowing whether the New Zealanders participating in New Zealand condition of the Hay et al. (2006)

experiment were influenced in any way by the presence of the label *New Zealand*, or if the divergence shown in responses to KIT and TRAP vowels was a result of participants in the Australian condition alone. Further, a control group would also allow us to determine how accurate participants were at the task without having to account for a priming effect. This would provide a more definitive baseline of responses, revealing any perceptual biases or idiosyncrasies in token selection that may influence results beyond the prime.

Each participant completed the task individually and in the same sound attenuated room. The room layout was identical for all participants and contained no other potential regional primes. Thus, the only difference between the conditions was the presence of the prime itself (or lack of prime in the control condition). All participants interacted with the same experimenter – a 29-year-old male, born in New Zealand who had moved to Australia at age 10 – therefore any potential priming effect from the experimenter was consistent for all participants. The experimenter was the author of the present thesis and therefore was aware of the hypotheses being tested.

3.4.2 Perception task

Consistent with Hay et al. (2006) and Hay and Drager (2010), participants were told the purpose of the experiment was to determine the accuracy of synthesised vowel sounds. Each participant was told they would be hearing a sentence of recorded human speech with an identified target word and then some synthesised vowel sounds. Participants were instructed to select the vowel sound they believed was the closest match to the vowel in the target word.

The perception task consisted of a familiarisation phase and an experimental phase. Prior to the experimental phase each participant was given three practice questions to familiarise them with the task. Practice questions presented in the same way as in the experimental phase but used a different recorded speaker. To reduce any potential priming effect from another voice, practice questions used recorded speech of the main experimenter referred to in Section 3.4.1.

Practice questions used non-target vowels: AusE /ɜː/ and /ɐ/ and contained no instances of KIT, DRESS or TRAP vowels. This would ensure participants had no exposure to either the speaker or recorded examples of the main target vowels prior to the experimental phase. The first practice question was used to guide participants through the task step-by-step. Participants were informed that they would hear a sentence, followed by six isolated synthesised vowel sounds, and then a selection screen would appear. Presentation of a further two practice questions was identical to that of the main task (outlined below in Figure 3.4). No feedback was given for the practice questions although participants were able to ask the experimenter, who was present and visible but in an adjacent room, for clarification on any part of the task. Participants were not provided with any information about the speaker in either the familiarisation or experimental phase.

In the experimental phase, participants were again informed on screen that for each question they would hear a sentence once, followed by the six isolated synthesised vowel sounds and then a selection screen would appear. Participants heard each sentence while it was simultaneously presented as text on-screen. In the written form, the phrase-final word was bold and underlined, identifying it as the target. The sentence remained visible for the duration of the recorded sentence plus an additional 2000ms. Immediately following, each 180ms continuum token was played with its corresponding number visible. The number remained on screen for an additional 820ms. Thus the entire continuum was 6000ms (6 tokens x 1000ms). Stimuli timing was piloted and feedback indicated the presentation of stimuli was suitable for the task. The experimental procedure is illustrated in Figure 3.4.

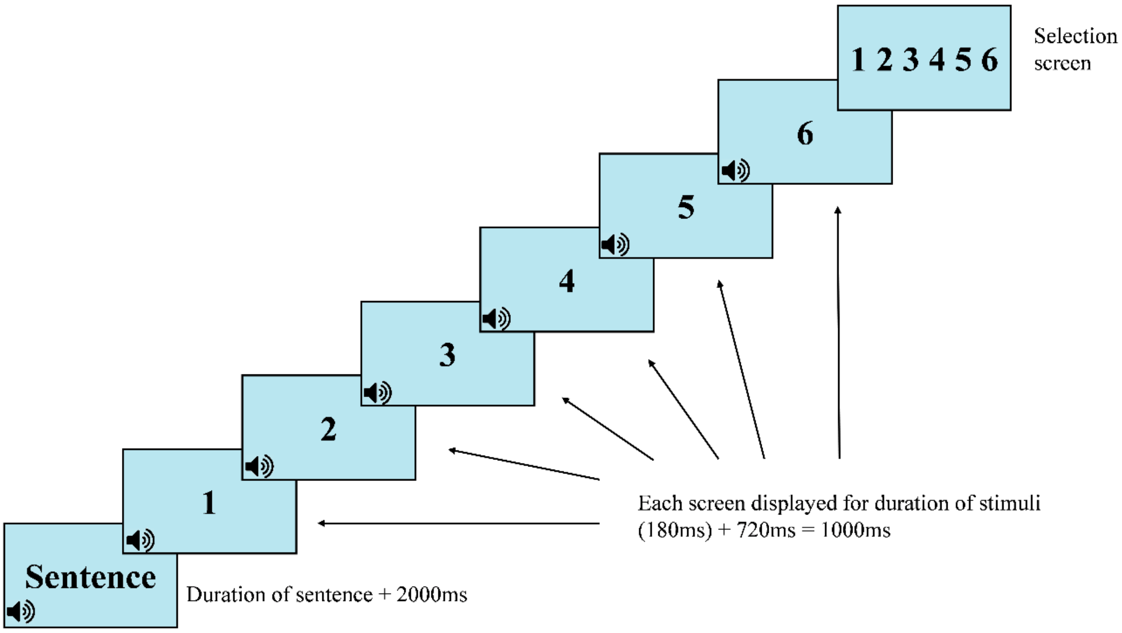


Figure 3.5. Illustration of experimental procedure.

Participants made their selection by pressing the corresponding number key on a keyboard upon the presentation of a selection screen, which displayed the numerals 1-6. Participants were not able to make a selection until after all six continuum tokens had been played and the selection screen was displayed. Any key press by the participant prior to the selection screen had no effect on the experiment. Reaction time was measured from the appearance of the selection screen (which signified when participants could make a selection), however, participants were not explicitly given any instructions regarding selection speed. Once a selection was made, the task continued with the next sentence. Presentation of the synthesised tokens with a visual numeric label differs slightly from Hay et al. (2006) and Hay and Drager (2010) who used spoken numbers preceding each token to label the steps. Visual labelling was preferred as it reduced any additional priming effect from the token labels. This was particularly important for label *six*, which contains a KIT vowel and pronunciation of the word is regarded as a highly salient identifier of NZE when compared with AusE. In Lawrence (2015), continuum tokens were represented by dots, rather than numbers. However, it is likely that this would complicate token identification.

Sentences were presented in two blocks. Each block contained all 30 stimuli sentences, therefore participants heard each target word once per block and twice in total. Sentences were presented once with the continuum tokens played in order from token 1 to token 6 (NZE-like to exaggerated AusE), and once with the token order reversed (exaggerated AusE to NZE-like). The reverse-order continua were still presented to the participant labelled from 1-6. In other words, items labelled 1 in the original order were the most NZE-like and those labelled 1 in the reverse order were the most AusE-like (exaggerated) tokens. This was intended to discourage any potential selection patterning by varying the position of the synthesised version of the speaker's vowel, as well as highly centralized or peripheral variants. In each block, half of the sentences were presented with their continua in the original order and half with their continua in reversed order. Those sentences presented with their continuum in the original order in the first block would be presented with their continuum in reverse-order in the second block, and vice versa. Although the two blocks were identical for all participants, presentation order of sentences was randomised. Full lists of continuum order are presented in Appendix B. Figure 3.5 illustrates the continuum for *hit* in the original presentation order and Figure 3.6 illustrates the continuum for *hit* when presented in the reversed order.

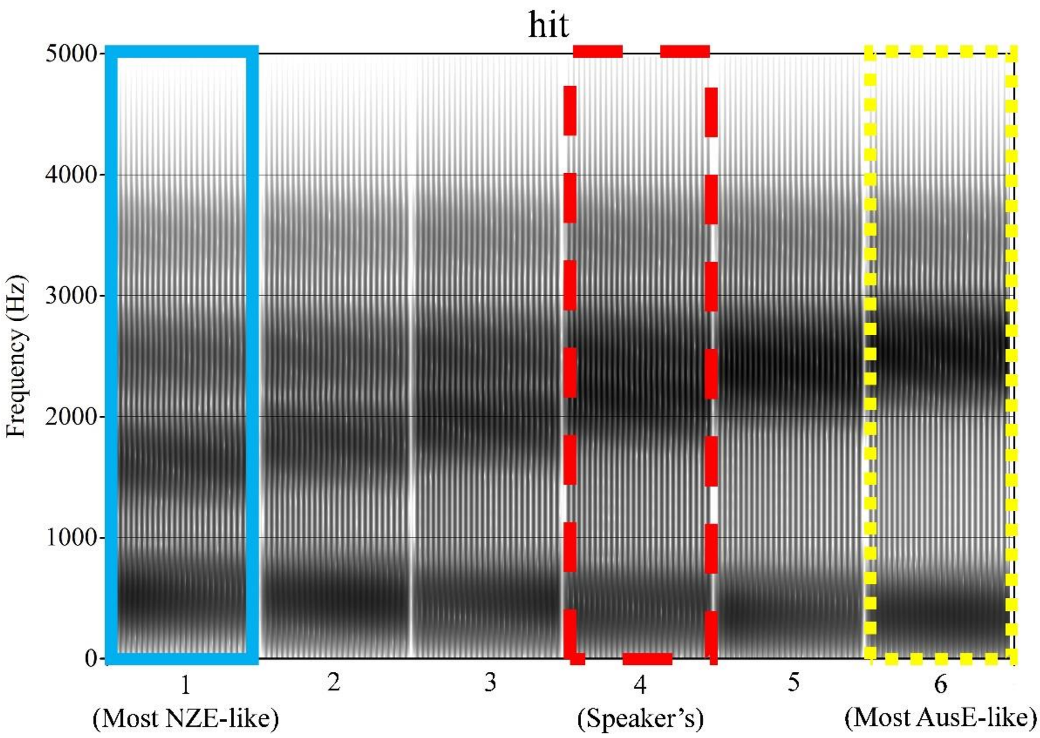


Figure 3.6. Spectrograms for continuum tokens for hit in original order. Blue solid box indicates the most NZE-like vowel. Red dashed box indicates the synthesised speaker's vowel. Yellow dotted box represents the most AusE-like (or exaggerated-AusE) vowel.

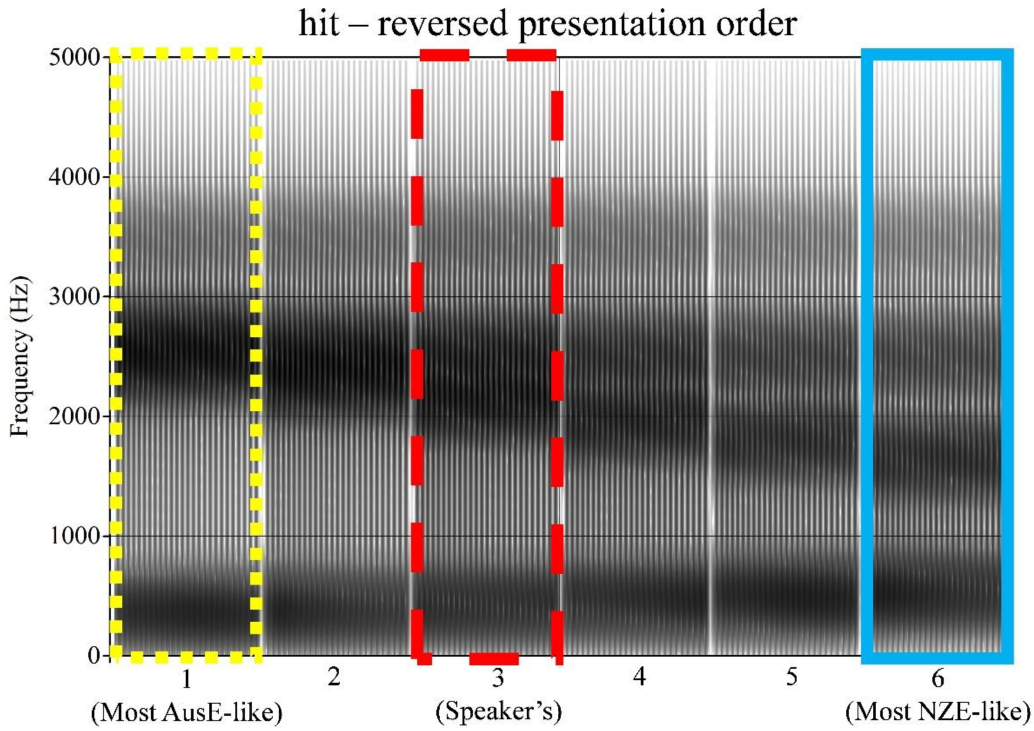


Figure 3.7. Spectrograms for continuum tokens for hit in reversed order. Yellow dotted box represents the most AusE-like (or exaggerated-AusE) vowel. Red dashed box indicates the synthesised speaker's vowel. Blue solid box indicates the most NZE-like vowel.

The perception task was presented on a Sony Vaio laptop using the E-Prime 2.0 software (Psychology Software Tools, 2012). All participants used the same pair of Sennheiser HD 461i closed over-ear headphones and were able to adjust volume to a comfortable level.

Following the perception task, participants completed two questionnaires, both modelled on those used in Hay et al. (2006) and Sanchez et al. (2015). The first concerned the participant's impressions of the speaker's age, nationality, occupation and education level. The second was designed to assess participant's level of exposure to New Zealand and NZE. Responses to this questionnaire were used to identify participants who had frequent exposure to NZE. Questionnaires are presented in Appendix C.

3.5 Data analysis

A total of 2580 token selections were recorded. This total represents 60 selections from each of the 43 participants. While participants were not given any explicit instruction regarding response time, we elected to exclude any selection that was three standard deviations longer than the mean response time for each participant. This was deemed necessary due to a small number of selections that took an unreasonably long time. In total, five selections, all having a reaction time of 9798ms or longer, were excluded. Reaction time for all remaining selections was 8264ms or less. This resulted in a total data set of 2575 selections to be used in analysis. Table 3.3 outlines the distribution of selections according to vowel and condition.

Table 3.3. Total selections recorded for KIT, DRESS and TRAP vowels across all conditions.

	KIT	DRESS	TRAP	Total
Australia	279	279	280	838
New Zealand	299	300	298	897
Control	280	280	280	840
Total	858	859	858	2575

3.5.1 Data coding

Token selections for all participants, including reaction time, were extracted from E-prime. Individual selections were represented by a number from 1-6, corresponding to the actual token selected by the participant. For those trials presented in reverse order, selections were re-coded (i.e. 1 => 6, 2 => 5, etc.). This meant that coding of selections was consistent according to the acoustic features of the selected token (i.e. F1 and F2), rather than its numeric label. Each selection was also coded according to whether it was in response to a continuum presented in the original order or reversed order.

Relevant demographic information about each participant was also collated. For the purposes of analysis, age and gender were not included in any models as all participants were female and aged 18-27.

Socioeconomic Index scores (SEI) for each participant were calculated from the Socioeconomic Index for Australia (McMillan, Beavis, & Jones, 2009). Due to participants' age, SEI scores represent an average of the participants' parents SEI score, which is a representation of either their occupation or highest level of education on a 0-100 scale. Both parents' occupations were used. Where no occupation was provided or the parent was not actively employed, highest level of education was used, as recommended in McMillan et al. (2009). The range of SEI scores for the participants in the present study was 29.9 to 90.75. Mean SEI was 61.83 with a standard deviation of 17.60.

Responses to the second questionnaire were used to create a NZE Exposure Score (NZES) intended to weight a participant's likely exposure to, or familiarity with, NZE. This score was custom designed for the present study. A maximum score of 13 points was possible. Points were allocated as follows:

Participant had been to NZ	2pts
For longer than one month	2pts
Since the beginning of 2015	2pts
Participant personally acquainted with New Zealanders	2pts
Participant came into personal contact with:	
1-5 New Zealanders per week	2pts
6-10 New Zealanders per week	3pts
More than 10 New Zealanders per week	4pts
Participant could name some NZ media	1pt

3.5.2 Statistical analysis

To test the influence of the priming condition on token selection we used a linear mixed effects model (Baayen, 2008). Token selection was the dependent variable. Random effects were speaker and word. Fixed effects included experimental condition (Australia, New Zealand and control), participant SEI, participant NZES and continuum presentation order (original and reversed). Collinearity was observed between the SEI and NZES. Although SEI was found to be a significant effect in Hay et al. (2006) and Hay and Drager (2010) it was suggested that SEI was a predictor of New Zealanders experience with AusE. As we created a metric to score our participants' exposure to NZE – the NZES – we elected to remove SEI from the model. In addition, it was necessary to include the NZES as our third hypothesis was that exposure to NZE would be a significant effect.

The model was fit to the data using the lme4 library (Bates, Meaechler, & Bolker, 2011) in R (R Core Team, 2016). Significance was calculated using Satterthwaite's (1946) approximations to degrees of freedom using the lmerTest library (Kuznetsova, Brockhoff, & Christensen, 2013).

We also fit a linear fixed effects regression model to our data, similar to the method reported in Hay et al. (2006) and Hay and Drager (2010). This allowed us to compare our results more directly with previous analyses. Predictor variables were experimental condition (Australia, New Zealand and control), participant NZES and continuum presentation order (original and reversed). This model does not treat subject and word as random effects.

4. Results

4.1 Introduction

This chapter presents the results from the matching task and post-task questionnaires. First, in Section 4.2 we will present descriptive results from the raw data. Responses to the post-task questionnaire will be provided in Section 4.3. Following, the results of the linear mixed effects model will be reported in Section 4.4. Statistical analysis allowed us to examine whether any observed trends in token selection were significant and establish whether there was any support for the hypotheses.

The hypotheses being tested were:

H1. The stuffed toys will influence participants' selections in the vowel matching task. Those AusE-speaking participants exposed to the New Zealand prime (stuffed toy kiwis) should select continuum tokens that represent more NZE-like vowels than those exposed to the Australian prime (stuffed toy koalas).

H2.Responses to KIT vowels will exhibit the most variability between the Australian and New Zealand conditions. However,

H3.The priming effect will be limited to those participants who indicate some level of familiarity to New Zealand and NZE.

4.2 Descriptive results

Table 4.1 shows mean token selection from all participants for KIT, DRESS and TRAP vowels in the Australian and New Zealand conditions only. These represent the two priming conditions which were the primary focus of this study. Table values represent mean token selections. A value of 1 represents the most NZE-like vowel, 4 represents the vowel synthesised from the F1 and F2 values of the speaker's actual vowel production, and 6 represents the most exaggerated AusE vowel. Although tokens 5 and 6 represent exaggerated AusE vowels, for clarity, higher token numbers will be referred to as more AusE-like and lower token numbers as more NZE-like for the remainder of this analysis.

