Dialect divergence and convergence in New Zealand English

Molly Babel
Department of Linguistics
University of British Columbia and University of California, Berkeley

INTRODUCTION

As people learn to speak, they acquire the language and dialect spoken around them. The exact forms of every level of linguistic representation – from syntax to lexical choice to pronunciation – are all determined by the patterns of the ambient language. This is done with great ease as young language learners. It has been documented, however, that after a certain age children are unable to fully acquire particularly complex phonological aspects of a second dialect (Payne 1980, Chambers 1992). Trudgill (1986:58) argues that this is due to the fact that accommodation of sound structure in dialect contact is "*phonetic* rather than phonological" (emphasis in original).

Despite the purported difficulty in acquiring a new dialect, research on adults exposed to new dialects demonstrates that ways of speaking do indeed change. For example, Munro, Derwing, and Flege (1999) found that accents of Canadians living in Alabama were perceived by listeners as sounding intermediate between those of Canadians living in Canada and native Alabamans. A series of studies have focused on longitudinal changes within the speech of a single speaker over a fifty year span. Harrington (2006, 2007) and Harrington, Palethrope, and Watson (2000a, b) have analyzed vowel changes in the speech of the Queen Elizabeth II. The recordings under analysis were her yearly Christmas broadcasts from 1952 through 2002. After accounting for maturational changes in the vocal tract, this work has demonstrated that the Queen's vowels shifted in the direction of Southern British English, away from Received

Pronunciation. Evans and Iverson (2007) examined changes in both perception and production in new dialect exposure. College students from a northern English dialect (Sheffield) were interviewed at the commencement of university studies, three months after the beginning of studies, after one year of study, and after two years of study at a southern English university where Standard Southern British English (SSBE) was spoken. SSBE has /ʌ/ in *bud* and *cud* and /a/ in *bath*. Northern varieties of English use the vowel /ʊ/ for *bud* and *cud* and /ɑ/ for *bath*. The authors found that both *bud* and *cud* became more centralized. In the northern dialects, *cud* and *could* are homophonous, both having the vowel /ʊ/. *Could* has this vowel in the southern dialects, but *cud* has /ʌ/. After their time at the university, the participants began to centralize the vowel for *could* as well. Participants were also rated on a 10-point scale from 'very northern' to 'very southern'. Overall participants were rated as sounding more southern after their time at the university. In terms of changes in perception, participants who were rated as having maintained a more northern accent did in fact choose more northern perceptual exemplars.

Delvaux and Soquet (2007) provide evidence demonstrating that speakers can shift from one dialect to another after very brief periods of exposure. Female speakers of the Mons dialect of Belgian French were exposed to a the voice of a female model talker from the Liège dialect of Belgian French. After auditory exposure to the model talker's voice, the Mons speakers modified their pronunciations of /o/ and /ɛ/ in a sentence production task.

In addition to these findings that speakers acquire new dialect patterns as adults, several laboratory-based studies have found that individuals acquire the phonetic characteristics of within-dialect model talkers in speech tasks. In what is known as a *shadowing task* or an *auditory naming task* where participants simply repeat single words after a model talker, Goldinger (1998) found that when participants repeat words, they acquire phonetic aspects of the model talker's voice. This general finding has been well replicated (Namy, Nygaard, & Sauerteig 2002, Goldinger & Azuma 2004). In the completion of a conversational map task, speakers were also found to converge phonetically on each other's productions; the direction of the convergence though was influenced by several social factors, namely gender and participant role (Pardo 2006). In Pardo's study, male dyads converged more than female dyads. As far as participant role, women were found to converge toward the speaker who was receiving instructions while men converged toward the speaker who was giving instructions. In the speech science literature, there has also been some investigation of what exactly can be or is being imitated. Shockley, Sabadini, and Fowler (2004) and Nielsen (2008) find that speakers imitate the voice onset time of aspirated stops. Recently, Babel (2009) found that phonetic accommodation of vowel formants is selective. In that study, the American English vowels /i æ ɑ o u/ were examined for imitative behaviors; only /æ/ and /ɑ/ exhibited shifts toward the model talker in an auditory naming task.

There is ample evidence demonstrating that phonetic accommodation occurs in language behavior. It is currently under great debate as to *why* this happens. Psycholinguists who study the mechanisms of speech behavior have developed models to account for accommodative behaviors in language. The interactive alignment model is a

popular model proposed by Pickering and Garrod (2004) to account for speech accommodation. In this model every level of linguistic representation – the situation model, semantics, syntax, the lexicon, phonology, and phonetics – is connected within an individual, and each level of representation between the listener and the talker is connected. An automatic priming process that percolates through the levels of representations of the interlocutors aligns speech. The use of a representation by a talker leads to the activation of that representation in the listener, and that activation leads to increased incidence of use.

Recently, sociolinguists have contributed to the discussion of why accommodation happens. Traditionally, research on dialect contact and acquisition has remained agnostic regarding the mechanisms and motivations for accommodation. Trudgill (2008:252) argues that "accommodation is not only a subconscious but also a deeply automatic process." He reaches this conclusion after reviewing four cases of European languages forming new varieties as a result of dialect contact and he discards the theory that new dialects arise as a result of the formation of new national identities:

> ...if a common identity is promoted through language, then this happens as a consequence of accommodation; it is not its driving force. Identity is not a powerful enough driving force to account for the emergence of new, mixed dialects by accommodation. It is parasitic upon accommodation, and is chronologically subsequent to it (Trudgill 2008:251).

It is striking to see Trudgill eschewing social factors as a palpable force in language change. Trudgill's claim has been criticized extensively (Bauer 2008, Mufwene 2008, Tuten 2008, Schneider 2008, Coupland 2008, Holmes & Kerswill 2008). Bauer, for example, points out there are instances where individuals have lived for a considerable amount of time in a new dialect area without displaying significant accommodation.

