Dialect divergence and convergence in New Zealand English

MOLLY BABEL

Department of Linguistics, Totem Fields Studios,
2613 West Mall, University of British Columbia, Vancouver,
British Columbia, Canada V6T 1Z4
mbabel@interchange.ubc.ca

ABSTRACT

Recent research has been concerned with whether speech accommodation is an automatic process or determined by social factors (e.g. Trudgill 2008). This paper investigates phonetic accommodation in New Zealand English when speakers of NZE are responding to an Australian talker in a speech production task. NZ participants were randomly assigned to either a Positive or Negative group, where they were either flattered or insulted by the Australian. Overall, the NZE speakers accommodated to the speech of the AuE speaker. The flattery/insult manipulation did not influence degree of accommodation, but accommodation was predicted by participants' scores on an Implicit Association Task that measured Australia and New Zealand biases. Participants who scored with a pro-Australia bias were more likely to accommodate to the speech of the AuE speaker. Social biases about how a participant feels about a speaker predicted the extent of accommodation. These biases are, crucially, simultaneously automatic and social. (Speech accommodation, phonetic convergence, New Zealand English, dialect contact)*

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 by 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 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 /ɑ/ in *bath*. Northern varieties of English use the vowel /ʊ/ for *bud* and *cud* and /a/ 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 ten-point scale from 'very northern' to 'very southern'. Overall, participants were rated as sounding more southern after their time at the university.

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 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.

Of course, dialect acquisition and dialect change are not the only ways in which an individual's speech fluctuates. Individuals regularly style-shift in language use; style-shifting is the process of changing ways of speaking according to social contexts (see Eckert & Rickford 2001 or Coupland 2007 for deeper treatments of style). In a landmark study, Coupland (1980) examined the speech patterns of a travel agent in Cardiff while she interacted with different participants (friends, clients, coworkers) about different topics and across different communication channels (in person or on the telephone). The frequency with which the travel agent used particular phonological variables was correlated with the contextual dynamic. Bell (1984) reports on similar data showing that four radio newsreaders in New Zealand varied their productions of intervocalic /t/ when reading on different radio stations. All newsreaders had higher incidence of flapping intervocalic /t/ on the community radio station, with lower flapping rates occurring on the national radio station. Bell's theory of audience design, which arose from his 1984 study, is a model that attempts to account for when stylistic shifting arises. In audience design, a speaker considers not only the addressee in the selection of a proper linguistic

variable for an utterance, but also the individuals who are assuming the role of auditor, overhearer, and eavesdropper. Bell (2001) expands the theory of audience design by adding a notion of REFEREE DESIGN as a mechanism parallel with AUDIENCE DESIGN. Referee design is the change in a speaker's style that affiliates the speaker with a particular social group as prompted by the audience (Bell 2001:163). Bell argues that a comprehensive theory of style-shifting must consider the two principal driving forces—audience and referee design—simultaneously. While a talker is crafting from a stylistic repertoire that heeds the needs of the audience, a speaker feels the pressure to maintain a level of faithfulness to their social network and larger social affiliations.

What determines a speaker's stylistic selection? This question is also addressed by Bell (1984:167), where he considers three possibilities: (i) speakers examine an addressee's personal traits and design their style accordingly; (ii) speakers gauge the general speech style of an addressee and design accordingly; or (iii) speakers listen for specific linguistic variables and select their stylistic variants to reflect their findings. Many of Bell's suggestions are both echoed in and echoes of COM-MUNICATION ACCOMMODATION THEORY (CAT; Giles 1973, Giles & Coupland 1991, Giles, Coupland, & Coupland 1991, Shepard, Giles, & LePoire 2001). In CAT, it is argued that linguistic choices made in interaction are a function of speakers creating and maintaining social distance. Linguistic accommodation or convergence lessens the social distance between interlocutors while divergence maximizes it. A speaker's speech style will shift in an interaction to modify the social distance between the speaker and the addressee, overhearers and eavesdroppers, and the referee. This connection between style-shifting and CAT is not novel. In fact, Giles (1973) claimed that, perhaps, all style-shifting in sociolinguistics interviews was the result of accommodation. Trudgill (1981), however, provided evidence from his interviews in Norwich to refute this claim. In his data, Trudgill found that he actually accommodated to the interviewees more than the other way around.