Table 4.1. Mean token selection for target vowels, Australian and New Zealand conditions.

Condition	**KIT**	**DRESS**	**TRAP**
Australia	5.21	3.15	3.39
New Zealand	5.25	3.09	3.51

Overall, selections do not appear to reflect a shift in perception consistent with the priming condition. It was predicted that participants in the New Zealand condition would select more NZE-like tokens, which would be represented by lower numbered tokens, than those in the Australian condition. However, the predicted effect only occurred for DRESS vowels. For both KIT and TRAP vowels, participants in the New Zealand priming condition selected more AusE-like tokens than those in the Australian condition. It was also predicted that, consistent with Hay et al. (2006) and Hay and Drager (2010), any observed shift would be most likely to occur for target words containing a KIT vowel. This effect is not evident in our data.

As previous experiments did not use a control, we had no expectations regarding whether token sections in the Australian condition should be different to a control (i.e. more AusE-like) or consistent with a control (i.e. the Australian priming condition having no effect on Australian participants). However, we could assume that selections in the control condition represent a baseline of responses made in an un-primed context. Thus those in the New Zealand condition should have selected lower numbered tokens than those in the control and those in the Australian condition may have selected higher numbered tokens. This prediction only holds for responses to KIT and DRESS vowels in the Australian condition where participants selected, on average, more AusE-like tokens than in the control. Mean token selections for all vowels in the New Zealand condition represent more AusE-like vowels than selections in the control condition. Table 4.2 outlines the mean token selections for all vowel contexts in each of the three conditions, as well as the mean token selection across each condition and vowel. Figure 4.1 illustrates how the mean token selections in each of the three conditions compare within each vowel context.

Table 4.2. Mean token selection – all conditions.

Condition	KIT	DRESS	TRAP	Total
Australia	5.21	3.15	3.39	3.91
New Zealand	5.25	3.09	3.51	3.95
Control	5.00	3.06	3.41	3.82
Total	5.15	3.10	3.44	3.90

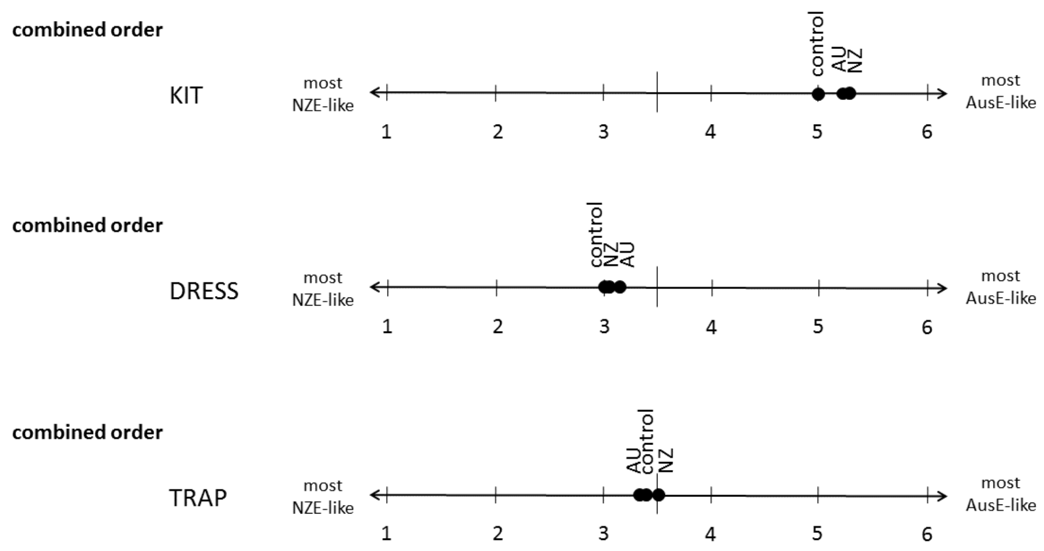


Figure 4.5. Mean token selections in each condition for KIT, DRESS and TRAP vowels.

On average, for each of the three target vowels, participants selected tokens that represent more phonetically raised and fronted vowels than the speaker's synthesised vowels. This explains why the mean token selection for KIT (5.15) was so different to the mean selections for DRESS (3.10) and TRAP (3.44). For KIT vowels, higher token numbers represent raised and fronted variants (more AusE-like) however, for the DRESS and TRAP vowels, lower token numbers represent raised and fronted variants (more NZE-like). Figure 4.2 provides a visual representation of the major trends in our data. This shows that token selections for KIT vowels are higher overall than for both DRESS and TRAP as well as showing that token selections made when the continuum was presented in reversed order are generally more AusE-like.

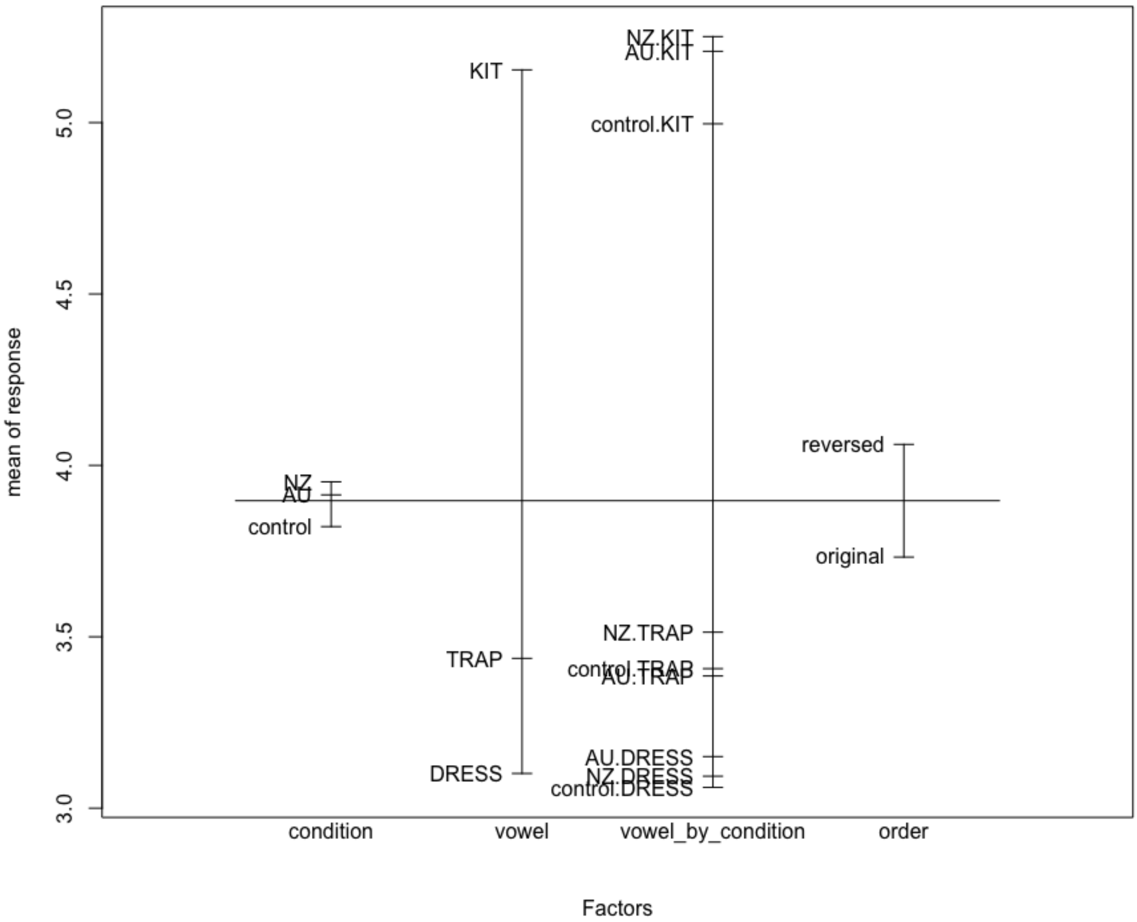


Figure 4.2. Major data trends. Vertical lines illustrate the levels for each factor. Grand mean for the data set is represented by the horizontal line. Vertical axis represents mean token selection.

4.2.1 Token selection

In no vowel context was the synthesised version of the speaker's recorded vowel (token 4) the statistical mode. Token 4 was the second most frequently selected token for DRESS (27.12%) and TRAP (31.00%) vowels, although for KIT vowels, token 4 was only selected 14.22% of the time across the three conditions. Overall, 72.78% of token selections were either token 3 (30.41%), 4 (24.12%), or 5 (18.25%) indicating that participants did typically select tokens that were acoustically similar to the speaker's vowel. Across all trials, there were only 42 selections (1.63%) of token 1, the most NZE-like token, which suggests this was not considered a plausible candidate for any of the three vowels. Similarly, for DRESS and TRAP vowels, token 6 received only 1.86% of selections. However, token 6 was considered a plausible candidate by participants for KIT vowels, receiving 39.74% of selections across the three conditions. Tables 4.3 and 4.4 outline the distribution of token selections for each vowel in the three conditions.

Table 4.3. Distribution of token selections by vowel and condition – selection totals.

	1	**2**	**3**	**4**	**5**	**6**	**Total**
KIT AU	0	0	8	44	109	118	279
KIT NZ	1	4	2	27	143	122	299
KIT C	0	5	18	51	105	101	280
DRESS AU	7	41	145	75	11	0	279
DRESS NZ	8	51	151	87	1	2	300
DRESS C	5	65	129	71	9	1	280
TRAP AU	6	45	108	85	28	8	280
TRAP NZ	7	40	106	94	40	11	298
TRAP C	8	35	116	87	24	10	280
Total	42	286	783	621	470	373	2575

Table 4.4. Distribution of token selections by vowel and condition – percentage.

	1	**2**	**3**	**4**	**5**	**6**	**Total**
KIT AU	0.00	0.00	2.87	15.77	39.07	42.29	100
KIT NZ	0.33	1.34	0.67	9.03	47.83	40.80	100
KIT C	0.00	1.79	6.43	18.21	37.50	36.07	100
DRESS AU	2.51	14.70	51.97	26.88	3.94	0.00	100
DRESS NZ	2.67	17.00	50.33	29.00	0.33	0.67	100
DRESS C	1.79	23.21	46.07	25.36	3.21	0.36	100
TRAP AU	2.14	16.07	38.57	30.36	10.00	2.86	100
TRAP NZ	2.35	13.42	35.57	31.54	13.42	3.69	100
TRAP C	2.86	12.50	41.43	31.07	8.57	3.57	100
Total	1.63	11.11	30.41	24.12	18.25	14.49	100

4.2.2 KIT vowels

For KIT vowels, selections favoured the two exaggerated AusE tokens, 5 and 6. These two tokens represent 73.57% of selections in the control condition, 81.36% in the Australian condition and 88.63% in the New Zealand condition. The more NZE-like tokens, 1 and 2, were only selected 10 times across all three conditions, representing 1.17% of selections.

Mode token selection did shift between conditions (token 6 in the Australian condition and token 5 in the New Zealand and control conditions). This suggests a slight shift towards more exaggerated AusE token selections in the Australian condition. Overall selections for KIT vowels are shown in Figure 4.3.

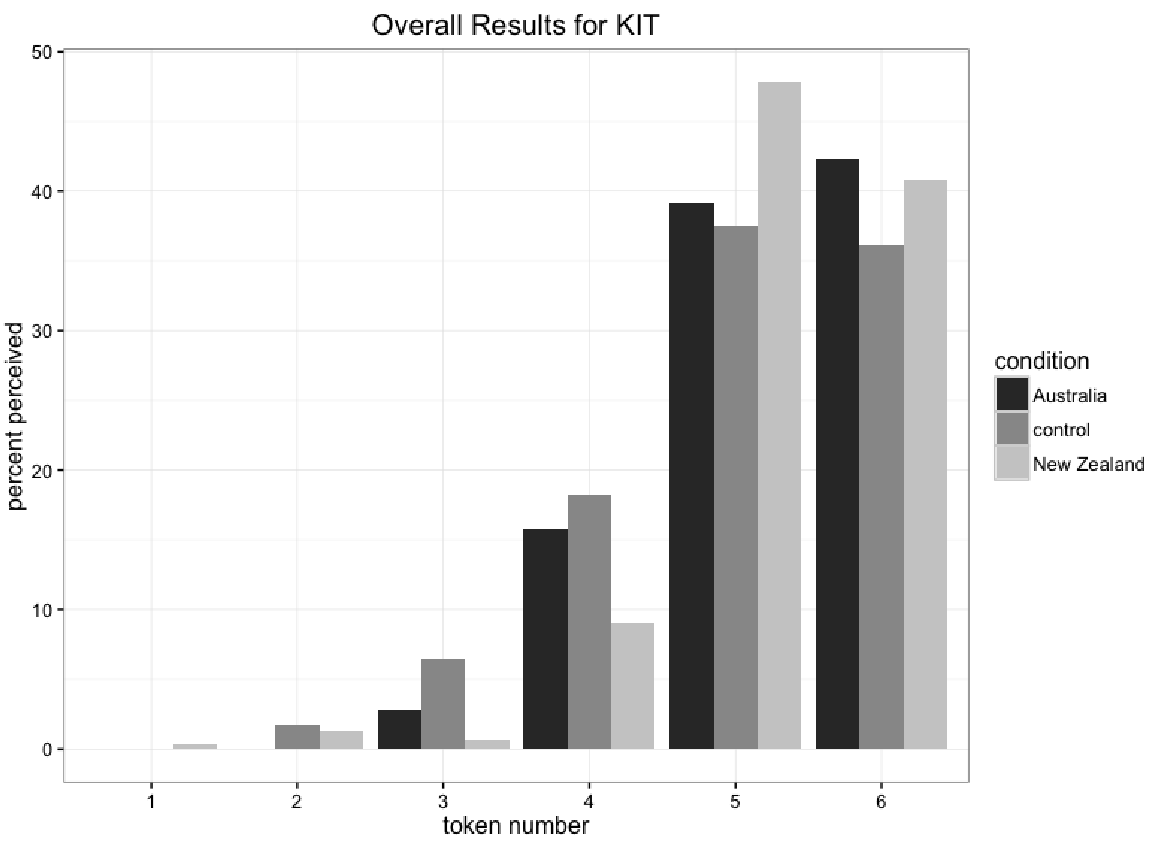


Figure 4.3. Token selections for KIT vowels – all conditions. Lower numbers represent more NZE-like tokens.

4.2.3 DRESS vowels

For DRESS vowels, selections favoured tokens 2, 3 and 4. Unlike selections made for KIT vowels, which were almost exclusively the exaggerated AusE tokens 5 and 6, for DRESS vowels, only 2.79% of selections were token 5 or 6.

The mode selection for DRESS vowels in all three conditions was token 3. Participants in the New Zealand condition did select slightly more NZE-like tokens than those in the Australian condition. However, participants in the control condition selected even more NZE-like tokens than those in the New Zealand condition. Overall selections for DRESS vowels are shown in Figure 4.4.

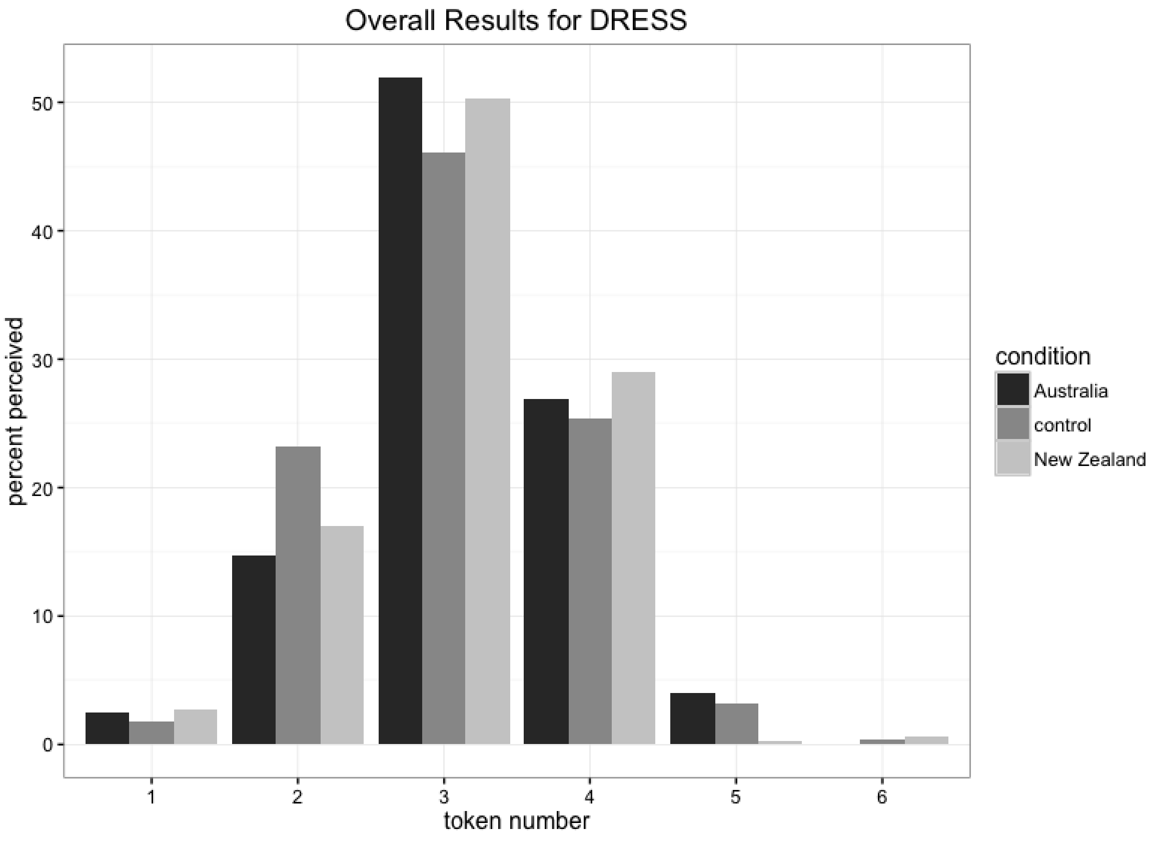


Figure 4.4. Token selections for DRESS vowels – all conditions. Lower numbers represent more NZE-like tokens.

4.2.4 TRAP vowels

Similar to DRESS vowels, selections for TRAP vowels favoured tokens 2, 3 and 4. The mode selection for TRAP vowels in all three conditions was token 3. Participants in the New Zealand condition selected slightly more AusE-like tokens than those in the Australian condition.

