From *fush* to *feesh*: Exemplar priming in speech perception[1]

JENNIFER HAY, AARON NOLAN and KATIE DRAGER

Abstract

Niedzielski (1999) reports on an experiment which demonstrates that individuals in Detroit 'hear' more Canadian Raising in the speech of a speaker when they think that speaker is Canadian. We describe an experiment designed to follow up on this result in a New Zealand context. Participants listened to a New Zealand English (NZE) speaker reading a list of sentences. Each sentence appeared on the answer-sheet, with a target word underlined. For each sentence, participants were asked to select from a synthesized vowel continuum the token that best matched the target vowel produced by the speaker. Half the participants had an answer-sheet with the word 'Australian' written on it, and half had an answer-sheet with 'New Zealander' written on it. Participants in the two conditions behaved significantly differently from one another. For example, they were more likely to hear a higher fronter /ɪ/ vowel when 'Australian' appeared on the answer sheet, and more likely to hear a centralized version when 'New Zealander' appeared – a trend which reflects production differences between the two dialects. This is despite the fact that nearly all participants reported that they knew they were listening to a New Zealander. We discuss the implication of these results, and argue that they support exemplar models of speech perception.

1. This article has benefited from the insights of Cynthia Clopper, Susanne Gahl, Andy Gibson, Keith Johnson, Margaret Maclagan, Rachel Matthews, Alice Murphy, Anna Pucilowski, Paul Warren, Alan Yu, an anonymous reviewer, and audiences at the University of Canterbury and the New Zealand Linguistics Society conference in Auckland, New Zealand. We would like to express our thanks to all the participants who took part in this study. This research was partially funded by a University of Canterbury research grant to the first author.

The Linguistic Review 23 (2006), 351–379
DOI 10.1515/TLR.2006.014
0167–6318/06/023-0351

1. Introduction

A large body of work demonstrates that individuals style-shift in production (see, e.g., the reviews in Schilling-Estes 2002 and Mendoza-Denton 1999). One common manifestation of style-shifting is the very well attested phenomenon of speech accommodation – in many contexts speakers shift their speech to more closely resemble that of their addressee (Giles et al. 1991). While it is well established that individuals may adjust their production depending on who they are talking to, much less is known about the degree to which they may adjust their speech perception. The speech signal is highly variable – both across individuals and within individuals. Variability in the signal has received attention largely in the context of investigations of how individuals might strip away the variability or normalize the signal, in order to get at its core 'underlying' properties.

Recently, however, a series of investigations has started to examine the possibility that individuals may actually exploit this variability in order to assist with speech perception. In particular, it has been demonstrated that individuals are highly skilled at exploiting socially governed variability. In cases where different groups in society vary in the production of certain variants, individuals can exploit this connection in speech perception. For example, individuals process the speech signal differently depending on the perceived gender of the speaker (Strand 1999; Johnson, Strand and D'Imperio 1999), perceived age (Hay et al. to appear; Drager 2005), perceived social class (Hay et al. to appear), and perceived dialect area (Niedzielski 1999). Niedzielski's (1999) study shows that stereotypes about a particular dialect area can affect the speech perception. In this paper, we report on a follow-up of Niedzielski's study in a slightly different context. The goal was to replicate the basic effect, as well as try and further understand its nature. After outlining the experimental methodology and its results, we turn to an interpretation of our results (and Niedzielski's) in the context of an exemplar account of speech perception.

2. Background

Bilingual speakers adjust their perceived boundaries between phonemes depending on which language they think they are listening to (Jansen and Schulman 1983). However, as mentioned above, recent work also suggests that listeners can make much more subtle within-language adjustments, depending on the social characteristics of the speaker they think they are listening to. For example Strand (1999) demonstrates that listeners hear the boundary between /s/ and /ʃ/ in different places, depending on whether they think they are listening to a male or female. Johnson et al. (1999) find a similar effect for the bound-

ary between the vowels in *hood* and *hud*. Remarkably, the strength of these effects is influenced by how prototypically 'male' or 'female' the voice or face is. They are also present when participants are not shown any visual cue, but simply asked to imagine a male or female face saying the words. Drager (2005) shows that the perceived age of a speaker can affect perceived vowel boundaries in the context of a chain-shift in progress. And Hay et al. (to appear) demonstrate that participants can more accurately distinguish between vowels undergoing merger when they think they are listening to an older speaker (who is less likely to have the merger). In all of these experiments listeners' expectations are manipulated, and they are asked to identify words in isolation (e.g., identify whether a particular token is an instance of *hood* or *hud*, or *ear* or *air*, etc).

Niedzielski (1999) approaches the question of whether social expectations can affect speech perception in quite a different way. In her experiment two groups of participants from Detroit, Michigan were played sentences spoken by an individual from Detroit. A continuum of the /aʊ/ diphthong was synthesized, including tokens representing the stigmatized, raised variant produced by both Canadians and Detroiters. On the response sheet, words were underlined for each sentence, and participants were asked to match the vowels from the underlined words with one from the continuum of synthesized vowels.

One of the two words 'Canadian' or 'Michigan' appeared at the top of the response sheet depending on the condition. Participants were more likely to respond with more raised synthesized tokens of /aʊ/ when 'Canadian' was written on the response sheet, even though both Canadians and Detroiters produce raised variants of the diphthong in natural speech.

Niedzielski (1999) attributes the apparent difference in perception to the expectations of the American participants who took part in the study. Detroiters believe that Canadians as a group produce the raised variant of the diphthong, while they appear to be unaware that the high variant also exists in their own speech. Similarly, for a variety of Northern Cities Chain Shift vowels, participants responded with more 'standard' or canonical variants when they thought the speaker was from Detroit.

In this article, we attempt to replicate Niedzielski's basic results in a new context. New Zealanders were asked to match three different vowels to synthesized continua. They were played the speech of a New Zealander, but half the participants had 'New Zealander' written on the top of the answer sheet, and half of them had 'Australian'.

A secondary manipulation was included in order to test for an effect of recency – do participants behave differently in this task if the word they are matching immediately precedes the synthesized continuum, as opposed to being phrase medial, with multiple words intervening (as was the case with most of Niedzielski's stimuli)? Matching an item to a vowel on a six-point contin-

uum is likely to be challenging at the best of times. Having the word embedded phrase-medially requires participants to hold the word in memory until the phrase is complete. It seems possible that this storage might exaggerate the effect, as the difficulty of the task may cause participants to fall back on their stereotypes. If this were true, then the results (while still fascinating) would relate less to speech perception itself than Niedzielski suggests. The manipulation of the recency of exposure to the target word was therefore included in order to further probe the locus of the effect.