Trudgill's hypothesis makes a prediction that is testable: if linguistic accommodation is automatic, it should not only be demonstrable in a laboratory, but everyone should do it. Some evidence may already exist to counter Trudgill's claim. Communication Accommodation Theory (Giles 1973, Giles & Coupland 1991, Giles, Coupland, & Coupland 1991, Shepard, Giles, & LePoire 2001) has long argued that language choices are a function of social distance. Linguistic convergence lessens the social distance between interlocutors while divergence maximizes it. In particular, one study poses a challenge for the automatic accommodation model. Bourhis and Giles (1977) is the hallmark example of dialect divergence and served as inspiration for the study presented in this paper. Bourhis and Giles examined accent convergence and divergence in two groups of Welsh-born adults. A group of Welsh adults who attended both Welsh language and Welsh culture classes were found to diverge from an out-group speaker of Received Pronunciation (RP) by adopting a Welsh-accented dialect. The RP interviewer had questioned the vitality and function of the Welsh language in modern times. The second group of Welsh adults participating in the experiment also attended Welsh language classes, but only on business time with the explicit goal of furthering their careers. These adults were found to converge with the RP interviewer. The

perceived changes in the accents of the adults were measured on an 11-point scale by two judges who were not linguistically trained and naive to the experiment. Crucially, however, the judgments of convergence and divergence came from instances of running speech. This means that the perceptual judgments could have been made based on lexical items, syntactic structure, or phonetic features. For example, one participant in the experiment who was judged to have diverged from the RP speaker was heard conjugating Welsh verbs. While this type of behavior is clearly divergent, it is distinct from a speaker using a slightly different pronunciation to socially distance themself.

The goal of this paper is to attempt a replication of Bourhis and Giles looking at phonetic convergence and divergence. When an individual wants to socially distance themself from another speaker, do they do it phonetically? To answer this question, a speech production task was designed using the voice of a male Australian as that of the model talker (like the RP speaker in the Bourhis and Giles study) and New Zealand participants. Australian and New Zealand Englishes (AuE and NZE) share many basic dialect features (Bauer, Warren, Bardsley, Kennedy, & Major 2007, Cox & Palethorpe 2007), but there are several key differences in the front vowel monophthongs (Watson, Harrington, & Evans 1998, Easton & Bauer 2000). Ongoing sound changes in NZE have made the front vowel space particularly distinct in the two dialects (Maclagan & Hay 2007). In NZE, for instance, the vowels DRESS and TRAP are raising.[1] The KIT vowel in AuE is raising, while in NZE it is centralizing. Not all of these sound changes are salient differences within the NZ community. Bayard (2000) states that KIT is the most salient difference between NZE and AuE. Hay, Nolan, and Drager (2006) examined NZE listeners' sensitivity to the salient differences in KIT, TRAP, and DRESS in NZE

[1] Following Wells (1982), lexical sets will be used to refer to the vowel categories in this paper.

and AuE; they found equally strong results for KIT and TRAP, but NZE listeners did not behave as though they were explicitly aware of the differences between DRESS in these two dialects. It is important to be aware of the salient differences across the two dialects as Trudgill (1981) argues that only socially salient variables are susceptible to accommodation. Following this, in interactions between NZE and AuE speakers, we predict that only TRAP and KIT would exhibit accommodation.

METHODOLOGY

In the following paragraphs I describe the speech production task used to explore phonetic convergence and divergence in NZE. Following that, I provide the methodology for the Implicit Association Task that measured each NZ participants' inherent bias toward Australia. Participants' scores on this task were used in the statistical analysis described in the ANALYSIS & RESULTS section.

Speech Production Task

Participants (females = 34, males = 8) from the Victoria University of Wellington community completed an auditory naming task. The task took approximately 30 minutes and participants were compensated with a $10 book voucher. The auditory stimuli were monosyllabic single word productions from a male talker who was born and raised in Melbourne, Australia. Monophthongal stimuli were words taken from the lexical sets KIT, DRESS, TRAP, BARN, STRUT, and THOUGHT. In addition to the target monophthongs, participants were also presented with diphthongs involved in the NEAR/SQUARE merger in New Zealand English (Hay, Warren, & Drager 2006). Only

the results from the monophthongs are presented in this paper. The full list of monophthongal words used in the task appears in the Appendix.

Participants were seated at a PC laptop and the experiment was presented using E-Prime 2.0 Experimental Software (Schneider, Eschman, & Zuccolotto 2002). Auditory stimuli were presented over AKG K271 headphones. Audio-recording was done directly in E-prime using an M-Audio USB audio device with a head-mounted AKG C520 microphone positioned three inches from the participant's mouth.

The task was designed as follows: Participants were randomly presented with hVd words (*hid*, *had*, *head*, etc.) which they were to read aloud. In the next block, participants were presented with the target word list which they were asked to produce aloud. The words in the list were presented in a different random order for each participant; this is referred to as the pre-task block. The purpose of the pre-task block is to obtain a baseline production of how a participant produces each word. The following block was the shadowing block where participants were exposed to the target word productions from the Australian model talker over headphones. Words were randomly presented twice through the course of the test block. Participants' instructions for this part of the task are to identify the word heard by saying it out loud. Finally, participants did a post-task reading of the wordlist; this block was identical to the pre-task block, except that the words were presented in a different random order. Finally, participants re-read the hVd words again. The methodology of the auditory naming task follows that of Goldinger (1998) and Namy, Nygaard, and Sauerteig (2002). In comparing pre-task and test-block productions, we can see how NZ participants modify productions as a result of exposure to the Australian model talker. In this task participants were assigned to one of

two conditions. In the POSITIVE CONDITION, participants were presented with the following text which was intended to make them view the talker and Australia as a whole in a positive light:

> The Australian talker you are about to hear was actually born in Auckland. At a young age, however, he and his parents moved to Melbourne where he has lived since. His grandparents and the rest of his extended family still live in New Zealand, so he visits frequently. In fact, he is currently looking for employment in New Zealand so that his children may live closer to their great-grandparents.

The other condition was a NEGATIVE CONDITION. The purpose of this condition was to inspire negative feelings toward the talker and Australia.

> The Australian talker you are about to hear was born in Sydney. Like many Australians, he has strong negative opinions of New Zealand. For one, he thinks that New Zealanders are rather stupid and that they lack culture. In addition, he finds the entire population backwards and naïve. In his mind, New Zealand is provincial and has a horrid cricket team. He never intends to visit New Zealand because of these views.

In both conditions participants were exposed to a screen with the assigned text immediately before beginning the test-block. After reading the Positive or Negative text

participants pressed a button that took them to the test-block. Male and female participants were evenly assigned to the two conditions.

Upon completion of the speech production task all participants took an Implicit Association Task. This task is described in the next section.

Implicit Association Task

Traditionally in sociolinguistics, attitudes are elicited explicitly through surveys or questionnaires. One goal of this project was to examine how implicit and subconscious attitudinal measures may predict speech behavior. To this end, an Implicit Association Task (IAT; Greenwald, McGhee, & Schwarz 1998) was used. The IAT is a standard social psychology tool that uses reaction time in category classification to measure implicit biases.