On average, across the three conditions, participants were also the most likely to select the synthesised variant of the speaker's token for TRAP vowels (31.00%). For KIT, this figure is 14.22%, and for DRESS, 27.12%. Overall selections for TRAP vowels are shown in Figure 4.5.

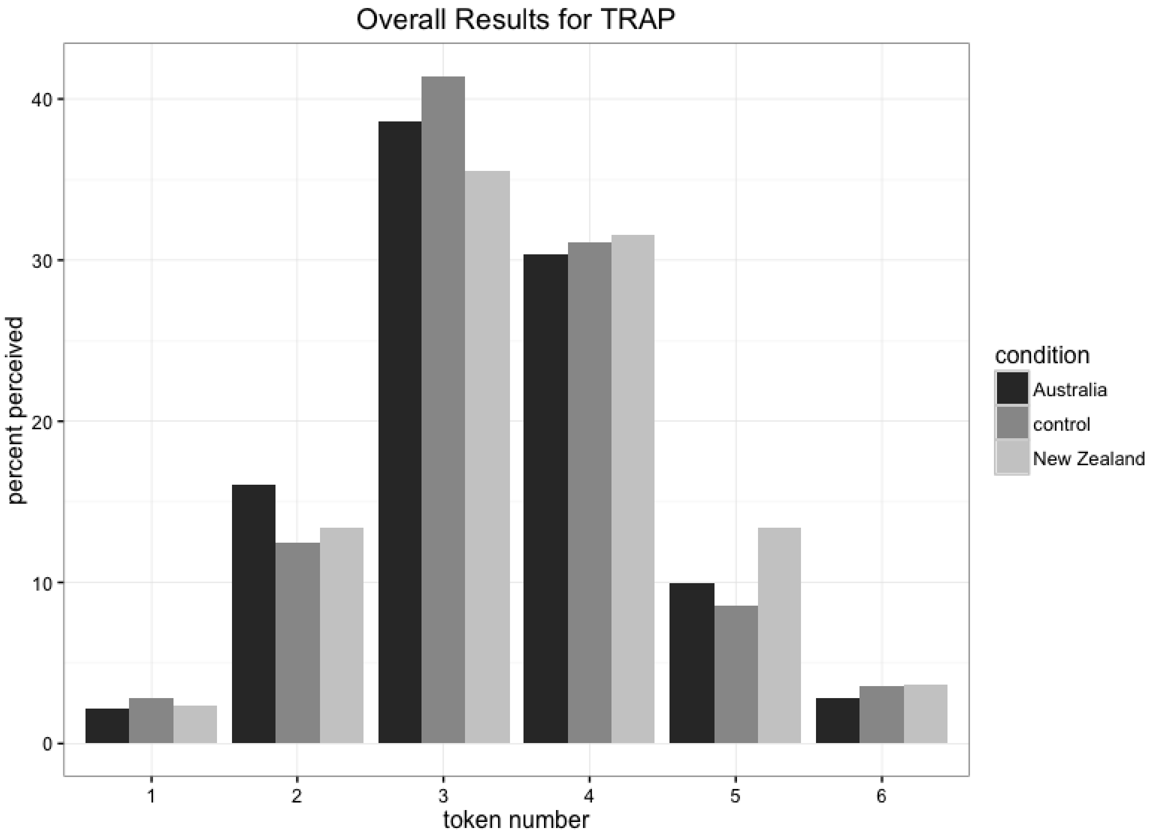


Figure 4.5. Token selections for TRAP vowels – all conditions. Lower numbers represent more NZE-like tokens.

4.2.5 Summary

These descriptive results indicate that the predicted shift in token selection did not occur between priming conditions for KIT and TRAP vowels. Participants in the New Zealand condition tended to select more AusE-like tokens than those in the Australian condition. For

DRESS vowels, it was found that participants in the New Zealand condition selected more NZE-like vowels. However, for all three vowel contexts, participants in the control condition selected more NZE-like tokens than those in the New Zealand condition.

4.3 Post-task questionnaires

In the post-task questionnaires (see Appendix C), participants were required to answer forced-choice questions relating to the speaker's gender, age, education level, and a free-response question about the speaker's nationality. All listeners were able to correctly identify the speaker as male. Only half (51.16%) of participants were able to correctly judge which age range the speaker belonged to (the speaker was 19). Participants were generally correct when asked about the speaker's highest level of education, with 71.11% selecting 'university degree'. Responses to the speaker age and education questions are outlined in Tables 4.5 and 4.6.

Table 4.5. Responses to forced-choice question regarding speaker's age – the speaker was 19.

Age range	Number selected	% Selected
18-25	22	51.16
26-35	17	39.53
36-45	4	9.30

Table 4.6. Responses to forced-choice question regarding speaker's highest level of education.

Education	Number selected	% Selected
Primary School	0	0.00
High School	1	2.32
Trade Certificate/TAFE	1	2.32
College Diploma	4	9.30
University Degree	32	71.11
Postgraduate Degree	5	11.63

4.3.1 Speaker origin

Some variation in participant response was observed in the free-choice question requiring participants to state where they believed the speaker was from. Participants were less likely to identify the speaker as 'Australian' in the New Zealand condition (73.33%) than in the Australian (92.86%) or control conditions (85.71%). However, as only one participant in the New Zealand condition responded New Zealand, it is unlikely that the presence of the kiwis had a direct influence on participants' belief about the speaker's nationality. In addition, one participant in the control condition also responded 'New Zealand'. It is possible that the kiwis lead participants to guess the speaker was not Australian, however, given the relatively small sample in each condition (n = 14-15) it would seem more likely that this is normal variance or uncertainty. Responses to speaker nationality are presented in Table 4.7.

Table 4.7. Nationality attributed to the speaker by participants.

Condition	**Response**				
	Australia	**New Zealand**	**Other**	**Total**	**% Australia**
Australia	13		1	14	92.86
New Zealand	11	1	3	15	73.33
Control	12	1	1	14	85.71

4.3.2 New Zealand Exposure Score

Responses to the post-task questionnaires were also used to assess participant's likely level of exposure to NZE. Overall, 30.23% of participants had been to New Zealand, and all had done so in the last ten years. 30.23% of participants answered that they were personally acquainted with New Zealanders (not all these participants had been to New Zealand). Almost half of the participants responded that they spoke with New Zealanders on a weekly basis, 44.19% spoke with 1-5, and 4.65% spoke with ten or more. However, 51.16% responded that they speak with no New Zealanders on a weekly basis. Only 16.28% of respondents could name any New Zealand media they had recently seen or heard.

Using the metric described in Section 3.4.1, each participant was given a New Zealand Exposure Score (NZES) between 0 and 13. Participants scored from 0 to 12 points. The mean score across all participants was 2.71 (*sd* = 3.01). Distribution of participant scores is outlined in Figure 4.6.

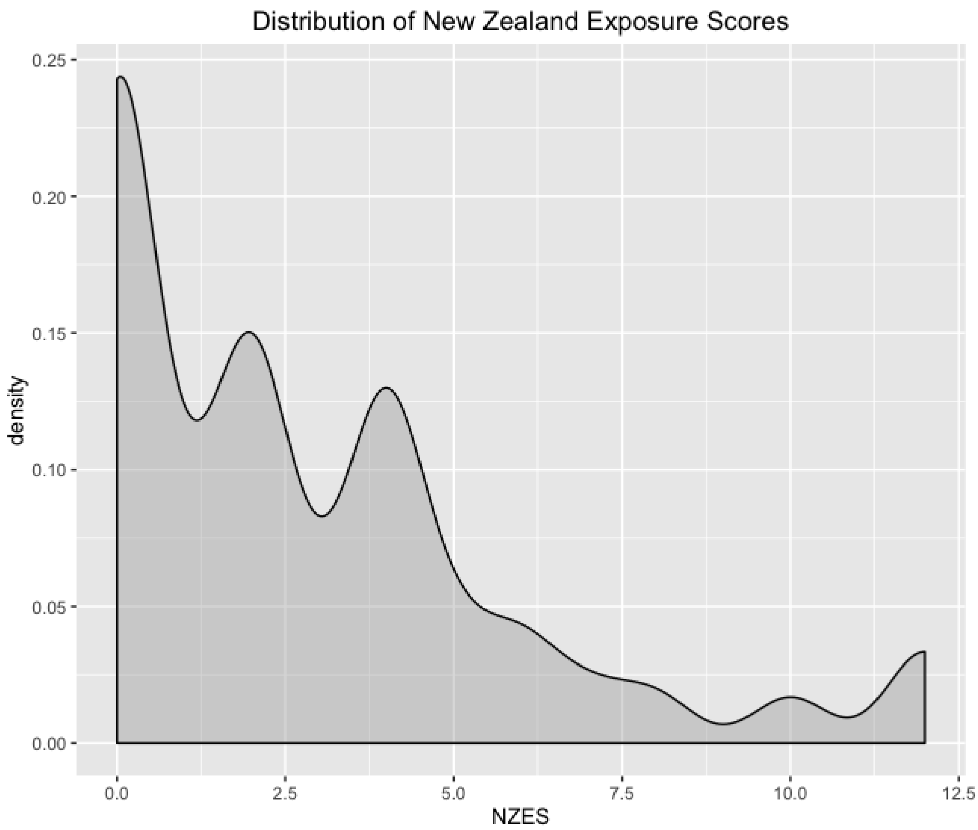


Figure 4.6. Distribution of New Zealand Exposure Scores.

4.4 Statistical Analysis

Model outputs from the linear mixed effects analysis are presented below in Tables 4.8-4.13. Regarding the first and second hypotheses – that token selection would differ between the Australian and New Zealand conditions and this would be most evident with KIT vowels – no main effect of condition was found. There was no significant difference between selections in the Australian and New Zealand conditions for any of the three target vowels (KIT, DRESS and TRAP). Further, no significant difference was found between token selections in the control condition and either primed condition for any of the three target vowels. Thus, the first and second hypotheses were not supported by our analysis. In addition, the third hypothesis was also not supported, as no effect of NZES was found.

Table 4.8. Fixed effects for KIT selections with Australia as default condition.

	Estimate	Std. error	df	t value	Pr(>|t|)
(Intercept)	4.93	0.13	41.90	37.06	< 0.001 ***
condition = control	-0.23	0.14	39.00	-1.59	0.12
condition = NZ	0.03	0.14	39.00	0.25	0.80
order = reversed	0.49	0.05	805.10	9.73	< 0.001 ***
NZES	0.01	0.02	39.00	0.81	0.42

Table 4.9. Fixed effects for KIT selections with control as default condition.

	Estimate	Std. error	df	t value	Pr(>|t|)
(Intercept)	4.70	0.14	44.00	33.64	< 0.001 ***
condition = NZ	0.26	0.14	39.00	1.88	0.07 .
condition = Aus	0.22	0.14	39.00	1.59	0.12
order = reversed	0.49	0.05	805.10	9.73	< 0.001 ***
NZES	0.01	0.02	39.00	0.81	0.42

Table 4.10. Fixed effects for DRESS selections with Australia as default condition.

	Estimate	Std. error	df	t value	Pr(>|t|)
(Intercept)	2.94	0.11	34.60	26.43	< 0.001 ***
condition = control	-0.10	0.11	39.00	-0.94	0.36
condition = NZ	-0.07	0.11	39.00	-0.61	0.55
order = reversed	0.36	0.05	806.00	7.17	< 0.001 ***
NZES	0.01	0.01	39.00	0.88	0.38

Table 4.11. Fixed effects for DRESS selections with control as default condition.

	Estimate	Std. error	df	t value	Pr(>|t|)
(Intercept)	2.84	0.11	37.40	24.43	< 0.001 ***
condition = NZ	0.04	0.11	39.00	0.35	0.73
condition = Aus	0.10	0.11	39.00	1.59	0.36
order = reversed	0.36	0.05	806.00	7.17	< 0.001 ***
NZES	0.01	0.01	39.00	0.88	0.38

Table 4.12. Fixed effects for TRAP selections with Australia as default condition.

	Estimate	Std. error	df	t value	Pr(>|t|)
(Intercept)	3.28	0.17	46.70	19.23	< 0.001 ***
condition = control	0.004	0.19	39.00	0.02	0.98
condition = NZ	0.12	0.19	39.00	0.62	0.54
order = reversed	0.13	0.06	805.00	2.13	0.03 *
NZES	0.02	0.02	39.00	0.72	0.48

Table 4.13. Fixed effects for TRAP selections with control as default condition.

	Estimate	Std. error	df	t value	Pr(>|t|)
(Intercept)	3.28	0.18	47.50	18.20	< 0.001 ***
condition = NZ	0.11	0.19	39.00	0.60	0.55
condition = Aus	-0.004	0.19	39.00	-0.02	0.99
order = reversed	0.13	0.06	805.00	2.13	0.03 *
NZES	0.02	0.02	39.00	0.72	0.42

However, as shown in the above tables we did find a significant effect of presentation order, with participants selecting higher numbered tokens when the continuum order was reversed for all vowels and conditions. That is, participants selected more AusE-like tokens when the continuum presented the most AusE-like token first, and the most NZE-like token last. Model predictions for the reversing effect from continuum order are presented in Figures 4.7-4.9.

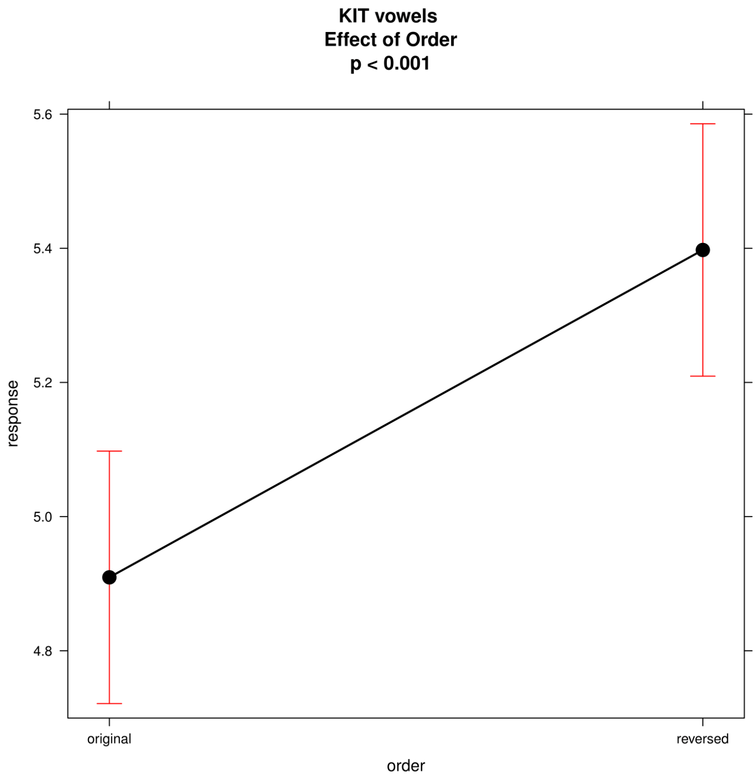


Figure 4.7. Model predictions for the effect of continuum order – KIT. Vertical axis represents mean token selection.

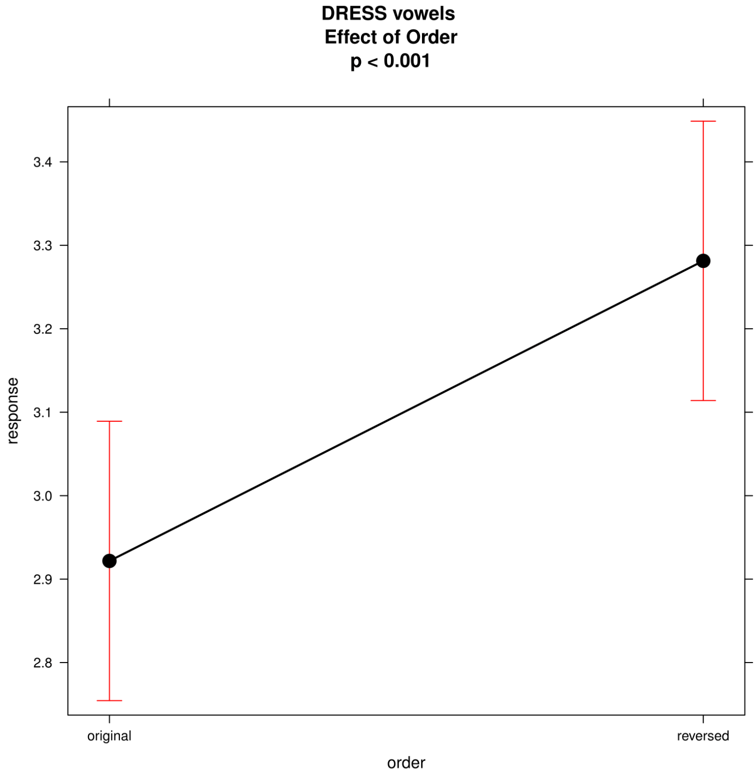


Figure 4.8. Model predictions for the effect of continuum order – DRESS. Vertical axis represents mean token selection.

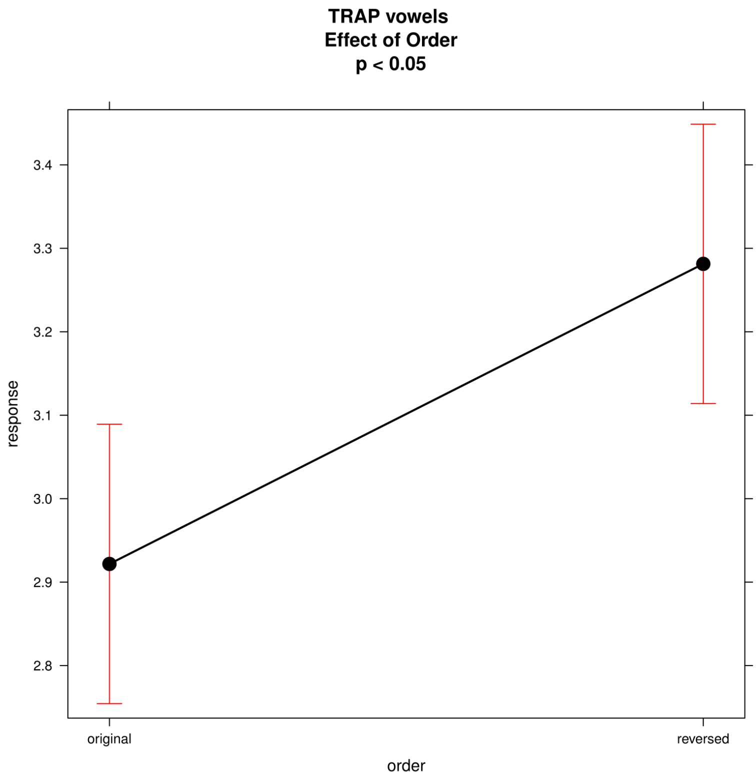


Figure 4.9. Model predictions for the effect of continuum order – TRAP. Vertical axis represents mean token selection.