3. New Zealand and Australian English

New Zealanders tend to have strong opinions about Australian English, generally considering it to be not only easily recognizable but also somewhat unpleasant (Bayard 1990, 1995). However, in an attitude study, Bayard (1990) found that participants from New Zealand (NZ) were actually relatively inaccurate at discriminating between NZ and Australian voices. This result was fairly speculative, and only based on a few voices. It nonetheless suggested that it may be possible to manipulate the perceptions of participants from NZ simply by writing 'New Zealander' or 'Australian' on their answer-sheets.

NZ and Australian English share a large number of features in common relative to other dialects of English. There are also, however, some key differences between the two different dialects, especially with respect to the realization of the short, front vowels. While /æ/ and /ɛ/ are raised in the speech of Australians as compared with other dialects of English, they are not as raised for Australians as they are in the speech of New Zealanders (Bauer and Warren 2004; Watson et al. 1998). There is an even more salient difference in the realization of the vowel, /ɪ/, which is raised in Australian English and is centralized in NZE (see, e.g., Watson et al. 1998). Differences in the pronunciation of /ɪ/ seem to be the most salient feature for both Australians and New Zealanders, and are a matter of considerable comment. Bayard (2000) cites this vowel as the most salient perceived feature by New Zealanders and Australians. Humorous Australian websites about NZE are dominated by examples containing this variable (e.g., *Bug hut: popular recording; day old chuck: very young poultry* …) The cross-Tasman production of the phrase *fish n chips* is particularly canonical and is the focus of mocking on both sides. This is the phrase most often cited when discussing differences between the accents (the NZ "fush and chups", vs. Australian "feesh and cheeps"). For example a displaced New Zealander recently wrote in the NZ magazine *The New Zealand Listener*:

> I've kept my accent, although when necessary I can say "feesh and cheeps". (Ivory 2005)

It is clear that this vowel is the most salient, most stereotyped difference between NZ and Australian English, and it would be reasonable to suggest that the phrase *fish and chips* is particularly iconic.

Differences between realizations of the other front, short vowels /æ/ and /ɛ/ are sometimes commented on, particularly by Australians, but the commentary is much less frequent. Indeed, the raised variants that occur in the speech of many New Zealanders seem to be non-stigmatized, suggesting that they are below the level of consciousness for most New Zealanders (Maclagan et al. 1999). Ludwig (2005) asked 24 New Zealanders to list ways in which they thought Australian English sounded different from NZE. Many of these participants commented on the vowels, in a non-specific way ("harsh vowels", "drawling vowels", "larger vowel sounds", …). Six explicitly mentioned the /ɪ/ vowel (e.g., " 'i', e.g., six as sex", "seednee", "Australian 'i' sounds like NZ 'e' " …). No-one mentioned /æ/, and one person may have been alluding to /ɛ/ when they wrote that the "e" is "drawn out".

It is important to note these variables differ from Niedzielski's in an important way. The Canadian Raising was present in both the speech of Canadians and Detroiters, but Detroiters only held a stereotype about it being present in Canada. In NZ, for /ɪ/, the realizations between NZ and Australian English are very different, and also reinforced by stereotype. For /æ/ and /ɛ/ the realizations are also different, but there is little (or no) stereotype associated with this difference.

4. Methodology

A male speaker of NZE was recorded reading a list of sentences in which target words containing the vowels /ɪ/, /æ/, and /ɛ/ occurred sentence-medially in one half of the sentences and sentence-finally in the other (matched) half. Example sentences are given in (1).

(1) a. During the colder months John and I both wanted to stay **fit** so we decided to join a yoga class.
b. During the colder months John and I both wanted to stay **fit**.

(2) a. Very quickly, I realized that this was actually a horribly uncomfortable **bed** and that I needed to go back and have it exchanged.
b. Very quickly, I realized that this was actually a horribly uncomfortable **bed**.

20 of the 40 stimulus sentences included the target vowel /ɪ/, which was the main focus of the experiment. The remaining 20 sentences were equally divided between sentences containing the target vowels /ɛ/ and /æ/. These addi-

Table 1. Continua formant values for /ɪ/, /ɛ/ and /æ/

	/ɪ/		/ɛ/		/æ/	
	F1	F2	F1	F2	F1	F2
Token 1 (AUS)	350	2350	450	2100	700	1750
Token 2	417	2167	400	2215	663.3	1867
Token 3	484	1984	350	2330	626.6	1983
Token 4 (actual)	550	1800	300	2450	590	2100
Token 5	618	1618	250	2565	553.3	2114
Token 6 (NZ)	685	1435	200	2680	516.6	2333

tional sentences were primarily included so that participants didn't get locked in to a particular pattern of response. The sentences were presented in pseudo-random order, and were not blocked. Corresponding sentences (i.e., the 'a' and 'b' forms from (1) and (2)) were never presented in immediate succession.

All sentences were recorded in a quiet room using Sonic Foundry Sound-Forge 6.0 on a Toshiba laptop linked via a USB port to a USB Pre 1.5 interface and an AKG C 444 PP condenser microphone.

A six-step continuum was created for each vowel using Klattworks v1.8 – a speech synthesis tool that allows a user to make slight changes in formant values (McMurray in preparation). The continuum was created by resynthesizing recordings of the vowels, produced by the same male speaker produced in an isolated h_d frame. By manipulating the values for the first and second formants (F1 and F2), the continua were designed to span a range from Australian-like variants to NZ-like variants, with equal steps between the values of each token of the continua.[2] The values set to F1 and F2 for the continua are presented in Table 1.

The spectrogram in Figure 1 displays the gradual shift in F1 and F2 values between the six vowels in the /ɪ/ continuum. The Australian-like variants of token 1 were based on values presented by Cox (2004), and token 4 was based on the actual values produced by the New Zealander in the h_d frame. Tokens 2 and 3 occurred in equal steps between tokens 1 and 4. Tokens 5 and 6 represent intervals of equivalent size, representing more centralized variants than the token on which they were based.