This research on dialect change and style-shifting demonstrates that speakers modify their speech patterns as a result of changes in social contexts and new cultural environments. In addition to these findings, several laboratory-based studies have found that individuals acquire the phonetic characteristics of model talkers in highly asocial speech production tasks. In what is known as a SHADOWING TASK or an AUDITORY NAMING TASK[1] 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, however, 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 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.

Thus, there is ample evidence demonstrating that phonetic and phonological accommodation occurs in language behavior both in ecologically valid social interactions and in more sterile laboratory contexts. There is currently considerable 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.

Sociolinguists have also contributed to the how-and-why discussion of accommodation. Traditionally, research on dialect contact and acquisition has remained agnostic regarding the mechanisms for accommodation, while the motivation is considered one of belonging, group identity, and social-network membership. But, recently, 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, Coupland 2008, Holmes & Kerswill 2008, Mufwene 2008, Schneider 2008, Tuten 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. One particular CAT study poses a clear 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 article. 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 eleven-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 article is twofold. The first goal 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, an auditory naming 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.[2] 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 DRESS and TRAP vowels are raising.[3] The KIT vowel in AuE is raising, while in NZE it is centralizing. Not all of these sound changes are considered 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 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 the DRESS vowel 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.

The second goal is to query the role of social factors in speech production and the role of social interaction in accommodation. The fact that speakers automatically accommodate in auditory naming tasks demonstrates that at least a small part of accommodation and style-shifting in social interactions may be unintentional and simply an automatic reflex of the language system. But, furthermore, demonstrating that social factors affect the degree of accommodation in asocial laboratory interactions would indicate that it is unlikely that language use lacks social conditioning and influence at any level.

METHODOLOGY

In the following paragraphs I describe the speech production task used to explore phonetic convergence and divergence in NZE. Next, I provide 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 AND 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 thirty 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, START, 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 article. 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 2007). 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 auditory naming 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 pretask block. The

purpose of the pretask 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. 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 posttask reading of the wordlist; this block was identical to the pretask block, except that the words were presented in a different random order. Finally, participants reread the hVd words again. The methodology of this auditory naming task follows that of Goldinger 1998 and Namy et al. 2002.

In comparing pretask 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, a POSITIVE CONDITION and a NEGATIVE CONDITION. In the Positive Condition, participants were presented with the following text that was intended to make them view the talker and Australia as a whole in a positive light:[4]

> 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.

Participants in the Positive Condition were expected to develop positive feelings toward the Australian talker. These positive feelings were predicted to create the desire to lessen the social distance, and thus following CAT, NZ participants were expected to converge toward the Australian model talker in this condition. The purpose of the Negative Condition was to inspire negative feelings toward the talker and Australia, which, according to CAT, would cause divergence:

> 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.

The texts associated with each condition were written in consultation with two Australians. In both conditions participants were exposed to a screen with the assigned text immediately before beginning the test-block. The Positive and Negative texts were presented under the guise that the purpose of the experiment was to examine how personal information about the model talker affected word recognition. 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 aspect of this project was to examine how implicit and subconscious attitudinal measures that participants may not admit forthright in an interview may predict speech behavior. To this end, an Implicit Association Task (IAT; Greenwald, McGhee, & Schwarz 1998) was used. While the IAT is not commonly used in sociolinguistics, it is a standard social psychology tool that uses reaction time in category classification to measure implicit biases.