4.4.1. Fixed effects model

Model outputs from the linear fixed effects regression analysis are presented below in Tables 4.14-4.19. This analysis did reveal a significant difference in token selections for KIT vowels between the control condition and the Australian and New Zealand conditions. Although this significant effect was found, this type of analysis does not account for the fact that responses come from different individuals and as such, may overstate the variance between conditions. Crucially, even this fixed effects model does not show a significant difference between the Australian and New Zealand conditions.

Interestingly, this model only found a significant effect of continuum presentation order with DRESS vowels, although there is a trend with TRAP vowels. A significant effect of NZES in KIT vowels was also found, with a higher NZES predicting more AusE-like token selections.

Table 4.14. Fixed effects for KIT selections with Australia as default condition – linear fixed effects regression model.

	Estimate	**Std. error**	**t value**	**Pr(>|t|)**
(Intercept)	4.93	0.06	80.65	< 0.001 ***
condition = control	-0.23	0.07	-3.18	0.001 **
condition = NZ	0.04	0.07	0.50	0.62
order = reversed	0.01	0.01	1.61	0.11
NZES	0.49	0.06	8.61	< 0.001 ***

Table 4.15. Fixed effects for KIT selections with control as default condition – linear fixed effects regression model.

	Estimate	**Std. error**	**t value**	**Pr(>|t|)**
(Intercept)	4.70	0.06	72.71	< 0.001 ***
condition = NZ	0.26	0.07	3.76	< 0.001 ***
condition = Aus	0.23	0.07	3.18	0.002 **
order = reversed	0.01	0.01	1.62	0.11
NZES	0.49	0.06	8.61	< 0.001 ***

Table 4.16. Fixed effects for DRESS selections with Australia as default condition – linear fixed effects regression model.

	Estimate	Std. error	t value	Pr(>|t|)
(Intercept)	2.94	0.06	50.14	< 0.001 ***
condition = control	-0.10	0.07	-1.50	0.14
condition = NZ	-0.06	0.07	-0.96	0.34
order = reversed	0.36	0.05	6.64	< 0.001 ***
NZES	0.01	0.01	1.45	0.15

Table 4.17. Fixed effects for DRESS selections with control as default condition – linear fixed effects regression model.

	Estimate	Std. error	t value	Pr(>|t|)
(Intercept)	2.84	0.06	45.74	< 0.001 ***
condition = NZ	0.04	0.07	0.57	0.57
condition = Aus	0.10	0.07	1.50	0.14
order = reversed	0.36	0.05	6.64	< 0.001 ***
NZES	0.01	0.01	1.45	0.15

Table 4.18. Fixed effects for TRAP selections with Australia as default condition – linear fixed effects regression model.

	Estimate	Std. error	t value	Pr(>|t|)
(Intercept)	3.28	0.08	42.84	< 0.001 ***
condition = control	0.004	0.09	0.04	0.97
condition = NZ	0.12	0.09	1.35	0.18
order = reversed	0.13	0.07	1.87	0.06 .
NZES	0.02	0.01	1.57	0.12

Table 4.19. Fixed effects for TRAP selections with control as default condition – linear fixed effects regression model.

	Estimate	Std. error	t value	Pr(>|t|)
(Intercept)	3.28	0.08	40.43	< 0.001 ***
condition = NZ	0.11	0.09	1.31	0.19
condition = Aus	-0.004	0.09	-0.04	0.97
order = reversed	0.13	0.07	1.87	0.06 .
NZES	0.18	0.01	1.57	0.12

4.5 Ordering effect

As the presentation order of continuum tokens was found to be a significant factor in token selection in our mixed effects model, this section provides a comparison of token selections in the original and reversed order (details about the two presentation orders can be found in Section 3.3.2). As a consequence of this ordering effect, it was deemed necessary to provide an analysis of selections when the continua was presented in the original order only (Section 4.5.1). This enabled a more appropriate comparison of findings with those of previous studies (Hay & Drager, 2010; Hay et al., 2006: Niedzielski, 1999; Lawrence, 2015) who presented all continua in a consistent order that is equivalent to the original order in the present experiment.

Tables 4.20 and 4.21 show how the mean token selection in each vowel condition differs between continua presented in the original order and those presented in reversed order. The selection values in Table 4.21 represent recoded data (recoding details are outlined in Section 3.4.1 and data in the remainder of this chapter is always recoded unless stated otherwise). Selection values in Table 4.22 represent the actual token selection made by participants. Recoding of selections for the continua in reversed order was necessary to allow for comparison of token selections that represent equivalent synthesised vowels.

Table 4.20. Mean token selection – all conditions, original order.

	KIT	DRESS	TRAP
Australia	4.99	3.02	3.39
New Zealand	5.02	2.89	3.44
Control	4.71	2.86	3.28
Total	4.91	2.92	3.37

Table 4.21. Mean token selection – all conditions, reverse order (recoded).

	KIT	DRESS	TRAP
Australia	5.43	3.28	3.38
New Zealand	5.48	3.30	3.59
Control	5.28	3.26	3.54
Total	5.40	3.28	3.50

Table 4.22. Mean token selection – all conditions, reverse order (actual selection).

	KIT	DRESS	TRAP
Australia	1.57	3.72	3.62
New Zealand	1.52	3.70	3.41
Control	1.72	3.74	3.46
Total	1.60	3.72	3.50

The values in Tables 4.20 and 4.21 indicate participants selected higher numbered tokens, representing more AusE-like vowels, when a continuum was presented in reversed order. This is true in all vowel contexts except TRAP vowels in the Australian condition.

These data suggest that if a participant selected, for example, token 5 as the best match to the target vowel in the sentence '*The new movie was a huge summer **hit***', when the continuum was presented in the original order, they were not necessarily selecting the equivalent token 2 when the continuum was presented in reverse order. Instead, participants were more likely to select the more AusE-like token 1 (i.e. token 6 in the original order). Table 4.23 outlines the distribution of token selections in original order only for each vowel in the three conditions and Table 4.24 outlines the distribution in reversed order only. Tables 4.25 and 4.26 display these distributions as percentages.

Table 4.23. Distribution of token selections in original order only by vowel and condition – selection totals.

	1	**2**	**3**	**4**	**5**	**6**	**Total**
KIT AU	0	0	6	31	61	41	139
KIT NZ	1	2	2	19	89	36	149
KIT C	0	2	14	37	56	31	140
DRESS AU	6	25	74	28	6	0	139
DRESS NZ	7	31	84	28	0	0	150
DRESS C	3	40	74	20	3	0	140
TRAP AU	2	22	49	55	10	2	140
TRAP NZ	4	20	50	59	14	2	149
TRAP C	7	14	67	41	7	4	140
Total	30	156	420	318	246	116	1286

Table 4.24. Distribution of token selections in reversed order only by vowel and condition – selection totals.

	1	2	3	4	5	6	Total
KIT AU	0	0	2	13	48	77	140
KIT NZ	0	2	0	8	54	86	150
KIT C	0	3	4	14	49	70	140
DRESS AU	1	16	71	47	5	0	140
DRESS NZ	1	20	67	59	1	2	150
DRESS C	2	25	55	51	6	1	140
TRAP AU	4	23	59	30	18	6	140
TRAP NZ	3	20	56	35	26	9	149
TRAP C	1	21	49	46	17	6	140
Total	12	130	363	303	224	257	1289

Table 4.25. Distribution of token selections in original order only by vowel and condition – percentage.

	1	2	3	4	5	6	Total
KIT AU	0.00	0.00	4.32	22.30	43.88	29.50	100
KIT NZ	0.67	1.34	1.34	12.75	59.73	24.16	100
KIT C	0.00	1.43	10.00	26.43	40.00	22.14	100
DRESS AU	4.32	17.99	53.24	20.14	4.32	0.00	100
DRESS NZ	4.67	20.67	56.00	18.67	0.00	0.00	100
DRESS C	2.14	28.57	52.86	14.29	2.14	0.00	100
TRAP AU	1.43	15.71	35.00	39.29	7.14	1.43	100
TRAP NZ	2.68	13.42	33.56	39.60	9.40	1.34	100
TRAP C	5.00	10.00	47.86	29.29	5.00	2.86	100
Total	2.33	12.13	32.66	24.73	19.13	9.02	100

Table 4.26. Distribution of token selections in reversed order only by vowel and condition – percentage.

	1	2	3	4	5	6	Total
KIT AU	0.00	0.00	1.43	9.29	34.29	55.00	100
KIT NZ	0.00	1.33	0.00	5.33	36.00	57.33	100
KIT C	0.00	2.14	2.86	10.00	35.00	50.00	100
DRESS AU	0.71	11.43	50.71	33.57	3.57	0.00	100
DRESS NZ	0.67	13.33	44.67	39.33	0.67	1.33	100
DRESS C	1.43	17.86	39.29	36.43	4.29	0.71	100
TRAP AU	2.86	16.43	42.14	21.43	12.86	4.29	100
TRAP NZ	2.01	13.42	37.58	23.49	17.45	6.04	100
TRAP C	0.71	15.00	35.00	32.86	12.14	4.29	100
Total	0.93	10.09	28.16	23.51	17.38	19.94	100

Figures 4.10 and 4.11 compare token selections between the original and reversed presentation orders for KIT vowels. For KIT vowels, the mode token selection shifts from token 5 to token 6 in all conditions. Token 6 was selected 25.23% of the time in the original order but 54.19% of the time in the reversed order. Selections for tokens other than 5 or 6 are 28.97% in original order and 10.70% in the reversed.

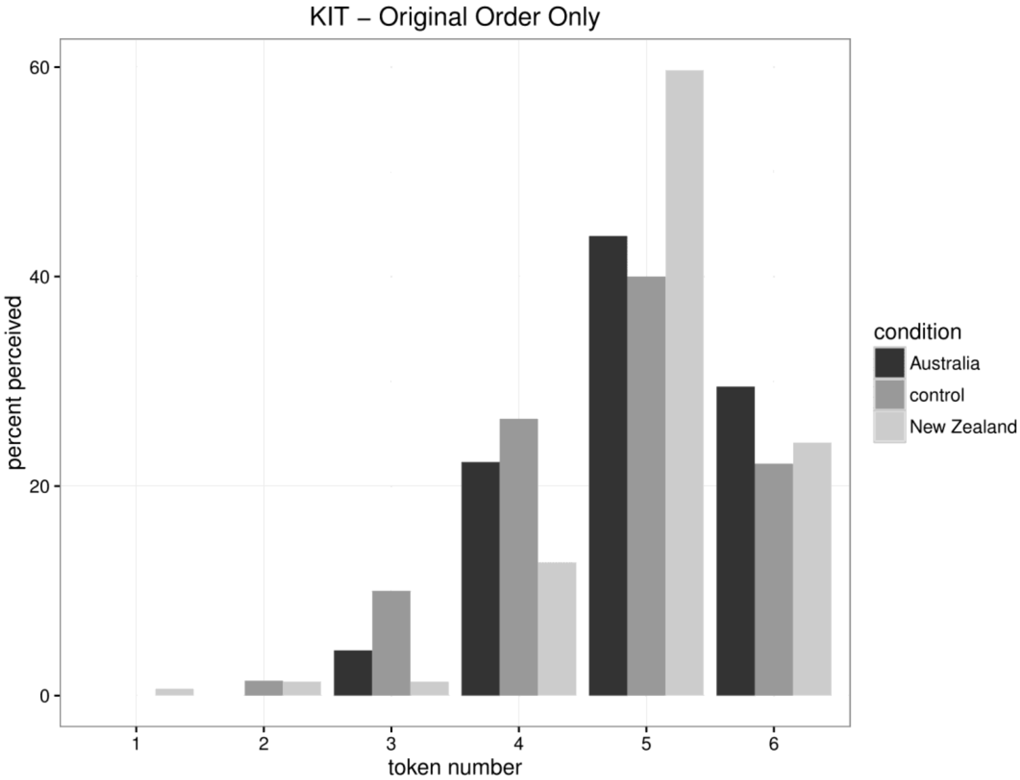


Figure 4.10. Token selections for KIT vowels – all conditions, original order only. Lower numbers represent more NZE-like tokens.

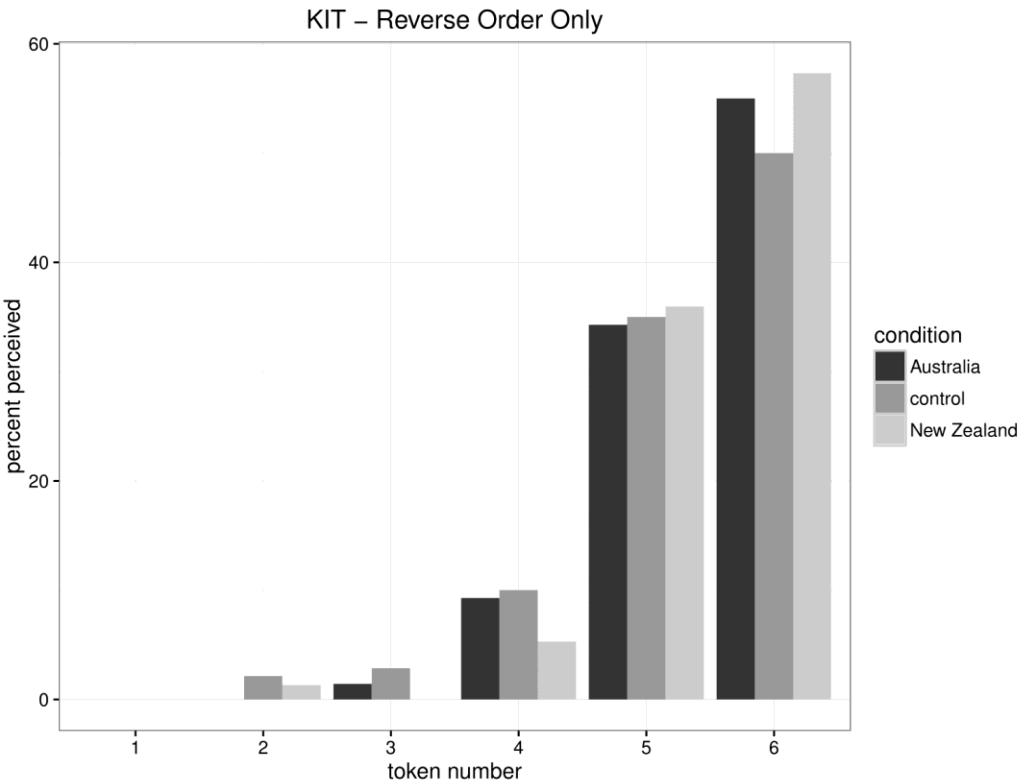


Figure 4.11. Token selections for KIT vowels – all conditions, reverse order only. Lower numbers represent more NZE-like tokens.

Figures 4.12 and 4.13 compare token selections between the original and reversed presentation orders for DRESS vowels. The most obvious change in selection preference is token 4, which accounts for 17.72% of selections in the original order and 36.51% in the reversed order.

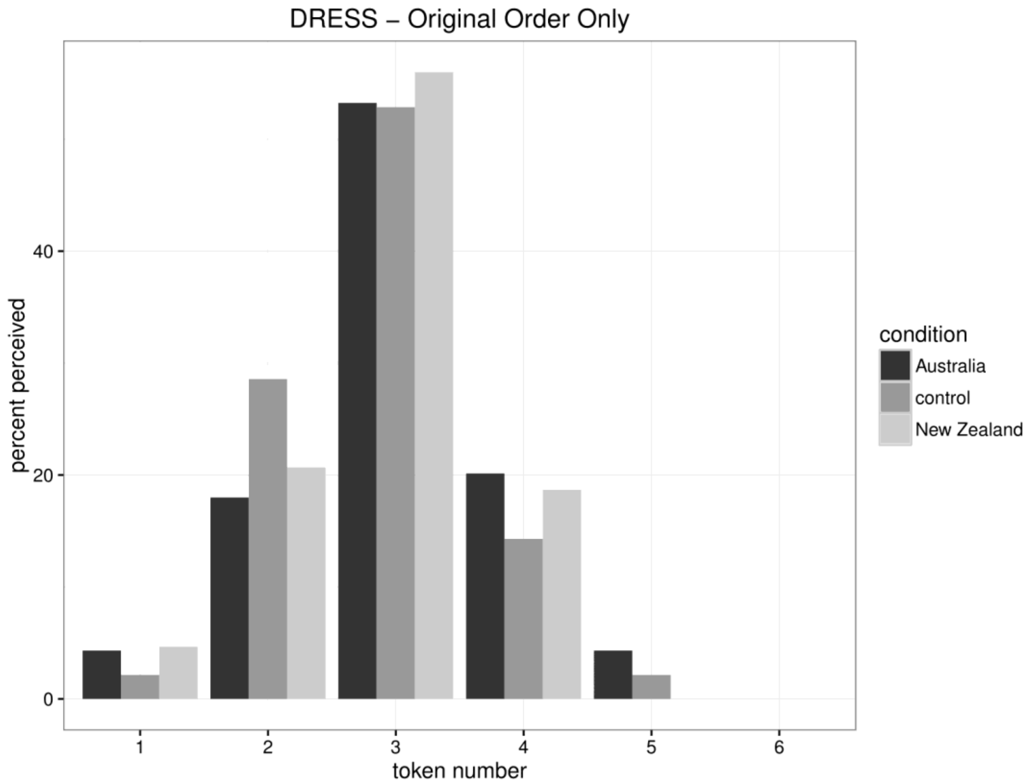


Figure 4.12. Token selections for DRESS vowels – all conditions, original order only. Lower numbers represent more NZE-like tokens.