Token 4 in each of the continuum was constructed to approximate the token produced in the h_d by the male speaker. However the actual tokens produced by him in the sentence context obviously showed some variation. The range of

2. Due to experimenter error, token 5 for /æ/ is not equidistant between the F2 values for tokens 4 and 6. Because most participants responded between 1 and 4, this is unlikely to have adversely affected the results.

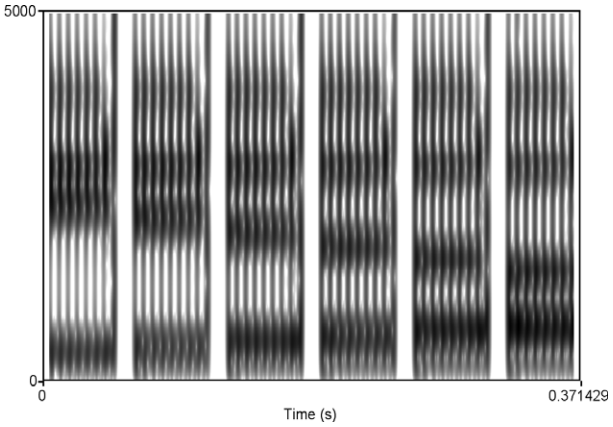

Figure 1. Spectrogram of the six vowels /ɪ/ tokens, ordered from most Australian to least Australian

formant values in the stimulus tokens is shown in Table 2. While the tokens are generally positioned near token 4 in the continua, some individual tokens are closer to tokens 3 or 5.

A total of 49 New Zealanders, 31 females and 18 males, took part in the experiment. The experiment was conducted in a quiet room over headphones. The participants were told that they were participating in an experiment testing the adequacy of synthesized voices. They were asked to both read and listen to the sentences, paying particular attention to the underlined word in each sentence. After each sentence, they were asked to listen to the synthesized vowels and to indicate on the response sheet the token that best matched the vowel in the underlined word. At the top of each response sheet was the word 'New Zealander' in one condition, and the word 'Australian' in the other. Participants' attention was not overtly drawn to this label. Participants were not explicitly told that the label was the nationality of the speaker, as we hypothesized that the mere presence of the label would bias perception. 26 participants took part in the NZ

Table 2. Range of formant frequencies in the stimulus sentences

	/ɪ/		/ɛ/		/æ/	
	F1	F2	F1	F2	F1	F2
Mean	556	1684	349	2373	576	2023
Min	504	1495	317	2246	502	1845
Max	642	1856	402	2552	630	2219

condition, and 23 participants were in the Australian condition. Each participant took part in just one of the two conditions.

The experiment was conducted in two phases, each involving slightly different methodology. Several participants in phase one remarked that the task was difficult, and so we made some small adjustments in an attempt to make the task less stressful. This methodology change also coincided with a change in the identity of the experimenter. The following differences existed between participants in 'phase 1' (20 participants – 10 in the Australian condition) and those in 'phase 2' (29 participants – 13 in the Australian condition).

1. The participants in phase 1 heard the sequence of vowels in quick succession, and had to choose which one was the best match. They heard the continuum twice. The participants in phase 2 heard the continuum only once, but the tokens in the continuum were numbered, in order to make them easier to track.
2. The participants in phase 2 encountered the stimuli in a different order from those in phase 1, and had slightly less time to answer.
3. While phase 1 was administered by a male New Zealander (the 2nd author), phase 2 was administered by a North American female (the 3rd author).
4. After responding to the sentences in the perception task, all 29 of the participants in phase 2 were asked to complete a short questionnaire about their intuitions regarding the person who read the sentences. Included in the questionnaire were questions about the speaker's age, education, occupation, and gender. Participants were also asked an open-response question about where the speaker was from, and they were asked to be as specific as possible.

The results from phase 1 and phase 2 were highly comparable, and are pooled in this discussion. In the regression models we report, which phase the participant was in was included as a potential predictor, in order to assess the degree to which these changes may have affected the results. Our major results hold within both subsets of the data.

All participants were volunteers and received chocolate in exchange for their time.

5. Results

Just as with Niedzielski's (1999) study, the results of this experiment demonstrate that apparent dialect area influences vowel perception. Participants were more likely to report a high variant of /ɪ/ if included in the experimental condition where the word 'Australian' appeared at the top of the response sheet. As shown in Figure 2, participants in the NZ condition were more likely to respond with a higher token number (or more central vowel) than participants in the Australian condition.

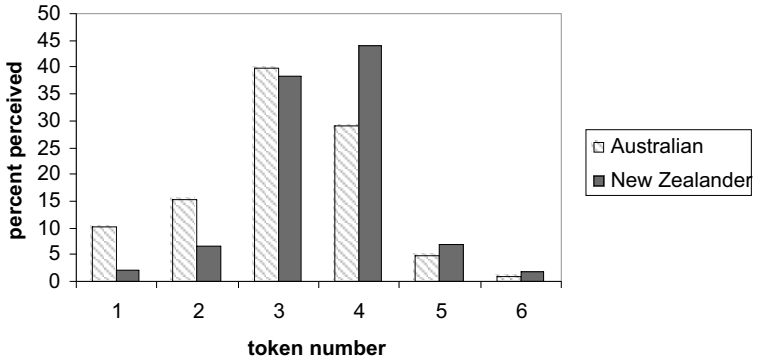

Figure 2. Relative frequency of participants' responses to the /ɪ/ vowels, separated by experimental regional label

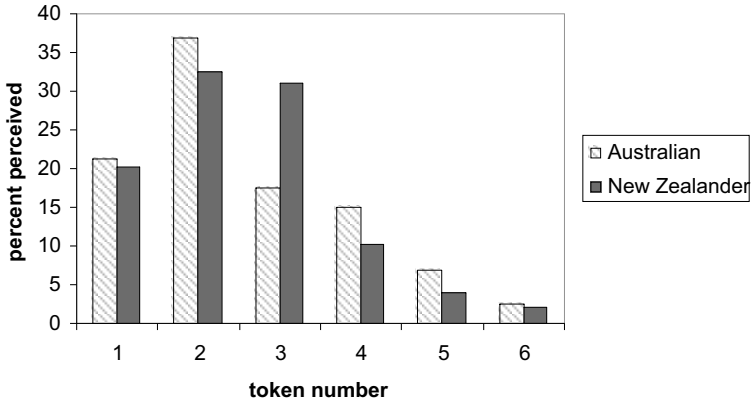

Figure 3. Relative frequency of female participants' responses to the /ɪ/ vowels, separated by experimental regional label

There is an overall tendency for subjects to respond towards the Australian end of the continuum compared to the speaker's average pronunciation, which probably relates to the fact that the male speaker recorded for the experiment has an extremely centralized /ɪ/, even for a New Zealander.