There are five blocks in this task. Block 1 is target-concept discrimination. The targets AUSTRALIA and NEW ZEALAND are 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.)—is 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. Block 2 is associated attribute discrimination. Here, the attributes *good* and *bad* are presented on opposite sides of the computer screen, and the concepts AUSTRALIA and NEW ZEALAND are no longer present on the screen. 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. Block 3 is a combined test block. Labels for the concept categories (AUSTRALIA vs. NEW ZEALAND) and word (*good* vs. *bad*) attributes are presented simultaneously at the top corners of the screen. In the center, either a concept (eliciting either an AUSTRALIA or NEW ZEALAND response) or a word (to be classified as *good* or *bad*) 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 illustrated as words (eg. Kevin Rudd and Helen Clark) 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 are presented on different sides of the screen (so, if AUSTRALIA is on the right-side of the screen in Block 1, it is on the left side in Block 4). Participants then categorize 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) match up with the original order of *good* and *bad* such that if AUSTRALIA is 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, & Banaji 2003. After eliminating responses with latencies over 10,000 milliseconds, 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, following the updated scoring algorithm described in Greenwald et al. 2003. Means were then recalculated 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. A negative IAT score is indicative of a pro-Australian bias and a positive IAT score indicates a pro-New Zealand bias. These values were used as predictors in the statistical model described below.

After completing the Implicit Association Task, participants were debriefed on the purposes of both the speech production task and the IAT. While most participants made no explicit comment about the task, a small number of participants assigned to the Negative Condition remarked that the text seemed a bit extreme.

ANALYSIS AND RESULTS

Analysis

Vowels from the participant productions were hand-marked. A Praat (Boersma & Weenink 2005) script calculated the average first and second formants over the middle 50% of the vowel. Data points that fell beyond the general vowel categories were identified as outliers and 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).

Euclidean distance is a way of measuring distance within a two-dimensional space; with vowels, the Euclidean distance calculates the distance between vowels by considering both the first and second formant. Using the normalized formant values, the Euclidean distance between each NZ word production and the same word from the Australian model talker was calculated. 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 posttask blocks were subtracted from the pretask block distances, creating a DIFFERENCE IN DISTANCE metric. This difference-in-distance measure is indicative of HOW MUCH a participant modified their vowel productions with respect to the speech

of the model talker. A negative value indicates that the distance between the NZ participant and the Australian model has shrunk, which would indicate convergence. A positive value signals that the phonetic distance has grown; this would mean there was divergence.

Results from the shadowed productions

Analyses were conducted separately on the task and posttask productions. First, the results from the task productions are discussed. A stepwise hierarchical linear regression analysis using only the data from the task 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 (DRESS, KIT, TRAP, START, 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 selected 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 pretask tokens.

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; rather, there was less imitation as word frequency was higher and IAT score was higher (which equals a pro-New Zealand bias). 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, posthoc 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

TABLE 1. *Results of a stepwise hierarchical linear regression used to predict difference-in-distance values from the shadowed 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

Symbols following the *t*-values indicate the associated *p*-value: ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$.

Negative Condition scored as more pro-New Zealand in the IAT than those partici-pants randomly assigned to the Positive Condition.

While the overall pattern was that of convergence, participants did not accom-modate toward all vowels to the same extent. Figure 2 presents difference-in-distan-ce values for each vowel averaged across all participants in the task block. Any value below zero—denoted 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 nonuniform manner. For example, it seems that DRESS elicited the most accommo-dation. Indeed, posthoc analyses found that the DRESS vowel was imitated more than STRUT ($p < 0.05$), THOUGHT ($p < 0.001$), and TRAP ($p < 0.01$). START was also imitated more than THOUGHT ($p < 0.05$).[5]