There are five blocks in this task. The first block is target-concept discrimination. The targets AUSTRALIA and NEW ZEALAND were presented on opposite sides of the monitor. A combination of Australian or New Zealand concepts – famous individuals, maps, and images (flags, native scenery, sports emblems, etc.) – were then randomly presented (e.g. an image of a kangaroo or an image of a kiwi) in the middle of the screen. A participant's task is to categorize the concepts as AUSTRALIA or NEW ZEALAND as quickly as possible. The second block is associated attribute discrimination. Here, the attributes *good* and *bad* are presented on opposite sides of the computer screen in place of AUSTRALIA and NEW ZEALAND. Attribute words are presented randomly (e.g. *rainbow* or *cancer*) in the middle of the screen. Participants categorize the words as semantically *good* or *bad* words. The third block is a combined test block. Labels for the

concept categories (AUSTRALIA vs. NZ) and word (*good* vs. *bad*) attributes are presented at the top corners of the screen. In the center, either a concept or a word are randomly presented and must be categorized. Participants are instructed to ignore the target-concept when categorizing words and ignore the attributes when categorizing concepts. Concepts that are words are presented in all capital letters and *good/bad* words are presented in all lowercase letters to facilitate the process. Block 4 is just like Block 1, except that the labels AUSTRALIA and NEW ZEALAND were presented on different sides of the screen (so, if AUSTRALIA was on the right-side of the screen in Block 1, it was on the left side in Block 4). Participants then categorized concepts as AUSTRALIA or NEW ZEALAND as they did in Block 1. Block 5 is the reversed combined task; the reversed order of the target-concepts (AUSTRALIA and NEW ZEALAND) are matched up with the original order of *good* and *bad* such that if AUSTRALIA was originally presented above *good* and NEW ZEALAND with *bad*, this pattern is reversed and Australia is presented with *bad* and New Zealand with *good*. The experiment was counterbalanced so that half of the participants were initially exposed to AUSTRALIA paired with *good* and NEW ZEALAND paired with *bad* while the other half were first presented with AUSTRALIA paired with *bad* and NEW ZEALAND paired with *good*.

Participants logged responses using assigned buttons on a computer keyboard. Responses were collected automatically using E-prime. Participants' scores were calculated using the updated methods described in Greenwald, Nosek, and Banaji (2003). After removing outliers, the mean reaction time was calculated for each participant based on correct responses for each block. One standard deviation was also calculated for each block. Then, each response error was replaced with the block mean and a 600 ms penalty.

Means were then re-calculated for each block and the difference between these two blocks was computed. Finally, to get the IAT score, the difference was divided by the standard deviation previously calculated. These values were used as predictors in the statistical model described below.

ANALYSIS & RESULTS

Analysis

Vowels from the participant productions were hand-marked. A Praat (Boersma & Weenink 2005) script calculated the average first and second formant over the middle 50% of the vowel. Obvious outliers were hand-corrected. In order to minimize physiological differences between participants so that the analysis can focus on speech production differences in dialect and style, it is necessary to normalize the vowel formants. As per Adank, Smits, and van Hout (2004), the Lobanov normalization method was used (Lobanov 1971).

With the normalized formant values, the Euclidean distance between each NZ word production and the same word from the Australian model talker was calculated. Euclidean distance is a way of measuring distance; essentially, it measures distance as the crow flies. This means that for each word from each block for each participant there is a single distance measure. To understand how NZ participants' productions changed as a result of exposure to the model talker, the distance for each word from the test-block and post-task blocks were subtracted from the pre-task block distances, creating a *difference in distance* metric. This *difference in distance* measure is indicative of *how much* a participant modified their phonetic distance to the speech of the model talker. A negative

value indicates the distance between the NZ participant and the Australian model has shrunk, indicating convergence. A positive value signals the phonetic distance has grown; this would mean there was divergence.

Results from the Shadowed Productions

A stepwise hierarchical linear regression analysis using only the data from the block of shadowed productions was run. With this type of analysis all potential main effects and interactions are evaluated against one another. The model automatically selects the predictor variables that account for larger proportions of variance in the data. The *difference in distance* metric was used as the dependent variable while vowel category (BARN, DRESS, KIT, TRAP, STRUT, THOUGHT), participant gender (male or female), experimental condition (Positive or Negative), participant age, word frequency, and IAT score were entered as independent predictors. Word frequency was determined by logarithmic frequency counts in CELEX (Baayen, Piepenbrock, & van Rijn 1993). The variables that were used by the model are shown in Table 1. The model selected Vowel, participants' IAT score, word frequency, and the Vowel by IAT interactions as predictors in the model [F (12, 2037) = 6.2, $p < 0.001$]. The negative intercept in Table 1 indicates that on average participants converged with the model talker; the *difference in distance* was less in the shadowed tokens than in the pre-task tokens.

	β	Standard Error	*t*-value
Intercept	-0.179455	0.025179	-7.127***
DRESS	-0.049902	0.027372	-1.823
KIT	0.040225	0.027291	1.474
STRUT	0.063596	0.024102	2.639**
THOUGHT	0.071848	0.025304	2.839**
TRAP	0.056860	0.023731	2.396*
IAT	0.089247	0.030549	2.921**
word frequency	0.045078	0.011709	3.850***
DRESS * IAT	0.060010	0.049759	1.206
KIT * IAT	-0.129701	0.049571	-2.616**
STRUT * IAT	-0.086119	0.043395	-1.985*
THOUGHT * IAT	0.007461	0.045661	0.163
TRAP * IAT	-0.059030	0.043033	-1.372

Table 1. Results of a stepwise hierarchical linear regression used to predict *difference in distance* values from the shadowed tokens. Symbols following the *t*-values indicate the associated *p*-value: '***' $p < 0.001$, '**' $p < 0.01$, and '*' $p < 0.05$.

Participants' IAT score and word frequency have positive *β* coefficients in Table 1. This means that the *difference in distance* value increased as these values increased. For word frequency, this is in accordance with previous work; lower frequency words exhibit more imitation than high frequency words (Goldinger 1998). In terms of the IAT score, the positive coefficient demonstrates that participants who scored as pro-Australian were more likely to accommodate toward the vowels of the Australian model talker. The Condition to which participants were assigned did not show up in the regression model. In addition, post-hoc analysis found no reliable effect of the Positive or Negative Conditions on *difference in distance* values. Figure 1 presents participants' IAT scores plotted against their average *difference in distance* value. Again, these results show that participants who scored as pro-Australian in the IAT were significantly more likely to spontaneously accommodate toward the vowels of the Australian talker [$t(40) = 2.2$, $p < 0.05$, Pearson's $R = 0.32$]. In Figure 1, each "positive" or "negative" data point represents one participant and indicates which Condition the participant was assigned to.