Interestingly, it is the female participants who carry the trend – the female participants in the NZ condition were considerably more likely to respond with a higher token number (or more central vowel) than female participants in the Australian condition. The results for the females only are shown in Figure 3.

While the trend of the female participants is in the expected direction, the tendency to respond to a particular token depending on regional label is ac-

tually in the opposite direction for the male participants. This should perhaps be treated with some caution, as only 7 male participants were included in the 'Australian' condition. Nonetheless, males who were included in the condition where 'New Zealander' appeared at the top of their response sheet were actually slightly more likely to respond with the lower token numbers that represent higher, more Australian-like variants. This difference between males and females is particularly surprising because Niedzielski (1999) reported that there were no gender differences in her results. Note, further, that females in general seem to be responding with more Australian variants overall. Comparing Figures 2 and 3 shows that the overall distribution for the female respondents is shifted somewhat leftwards, relative to the participants as a whole.

Because males and females appear to behave differently from one another during this perception task, it is important that a gender difference is taken into account when analyzing the data.

We also wanted to investigate the potential role of recency – did sentence-final tokens elicit different responses from sentence-medial tokens? As there was the potential for a variety of factors to influence individual responses, we fit a statistical model which attempted to take all of these into account.

In order to investigate the effect of apparent dialect area within the context of other potential influences on perception, an ordinary least squares linear model was fit to the data by hand using Harrell's (2001) Design library for the statistical tool R (R Development Core Team 2004).

The following factors were considered:

– Whether the participant was in the NZ or the Australian condition;
– Whether the participant was in phase 1 (run by the NZ experimenter) or phase 2 (run by the US experimenter);
– The participant's age, social class index score, and gender;
– The sentence number – which indicates how far through the experiment the response comes from;
– The actual F1 and F2 of the token the participant is trying to match;
– Whether the sentence contained the word *fish* – two sentences contained this word (one phrase medially and one phrase finally), and we speculated that the effect may be stronger on this word than any other;
– Whether the response was to a phrase-medial vs. phrase-final token.

All two way interactions were checked, and all three way interactions which seemed to us to be plausible. An Analysis of Variance table for the resulting linear model is presented in Table 3, and the associated co-efficients are shown in Table 4. The coefficients can range from negative infinity to positive infinity, and show the predicted extent and direction of each effect, with negative values indicating more Australian-like tokens and positive tokens indicating more NZ-like tokens. In the case of categorical variables, the variant which is not listed in the coefficient table has a coefficient of zero (by definition).

Table 3. Analysis of variance for /ɪ/ model (based on 980 responses from 49 speakers)

Factor	d.f.	F	P
regional label (factor+higher order factors)	4	16.87	< .0001
sex (factor+higher order factors)	2	21.98	< .0001
Sentence type	1	11.7	0.0007
fish (factor+higher order factors)	2	3.17	0.0426
social class (factor+higher order factors)	2	4.64	0.0099
experimenter (factor+higher order factors)	2	2.69	0.0682
number (factor+higher order factors)	2	2.18	0.114
regional label * sex (factor+higher order factors)	1	42.88	< .0001
regional label * fish (factor+higher order factors)	1	6.32	0.0121
regional label * social class (factor+higher order factors)	1	8.62	0.0034
experimenter * number (factor+higher order factors)	1	3.84	0.0502

Table 4. Coefficients for /ɪ/ model (based on 980 responses from 49 speakers)

Factor	Coefficient
Intercept	3.949
Regional label = New Zealander	−0.262
sex = M	0.529
sentence type = phrase-final	−0.205
fish = y	−0.254
social class	−0.009
Experimenter = US * number	−0.289
Number	−0.008
Regional label = New Zealander * sex = M	−0.836
Regional label = New Zealander * fish = y	0.502
Regional label = New Zealander * social class	0.012
Experimenter = US * number	0.010

Participants who were administered the condition with the 'New Zealander' regional label at the top of the response sheet responded with higher token numbers (more central vowels) than participants in the 'Australian' condition. In other words, subjects who were exposed to the 'New Zealander' regional label were more likely to respond with a higher token that represents a central, more NZ-like variant, and subjects who had response sheets with the 'Australian' regional label were more likely to respond with lower token numbers that represent higher, more Australian-like variants.[3] Like Niedzielski's (1999)

3. Note that the coefficient for 'New Zealander' in Table 3 is actually negative indicating lower token numbers (relative to the Australian label), and so more raised variants. This is because

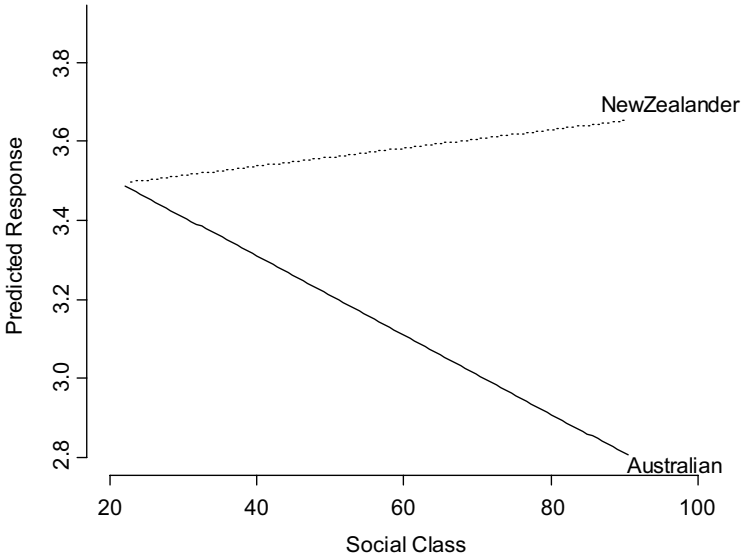

Figure 4. Response predicted by the model, as a function of social class and regional label

study, participants' vowel perception seems to have been influenced by their expectations regarding the dialect of the speaker. This interacts with three other factors: social class, whether the sentence contained *fish*, and participant sex.