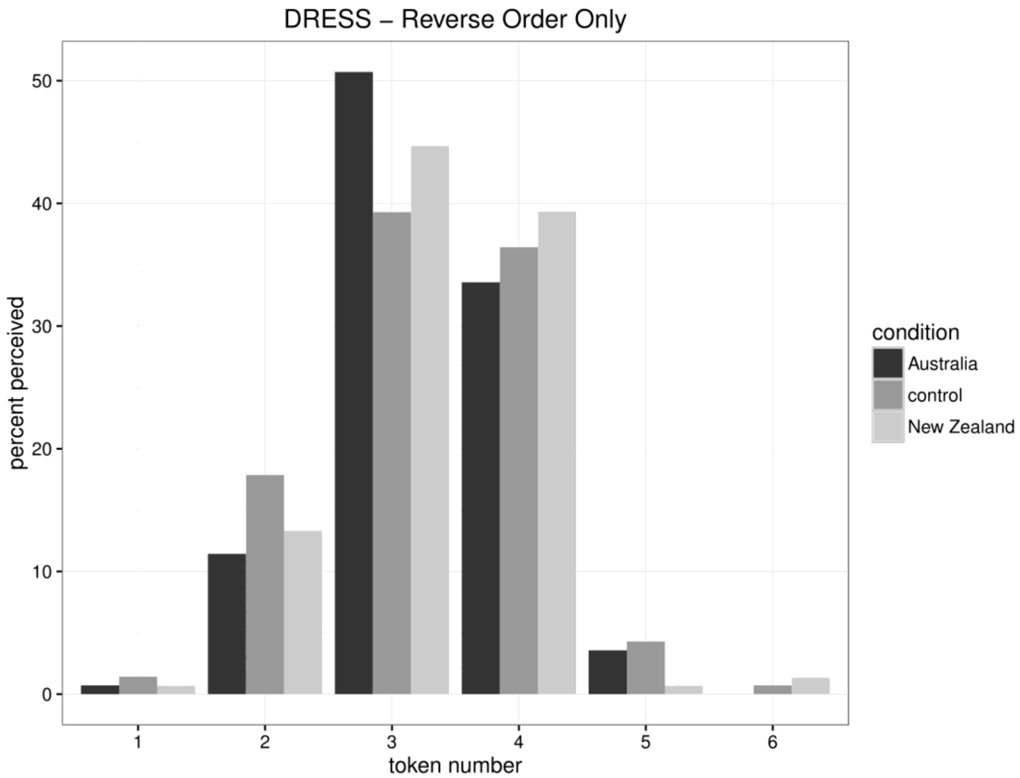


Figure 4.13. Token selections for DRESS vowels – all conditions, reverse order only. Lower numbers represent more NZE-like tokens.

Figures 4.14 and 4.15 compare token selections between the original and reversed presentation orders for TRAP vowels. For the two priming conditions (Australia and New Zealand) the mode token selection shifts from token 4 in the original order to token 3 in the reversed. However, for selections made in the control condition, the mode is token three for both presentation orders but the proportion of selection lowers from 47.86% to 35.00%.

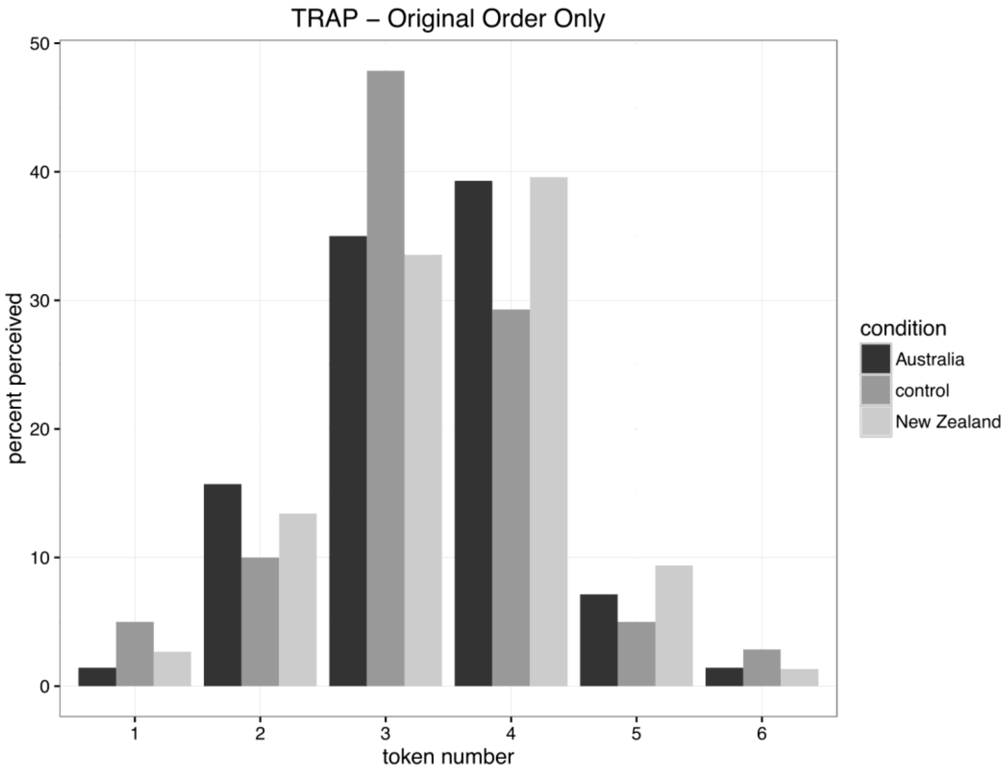


Figure 4.14. Token selections for TRAP vowels – all conditions, original order only. Lower numbers represent more NZE-like tokens.

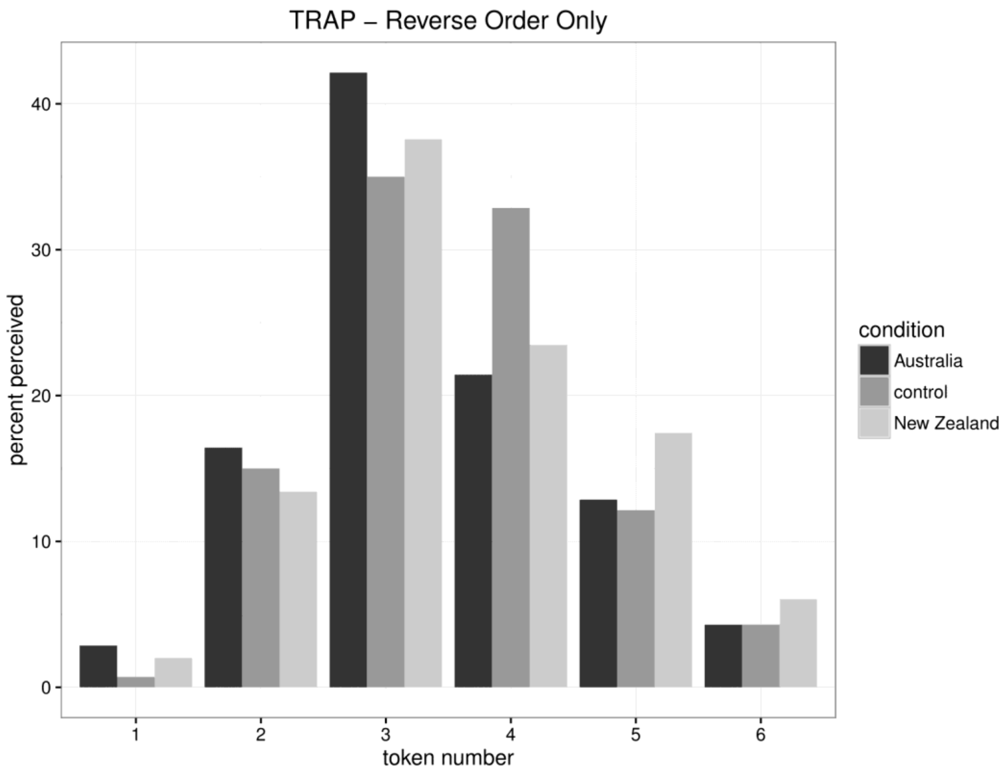


Figure 4.15. Token selections for KIT vowels – all conditions, reverse order only. Lower numbers represent more NZE-like tokens.

4.6 Analysis of results from original order

We fit the same linear mixed effects model used for the main analysis to the original order data only with presentation order removed from the fixed effects. Model outputs are presented below in Tables 4.27-4.32. Consistent with the initial analysis, no main effect of condition was found. There was no significant difference between selections in the Australian and New Zealand conditions for KIT, DRESS or TRAP. In addition, no significant difference was found between token selections in the control condition and either primed condition for any of the three target vowels. This adds further weight to the rejection of our hypotheses. Regarding NZE exposure, a trend can be observed for DRESS vowels where a higher NZES results in more AusE-like token selections. However, no trend is present in KIT and TRAP vowels.

Table 4.27. Fixed effects for KIT selections in original order only with Australia as default condition.

	Estimate	Std. error	df	t value	Pr(>|t|)
(Intercept)	4.93	0.17	40.54	28.66	< 0.001 ***
condition = control	-0.29	0.19	38.98	-1.53	0.13
condition = NZ	0.02	0.19	39.04	0.11	0.91
NZES	0.02	0.02	38.93	0.86	0.40

Table 4.28. Fixed effects for KIT selections in original order only with control as default condition.

	Estimate	Std. error	df	t value	Pr(>|t|)
(Intercept)	4.64	0.18	42.45	25.55	< 0.001 ***
condition = NZ	0.31	0.19	38.98	1.68	0.10
condition = Aus	0.29	0.19	38.98	1.53	0.13
NZES	0.02	0.02	38.93	0.86	0.40

Table 4.29. Fixed effects for DRESS selections in original order only with Australia as default condition.

	Estimate	Std. error	df	t value	Pr(>|t|)
(Intercept)	2.95	0.12	29.21	24.80	< 0.001 ***
condition = control	-0.20	0.12	39.09	-1.60	0.12
condition = NZ	-0.15	0.12	39.10	-0.28	0.21
NZES	0.03	0.02	39.02	1.90	0.06 .

Table 4.30. Fixed effects for DRESS selections in original order only with control as default condition.

	Estimate	Std. error	df	t value	Pr(>|t|)
(Intercept)	2.76	0.12	31.94	22.13	< 0.001 ***
condition = NZ	0.04	0.12	38.97	0.35	0.73
condition = Aus	0.20	0.12	39.09	1.60	0.12
NZES	0.03	0.02	39.02	1.90	0.06 .

Table 4.31. Fixed effects for TRAP selections in original order only with Australia as default condition.

	Estimate	Std. error	df	t value	Pr(>|t|)
(Intercept)	3.31	0.18	41.17	18.24	< 0.001 ***
condition = control	-0.15	0.20	39.00	-0.74	0.47
condition = NZ	0.02	0.20	39.06	0.12	0.91
NZES	0.04	0.03	39.00	1.40	0.17

Table 4.32. Fixed effects for TRAP selections in original order only with control as default condition.

	Estimate	Std. error	df	t value	Pr(>|t|)
(Intercept)	3.16	0.19	42.89	16.46	< 0.001 ***
condition = NZ	0.17	0.20	39.06	0.87	0.39
condition = Aus	0.15	0.20	39.00	0.74	0.47
NZES	0.04	0.03	39.00	1.40	0.17

4.6.1 Fixed effects analysis of results from original order

We also fit the linear fixed effects regression model to original order data only. Model outputs are presented below in Tables 4.33-4.38. A significant difference between token selections in the control condition and the two priming conditions was found for KIT vowels and between the Australian and control conditions in DRESS vowels. For each of these, token selections in the control condition represent more NZE-like vowels. Again, no significant difference was found between selection made the Australian and New Zealand conditions in any vowel context. A significant effect is observed for NZES in the DRESS and TRAP vowel contexts where a higher NZES results in more AusE-like token selections.

Table 4.33. Fixed effects for KIT selections in original order only with Australia as default condition – linear fixed effects regression model.

	Estimate	Std. error	t value	Pr(>|t|)
(Intercept)	4.93	0.08	61.44	< 0.001 ***
condition = control	-0.29	0.11	-2.78	0.006 **
condition = NZ	0.02	0.10	0.22	0.83
NZES	0.02	0.01	1.56	0.12

Table 4.34. Fixed effects for KIT selections in original order only with control as default condition – linear fixed effects regression model.

	Estimate	Std. error	t value	Pr(>|t|)
(Intercept)	4.64	0.09	53.85	< 0.001 ***
condition = NZ	0.31	0.10	3.07	0.002 **
condition = Aus	0.29	0.11	2.78	0.006 **
NZES	0.02	0.01	1.56	0.12

Table 4.35. Fixed effects for DRESS selections in original order only with Australia as default condition – linear fixed effects regression model.

	Estimate	Std. error	t value	Pr(>|t|)
(Intercept)	2.95	0.07	40.59	< 0.001 ***
condition = control	-0.19	0.10	-2.03	0.04 *
condition = NZ	-0.15	0.09	-1.63	0.10
NZES	0.03	0.01	2.46	0.01 *

Table 4.36. Fixed effects for DRESS selections in original order only with control as default condition – linear fixed effects regression model.

	Estimate	Std. error	t value	Pr(>|t|)
(Intercept)	2.76	0.08	35.36	< 0.001 ***
condition = NZ	0.04	0.09	0.46	0.65
condition = Aus	0.19	0.10	2.04	0.04 *
NZES	0.03	0.01	2.46	0.01 *

Table 4.37. Fixed effects for TRAP selections in original order only with Australia as default condition – linear fixed effects regression model.

	Estimate	Std. error	t value	Pr(>|t|)
(Intercept)	3.31	0.09	37.42	< 0.001 ***
condition = control	-0.15	0.12	-1.30	0.19
condition = NZ	0.02	0.11	0.20	0.84
NZES	0.04	0.01	2.475	0.01 *

Table 4.38. Fixed effects for TRAP selections in original order only with control as default condition – linear fixed effects regression model.

	Estimate	Std. error	t value	Pr(>|t|)
(Intercept)	3.16	0.10	33.16	< 0.001 ***
condition = NZ	0.17	0.11	1.53	0.13
condition = Aus	0.15	0.12	1.30	0.19
NZES	0.04	0.01	2.48	0.01 *

4.7 Summary of results

In this chapter we have presented the results of our descriptive analysis as well as the linear mixed effects and linear fixed effects regression analyses. The main findings are as follows:

- Token selections are generally accurate in that the most frequently selected tokens overall were 3, 4, and 5.
- Participants tended to select tokens that represent more raised and fronted vowels than the vowel produced by the speaker. This is reflected in the differences in selection between KIT vowels and DRESS and TRAP.
- The linear mixed effects model revealed no significant differences in token selection between the Australian and New Zealand conditions.
- The linear mixed effects model revealed no significant differences between token selections in the control condition and the two priming conditions. However,
- The linear fixed effects regression model did show that token selections in the control condition were significantly more NZE-like than in either priming condition for KIT vowels. However, as individual participant effects are not taken into account in this model, it is likely that the significance is overestimated
- Similarly, NZES was not found to be a significant predictor of selections in the mixed effects model but was in the linear effects regression, with participants allocated a higher NZES selecting more AusE-like tokens on average.

- Token selections did differ significantly between the two presentation orders. Overall, more AusE-like tokens were selected when the continuum presentation order was reversed.
- The linear mixed effects model was fit to original order data only but no significant differences in token selection between any of the conditions was found.

5. Discussion

5.1 Introduction

This thesis reports on a pseudo-replication of Hay and Drager (2010). Our aim was to establish whether the observed regional priming effect could be reproduced with an AusE-speaking sample. Using priming conditions similar to those in Hay and Drager (2010), we predicted that the presence of culturally significant stuffed toys would influence participants' selections in a vowel matching task. Based on this prediction, we hypothesised that participants exposed to the New Zealand prime (stuffed toy kiwis) would select continuum tokens that represent more NZE-like vowels than those exposed to the Australian prime (stuffed toy koalas). Following the results of Hay and Drager (2010) and Hay et al. (2006), we also hypothesised that selections made for KIT vowels would exhibit the most variability between the Australian and New Zealand conditions.

No support was found for these two hypotheses. Token selections in the matching task did not differ significantly between the Australian and New Zealand conditions. Although variance was observed between selections in the control condition and both priming conditions, this variance was not found to be significant using a linear mixed effects model. Further, the majority of variance that was observed between conditions was in the opposite direction to that predicted.

Our third hypothesis was that exposure to NZE would be a significant predictor for response variability between the conditions. Participants in the New Zealand condition who had more exposure to NZE (resulting in a higher NZES) were expected to be more likely to show the priming effect in their token selections. This hypothesis was also not supported by our analysis. Participants exposed to the New Zealand condition with a higher NZES did not show the predicted priming effect.

Analysis did reveal an unexpected effect of continuum presentation order. Selections made in response to continua played in the reversed order (AusE-like to NZE-like) reflect more AusE-like tokens than selections made in response to the continua presented in the original order (NZE-like to AusE-like). We elected to analyse those data from original order selections only as a secondary analysis. These token selections were not found to differ significantly between conditions. This allows us to rule out the possibility that the effect of presentation order was masking any regional priming effect.

The remainder of this chapter will be structured as follows. First, in Section 5.2, the main findings will be discussed and comparisons will be made to the previous studies. Following this, Section 5.3 will consider the extent to which the participants' level of NZE exposure may explain the reason for our lack of support for the hypotheses. In Section 5.4 we will consider the implications of our findings for exemplar theory and the claim that indexical information plays a role in speech perception. In Section 5.5, the effect of presentation order and the implications of task design will be considered. Finally, the limitations of the study will be discussed (Section 5.6).

5.2 Priming effect and discussion of previous research

Hay and Drager (2010) suggest that the observed regional priming effect illustrates "how subtle differences in experimental environment can influence subjects' responses" (p. 883). This idea extends the evidence from a large body of literature, reviewed in Chapter 2, indicating that indexical information plays a role in speech perception (e.g. Drager, 2011; Strand & Johnson, 1996). In the present study, some variability consistent with the priming condition was observed, for example participants in the Australian condition selected more AusE DRESS vowels than those in the New Zealand condition. However, this variability did not represent broader trends in the data. Overall, there is little evidence to suggest that the priming condition influenced participants' token selection in the present experiment.

Results from the linear mixed effects model confirmed that token selections were not significantly different between the three conditions. This was consistent for each of the three vowel contexts. Interestingly, when we fit a linear fixed effects regression model, similar to the method reported in Hay and Drager (2010), a significant difference in token selections for KIT vowels was found between the control condition and the two priming conditions. As a fixed effects model cannot account for individual variation amongst participants, it is possible that the significance of condition as a predicting factor was overstated by this model. It is now widely accepted that mixed effects models, with the inclusion of random effects (which were speaker and word in our analyses), present a more statistically robust presentation of trends in a dataset (Drager & Hay, 2012). Further, the fixed effects model did not find a significant difference between selections in the Australian and New Zealand conditions.