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' pretask pro-ductions. The smaller words in black italics are the shadowed productions and those in small gray italics are productions from the posttask block. These figures depict the direction of the accommodation. For female participants (Fig. 3), all vowels in the test block shifted in the direction of the Australian talker. Male par-ticipants (Fig. 4), however, display this same pattern of drift toward the Australian talker for DRESS, START, STRUT, and THOUGHT, but demonstrate divergence or maintenance for KIT and TRAP.
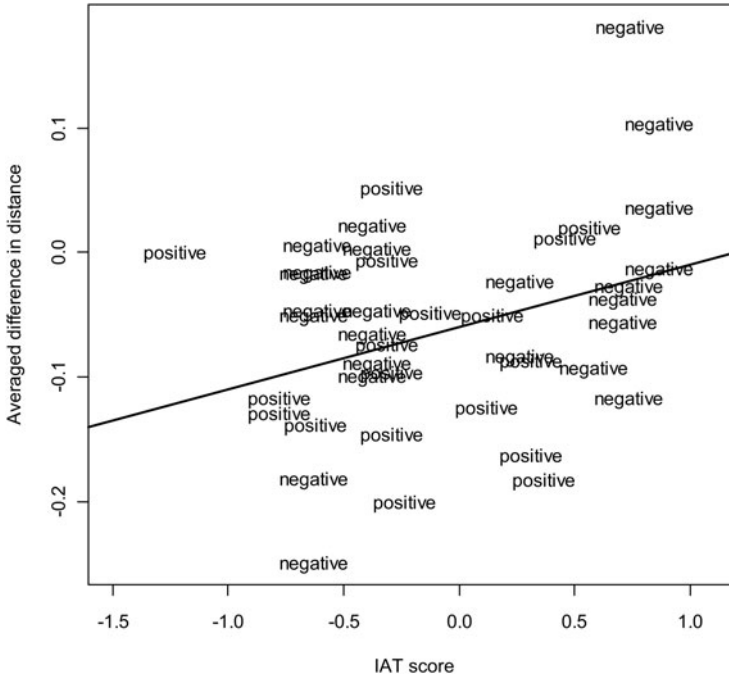
FIGURE 1. Each participant's averaged difference-in-distance plotted against their IAT score. The difference-in-distance measure on the *y*-axis indicates the amount of phonetic accommodation. A negative value demonstrates phonetic convergence while a positive value demonstrates phonetic divergence. 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.

Results from the posttask productions

For the posttask productions, another stepwise hierarchical linear regression analysis was used. Again, the difference-in-distance metric was used as the dependent variable, while vowel category (DRESS, KIT, TRAP, START, 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 posttask block were smaller than those in the pretask. After exposure to the

448   *Language in Society* **39**:4 (2010)
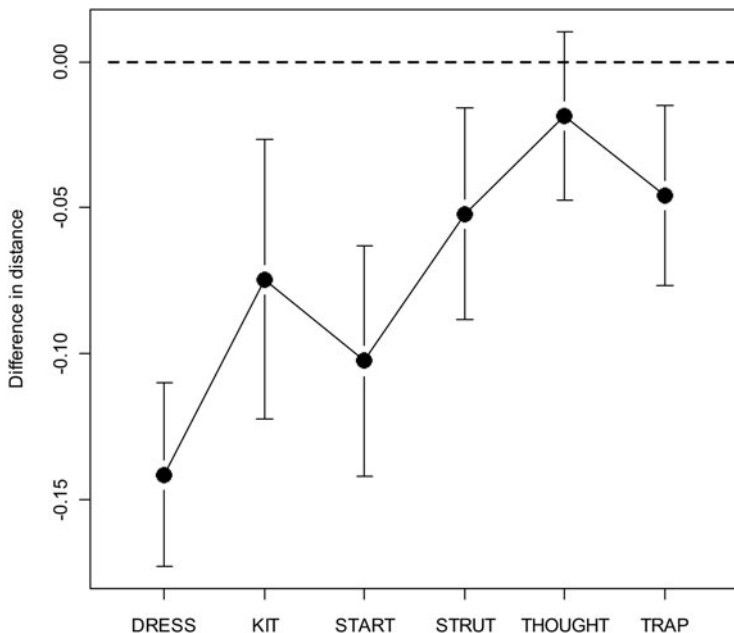
FIGURE 2. Amount of phonetic accommodation to the vowels averaged across all participants in both the Positive and Negative Conditions. The difference-in-distance measure on the *y*-axis indicates the amount of phonetic accommodation. A negative value demonstrates phonetic convergence while a positive 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.