The regional label interacts with social class. "Social class" is defined as the combined (averaged) parents' scores on the NZ Socioeconomic Index (NZSEI) – a NZ occupation-based social class indexing system devised by social scientists (Davis et al. 2003; Davis et al. 1997). Higher scores reflect occupations which tend to be more prestigious, associate with higher incomes, and require more tertiary education. We use parents' scores, because many of our participants are students, and so assigning them an occupation-based measure based on their own occupation is not possible. Figure 4 shows the nature of the social class effect – this shows the predicted effect of social class for the two regional labels when all other factors are held constant in the model. What this shows is that while there was not much effect of social class amongst the participants

of the presence in the model of the interaction between condition and social class. The lowest social class assigned to any of our participants is 22. Thus, if someone is in the New Zealand condition, the 'New Zealand' coefficient of $-.26$ is immediately counteracted by a minimum coefficient from this interaction of $(22 \times .012) = .264$. That is, the negative coefficient for New Zealander assumes a social class of zero, which is never the case. Figure 4 illustrates the social class/condition interaction in a way that clarifies the true direction of the effect.

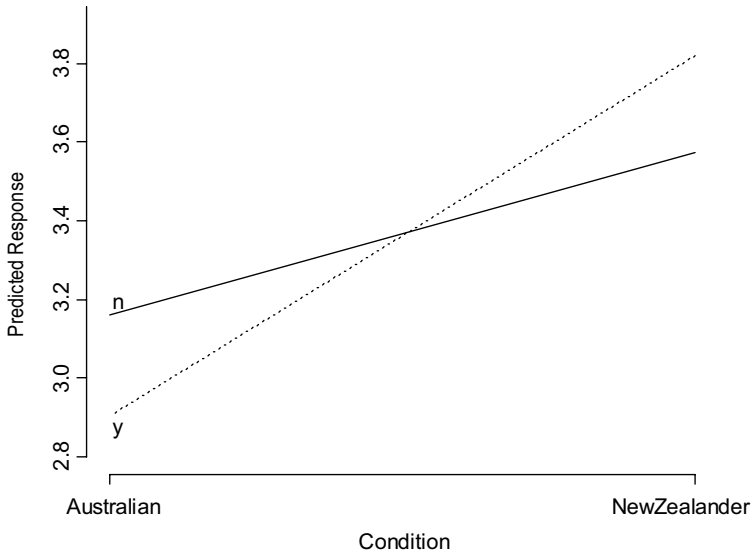

Figure 5. Response predicted by the model, as a function of regional label, and whether the sentence did ("y") or did not ("n") contain the word *fish*

in the 'New Zealander' condition, the participants in the 'Australian' condition were increasingly likely to have responded with an Australian-type variant as their social class index score increased. The higher the social class index score, the more the individual is influenced by the 'Australian' label.

The regional label also interacted with whether the sentence contained the word *fish*. As we had hypothesized, the effect of the regional labels was stronger for sentences containing this word than for the other sentences. This interaction is shown in Figure 5. Note that these tokens were not produced with extreme vowels. Indeed, the minor variation across the different lexical items played no role in participants' responses, as evidenced by the fact that the token formant values did not approach significance in this model. There is nothing about the *fish* sentences that makes them special, except for the very fact that they contain this word, which is iconic of the stereotype, and so is often used to exemplify the differences between NZ and Australian speech.

Figure 6 further illustrates this phenomenon by repeating Figure 3 (the results for the female participants), this time for the *fish* sentences only. The effect is very clear.

There is also an interaction between the sex of the subject and the regional label condition in which the subject participated, confirming the trend noted

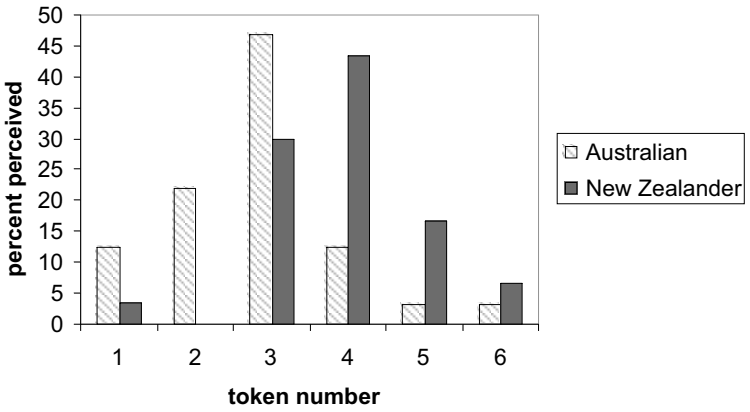

Figure 6. Relative frequency of female participants' responses to the two sentences in which *fish* was the target word, separated by experimental regional label

more informally above. The true magnitude of the interaction is difficult to assess, as only seven males participated in the 'Australian' condition. Continued recruitment of participants is required in order to determine whether or not the perception of male participants follows a trend in the opposite direction as that of the females or whether their perception is not significantly affected by the presence of either 'Australian' or 'New Zealander' on the top of their response sheet. The main effect of the regional label remains significant regardless of whether the interaction is included in the model. Also consistent with the trend observed above, there is a significant overall tendency for males to respond with more NZ-like tokens than the females – i.e., across both conditions.

The sentence type had a significant effect. This is the manipulation in which identical variants were included in the middle of a phrase, or phrase finally. In the phrase-final condition, participants matched the token to a step on the continuum immediately after hearing it. We included this because the task as Niedzielski framed it seemed very difficult. Participants were required to hold the token in memory until the end of the phrase. We speculated that this might be very difficult for them, and so they would be forced to rely on their stereotypes alone when making their choice. If this were so, then Niedzielski's result, while still fascinating, may not relate to speech perception at all.

The results show that there is an effect of how recently the token was played, but that this did not interact with the regional label. This lack of interaction indicates that the overall effect of the regional label condition is not an artifact of the difficulty of the task. There is simply an overall result that when the item

is phrase-final, all participants respond with a slightly more Australian variant. This is curious – the participants are actually less accurate when they have just heard the token than when they have had to hold it in memory.