This method of presenting the data also reveals the trend – albeit insignificant – that participants in the Negative Condition scored as more pro-New Zealand in the IAT than those participants randomly assigned to the Positive Condition.

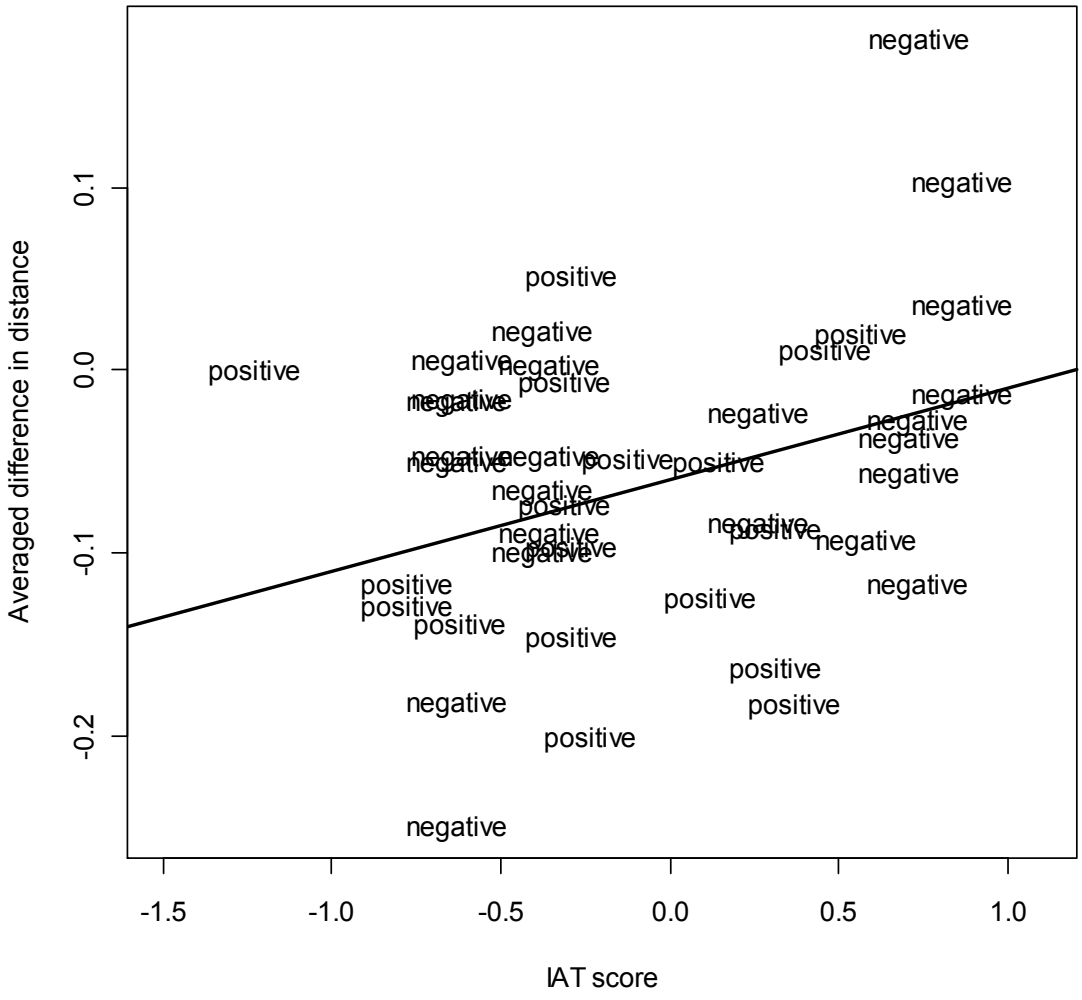

Figure 1: Each participant's averaged *difference in distance* plotted against their IAT score. The *Difference in distance* measure on the *y*-axis indicates the amount of phonetic accommodation. A negative value demonstrates phonetic convergence while a positive value demonstrates phonetic divergence. On the *x*-axis, a negative IAT score indicates a pro-Australian bias and a positive score a pro-New Zealand bias. The negatively skewed regression line in the figure demonstrates that a participant was more likely to phonetically accommodate to the Australian talker when they scored with a pro-Australian bias on the IAT.