Australian model talker ceased, NZ participants retained the more Australian-like vowels they had acquired during the shadowing block.

Only the 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 posttask reading. The correlation between a participant's average difference-in-distance and their IAT score was slightly stronger with the posttask productions than it had been with the task 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, posthoc analysis found no difference in the amount of accommodation across the Positive and Negative Conditions in the posttask productions. Again, like in the shadowed productions, lower-frequency words elicited more accommodation (Goldinger 1998). Participant age was selected by the model, but it did not make a significant contribution in accounting for the 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
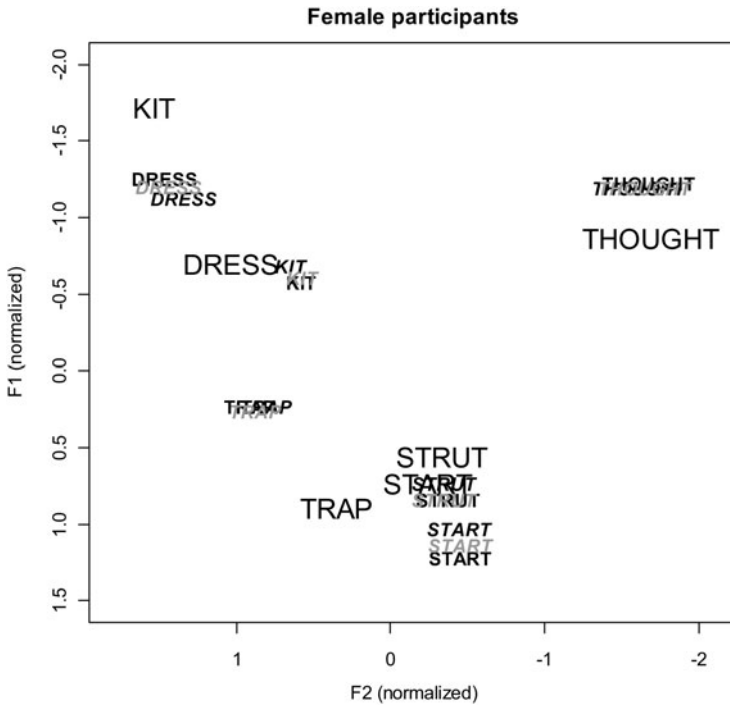
449

Female participants

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' pretask productions are in a small black font, shadowed productions are in black italics, and posttask productions are in gray italics. The shadowed productions move in the direction of the model talker's tokens and the posttask productions generally lie somewhere between the pretask and shadowed productions.

TABLE 2. *Results of a stepwise hierarchical linear regression used to predict difference-in-distance values from the posttask tokens.*