Lawrence's (2015) replication of Niedzielski (1999), using BATH and STRUT vowels with Southern British English listeners, also failed to find support for the claim that priming listeners towards another dialect could influence their performance in a matching task. Although results appeared to indicate some shift consistent with the priming condition, this was not found to be significant. This was a surprising result as the target vowels, BATH and STRUT, are highly salient identifiers of Northern and Southern British English. Lawrence (2015) argues that either the influence of a regional prime is more limited than was previously suggested or the influence exists but is highly contextually specific. Similarly, Squires (2013) argues in favour of a more limited expectation of bidirectionality between linguistic and indexical influences on perception. In a morphosyntactic paradigm, Squires (2013) found that linguistic information influenced impressions about a speaker's social status but socioeconomic cues did not influence perception of non-standard speech. While the results reported in the present study might be seen to add support to these conclusions, there are a number of other possible explanations for the result observed here.

One possible explanation for the lack of support for our hypotheses is an ineffective priming condition. This explanation echoes that of Squires (2013) who concedes that the lack of a priming effect may have been due to the primes not working as expected. If the effect reported in Hay and Drager (2010) relies on a strong association between the toy, its elicited dialect and relevant dialectal variation, then it is possible that the toy kiwis were not culturally significant enough to activate 'New Zealand' for our participants. In other words, there is no way of knowing whether those participants in the New Zealand condition recognised the toys as kiwis and in turn associated them with New Zealand. It may be that the kangaroos and koalas used in Hay and Drager's (2010) experiment are more salient signifiers of Australia for New Zealanders than our kiwis are signifiers of New Zealand for Australians. Further, although the kiwi is a distinctive emblem of New Zealand, our toys may not have been good enough representations of kiwis for them to be easily identified (see image, Section 3.4.1). One way to overcome this problem in the future would be to test the identification or recognition of the kiwis before their use in the experiment. Further, if participants had been asked to identify the toys used in the present study post-task, this would provide some clarification regarding prime identification.

5.3 Exposure and sensitivity to NZE

Another possible explanation for our lack of result is that the participants did not have the level of NZE exposure required to complete the matching task as predicted. Questionnaire responses indicated that 11 of the 43 participants had never been to New Zealand, didn't speak with, or know, any New Zealanders and couldn't name any New Zealand media. Although these individuals might be generally aware of NZE and how it differs from AusE, it is possible that they are not. If this is the case, the lack support for Hay and Drager (2010) is not surprising, at least from these participants.

Supporting the lack of exposure explanation, it has been shown that dialect recognition and feature identification is facilitated by a listener's familiarity with that dialect (Sumner & Samuel, 2009). Listeners are able to make use of social and phonological information to build perceptual categories for regional dialects, consistent with the claims made by exemplar theory (Clopper & Pisoni, 2007). Thus those listeners who are experienced with another dialect are more likely to have accurately labelled categories within their exemplar space, relevant to that dialect, facilitating the categorisation of new input. Peterson and Barney (1952) found that identification of vowels was dependent on a listener's previous language experience, as well as finding that peripheral vowels were more likely to be correctly identified than other vowels. Labov (2010) found a 'significant local advantage' whereby listeners from three dialect areas of North America were better at identifying vowels produced by a speaker from their own area than from one of the other two areas.

Expecting that exposure to NZE would improve a listener's representation of a NZE vowel space, as well as increase the likelihood of prime recognition, we predicted that, in the New Zealand condition, participants with frequent exposure to NZE should be more likely to show the priming effect. More than half (n = 26) of the participants had travelled to New Zealand or spoke with New Zealanders on a regular basis. In addition, six of these participants had a parent, partner or close friends from New Zealand. Yet, participants with higher NZES scores based on contact with NZE did not demonstrate the priming effect. This suggests that our result cannot be attributed to a lack of exposure to NZE alone.

It could be that the level of exposure assumed to our participants was not enough to allow them the requisite fine phonetic knowledge that would result in meaningful influence from the priming condition. However, as discussed in Section 2.5.2, there is evidence to support the idea that Australians are generally sensitive to the differences between NZE and AusE and likely do have an understanding of the most salient differences between the dialects. Australians believe NZE to be identifiable (Weatherall et al., 1998), and have been shown to accurately identify NZE vowels in forced-choice tasks (Bayard et al., 1998; Ludwig, 2007; Weatherall et al., 1998). In Ludwig (2007), Australians were able to distinguish between AusE and NZE variants of synthesised words 67% of the time. NZE KIT vowels were correctly identified in 90% of instances, although the AusE KIT was typically perceived as more NZE-like. Australians were also found to be more accurate at correctly identifying DRESS and TRAP vowels than New Zealanders. This suggests that if New Zealanders showed the priming effect, it is reasonable to expect that Australians would too.

Further, even naïve listeners have been shown to make use of reliable acoustic-phonetic properties to identify dialects of American English (Clopper & Pisoni, 2004; 2007; Williams, Garrett, & Coupland, 1999) as well as variation in Dutch (Van Bezooijen & Gooskens, 1999; Van Bezooijen & Ytsma, 1999). Tamasi (2004) and Clopper and Pisoni (2007) showed that listeners were able to identify dialects of American English in a free classification task using fine-grained phonetic distinctions. For our study, the salience of the KIT vowel (Maclagan et al., 1999; Watson et al., 1998) means it should have reliable acoustic-phonetic properties that differentiate AusE and NZE, even for those participants who may otherwise lack frequent exposure to NZE.

It would have been useful to have participants in the New Zealand condition assessed on their ability to identify AusE and NZE in a post-experiment task. If participants could identify a NZE speaker on the basis of KIT, DRESS and TRAP vowels and still failed to show any influence of the prime, this would strengthen the null result.

5.4 Implications for exemplar theory

The lack of result in the present study does not necessarily contradict Hay et al. (2006) and Hay and Drager's (2010) support for an exemplar-based model of speech perception. According to exemplar models, speech input is categorised by comparing the relative activation levels for each candidate category (Hay et al. 2006; Pierrehumbert, 2001). As discussed in Sections 2.2 and 2.3, there is evidence to suggest that socially salient cues may bias categorisation in favour of acoustic variants associated with that indexical information (Foulkes & Docherty, 2006). The system may even be directed toward a particular categorisation when the acoustic information does not match (Niedzielski, 1999). In Hay and Drager (2010), it is argued that, for New Zealanders, stuffed toy koalas and kangaroos raise the activation level of exemplars indexed as 'Australian'. Phonetic input is then more likely to be classified along with those raised exemplars, because the activated portion of the category distribution is centred around Australian exemplars (Hay et al., 2006: p. 24). Hence, the bias towards more AusE-like vowels. It may be that for our experiment, any existing New Zealand or NZE exemplars were not activated enough (if at all) by the kiwi to compete with the resting activation level of AusE exemplars, even for those Australians with exposure to NZE. In order to produce a shift in token selection towards NZE-like variants and further test the claims of exemplar theory, a more overt prime might have been required to activate the relevant indexical and linguistic categories.

5.5 Token selection

Consistent across all three conditions and vowel contexts, selections favoured continuum tokens that represent more raised and fronted vowels than the speaker's production. This explains why the mean token selection in KIT vowels is different from DRESS and TRAP. For KIT vowels, the fronted and raised tokens were numerically higher tokens (5 and 6) and in DRESS and TRAP these were numerically lower tokens (typically 3). For both KIT and DRESS, the preferences shown by participants in token selection lend support to a peripherality bias in vowel perception. As discussed in Polka & Bohn (2003; 2011), directional asymmetries in vowel discrimination hint at underlying perceptual biases that favour more peripheral vowels in an F1/F2 defined vowel space. This suggests listeners may use peripheral vowels as perceptual reference points. These reference points may have influenced participants' representation of the target vowel as the continuum tokens are played, resulting in the selection of the more peripheral tokens. However, this explanation does not account for the selection preferences for TRAP vowels, where less peripheral raised tokens, rather than more peripheral lowered tokens, were preferred. Further, it is unclear whether a peripherality bias would be expected to influence a matching task requiring listeners to compare synthesised stimuli with recorded speech. Investigation of this explanation would require more targeted experimentation.

5.5.1 Working memory and paradigm issues

As has been discussed extensively in this thesis, aside from Lawrence (2015), the previous studies testing the regional priming paradigm with a matching task were able to successfully show a significant difference in token selections between two priming conditions. However, the task assumes participants are able to give equal consideration to all six tokens and then assess the best match. Research into the limitations of working memory suggest the requirements of the task may be beyond the capabilities of untrained listeners, raising concerns about participant performance in the previous studies.

Miller (1956) proposed that working memory was limited to lists of seven items, plus or minus two, a conclusion supported by Kinsbourne and Cohen (1971). However, additional studies investigating working memory capacity seem to support the even more modest limit of four items (Cowan, 2001; Luck & Vogel, 1997; Sperling, 1960). As the present study required listeners to hold seven items (six continuum tokens and the target word) it seems likely that this is beyond the capacity of working memory. This could conceivably result in accuracy issues with the task which, given the subtleties of the effect overall, are troubling. Hay et al. (2006) themselves acknowledged the difficulties of the task outlined by Niedzielski (1999), and designed a simplified procedure for their replication. A number of changes were also made to the procedure in the present study to further simplify the task, although the six continuum tokens plus target vowel has remained constant.

Baddeley (1992; 2010) has suggested that a phonological similarity effect impairs recall of words that are similar in sound so it can be assumed this extends to the continuum tokens used in this study. Li, Cowan, and Saults (2012) also found that listeners struggled to retain more than four tones in memory. It is worth noting however, that none of the pilot participants used for the present study indicated that the task was difficult. Although, one pilot participant did mention that she had trouble earlier on in the task until she started repeating the target word to herself between continuum tokens. This strategy might be problematic for the task, as the mental representation of the speaker's target word could be influenced by the listener's repetitions.

Nonetheless, the concerns relating to working memory should be taken on board with future research design in this paradigm. Future matching tasks testing a regional priming effect should minimise the requirements of working memory. One possibility could be to use either a shorter sentence or an isolated target word played before each token on the continuum. The participant would then rate each token against the preceding target word. It would be interesting to see whether a replication using a method such as this would produce the same effect in Detroit or New Zealand.

5.5.2 The reversing effect

Analysis revealed a significant difference between token selections in the two continuum presentation orders for all three vowel contexts. We elected to vary the presentation order of continuum tokens to determine whether selections were based on acoustic characteristics of the synthetic vowel or were a consequence of any potential response patterning. That the majority of token selections were within one continuum step of the synthesised speaker's vowel could be evidence for participants selecting tokens on the basis of acoustic information. However, this does not preclude a centring bias or other form of response patterning as both of these explanations might conceivably produce similar trends in selection.

The observed reversing effect may be further evidence that participants' ability to perform accurately in the task is undermined by the limitations of working memory. It would be expected that token selections should be consistent between presentation orders. However, it was found that participants selected more AusE-like tokens when the presentation order was reversed. For KIT vowels, when continua were presented in the original order, the mode token selection was token 5. If participants performed as expected, the mode token selection in the reversed order would be token 2 (which is recoded as token 5). However, the mode selection in the reversed order was token 1 (which is recoded as token 6).

As discussed, it is possible that listeners do not have the working memory capacity to hold six acoustically similar synthetic tokens equally in memory and make an accurate comparison to a target vowel. It is likely that participants were making the decision between tokens 5 and 6 (as presented in the original order). In reversed order these tokens would appear as the first two in the continuum and participants would have heard the other four tokens in the continuum before being able to make their selection. We speculate that the last four tokens in the reversed continuum may have acted as distractors. Participants, unsure of which of the first two tokens they preferred, selected token 1 as a more certain option. However, this explanation does not fit as well for DRESS and TRAP vowels, where selection differences, while significant, tended to vary between tokens 3 and 4. It would be worthwhile designing a matching task to test the reversing effect specifically in order to determine whether this is a problem inherent to the task or a generalizable phenomenon.

5.5.3 The control condition

Our decision to include a control group was motivated by uncertainty about the extent to which both priming conditions might show an effect, as well as participants' ability to perform the task. Without a control condition in any of the previous studies, there is no way of knowing whether the 'local' priming condition resulted in any shift in selection from what a participant would otherwise choose. For example, it is unclear whether there was any priming effect from the New Zealand condition on the NZE-speaking participants in Hay et al. (2006) and Hay and Drager (2010). It would be interesting to see whether participants exposed to the New Zealand condition in these studies would have selected more NZE-like tokens than a control group. As our results indicate, for KIT and DRESS vowels, participants in the Australian condition did select more AusE-like tokens than those in the control, although this did not occur with TRAP vowels. Further, selection differences between groups did not reach significance in the linear mixed effects model.

5.6 Limitations

Unfortunately, we did not have enough male participants to test the gender effect observed in both Hay et al. (2006) and Hay and Drager (2010). In both New Zealand studies, males showed the opposite effect to females, selecting more NZE-like tokens in the Australian condition. Hay and Drager (2010) argued that many New Zealand males have an inherently negative association towards Australia due to a sporting rivalry between the two countries. This resulted in males displaying a divergence effect in the Australian condition. Females' associations are more likely to be positive. It would not be surprising if this sporting rivalry has a reciprocal influence on Australian males' attitudes towards New Zealand. It would be interesting to see whether the divergence between male and female selections would have been replicated with an Australian sample.

Related to the procedure, there were some difficulties establishing the two priming conditions with the stuffed toys. The intention was to introduce the prime without it being obviously related to the task. This was achieved by the experimenter 'finding' the headphones used for the task under the prime toys, in a drawer. Both the headphones and toys were then removed from the drawer and casually placed on the desk. For consistency, the toys were deliberately placed beside the laptop, facing the participant to allow for easier identification. However, introduction of the prime in this way did not guarantee the participant would be able to identify the toys or even give them any significant attention. Further, the toys did occasionally fall over or behind the laptop, and in one trial a toy fell from the table completely. The experimenter was hesitant to fuss over placement too much as any excessive attention placed on the toy was thought to potentially lead the participant to suspect that the toy was relevant to the experiment.

It might be that drawing explicit attention to the prime is not problematic to the purpose of the task. However, explicit attention would detract from the purpose of Hay and Drager's (2010) priming condition, that being, the incidental presence of a regionally significant token in the experimental environment would produce an effect. However, given the nature of the results observed in the present study, perhaps a slightly more overt and consistent priming condition would be worth testing. This may improve our understanding of what level of prime is required to produce a measurable effect.

As we elected to restrict the target words used in the experiment to the form /CVt/ or /CCVt/, we were unable to control for lexical frequency. Although the constraints on our target words resulted in a more tightly controlled set of stimuli than those used in previous experiments, we were forced to use some less frequent lexical items (such as *grit*). It is possible that lexical frequency could influence the experiment and further replications should take this into account, either in experimental design or analysis. Finally, we did not account for just noticeable difference in our continua design. As a result, we cannot be certain that adjacent continuum tokens were able to be easily and consistently disambiguated by participants. More extensive piloting of the stimuli would have been required to ensure continuum tokens were suitable for the task.

6. Conclusions

Our aim was to complete a replication of the experiment reported in Hay and Drager (2010) in an Australian context. The replication described in this thesis also considered similar studies reported in Niedzielski (1999), Hay et al. (2006) and Jannedy et al. (2011) all of which showed that reference to another dialect could influence speech perception. We took into account the study by Lawrence (2015) which found no evidence of a priming effect. In our experiment, speakers of AusE completed a matching task in one of three priming conditions intended to influence their token selections. Conditions were established with either stuffed toy koalas (Australian), stuffed toy kiwis (New Zealand) or no toys (control).

We found no support for our three hypotheses. The stuffed toy priming effect observed in Hay and Drager (2010) was not replicated, token selections did not vary significantly according to the priming condition. This is consistent for each of the three vowels: KIT, DRESS and TRAP. In addition, participants who had frequent contact with NZE did not show any more sensitivity to the New Zealand prime than those without frequent contact with NZE. The results from this study suggest that, although there is evidence that New Zealanders can be influenced in their perception of vowels by the presence of cultural references to Australia, the reverse is not true. This may be a result of cultural asymmetry in recognition and familiarity between New Zealanders and Australians, it may simply be that the effect is limited to highly contextually specific situations or the effect may not exist.

In addressing the lack of result we considered potential issues with the priming condition and sensitivity to NZE. We attribute the effect of continuum presentation order to the inherent difficulty of the task and the limitations of a listener’s working memory. These issues highlight the need for carefully considered experimental design, particularly when investigating variation at a fine phonetic level.

6.1 Recommendations for further research

Further experimentation in this paradigm may be able to establish whether the results reported in the previous studies are limited to highly contextually specific situations or whether the participants tested in the present experiment and in Lawrence (2015) simply did not have sufficient exposure to the primed dialect.

Throughout this discussion we have offered numerous potential ideas for continuation of research in this paradigm. Despite the findings of the present study supporting the null result reported in Lawrence (2015), the extent to which regional priming influences speech perception warrants ongoing experimentation. Given the methodological and procedural issues identified in this analysis, there are two main areas that would require modification: presentation of stimuli and introduction of the priming condition.

Regarding concerns related to participants' working memory, the task might be best suited to four token continua with participants required to make a comparison to a vowel in an isolated target word. Alternatively, each token could be immediately preceded by the target word with participants required to rate the similarity of the token to that target word. This design would have the advantage of not having to consider the ordering or labelling of tokens. In addition, the presentation order of continuum tokens could be randomised.

The difficulties of establishing the priming condition demand a need for a priming condition that is able to be introduced with a greater level of consistency for all participants. Further, incorporation of additional post-task questionnaires requiring participants to identify the prime, state whether it was noticed and identify features of the primed dialect would assist evaluation of the success of the priming condition and its effect.

References

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press.

Babel, M. (2010). Dialect divergence and convergence in New Zealand English. *Language in Society*, *39*(04), 437–456. doi: 10.1017/S0047404510000400

Babel, M., & Russell, J. (2015). Expectations and speech intelligibility. *The Journal of the Acoustical Society of America*, *137*(5), 2823–2833. doi: 10.1121/1.4919317

Baddeley, A. (1992). Working Memory. *Science*, *255*(5044), 556–559.

Baddeley, A. (2010). Working memory. *Current Biology*, *20*(4), R136–R140.

Bates, D., Maechler, M., & Bolker, B. (2011). *Linear mixed-effects models using S4 classes* (R-Version 0.999375–41).

Bauer, L., & Warren, P. (2008). New Zealand English: phonology. *Varieties of English*, *3,* 39-63.