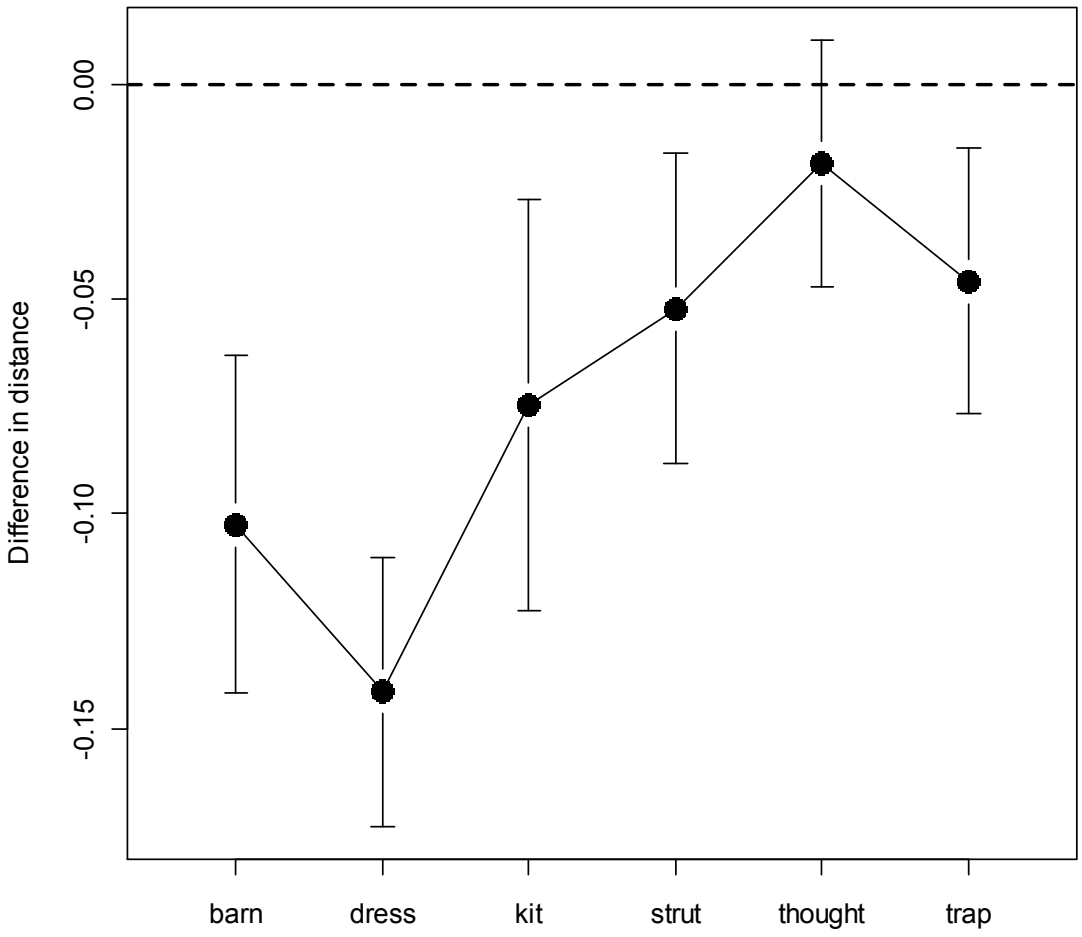

Figure 2. Amount of phonetic accommodation to the vowels averaged across all participants in both the Positive and Negative Conditions. The *difference in distance* measure on the *y*-axis indicates the amount of phonetic accommodation. A negative value demonstrates phonetic convergence while a positive value demonstrates phonetic divergence. The dashed-line marks the cut-off point for accommodation; all data below the line is indicative of phonetic convergence. The error bars represent 95% confidence intervals.

While the overall pattern showed convergence, participants did not accommodate toward all vowels to the same extent. Figure 2 presents *difference in distance* values for each vowel averaged across all participants. Any value below zero – marked by the

dashed line in the figure – indicates some degree of phonetic accommodation toward the vowels of the AuE talker. Negative *difference in distance* values for all vowels show that there was convergence toward the model talker and no patterns of divergence in the data set as a whole. Looking at Figure 2, it is also apparent that participants respond to vowels in a non-uniform manner. For example, it seems that DRESS elicited the most accommodation. Indeed, post-hoc analyses found that the DRESS vowel was imitated more than STRUT ($p > 0.05$), THOUGHT ($p > 0.001$), and TRAP ($p > 0.01$). BARN was also imitated more than THOUGHT ($p > 0.05$).[2]

Figures 3 and 4 plot the averaged normalized vowel formants for female and male participants, respectively. In both figures the words printed in the larger font represent the averaged vowel formants for the lexical sets from the Australian talker. The smaller words in black are averaged from participants' pre-task productions. The smaller words in black italics are the shadowed productions and those in small gray italics are productions from the post-task block. These figures depict the direction of the accommodation. They also demonstrate trends toward different behaviors between men and women that did not come out as significant in the model, but are interesting nevertheless. For female participants (Figure 3), all vowels in the test block shifted in the direction of the Australian talker. Male participants (Figure 4), however, display this same pattern of drift toward the Australian talker for DRESS, BARN, STRUT, and THOUGHT, but demonstrate divergence or maintenance for KIT and TRAP. TRAP in particular seems to be diverging from that of the AuE model speaker in this figure. The lowering of the first formant in the test block is a shift in the direction of the ongoing sound change in NZE (Maclagan & Hay 2007). This data is only based on the speech of

[2] Post-hoc analyses were Tukey's Honest Significant Difference tests.

eight males, but it is important and interesting to observe the pattern for TRAP and KIT as these are the vowel sets documented as the most salient distinguishers of NZ and Au Englishes.