	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

Symbols following the *t*-values indicate the associated *p*-value: ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$.
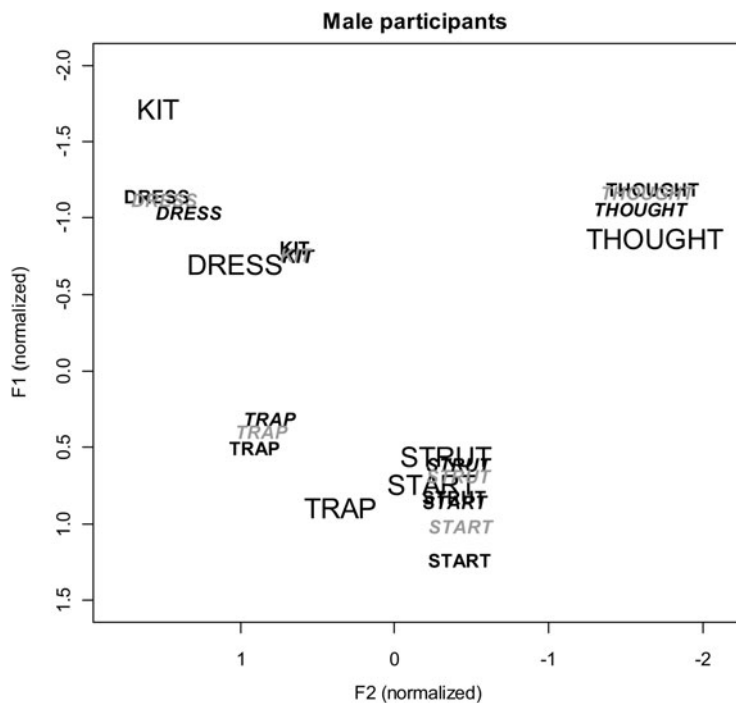
FIGURE 4. Formant plot displaying phonetic accommodation for the male participants. The Australian model talker's mean vowels are plotted in the larger black typeface. Male participants' pretask productions are in a small black font, shadowed productions are in black italics, and posttask productions are in gray italics. The shadowed productions move in the direction of the model talker's tokens and the posttask productions generally lie somewhere between the pretask and shadowed productions.

participants. There were no differences in degree of imitation between vowels in the posttask productions. The general pattern of accommodative behavior can be seen in Figs. 3 and 4. The posttask productions (small gray italics) are intermediate between the pretask productions (small black regular font) and the shadowed productions (small black italics).

DISCUSSION AND 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 & Giles 1977, one group of NZ participants was insulted (the Negative Condition) and the other was flattered (the Positive Condition) through a short story about the beliefs of the AuE speaker. The first and second

formants of monosyllabic words containing the vowels KIT, DRESS, TRAP, START, STRUT, and THOUGHT were acoustically analyzed for acoustic convergence and divergence. 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 posttask block. This type of model automatically selects predictors that best account for the data and ignores others. For the model using the shadowed task 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 posttask 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 posttask 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 incited by the task design and their assignment to the Positive or Negative Conditions. Differences in accommodation were found, however, based on participants' preexisting sentiments toward Australia. The more positive participants' feelings were toward Australia, the more likely they were to converge on the model speaker's vowels. Following CAT, participants with a positive perspective of Australia can be interpreted as having decreased the social distance between themselves and the model talker by accommodating to a greater degree than those participants who viewed Australia negatively. In terms of referee design (Bell 2001), participants with positive views of Australia would be interested in affiliating themselves with that group through convergent speech acts. As far as actual replication of Bourhis & Giles (1977) goes, if IAT scores are a means of participants self-organizing themselves into "positive" and "negative" groups (like those of Bourhis & Giles), the finding here is then similar. While positive feelings lead to greater likelihood to accommodate, a key result in this experiment is that ALL participants accommodated to some extent. What this potentially means is that the default behavior is for accommodation and convergence, and this may be due to psychological mechanisms stemming from the organization of language systems. Then, sociolinguistic factors such as social biases, the

desire to create social distance, and positive affiliation may affect the DEGREE of accommodation. Importantly, if implicit biases (eg. the IAT) mediate spoken language outside of socially meaningful interaction, it can only be expected that social biases would have an even greater effect in real-world language use.

An interesting difference revealed by the analysis of the task and posttask productions relates to the vowel effects. In the task productions, participants accommodated toward the vowels to different extents. Participants accommodated to the DRESS vowel considerably, but, for example, to the THOUGHT vowel much less so. In the posttask productions, there were no specific vowels effects, but simply a general tendency for the NZ participants to pronounce their vowels more like those of the Australian model talker. The carry-over effects into the posttask productions once exposure to the model talker has ceased indicates that there is some long-term change to the word-level representations such that they are more Australian-like; this effect influences all the vowels roughly to the same extent. In the task productions, 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 [dɹis̈], AuE [dɹes̈]), it is not considered a particularly salient difference. Hay and colleagues (2006) report that the largest vowel differences between NZE and AuE are for the KIT, DRESS, and TRAP vowels. They found that NZE speakers are not particularly sensitive to the difference between the DRESS vowels across the dialects, but they are aware of the dialect difference for the KIT and TRAP vowels. This awareness of the KIT and TRAP vowels may have prevented larger accommodation results for those vowels. For the DRESS vowel, however, with a large amount of phonetic space between the NZE and AuE categories, there is ample room for a New Zealander to accommodate, and the lack of salience for this category may facilitate the accommodation process.