Bauer, L., Warren, P., Bardsley, D., Kennedy, M., & Major, G. (2007). New Zealand English. *Journal of the International Phonetic Association*, *37*(1), 97–102.

Bayard, D., Weatherall, A., Gallois, C., & Pittam, J. (2001). Pax Americana? Accent attitudinal evaluations in New Zealand, Australia and America. *Journal of Sociolinguistics*, *5*(1), 22–49. doi: 10.1111/1467-9481.00136

Bell, A. (1984). Language style as audience design. *Language in Society, 13*(02), 145–204. doi: 10.1017/S004740450001037X

Boersma, P. & Weenink, D. (2014). *Praat: doing phonetics by computer* (Version 5.4). Retrieved 16 October 2014 from http://www.praat.org/

Bourhis, R. Y., & Giles, H. (1977). The language of intergroup distinctiveness. In H. Giles (Ed.), *Language, Ethnicity, and Intergroup Relations* (pp. 119–36). London: Academic Press.

Bybee, J. (2006). From usage to grammar: The mind's response to repetition. *Language, 82*(4), 711–733.

Clopper, C. G., & Pisoni, D. B. (2004). Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of Phonetics*, *32*(1), 111–140.

Clopper, C. G., & Pisoni, D. B. (2007). Free classification of regional dialects of American English. *Journal of Phonetics, 35*(3), 421–438. doi: 10.1016/j.wocn.2006.06.001

Cowan, N. (2001). The magical number four in short-term memory: A reconsideration of mental storage capacity. *The Behavioral and Brain Sciences*, *24*, 87–114.

Cox, F. (1999). Vowel change in Australian English. *Phonetics*, *56*, 1–27.

Cox, F. (2006). The acoustic characteristics of /hVd/ vowels in the speech of some Australian teenagers. *Australian Journal of Linguistics*, *26*, 147-179.

Cox, F., & Palethorpe, S. (2007). Australian English. *Journal of the International Phonetic Association*, *37*(3), 341–350. doi: 10.1017/S0025100307003192

Cox, F., & Palethorpe, S. (2008). Reversal of short front vowel raising in Australian English. In *INTERSPEECH*, pp. 342-345.

Cox, F., Palethorpe, S., Miles, K. & Davies, B. (2014, December). *Is there evidence for region specific vowel variation in /hVd/ word list data from AusTalk?* Paper presented at the Australian Linguistic Society Conference, Newcastle.

Docherty, G. J., & Foulkes, P. (2014). An evaluation of usage-based approaches to the modelling of sociophonetic variability. *Lingua*, *142*, 42–56. doi: 10.1016/j.lingua.2013.01.011

Drager, K. (2011). Speaker Age and Vowel Perception. *Language & Speech*, *54*(1), 99–121. doi: 10.1177/0023830910388017

Drager, K., & Hay, J. (2012). Exploiting random intercepts: Two case studies in sociophonetics. *Language Variation and Change*, *24*(1), 59–78. doi: 10.1017/S0954394512000014

Drager, K., Hay, J., & Walker, A. (2010). Pronounced Rivalries: Attitudes and Speech Production. *Te Reo*, *53*, 27–53.

Easton, A., & Bauer, L. (2000). An Acoustic Study of the Vowels of New Zealand English. *Australian Journal of Linguistics*, *20*(2), 93–117. doi: 10.1080/07268600020006021

Foulkes, P., & Docherty, G. (2006). The social life of phonetics and phonology. *Journal of Phonetics*, *34*(4), 409–438. doi: 10.1016/j.wocn.2005.08.002

Foulkes, P., Scobbie, J. M., & Watt, D. (2010). Sociophonetics. *The Handbook of Phonetic Sciences, Second Edition*, 703–754.

Giles, H., & Ogay, T. (2006). Communication Accommodation Theory. In B. B. Whaley, & W. Samter (Eds.), *Explaining Communication: Contemporary Theories* (pp. 325-341). Routledge.

Giles, H. & Smith, P. (1979). Accommodation Theory: Optimal Levels of Convergence. In H. Giles, & R. N. St. Clair (Eds.), *Language and Social Psychology* (pp. 45-65). Baltimore: Basil Blackwell.

Goldinger, S. 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. doi: 10.1037/0278-7393.22.5.1166

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, *1052*, 51–79.

Greenwald, A. G., McGhee, D. E. & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *7*, 1464–1480.

Hintzman, D. L. (1986). Schema abstraction in a multiple-trace memory model. *Psychological Review*, *93*, 411–428.

Harrington, J., Cox, F., & Evans, Z. (1997). An acoustic phonetic study of broad, general, and cultivated Australian English vowels. *Australian Journal of Linguistics, 17*, 155–184.

Hay, J., & Drager, K. (2010). Stuffed toys and speech perception. *Linguistics*, *48*(4), 865–892. doi: 10.1515/LING.2010.027

Hay, J., Nolan, A., & Drager, K. (2006). From fush to feesh: Exemplar priming in speech perception. *The linguistic review*, *23*(3), 351–379.

Hay, J., & Foulkes, P. (2016). The evolution of medial /t/ over real and remembered time. *Language*, *92*(2), 298–330. doi:10.1353/lan.2016.0036

Hay, J., Warren, P., & Drager, K. (2006). Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics, 34*(4), 458–484.

Holmes, J. (1997). Setting new standards: sound changes and gender in New Zealand English. *English World-wide, 18*(1), 107–142.

Jannedy, S., Weirich, M., & Brunner, J. (2011). The effect of inferences on the perceptual categorization of Berlin German fricatives. In *Proceedings of the International Congress of Phonetic Sciences, ICPHS 2011*, 962–965.

Johnson, K. (1997). Speech perception without speaker normalization. In K. Johnson & J. W. Mullennix (Eds.), *Talker Variability in Speech Processing* (pp. 145–166). NY: Academic Press.

Johnson, K., Strand, E. A., & D'Imperio, M. (1999). Auditory–visual integration of talker gender in vowel perception. *Journal of Phonetics*, *27*(4), 359–384. doi: 10.1006/jpho.1999.0100

Kinsbourne, M., & Cohen, V. (1971). English and Hebrew consonant memory span related to the structure of the written language. *Acta Psychologica*, *35*(5), 347–351. doi: 10.1016/0001-6918(71)90009-6

Koops, C., Gentry, E., & Pantos, A. (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.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2015). Package 'lmerTest'. *R package version*, 2-0.

Labov, W. (1966). *The Social Stratification of English in New York City*. Washington, DC: Cent. Appl. Ling.

Labov, W. (1972). The social motivation of a sound change. In *Sociolinguistic Patterns* (pp. 1–42). Philadelphia: University of Pennsylvania Press.

Labov, W. (1991). The three dialects of English. *New ways of analyzing sound change*, *5*, 1–44.

Labov, W. (2010). A Controlled Experiment on Vowel Identification. In *Principles of Linguistic Change* (pp. 48–58). Wiley-Blackwell.

Lacerda, F. (1997). Distributed memory representations generate the perceptual-magnet effect. *Ms., Institute of Linguistics, Stockholm University*.

Lawrence, D. (2015). Limited evidence for social priming in the perception of the bath and strut vowels. In The Scottish Consortium for ICPhS 2015 (Ed.), Proceedings of the 18th International Congress of Phonetic Sciences. Glasgow, UK: the University of Glasgow.

Li, D., Cowan, N., & Saults, J. S. (2012). Estimating working memory capacity for lists of nonverbal sounds. *Attention, Perception, & Psychophysics*, *75*(1), 145–160. doi: 10.3758/s13414-012-0383-z

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*, 279–281.

Ludwig, I. (2007). *Identification of New Zealand English and Australian English based on stereotypical accent markers.* Unpublished master's thesis, University of Canterbury.

Maclagan, M. A., & Gordon, E. (1996). Out of the AIR and into the EAR: Another view of the New Zealand diphthong merger. *Language Variation and Change*, *8*, 125–147.

Maclagan, M. A., Gordon, E., & Lewis, G. (1999). Women and sound change: Conservative and innovative behavior by the same speakers. *Language Variation and Change*, *11*(1), 19–41.

Maclagan, M. A., & Hay, J. (2004, December). The rise and rise of New Zealand English DRESS. *Proceedings of the 10th Australian speech science and technology conference*, 183–188.

Maclagan, M. A., & Hay, J. (2007). Getting fed up with our feet: Contrast maintenance and the New Zealand English front vowel shift. *Language Variation and Change*, *19*(1), 1–25. doi: 10.1017/S0954394507070020

May, J. (1976). Vocal tract normalization for /s/ and /ʃ/. *The Journal of the Acoustical Society of America*, *59*(S1), S25–S25. doi: 10.1121/1.2002554

McLennan, C. T. (2007). Challenges facing a complementary-systems approach to abstract and episodic speech perception. *Proceedings of the 17th International Congress of Phonetic Sciences*.

McMillan, J., Beavis, A., & Jones, F. L. (2009). The AUSEI06 A new socioeconomic index for Australia. *Journal of Sociology*, *45*(2), 123–149. doi: 10.1177/1440783309103342

Miller, G. A. (1956). The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81–97.

Niedzielski, N. A. (1997). *The effect of social information on the phonetic perception of sociolinguistic variables* (Unpublished doctoral dissertation). University of California, Santa Barbara.

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

Peterson, G. E., & Barney, H. L. (1952). Control Methods Used in a Study of the Vowels. *The Journal of the Acoustical Society of America*, *24*(2), 175–184. doi: 10.1121/1.1906875

Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In J. Bybee & P. J. Hopper (Eds.), *Frequency effects and emergent grammar* (pp. 137-158). Amsterdam & Philadelphia: John Benjamins.

Pierrehumbert, J. B. (2006). The next toolkit. *Journal of Phonetics, 34*(4), 516-530. doi: 10.1016/j.wocn.2006.06.003

Pierrehumbert, J. B. (2016). Phonological Representation: Beyond Abstract Versus Episodic. *Annual Review of Linguistics*, *2*(1), 33–52. doi: 10.1146/annurev-linguistics-030514-125050

Polka, L., & Bohn, O.-S. (2003). Asymmetries in vowel perception. *Speech Communication, 41*(1), 221–231. doi: 10.1016/S0167-6393(02)00105-X

Polka, L., & Bohn, O.-S. (2011). Natural Referent Vowel (NRV) framework: An emerging view of early phonetic development. *Journal of Phonetics*, *39*(4), 467–478. doi: 10.1016/j.wocn.2010.08.007

Preston, D. (1993). Folk dialectology. In D. Preston (Ed.), *American dialect research* (pp. 333 – 378). Philadelphia: John Benjamins.

Psychology Software Tools, Inc. (2012) E-Prime 2.0. Retrieved from http://www.pstnet.com.

R Core Team. (2016). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rubin, D. L. (1992). Nonlanguage factors affecting undergraduates’ judgments of nonnative English-speaking teaching assistants. *Research in Higher Education, 33*(4), 511–531.

Sanchez, K., Hay, J., & Nilson, E. (2015). Contextual activation of Australia can affect New Zealanders’ vowel productions. *Journal of Phonetics*, *48*, 76–95. doi: 0.1016/j.wocn.2014.10.004

Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics bulletin, 2*(6), 110-114.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, *74,* 1–29.

Squires, L. (2013). It don’t go both ways: Limited bidirectionality in sociolinguistic perception. *Journal of Sociolinguistics*, *17*(2), 200–237. doi: 10.1111/josl.12025

Strand, E. A., & Johnson, K. (1996). Gradient and visual speaker normalization in the perception of fricatives. In D. Gibbon (Ed.), *Natural language processing and speech technology: Results of the 3rd KONVENS conference, Bielefeld, October 1996* (pp. 14-26). Berlin: Mouton.

Sumner, M., & Samuel, A. G. (2009). The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language*, *60*(4), 487–501. doi: 10.1016/j.jml.2009.01.001

Tamasi, S. L. (2003). *Cognitive patterns of linguistic perceptions* (Unpublished doctoral dissertation). University of Georgia, Athens.

Thomas, E. R. (2002). Sociophonetic applications of speech perception experiments. *American Speech, 77,* 115–47.

Van Bezooijen, R., & Gooskens, C. (1999). Identification of language varieties: The contribution of different linguistic levels. *Journal of Language and Social Psychology, 18*, 31–48.

Van Bezooijen, R., & Ytsma, J. (1999). Accents of Dutch: Personality impression, divergence, and identifiability. *Belgian Journal of Linguistics*, *13*, 105–129.

Walsh, M., Möbius, B., Wade, T., & Schütze, H. (2010). Multilevel exemplar theory. *Cognitive Science, 34*(4), 537-582. doi: 10.1111/j.1551-6709.2010.01099.x

Watson, C. I., Harrington, J., & Evans, Z. (1998). An acoustic comparison between New Zealand and Australian English vowels. *Australian Journal of Linguistics*, *18*(2), 185–207.

Watson, C. I., Maclagan, M., & Harrington, J. (2000). Acoustic evidence for vowel change in New Zealand English. *Language Variation and Change*, *12*(1), 51–68.

Weatherall, A., Gallois, C., & Pitam, J. (1998). Australasians Identifying Australasian Accents. *Te Reo, 41*, 153–162.

Wedel, A. B. (2006). Exemplar models, evolution and language change. *Linguistic Review*, *23*(3), 247–274. doi: 10.1515/TLR.2006.010

Wells, J. C. (1982). *Accents of English*. Cambridge: Cambridge University Press.

Williams, A., Garrett, P., & Coupland, N. (1999). Dialect recognition. In D. R. Preston (Ed.), *Handbook of perceptual dialectology* (pp. 345–358). Philadelphia: John Benjamins.

Zwicker, E. (1961). Subdivision of the Audible Frequency Range into Critical Bands (Frequenzgruppen). *The Journal of the Acoustical Society of America*, *33*(2), 248–248. doi: 10.1121/1.1908630

Appendix A

Full sentence list with identified target words:

1. I wanted to go horse riding with my brother but I couldn't find my **bit**
2. During the colder months John and I both wanted to stay **fit**
3. I was enjoying my lunch until I took a bite and got a piece of **grit**
4. The new movie was a huge summer **hit**
5. Today I asked my wife what I'd done with my tool **kit**
6. For my birthday my mother gave me a woollen **knit**
7. Before removing the hot tray, I put on an oven **mitt**
8. We worked all afternoon to widen the **pit**
9. The comedian made me laugh with her **skit**
10. I charmed her with my good looks, smiling blue eyes and striking **wit**
11. The morning after, I honestly couldn't recall making the **bet**
12. Tonight I am meeting up with **Brett**
13. The government should do something about student **debt**
14. The plane that flew over us was a **jet**
15. We caught a lobster in our **net**
16. For as long as I could recall I'd wanted a **pet**
17. They wanted her mum to buy a new tea **set**
18. The team they were playing last week was their strongest **threat**
19. She is studying to become a **vet**
20. Her mother told her to keep out of the **wet**
21. Today was the first time Johnny tried to use a baseball **bat**
22. There was no way Warren was old enough to look after a **cat**
23. She called her mum for a short **chat**
24. Over the summer he was surprised to find he'd gotten quite **fat**
25. The bug looked like it was a **gnat**
26. Corrine wore her favourite **hat**
27. We walked all afternoon furiously looking for **Matt**
28. She hated the thought of the tame **rat**
29. The screen was defaced by a broken **slat**
30. She told us the drum was called a **vat**

Appendix B

Sentences in each block of the experiment with their continuum presentation order. Sentence order was randomised for each participant.

Block One		**Block Two**	
Target word	**Presentation order**	**Target word**	**Presentation order**
bit	original	knit	original
fit	original	mitt	original
grit	original	pit	original
hit	original	skit	original
kit	original	wit	original
bet	original	pet	original
Brett	original	set	original
debt	original	threat	original
jet	original	vet	original
net	original	wet	original
bat	original	hat	original
cat	original	Matt	original
chat	original	rat	original
fat	original	slat	original
gnat	original	vat	original
knit	reversed	bit	reversed
mitt	reversed	fit	reversed
pit	reversed	grit	reversed
skit	reversed	hit	reversed
wit	reversed	kit	reversed
pet	reversed	bet	reversed
set	reversed	Brett	reversed
threat	reversed	debt	reversed
vet	reversed	jet	reversed
wet	reversed	net	reversed
hat	reversed	bat	reversed
Matt	reversed	cat	reversed
rat	reversed	chat	reversed
slat	reversed	fat	reversed
vat	reversed	gnat	reversed

Appendix C

Questionnaires.

Questionnaire One

1. Participant Number:

2. Gender (Please circle one): **M** **F**

3. Date of birth:

4. Place of birth:

5. Mother's place of birth and first language:

6. Father's place of birth and first language:

7. Describe the occupation of your parents/guardians:

Female:

Male:

8. Which category best describes these occupations (Please circle one):

Female

a. Unskilled labourer, e.g. fast food employee

b. Skilled labourer, e.g. machine operator

c. Clerical worker, e.g. bookkeeper

d. Manager, e.g. business executive

e. Professional, e.g. lawyer

Male

a. Unskilled labourer, e.g. fast food employee

b. Skilled labourer, e.g. machine operator

c. Clerical worker, e.g. bookkeeper

d. Manager, e.g. business executive

e. Professional, e.g. lawyer

Which category best describes the education level of your parents/guardians:

Female

a. Primary school

b. High school (before Year 10)

c. High school (Year 10)

d. High school (Year 12)

e. Trade certificate or TAFE

f. College Diploma

g. University Degree

h. Postgraduate Degree

Male

a. Primary school

b. High school (before Year 10)

c. High school (Year 10)

d. High school (Year 12)

e. Trade certificate or TAFE

f. College Diploma

g. University Degree

h. Postgraduate Degree

Questionnaire Two

Regarding the sentences you heard in the previous listening exercise, please answer the following questions about the speaker. There is no right or wrong answer, just select what you think is the most accurate.

1. Would you say the speaker is male or female? (please circle one)

M F

2. To which age group would you guess the speaker belongs?

a. 18-25 b. 26-35 c. 36-45 d. 46-55 e. 56-65 f. 65+

3. Where would you guess the speaker is from? (please be as specific as possible)

4. What would you guess is the occupation of the speaker?

a. Unskilled labourer, e.g. fast food employee

b. Skilled labourer, e.g. machine operator

c. Clerical worker, e.g. bookkeeper

d. Manager, e.g. business executive

e. Professional, e.g. lawyer

5. What would you guess is the highest education level achieved by the speaker?

a. Primary school

b. High school

c. Trade certificate or TAFE

d. College Diploma

e. University Degree

f. Postgraduate Degree

1. Participant Number:

2. Have you ever been to New Zealand?

No Yes

For how long total?