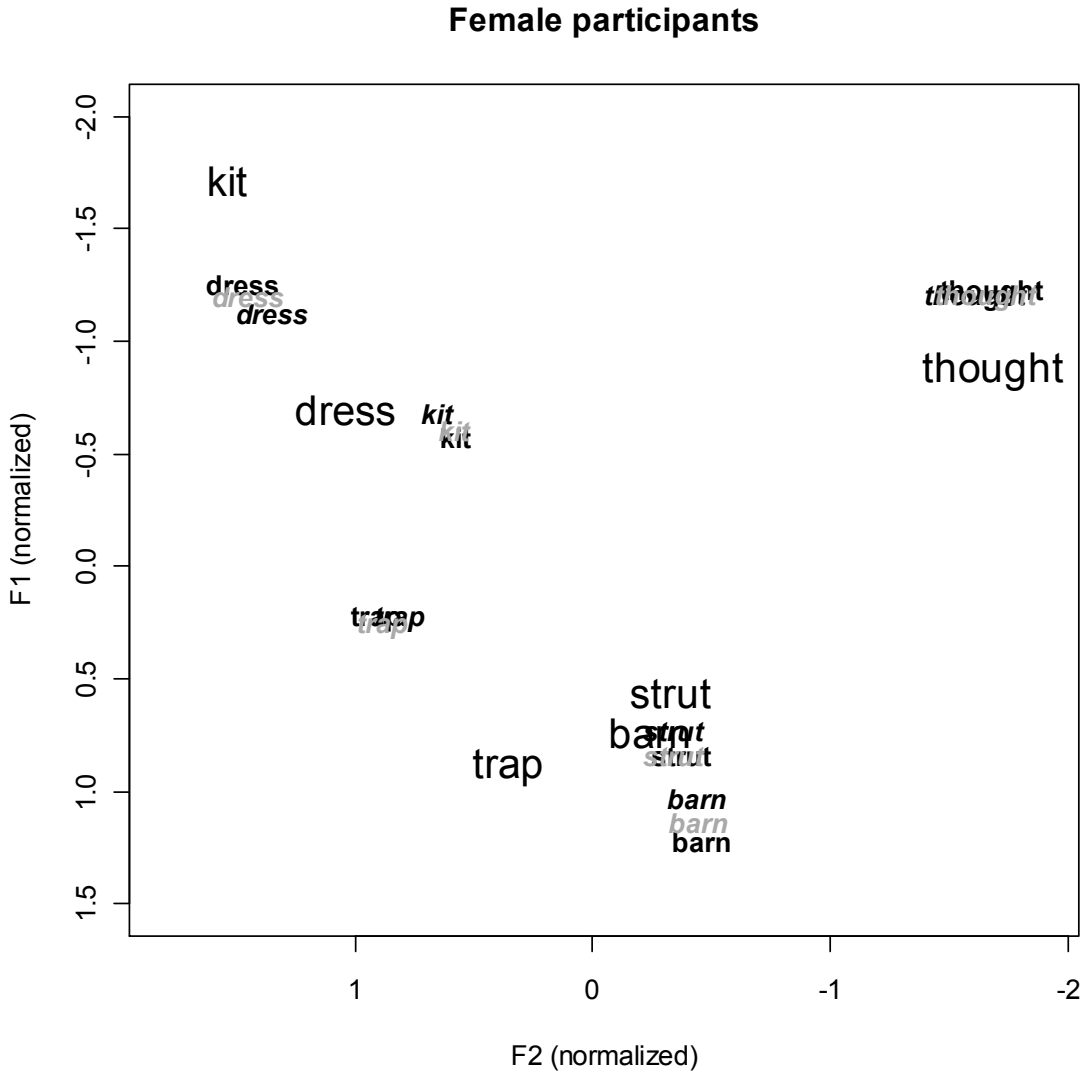

Figure 3: Formant plot displaying phonetic accommodation for the female participants. The Australian model talker's mean vowels are plotted in the larger black typeface. Female participants' pre-task productions are in a small black font, shadowed productions are in black italics, and post-task productions and in gray italics. The shadowed productions move in the direction of the model talker's tokens and the post-task productions generally lie somewhere between the pre-task and shadowed productions.

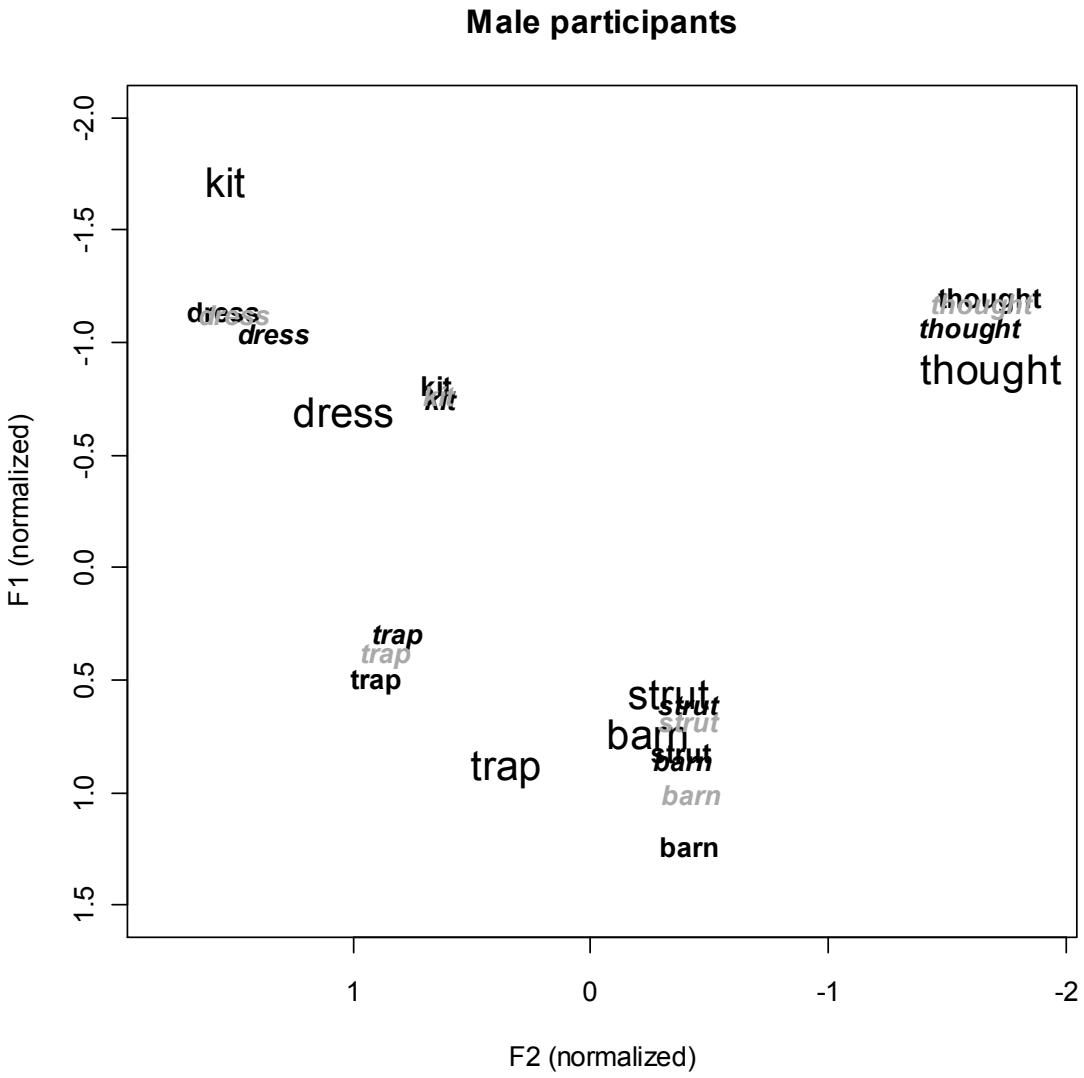

Figure 4. Formant plot displaying phonetic accommodation for the male participants. The Australian model talker's mean vowels are plotted in the larger black typeface. Male participants' pre-task productions are in a small black font, shadowed productions are in black italics, and post-task productions and in gray italics. The shadowed productions move in the direction of the model talker's tokens and the post-task productions generally lie somewhere between the pre-task and shadowed productions.

Results from the Post-task Productions

For the post-task productions, another stepwise hierarchical linear regression analysis was used. Again, the *difference in distance* metric was used as the dependent variable while vowel category (BARN, DRESS, KIT, TRAP, STRUT, THOUGHT), participant gender (male or female), experimental condition (Positive or Negative), participant age, word frequency, and IAT score were entered as the predictor variables. The variables chosen by the model are summarized in Table 2. The model selected IAT scores, Experiment Condition (with Positive as the model's default reference variable), word frequency, and participant age [F (4, 1039) = 4.3, $p < 0.05$]. Like what was found for the shadowed productions, the negative intercept indicates that overall the *difference in distance* values in the post-task block were smaller than those in the pre-task. After exposure to the Australian model talker ceased, NZ participants retained the more Australian-like vowels they had acquired during the shadowing block.