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 as well. 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 study, 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 social biases about how a participant feels about a speaker strongly influence the extent of accommodation. These biases are automatic in that they are subconscious and not derived by explicit and strategic decision making (Djiksterhuis & Bargh 2001), but they, 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, much like what is predicted in both audience and referee design (Bell 1984, 2001)

and CAT (Giles et al. 1991). Additionally, these results demonstrate that social factors overlay even very basic and low-level aspects of speech production.

NOTES

[1]In this type of task participants are presented with prerecorded tokens of words. These words are presented to participants one at a time. Upon hearing the words, participants are instructed to identify (name) the word (the auditory object) by saying it aloud; hence the term "auditory naming."

[2]A reviewer states that the relationship between Australian and New Zealand English is not like that of RP in Wales. This is true. But, while the social dynamic between Australia and New Zealand speakers is not identical to that of RP and Welsh speakers, it provides a parallel environment of cultural and linguistic competition.

[3]Following Wells 1982, lexical sets will be used to refer to the vowel categories in this article.

[4]One reviewer points out that the Positive Condition is "in-group" and could trigger perceptions of New Zealand English in the Australian talker. If this were the case then we would predict that participants in the Positive Condition would not accommodate to the model talker. As is shown below, this was not the case.

[5]Posthoc analyses were Tukey's Honest Significant Difference tests.

REFERENCES

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

Baayen, R. Harald; Piepenbrock, Richard; & van Rijn, H. (1993). *The CELEX lexical database* (CD-ROM). Linguistic Data Consortium, University of Pennsylvania.

Babel, Molly (2009). *Phonetic and social selectivity in speech accommodation*. Berkeley: University of California, Berkeley dissertation.

Bauer, Laurie (2008). A question of identity: A response to Trudgill. *Language in Society* 37:272–73.

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

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. Amsterdam: John Benjamins.

Bell, Allan (1984). Language style as audience design. *Language in Society* 13:145–204.

——— (2001). Back in style: Reworking audience design. In Penelope Eckert & John R. Rickford (eds.), *Style and sociolinguistic variation*, 139–69. Cambridge: Cambridge University Press.

Boersma, Paul, & Weenink, David (2005). Praat: Doing phonetics by computer, version 4.3.29. Online: http://www.praat.org/.

Bourhis, Richard Y., & Giles, Howard (1977). The language of intergroup distinctiveness. In Howard Giles (ed.), *Language, ethnicity, and intergroup relations*, 119–36. London: Academic Press.

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

Coupland, Nikolaus (1980). Style-shifting in a Cardiff work-setting. *Language in Society* 9:1–12.

——— (2007). *Style: Language variation and identify*. Cambridge: Cambridge University Press.

——— (2008). The delicate constitution of identity in face-to-face accommodation: A response to Trudgill. *Language in Society* 37:267–70.

Cox, Felicity, & Palethorpe, Sallyanne (2007). An illustration of the IPA: Australian English. *Journal of the International Phonetic Association* 37:341–50.

Delvaux, Véronique, & Soquet, Alain (2007). The influence of ambient speech on adult speech productions through unintentional imitation. *Phonetica* 64:145–73.

Dijksterhuis, Ap, & Bargh, John A. (2001). The perception-behavior expressway: Automatic effects of social perception on social behavior. In Mark Zanna (ed.), *Advances in experimental social psychology*, vol. 33, 1–40. San Diego: Academic Press.