When were you last there?

3. Are you personally acquainted with many New Zealanders?

No Yes

Give relationship:

4. How many New Zealanders would you come into personal contact with during an average week?

None 1–5 5–10 10+

5. Which of the following sentences best describes you:

I never watch TV.

I sometimes watch TV.

I watch a lot of TV.

6. Can you name any TV shows, movies or online media you have watched recently that were made in New Zealand?

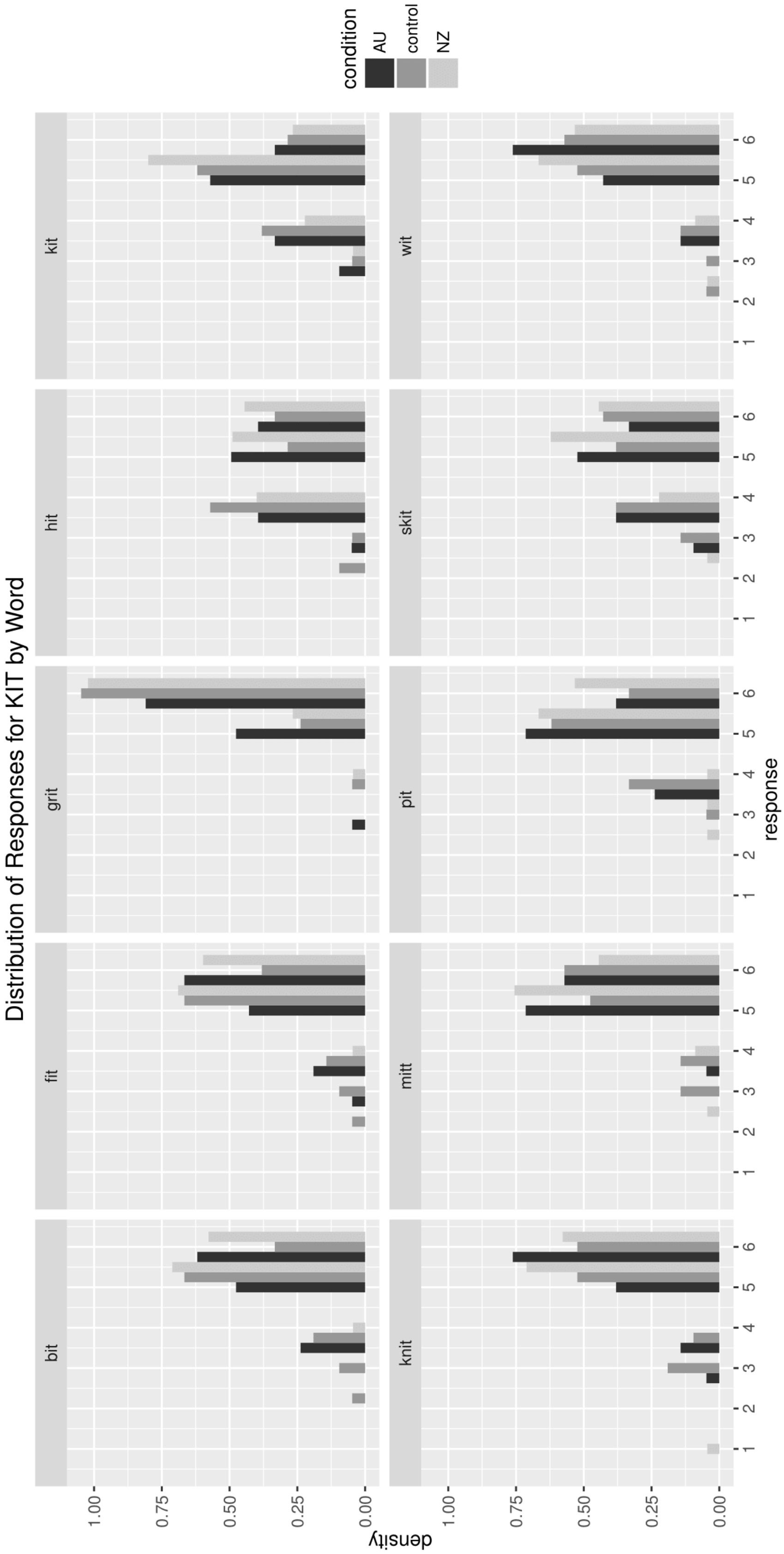
7. Do you follow international sport?

No Yes

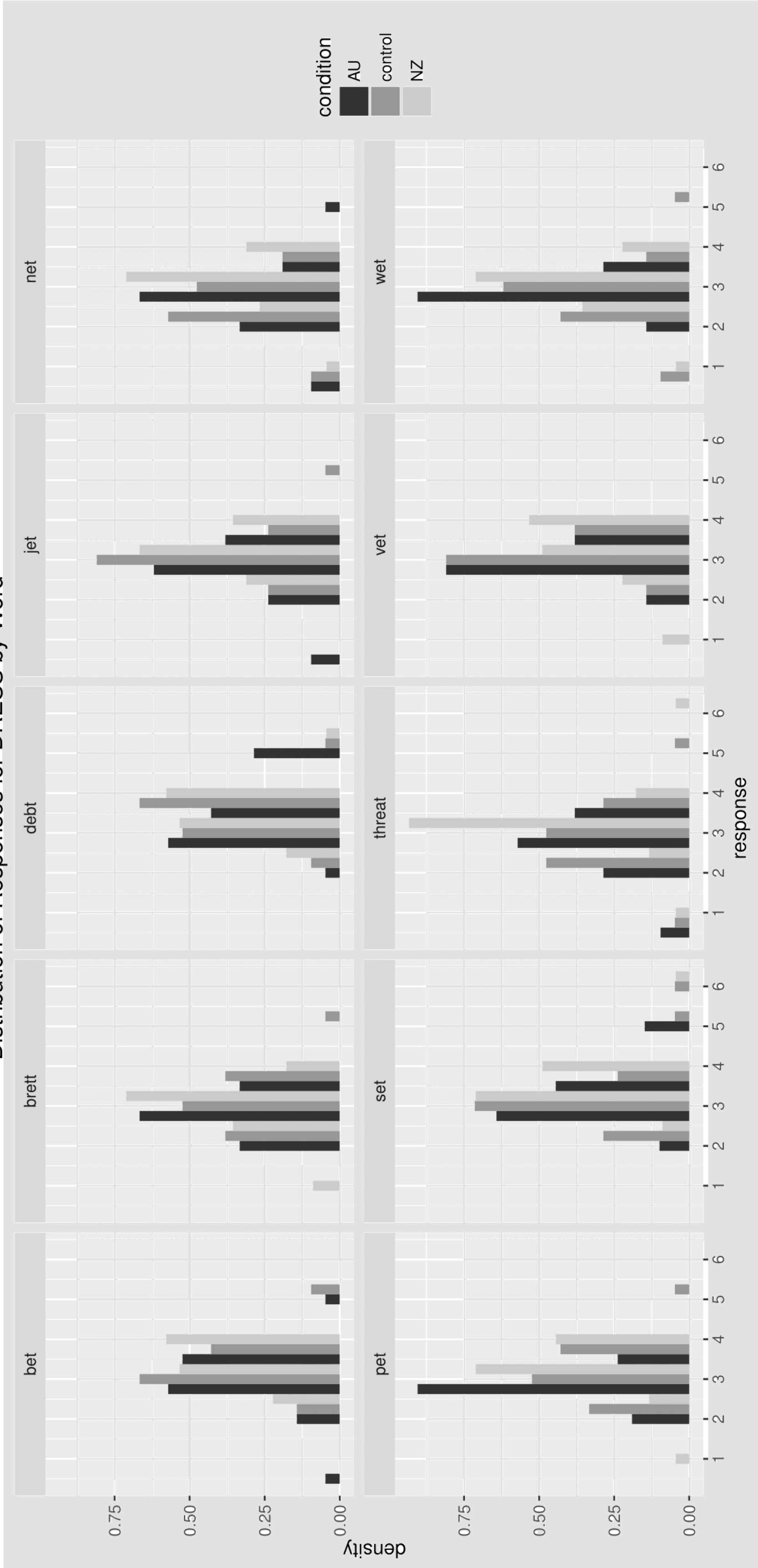
Which sports?

Appendix D

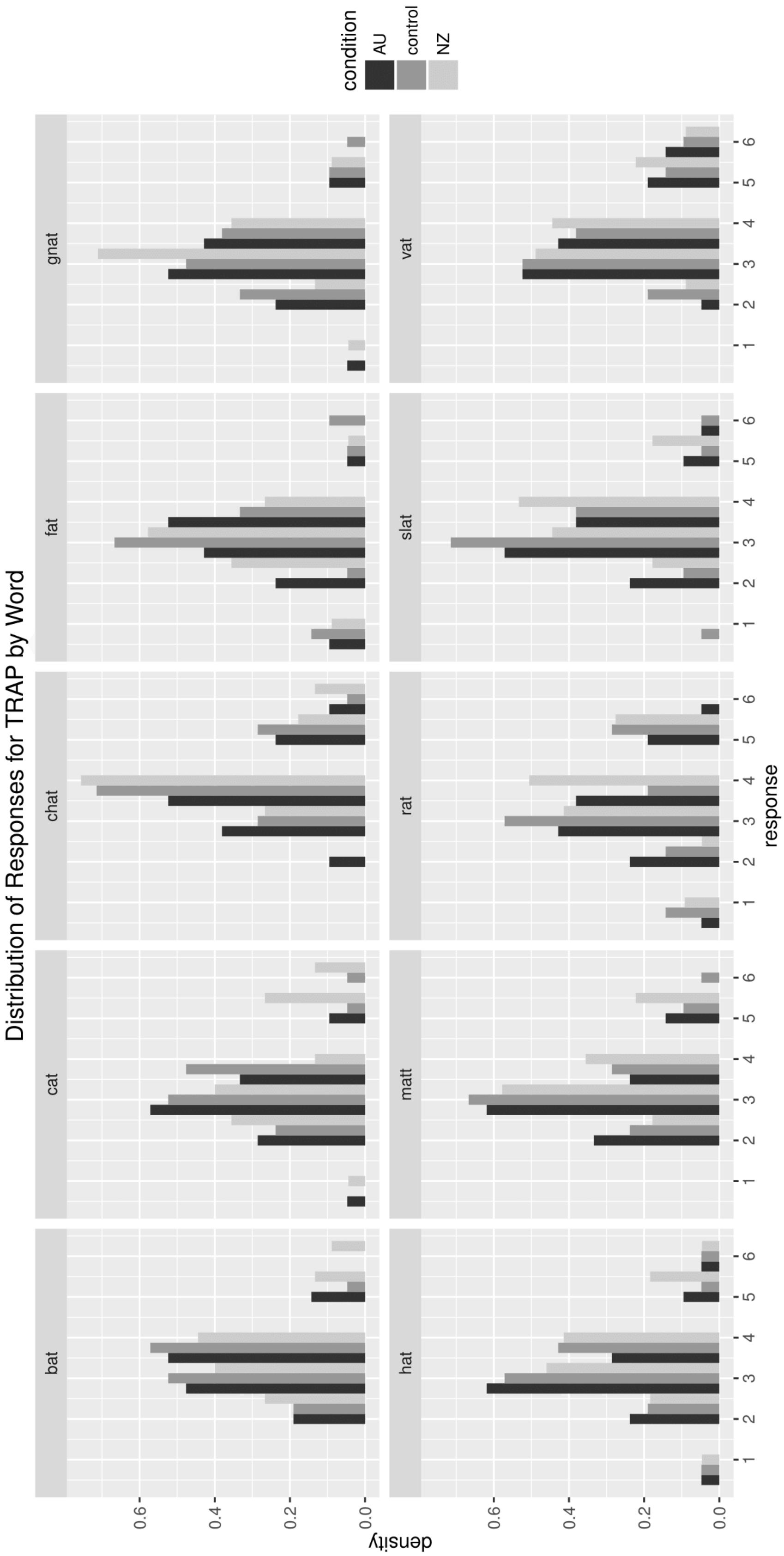
By-word distribution of response plots, KIT target words.



By-word distribution of response plots, DRESS target words.



By-word distribution of response plots, TRAP target words.



Appendix E

Dear Associate Professor Cox,

Re: "The effect of cultural reference on speech perception" (5201600056)

Thank you very much for your response. Your response has addressed the issues raised by the Faculty of Human Sciences Human Research Ethics Sub-Committee and approval has been granted, effective 9th March 2016. This email constitutes ethical approval only.

This research meets the requirements of the National Statement on Ethical Conduct in Human Research (2007). The National Statement is available at the following web site:

http://www.nhmrc.gov.au/_files_nhmrc/publications/attachments/e72.pdf.

The following personnel are authorised to conduct this research:

Associate Professor Felicity Cox
Dr Anita Szakay
Mr Michael Stuart Walker

Please note the following standard requirements of approval:

1. The approval of this project is conditional upon your continuing compliance with the National Statement on Ethical Conduct in Human Research (2007).

2. Approval will be for a period of five (5) years subject to the provision of annual reports.

Progress Report 1 Due: 9th March 2017
Progress Report 2 Due: 9th March 2018
Progress Report 3 Due: 9th March 2019
Progress Report 4 Due: 9th March 2020
Final Report Due: 9th March 2021

NB. If you complete the work earlier than you had planned you must submit a Final Report as soon as the work is completed. If the project has been discontinued or not commenced for any reason, you are also required to submit a Final Report for the project.

Progress reports and Final Reports are available at the following website:

http://www.research.mq.edu.au/current_research_staff/human_research_ethics/application_resources

3. If the project has run for more than five (5) years you cannot renew approval for the project. You will need to complete and submit a Final Report and submit a new application for the project. (The five year limit on renewal of approvals allows the Sub-Committee to fully re-review research in an environment where legislation, guidelines and requirements are continually changing, for example, new child protection and privacy laws).

4. All amendments to the project must be reviewed and approved by the Sub-Committee before implementation. Please complete and submit a Request for Amendment Form available at the following website:

http://www.research.mq.edu.au/current_research_staff/human_research_ethics/managing_approved_research_projects

5. Please notify the Sub-Committee immediately in the event of any adverse effects on participants or of any unforeseen events that affect the continued ethical acceptability of the project.

6. At all times you are responsible for the ethical conduct of your research in accordance with the guidelines established by the University. This information is available at the following websites:

http://www.mq.edu.au/policy

http://www.research.mq.edu.au/for/researchers/how_to_obtain_ethics_approval/human_research_ethics/policy

If you will be applying for or have applied for internal or external funding for the above project it is your responsibility to provide the Macquarie University's Research Grants Management Assistant with a copy of this email as soon as possible. Internal and External funding agencies will not be informed that you have approval for your project and funds will not be released until the Research Grants Management Assistant has received a copy of this email.

If you need to provide a hard copy letter of approval to an external organisation as evidence that you have approval, please do not hesitate to contact the Ethics Secretariat at the address below.

Please retain a copy of this email as this is your official notification of ethics approval.

Yours sincerely,

Dr Anthony Miller
Chair
Faculty of Human Sciences

Human Research Ethics Sub-Committee

Faculty of Human Sciences - Ethics
Research Office
Level 3, Research HUB, Building C5C
Macquarie University
NSW 2109

Ph: +61 2 9850 4197
Email: fhs.ethics@mq.edu.au
http://www.research.mq.edu.au/

Department of Linguistics
Faculty of Human Sciences
MACQUARIE UNIVERSITY NSW 2109

Phone: +61 (0)2 9850 8740
Fax: +61 (0)2 9850 9199
Email: lingadmin@mq.edu.au

Chief Investigator's / Supervisor: Associate Professor Felicity Cox

Participant Information and Consent Form

Name of Project: Perception of artificial speech.

You are invited to participate in a study on synthesised speech. The purpose of the study is to determine the accuracy of artificial speech when compared with recorded human speech. If you decide to participate, your voice will be recorded while you read a set of words and sentences from a computer screen. You will be asked to participate in a perception task where you will match recorded and artificial speech samples. You will also be asked to complete two short background questionnaires. All tasks will be explained before you start.

The recording and perception session will take about 1 hour. For your participation you will receive 5% research credit.

This study is being conducted by:
Michael Walker (04 8899 3551)
Department of Linguistics
michael.walker@students.mq.edu.au

as being conducted to meet the requirements of the Masters of Research under the supervision of:

Associate Professor Felicity Cox (02 9850 8767)
Centre for Language Sciences
Department of Linguistics
felicity.cox@mq.edu.au

Dr Anita Szakay (02 9850 4155)
Centre for Language Sciences
Department of Linguistics
anita.szakay@mq.edu.au

Any information or personal details gathered in the course of the study are confidential, except as required by law. No individual will be identified in any publication of the results. The material collected will be used for research and teaching purposes and any information collected during this experiment will remain confidential and no personal identifying information will be used on records or in publication. Excerpts from your audio recordings may be used for illustrative purposes in teaching, in conference presentations or in various relevant electronic media but you will not be identified if excerpts from your recordings are used for any of these purposes. The data may be used in the future for other speech research projects and included in the Australian National Corpus (http://www.ausnc.org.au/) but you will never be identified by name to anyone except the researchers directly involved with this project.

A summary of the results of the data can be made available to you on request by contacting Michael Walker or if you have any questions about this study please contact me and I will be happy to discuss it further with you.

Participation in this study is entirely voluntary: you are not obliged to participate and if you decide to participate, you are free to withdraw at any time without having to give a reason and without consequence.

I, have and understand the information above and any questions I have asked have been answered to my satisfaction. I agree to participate in this research, knowing that I can withdraw from further participation in the research at any time without penalty or loss of research credit. I understand that excerpts of the audio recordings may be used for illustrative purposes by the researchers in teaching, in conference presentations and in various electronic media. I have been given a copy of this form to keep.

Participant's Name:____
(Block letters)

Participant's Signature: ______________________ Date:__________________

Investigator's Name:____
(Block letters)

Investigator's Signature:___________________ __ Date:__________________

The ethical aspects of this study have been approved by the Macquarie University Human Research Ethics Committee. If you have any complaints or reservations about any ethical aspect of your participation in this research, you may contact the Committee through the Director, Research Ethics & Integrity (telephone (02) 9850 7854; email ethics@mq.edu.au). Any complaint you make will be treated in confidence and investigated, and you will be informed of the outcome.

(INVESTIGATOR'S [OR PARTICIPANT'S] COPY)