The recency with which the participant has encountered the item matters. Exactly how to interpret the effect is not so clear. One clue comes from the fact that even in the NZ condition, participants' responses are more front and higher than the actual token produced. Note in Figures 2 and 3, that while many responses across both conditions are more front than token 4, very few are more central than it. It appears that the label 'New Zealander' is responded to with the 'canonical standard' NZE variant – a variant which is more front than that produced by most New Zealanders. Just as Niedzielski's participants perceived vowels as being conservative with respect to the Northern Cities Chain-Shift, so our participants perceived /ɪ/ conservatively in the context of the NZ short front vowel shift. Time holding the variant in memory, then, may have the effect of shifting it more in the direction of the participants' own speech. A second possible interpretation of this effect lies in the fact that our stimuli in fact confound recency with sentence length.[4] The sentence-medial tokens (which were responded to as more NZ like) have further linguistic material following them. The sentence-final tokens do not. It is possible that the following material contains phonetic cues which are associated with NZ – further raising the activation level of NZ indexed exemplars. An experiment which separately manipulated recency and sentence length would be required to untangle these two explanations.

Finally, there was a near-significant interaction between which phase of the experiment the participants were in, and how far through the experiment they were. Toward the beginning of the experiment, participants in the second phase of the experiment were more likely to respond with a more Australian variant than participants in the first phase of the experiment. Our tentative explanation for this trend relates to the different dialects of the experimenters. In past work, we have observed that being met by experimenters speaking different dialects of English can affect participants' behavior in speech perception tasks (Hay et al. to appear; Hay et al. 2006). The US experimenter's /ɪ/ vowel more closely resembles the 'Australian' end of the continuum than the NZ experimenter's does, and so this could reflect short term accommodation on the behalf of the participants. This effect does not quite reach significance, and so one should be careful not to over-interpret it. However we choose to leave it in the model to be sure that we are statistically 'subtracting' this potential methodological effect from the rest of the results. All other results reported here remain significant even if these two interacting factors are removed from the model.

4. We are grateful to Cynthia Clopper for pointing out this possible interpretation of this effect.

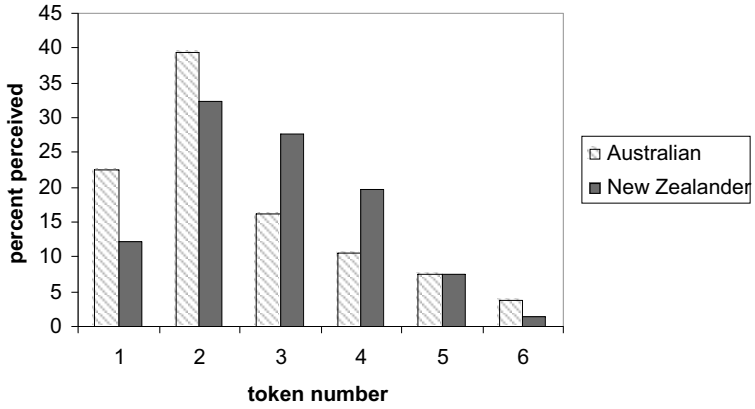

Figure 7. Relative frequency of female participants' responses to the /æ/ vowels, separated by experimental regional label

We now briefly examine the effect, if any, that the regional label had on the less salient filler variants included in the experiment, namely the vowels /æ/ and /ɛ/. Due to the smaller number of stimulus sentences containing these vowels, the results are not as robust as they are with /ɪ/ and should be viewed more tentatively.

The primary result regarding the effect of the two regional labels on the perception of /æ/ replicates the result for /ɪ/. Participants were more likely to respond with a more Australian-like token if they participated in the 'Australian' condition. As with /ɪ/ this result also interacts with the participant's sex. As with /ɪ/, females and males appear to behave differently, and for this variant, it is even more clearly the female participants who carry the trend in the expected direction. The graph for the female participants is shown in Figure 7.

In order to investigate the significance level of this trend, an ordinary least squares linear model (displayed in Tables 5 and 6) was fit to the data by hand. Included as factors in the final model was the combined NZSEI of the participant's parents, and an interaction between the regional label and whether a participant was male or female.

Overall, males were more likely to respond with a token toward the NZ end of the continuum. This difference between males and females reaches significance ($p < 0.01$). The condition in which a subject participated was a significant predictor of perception, and there is also a significant interaction between regional label and whether the participant was male or female ($p < 0.001$). Overall, subjects were significantly more likely to respond with a more NZ-like token if included in the NZ condition than when included in the Australian condition ($p < 0.01$), and as with results from /ɪ/, it was the female partic-

Table 5. Analysis of variance for /æ/ model (based on 490 responses from 49 participants)

Factor	d.f.	F	P
social class	1	6.69	0.01
regional label (factor+higher order factors)	2	5.58	0.004
sex (factor+higher order factors)	2	6.06	0.0025
regional label * sex (factor+higher order factors)	1	11.14	0.0009

Table 6. Coefficients for /æ/ model (based on 490 responses from 49 participants)

Factor	Coefficient
Intercept	3.061
social class	−0.009
Regional label = New Zealander	0.295
sex = M	0.562
Regional label = New Zealander * sex = M	−0.793

ipants who carried the trend. In other words, female participants in the NZ condition were more likely to respond with a raised variant representative of the NZ dialect. Males on the other hand were actually slightly more likely to respond with a raised variant if taking part in the Australian condition. The direction of the trend with /æ/ perception and the difference between male and female participants is comparable to the results from /ɪ/, despite the fact that the raised variant of /æ/ produced by New Zealanders is much less stereotyped or salient.

A participant's social class also influences which variant each participant perceived. Subjects whose parents' occupation received a lower NZSEI score were more likely to respond with a higher (more NZ-like) token number, indicating that they perceived the target vowel as more raised. This trend seems to reflect the variants produced by individuals from different social classes. New Zealanders from higher social classes tend to produce lower, more conservative variants of /æ/, and it appears that they also perceive lower, more conservative variants of /æ/ (a similar effect of social class on the perception of /æ/ was also reported by Drager 2005).

Variants of /ɛ/, which are also said to be below the level of consciousness, also revealed a link between vowel perception and social class of the perceiver. For this vowel, however, there is no difference in perception between participants in the 'New Zealander' and 'Australian' conditions. The significance levels for each of the factors included in the final model are presented in Table 7, and the coefficients are in Table 8.