	β	Standard Error	t-value
Intercept	-0.152895	0.048217	-3.171**
IAT	0.050129	0.017045	2.941**
Positive Condition	0.032860	0.018559	1.771
word frequency	0.021871	0.013836	1.581
participant age	0.002761	0.001794	1.539

Table 2. Results of a stepwise hierarchical linear regression used to predict *difference in distance* values from the post-task tokens. Symbols following the t-values indicate the associated p-value: '***' $p < 0.001$, '**' $p < 0.01$, and '*' $p < 0.05$.

Only IAT score contributed significantly to this model. The positive β coefficient for IAT in Table 2 means that participants who scored as pro-Australian were more likely to continue their accommodative vowel behavior into the post-task reading. The correlation between a participant's average *difference in distance* and their IAT score was

slightly stronger with the post-task productions than it had been with the shadowed productions [t (40) = 2.5, $p < 0.05$, Pearson's $R = 0.37$]. While Positive Condition was used as a predictor in the regression model, its contribution was not significant. Moreover, post-hoc analysis found no difference in the amount of accommodation across the Positive and Negative Conditions in the post-task productions. Again, like in the shadowed productions, lower frequency words elicited more accommodation (Goldinger 1998). Participant age was used in the model, but it did not make a significant contribution in account for variance in the data. Still, it is interesting that this predictor has a positive β coefficient. This indicates a trend that younger participants accommodated more than older participants. There were no differences in degree of imitation between vowels in the post-task productions. The general pattern of accommodative behavior can be seen in Figure 3 and 4. The post-task productions (small gray italics) are intermediate between the pre-task productions (small black regular font) and the shadowed productions (small black italics).

CONCLUSION

In this study New Zealand participants completed an auditory word-naming speech production task where the model voice was a speaker of Australian English. In an attempt to replicate Bourhis and Giles (1977), one group of NZ participants was insulted (the Negative Condition) and the other was flattered (the Positive Condition) through a story about the AuE speaker. The first and second formants of monosyllabic words containing the vowels KIT, DRESS, TRAP, BARN, STRUT, and THOUGHT were acoustically analyzed for acoustic convergence and divergence. This was determined by

calculating the phonetic distance of responses in the shadowed block and post-task block to NZ productions made in the pre-task block prior to any knowledge that the task involved an Australian. In addition to completing a speech production task, NZ participants took an IAT designed to explore implicit biases to Australia and New Zealand. Stepwise hierarchical linear regression models were used to analyze the tokens from the shadowed block and the post-task block. This type of model automatically selects predictors that best account for the data and ignores others. For the model using the shadowed productions, each of the vowels were selected by the model along with word frequency and participants' IAT scores. The results were vowel-specific such that not all vowels were imitated to the same extent. In terms of word frequency, participants accommodated to lower frequency words more than higher frequency words (Goldinger 1998). For IAT scores, participants who scored as pro-Australian on the task were more likely to accommodate to the AuE model talker than those who scored with a pro-New Zealand bias. Crucially, scores on this task were selected by the model as a predictor while assignment to the Positive or Negative Condition was not. For the post-task productions, IAT scores and word frequency had the similar effect as they did for the shadowed model. Only IAT scores contributed significantly to this second model; pro-Australian IAT scores resulted in more accommodation. In the post-task model, lower word frequency, younger participants, and assignment to the Positive Condition all tended toward more accommodation, but not significantly so.

Bourhis and Giles (1977) found convergence in a group of Welsh participants when they were meant to feel solidarity with a speaker of RP and divergence in a group who disagreed with a view expressed by the RP speaker. In general in this study, NZ

participants converged on the spectral characteristics of an AuE speaker regardless of the feelings presumably incited by the task design. Differences in accommodation were found, however, based on participants' pre-existing sentiments toward Australia. The more positive participants' feelings were toward Australia, the more likely they were to converge on the model speaker's vowels. If IAT scores are a means of participants self-organizing themselves into "positive" and "negative" groups (like those of Bourhis & Giles), the finding is similar to that of Bourhis & Giles. Positive feelings lead to greater likelihood to accommodate, but, crucially, the key result in this experiment is that *all* participants accommodated. Social biases inhibited the degree of accommodation, but the default behavior was for vowel convergence.

The vowel-specific findings also merit discussion. The DRESS vowel was imitated to a greater extent than all other vowels used in the task. While this vowel is produced very differently in the two dialects (compare NZE [dɹɪs], AuE [dɹɐ̟̈s]), it is not considered a particularly salient difference (Hay, Nolan, & Drager 2006). In this study, the most salient dialect differences were not the most imitated; this finding is contra Trudgill (1981). The vowel-specific results seem to contradict Trudgill (2008) to some extent. Trudgill's recent work claims accommodation is automatic and that social identities play no role in whether it happens, but that social ties are fostered as a result of accommodation. In this project, it was found that vocalic convergence is automatic in the sense that participants are not aware they are doing it, but it is not automatic in the sense of being a process that happens at all times. Some vowels are targeted more than others (see also Babel 2009). Moreover, with respect to social identities, implicit socio-cognitive biases about how a participant feels about a speaker determine the extent of

accommodation. These biases are automatic (Djiksterhuis & Bargh 2001), but, crucially, exist prior to the interaction that elicits convergent speech behavior. This result leads to a nuanced view reminiscent to that of Trudgill (2008): speakers of language cannot help accommodating, but group identity attitudes modulate this automatic process.

APPENDIX

Word list for speech production task.

hint	beg	bag	bug	path
bid	ban	bore	bud	bark
hit	bad	bored	buck	laugh
bet	bat	born	but	BARN
bed	back	bought	bun	dance

ACKNOWLEDGEMENTS

This project has been supported by funds from the National Science Foundation EAPSI program and the Abigail Hodgen Publication Fund at University of California, Berkeley. Thanks to Dr. Paul Warren for hosting me at Victoria University of Wellington and allowing me to conduct this research. This work has been improved by Keith Johnson and Andrew Garrett. Many thanks to the participants who volunteered for this project. All errors are, of course, my own.

REFERENCES

Adank, Patti, Roel Smits, & Roeland van Hout. (2004). A comparison of vowel normalization procedures for language variation research. *Journal of the Acoustical Society of America*, 116:3, 99-107.

Baayen, R.H., R. Piepenbrock, and H. van Rijn. (1993). *The CELEX lexical database* (CDROM). Linguistic Data Consortium, University of Pennsylvania.

Babel, Molly. (2009). *Phonetic and Social Selectivity in Speech Accommodation.* Doctoral Dissertation, University of California, Berkeley.