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

Eckert, Penelope, & Rickford, John R. (2001). *Style and sociolinguistic variation*. Cambridge: Cambridge University Press.

Evans, Bronwen G., & Iverson, Paul (2007). Plasticity in vowel perception and production: A study of accent change in young adults. *Journal of the Acoustical Society of America* 121:3814–26.

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

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

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

Goldinger, Stephen D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105:251–79.

———, & Azuma, Tamiko (2004). Episodic memory in printed word naming. *Psychonomic Bulletin & Review* 11:716–22.

Greenwald, Anthony G.; McGhee, Debbie E.; & Schwartz, Jordan L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology* 74:1464–80.

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

Harrington, Jonathan (2006). An acoustic analysis of 'happy-tensing' in the Queen's Christmas broadcasts. *Journal of Phonetics* 34:439–57.

——— (2007). Evidence for a relationship between synchronic variability and diachronic change in the Queen's annual Christmas broadcasts. In Jennifer Cole & José I. Hualde (eds.), *Laboratory phonology 9: Phonetics and phonology*, 125–44. Berlin: Walter de Gruyter.

———; Palethrope, Sallyanne; & Watson, Catherine I. (2000a). Does the Queen speak the Queen's English? *Nature* 408:927–28.

———; ———; & ——— (2000b.) Monophthongal vowel changes in Received Pronunciation: An acoustic analysis of the Queen's Christmas broadcasts. *Journal of the International Phonetic Association* 30:63–78.

Hay, Jennifer; Nolan, Aaron; & Drager, Katie (2006). From *fush* to *feesh*: Exemplar priming in speech perception. *The Linguistic Review* 23:351–79.

———; Warren, Paul; & ——— (2006). Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics* 34:458–84.

Holmes, Janet, & Kerswill, Paul (2008). Contact is not enough: A response to Trudgill. *Language in Society* 37:273–77.

Lobanov, B. (1971) Classification of Russian vowels spoken by different listeners. *Journal of the Acoustical Society of America* 49:606–8.

Maclagan, Margaret, & Hay, Jennifer (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* 37:254–58.

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

Namy, Laura L.; Nygaard, Lynne C.; & Sauerteig, Denise (2002). Gender differences in vocal accommodation: The role of perception. *Journal of Language and Social Psychology* 21:422–32.

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

Pardo, Jennifer S. (2006). On phonetic convergence during conversational interaction. *Journal of the Acoustical Society of America* 119:2382–93.

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. New York: Academic Press.

Pickering, Martin J., & Garrod, Simon (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* 37:262–67.

Schneider, Walter; Eschman, Amy; & Zuccolotto, Anthony (2007). *E-Prime: User's Guide*, version 2.0. Psychology Software Tools.

Shepard, Carolyn A.; Giles, Howard; & LePoire, Beth A. (2001). Communication accommodation theory. In W. Peter Robinson & Howard Giles (eds.), *The new handbook of language and social psychology*, 33–56. Chichester: John Wiley & Sons.

Shockley, Kevin; Sabadini, Laura; & Fowler, Carol A. (2004). Imitation in shadowing words. *Perception & Psychophysics* 66:422–29.

Trudgill, Peter (1981). Linguistic accommodation: Sociolinguistic observations on a sociopsychological theory. In Roberta Hendrick, Carrie Masek, & Mary Frances Miller (eds.), *Papers from the Parasession on Language and Behavior*, 218–37. Chicago: Chicago Linguistics Society.

——— (1986). *Dialects in contact*. Oxford: Blackwell.

——— (2008). Colonial dialect contact in the history of European languages: On the irrelevance of identity in new-dialect formation. *Language in Society* 37:241–80.

Tuten, Donald N. (2008). Identity formation and accommodation: Sequential and simultaneous relations. *Language in Society* 37:259–62.

Watson, Catherine; Harrington, Jonathan; & Evans, Zoe (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: Cambridge University Press.

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