Table 7. Analysis of variance for /ɛ/ model

Factor	d.f.	F	P
social class	1	18.20	< .0001
Experimenter (factor+higher order factors)	2	7.51	0.0006
sex (factor+higher order factors)	2	8.01	0.0002
Experimenter * sex (factor + higher order factors)	1	14.00	0.0001

Table 8. Coefficients for /ɛ/ model

Factor	Coefficient
Intercept	3.093
social class	−0.015
Experimenter = US	0.436
sex = M	0.304
Experimenter = US * sex = M	−0.829

As shown in the coefficient table in Table 8, the social class of the participant is linked with perception, so that participants with lower social class indices are more likely to respond as perceiving a higher token number, or more NZ-like /ɛ/ vowel. This trend is in the expected direction. Individuals from higher social classes tend to produce lower variants of /ɛ/ (see, e.g., Maclagan et al. 1999; Maclagan and Hay 2007), and here we again see that perception patterns are aligned with production patterns.

There is also a curious interaction between whether the participant was male or female and the experimenter who administered the experiment. As shown by the coefficients in Table 8, female participants were more likely to perceive a higher token number (or more NZ-like tokens) when met by the female American experimenter (in phase 2 of the experiment) than when met by the male New Zealander (phase 1). Male participants on the other hand were more likely to perceive a higher token number when met by the male experimenter from NZ. This interaction is significant and intriguing. It is difficult to say whether the difference in perception between the two groups is due to the nationality or gender of the experimenter, or to some difference caused by the subtly different methodologies used in the two parts. We have no good explanation, as the US experimenter's /ɛ/ vowel in fact resembles the Australian variant more than the NZ variant. We could perhaps speculate that the males who met the female experimenter responded to her more as an American (corresponding to lower tokens), and the females responded more to her as a woman (corresponding to higher tokens). Certainly in past work we have seen that females are more sensitive to gender-specific differences in speech perception than men are (Drager

2005; Hay et al. to appear). However this is highly speculative, and as this effect is not central to our main research questions, we leave it as an intriguing pointer towards further required work.

6. Post-task interview

Informal discussions with the first group of participants who participated in the experiment indicated that they did not, in fact, believe that the voice belonged to an Australian. In phase two this was tested more formally. Participants were administered a short survey asking about the participants' intuitions concerning the voice in the experiment. They were not exposed to further recordings but were expected to rely on their memory of the voice. Using a multiple choice format, the survey collected information on the perceived gender, age, education, and occupation of the speaker who read the sentences. Participants were also asked to indicate the likely nationality of the speaker, and they were asked to be as specific as possible.

All participants accurately indicated that the speaker was male and under the age of 45. There was a wide range of occupations and education levels attributed to the speaker. All but one of these participants indicated that the speaker was a New Zealander. This includes those who were in the Australian condition. There is a sense in which our manipulation failed, then, as participants with 'Australian' written on their answer sheets did not believe that they were listening to an Australian. Despite this 'failed' manipulation, however, having this word on their answer-sheet did affect participants' behavior. Participants seem to have been influenced by the "apparent dialect" of the voice despite being explicitly aware that the voice was in fact a New Zealander.

7. Discussion

In both Niedzielski's study and our study, a regional label at the top of an answer sheet affected participants' behavior in a speech perception task. Niedzielski (1997) argues that her results are a reflection of a speech perception strategy involving a 'filter'. A filter will reflect an individual's stereotypes and expectations about the person s/he is listening to, and characteristics not matching this filter do not pass through.

There are two important differences between Niedzielski's study and ours. First, Niedzielski verbally explained to each group where the speaker was from. Thus the participants had two pieces of information about the speaker origin – the verbal explanation, and the written label on the answer-sheet. In our study, the speaker origin was not discussed, and the written label was the only cue to speaker origin. In fact, one participant asked the experimenter if the label

implied that the speaker was going to be Australian, and (in order to be as consistent as possible with information provided to the other participants) the experimenter replied that she didn't know.

Second, Niedzielski's participants seem to have (for the most part) believed the manipulation. However, when our participants were questioned at the end of the study, all bar one indicated that they believed that the speaker was a New Zealander. Thus, either they didn't notice the speaker label at the top of the answer-sheet, or they noticed it, but then discounted it when they heard the voice of the speaker. Given that there were significant differences between the participants' responses in our two conditions, the former is unlikely to be true. Thus, we can assume that the participants were actually quite accurate at identifying the origin of the NZE speaker, and so ignored, or did not believe, the label on the answer-sheet.

Given that the participants did not think that they were listening to an Australian, it is difficult to situate these results within an analysis involving a filter during speech perception. This is especially so as the regional label did not interact with a token's position in the experiment. That is – the difference between the two regional labels is stable throughout the entire experiment. If there was such a filter, the 'New Zealandness' clearly passed through the filter. Once it did, then, why was the filter not removed or altered? An analysis which does not rely so centrally on overt beliefs and stereotypes is likely to be more effective at explaining these effects. We believe that exemplar (or 'multiple-trace') theory is a good candidate.

Exemplar models of lexical representation assume that individual speech utterances are stored in the mind as separate exemplars, and these exemplars are activated during both the production and the perception of speech (Pierrehumbert 2001; Johnson 1997). Lexical representations therefore consist of distributions of remembered exemplars, complete with phonetic detail. (Pierrehumbert 2001, 2002; Goldinger 1997; Johnson 1997, to appear; Foulkes and Docherty to appear; Bybee 2001; Scobbie 2005).

A range of information is indexed to each exemplar, including social information regarding the person who produced it (Johnson et al. 1999; Foulkes and Docherty to appear; Hay et al. to appear; Drager 2005). Thus, individual exemplars may be indexed to the identity of the speaker (e.g., "mum"), regional labels, gender labels as well as any number of other potential factors, including contextual information. During speech perception, exemplars are activated depending on their acoustic similarity to the incoming utterance. Exemplars which are indexed to appropriate contexts and social factors also receive activation. Words are identified by attending to the relative activation levels in the candidate distributions.

Many authors have recently argued that exemplar approaches to speech perception and production may offer a unified account of a range of phenom-

ena which would otherwise appear unrelated. These effects are wide-ranging, and include the role of frequency in production, perception and sound change, effects of speech accommodation and style-shifting, the evolution of certain phonological phenomena, acquisition processes, the gradience of well-formedness judgements, listener sensitivity to individual voices, and changes in individuals' phonology during adulthood (see, e.g., Docherty and Foulkes 2000; Bybee 2001; Smith 2003; Mendoza-Denton et al. 2003; Hawkins 2003; Scobbie 2005; Hay and Sudbury 2005; Foulkes and Docherty to appear; Harrington to appear; Wedel this volume).