Bayard, Donn. (2000). The cultural cringe revisited: changes through time in Kiwi attitudes toward accents. In Allan Bell & Koenraad Kuiper (eds.) *New Zealand English*, 297-324. John Benjamins Publishing Company.

Bauer, Laurie, Paul Warren, Dianne Bardsley, Marianna Kennedy, & George Major. (2007). An illustration of the IPA: New Zealand English. *Journal of the International Phonetics Association*, 37(1). 97–102.

Bauer, Laurie. (2008). A question of identity: A response to Trudgill. *Language in Society*, 37:272–273.

Boersma, Paul & David Weenink. (2005). Praat: doing phonetics by computer (version 4.3.29) [computer program], from http://www.praat.org/. Institute of Phonetic Sciences, Amsterdam.

Bourhis, Richard Y. & Howard Giles. (1977). The language of intergroup distinctiveness. In Howard Giles (ed.) *Language, ethnicity, and intergroup relations*, 119–136.

Chambers, Jack. (1992). Dialect acquisition. *Language* 68:673–705.

Coupland, Nikolaus. (2008). The delicate constitution of identity in face-to-face accommodation: A response to Trudgill. *Language in Society*, 37:267–270.

Cox, Felicity & Sallyanne Palethorpe (2007). An illustration of the IPA: Australian English. *Journal of the International Phonetic Association*, 37, 341-350.

Delvaux, Véronique & Alain Soquet. (2007). The influence of ambient speech on adult speech productions through unintentional imitation. *Phonetica*, 64:145–173.

Dijksterhuis, Ap & John A. Bargh. (2001). The perception-behavior expressway: Automatic effects of social perception on social behavior. In Mark Zanna (ed), *Advances in Experimental Social Psychology*, 1–40.

Easton, Anita & Bauer, Laurie. (2000). An acoustic study of the vowels of New Zealand English. *Australian Journal of Linguistics* 20 (2), 93-117.

Evans, Bronwen G. & Paul Iverson. (2007). Plasticity in vowel perception and production: A study of accent change in young adults. *Journal of the Acoustical Society of America*, 121:3814–3826.

Holmes, Janet & Paul Kerswill. (2008). Contact is not enough: A response to Trudgill. *Language in Society*, 37:273–277.

Giles, Howard. (1973). Accent mobility: a model and some data. *Anthropological Linguistics*, 15:87–105.

Giles, Howard, & Nikolas Coupland. (1991). *Language: Contexts and consequences.* Milton Keynes: Open University Press.

Giles, Howard, Nikolas Coupland, & Justine Coupland. (1991). Accommodation theory: Communication, context, and consequence. In Howard Giles, Justine Coupland, and Nikolas Coupland (eds.) *Contexts of Accommodation*, 1–68.

Goldinger, Stephen D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, 105:251–279.

Goldinger, Stephen D. & Tamiko Azuma. (2004). Episodic memory in printed word naming. *Psychonomic Bulletin & Review*, 11:716–722.

Greenwald, Anthony G., Debbie E. McGhee, & Jordan L. K. Schwartz. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74:1464–1480.

Greenwald, Anthony G., Brian A. Nosek, & Mahzarin R. Banaji. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, 85:197–216.

Hay, Jennifer, Aaron Nolan, & Katie Drager. (2006). From *fush* to *feesh*: Exemplar priming in speech perception. *The Linguistic Review*, 23, 351-379.

Hay, Jennifer, Paul Warren, & Katie Drager. (2006). Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics*, 34, 458-484.

Lobanov, B. (1971) Classification of Russian vowels spoken by different listeners. *Journal of the Acoustical Society of America*, 49, 606-08.

Maclagan, Margaret & Jennifer Hay. (2007). Getting fed up with our feet: Contrast Maintenance and the New Zealand English "short" front vowel shift. *Language Variation and Change*, 19, 1-25.

Mufwene, Salikoko S. (2008). Colonization, population contacts, and the emergence of new language varieties: A response to Trudgill. *Language in Society*, 37:254–258.

Munro, Murray J., Tracey M. Derwing, & James E. Flege. (1999). Canadians in Alabama: A perceptual study of dialect acquisition in adults. Journal of Phonetics 27:385–403

Namy, Laura L., Lynne C. Nygaard, & Denise Sauerteig. (2002). Gender differences in vocal accommodation: the role of perception. *Journal of Language and Social Psychology*, 21:422–432.

Nielsen, Kuniko. (2008). *The specificity of allophonic variability and its implications for accounts of speech perception*. Doctoral Dissertation, University of California, Los Angeles.

Pardo, Jennifer S. (2006). On phonetic convergence during conversational interaction. *Journal of the Acoustical Society of America*, 119:2382–2393.

Payne, Arvilla C. (1980). Factors controlling the acquisition of the Philadelphia dialect by out-of-state children. In William Labov (ed.) *Locating language in time and space*, 179–218. Academic Press: New York.

Pickering, Martin J. & Simon Garrod. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27:169–226.

Schneider, Edgar W. (2008). Accommodation versus identity?: A response to Trudgill. *Language in Society*, 37:262–267.

Schneider, Walter, Amy Eschman, & Anthony Zuccolotto. (2002). *E-Prime: User's Guide*, version 1.0. Psychology Software Tools.

Shepard, Carolyn A., Howard Giles, & Beth A. LePoire. (2001). Communication accommodation theory. In W. Peter Robinson and Howard Giles (eds.) *The New Handbook of Language and Social Psychology*, 33–56. John Wiley & Sons Ltd.

Shockley, Kevin, Laura Sabadini, & Carol A. Fowler. (2004). Imitation in shadowing words. *Perception & Psychophysics*, 66:422–429.

Trudgill, Peter. (1981). Linguistic accommodation: Sociolinguistic observations on a sociopsychological theory. In Roberta Hendrick, Carrie Masek, and Mary Frances Miller (eds.), *Papers from the Parasession on Language and Behavior*, 218–237. Chicago: Chicago Linguistics Society.

Trudgill, Peter. (1986). *Dialects in contact*. Blackwell Publishing.

Trudgill, Peter. (2008). Colonial dialect contact in the history of European languages: On the irrelevance of identity in new-dialect formation. *Language in Society*, 37:241–280.

Tuten, Donald N. (2008). Identity formation and accommodation: Sequential and simultaneous relations. *Language in Society*, 37:259–262.

Watson, Catherine, Jonathan Harrington & Zoe Evans. (1998). An acoustic comparison between New Zealand and Australian English vowels. *Australian Journal of Linguistics*, 18, 185–207.

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