In speech perception, those exemplars which are activated on the basis of both social and acoustic information reach full activation faster. This can account for a range of social effects on speech perception, such as the fact that consonant and vowel boundaries are perceived differently for male and female voices, as well as for stereotypical vs. non-stereotypical speakers (Johnson et al. 1999); that individuals can more accurately distinguish between phonemes participating in a merger in progress when they think they are listening to an older speaker (Hay et al. to appear), and that in the context of a chain-shift, individuals perceive vowel boundaries in different places depending on the perceived social characteristics of the speaker (Drager 2005, 2006). The automatic activation of socially appropriate exemplars explains these results without the need for assuming that the listener makes overt choices, invokes stereotypes, or utilizes filters.

This link between 'social labels' and individual, phonetically rich exemplars provides a natural explanation for such effects. The acoustic activation of exemplars will automatically activate the social categories they are associated with. This enables us, for example, to state with some confidence whether a voice is male or female, American or British, or someone we know well vs. someone we don't. Likewise, activating a social category raises the activation level of exemplars associated with that category. This raised activation effectively acts as a bias in perception (or indeed, in production – hence effects of speech accommodation). This bias is usually highly beneficial – making us more likely to identify words on the basis of exemplars which are appropriate to the context, and enabling us to identify quite different acoustic signals as being examples of the same 'word' when spoken by different speakers. Thus, activation of social concepts can effectively 'prime' the phonetic memories which are associated with them.

Individual exemplars may have different resting activation levels, depending on the social weight we place on them, as well as their recency. Exemplar memories decay over time. Links between social labels and exemplars may also have differing strengths depending on the importance of the link, and these may also decay, especially if the informativeness of the category is not seen as relevant. One might store, for example, the fact that a speaker was wearing a

green dress. But this link is likely to decay rapidly due to the lack of distributional reinforcement from the rest of the lexicon indicating that this is a relevant category. The degree to which specific exemplars can be "primed" with social information, then, is likely to be proportional to the strength of the association between those exemplars and the relevant social category.

In our very artificial task, participants are asked not to identify which distribution is most activated (which is the usual task in speech perception), but rather which *part* of a distribution is most activated. Here, then, when exposure to the concept 'Australia' primes the exemplars which are associated with Australia, this shifts the center of the activated portion of the distribution towards the more 'Australian' end.

An exemplar account therefore seems a promising avenue for accounting for our results. However at first glance, an exemplar approach actually struggles with Niedzielski's result. This is because Detroiters' accumulated experience of 'Canadian' words containing /au/, and Detroit words containing /au/ do not actually differ. The realization of the vowel is reportedly the same for both sets of speakers. Thus, activating the 'Canadian' exemplars and activating the 'Michigan' exemplars should lead to the same kind of bias.

Niedzielski's participants matched the vowels of their fellow Detroiter to more standard vowels than the speaker actually produced. Niedzielski discusses how the participants regard themselves to be speakers of 'Standard American English', and – when asked where in the US the most standard English is to be heard, many of them answered "right here". The clue, then, may come from the fact that Detroiters do not consider 'Detroit' a relevant social characteristic for speech. When encountering speech from Detroit speakers, these may be indexed to 'Standard American English', along with other tokens of Standard American English, which have been encountered from speakers from elsewhere and from the media. The label 'Michigan', then, activates the 'Standard American English' exemplars. While these include the Canadian-raised Detroit variants, they also include many other non-raised variants – which shift the center of the activated distribution to be lower. When 'Canadian' is activated, however, the majority of the relevant exemplars contain Canadian-raising, and so the match between the acoustic signal and the exemplar cloud is more accurate. Thus the 'stereotype' affects speech perception precisely because when Detroiters encounter other Detroit speech, the fact that they come from Detroit is regarded as incidental. Exemplars from Detroit speakers are indexed as 'Standard American'. It is possible that they are also indexed as 'Detroit', but that this indexing rapidly decays as it is perceived as irrelevant.

In the context of this interpretation of Niedzielski's results, we can now return to our own results for /ɪ/. Our participants have a /ɪ/ distribution which contains a range of phonetic variants. These variants will include front and

raised variants produced by Australian speakers, and central variants produced by New Zealanders. These exemplars will be indexed with social information, including – when it's known – the regional origin of the speaker.

Within an exemplar interpretation of the task, the task proceeds as follows.

The participant sees the label on the answer-sheet. When this is 'Australian', it raises the activation level of exemplars indexed as Australian. When it is 'New Zealander' it raises the activation level of all exemplars indexed as NZ. During the experiment, exemplars which are phonetically similar to the acoustic signal are activated. If the 'Australian' exemplars have previously been primed, this priming has the effect of shifting the overall center of the activated distribution towards more 'Australian' variants.

When the token is phrase-final, the participant responds immediately with the step on the continuum that best matches the region of most activation in the exemplar cloud. When the token is phrase-medial, more NZ-like variants are selected. This may be because additional NZE cues following the token may further activate NZ exemplars. Or it may be because the participant must hold the token in memory until the end of the phrase, and activation continues to spread during this time, shifting the distribution towards the participant's own pronunciation – i.e., toward a more NZ exemplar.

The raising of /ɛ/ and /æ/ and the centralization of /ɪ/ are sound changes which are lead by speakers from lower social classes in NZ (Maclagan et al. 1999).

Participants with lower social class index scores are therefore more likely to associate with speakers who have higher /ɛ/ and /æ/, and more centralized /ɪ/. Although participants were not asked specifically about travel to Australia, it seems likely that participants with lower social class indices are also less likely to have traveled to Australia, or to have received extensive exposure to Australian English (and would have fewer tokens indexed as 'Australian'). We appear to see reflexes of both of these effects in the data. For /ɪ/, we see that the effect of the regional label increases with increasing social class index scores, presumably reflecting a greater degree of exposure to Australian English. For /ɛ/ and /æ/, we find an overall effect in which participants with lower social class indices respond with higher variants, reflecting their own productions and those of their communities.

That the lexical distributions are 'lexical' (rather than being distributions of vowels) is evidenced by the fact that the effect is strongest for the word *fish* – the word which is most often used to 'perform' the difference between Australian and NZ accents. When New Zealanders characterize Australians (and vice-versa), the *fush and chups* or *feesh and cheeps* performances are actually more extreme than would tend to be produced naturally. However it is nonetheless likely that these performed forms still get stored, and in fact their association with 'New Zealandness' or 'Australianness' will be even more robust than

forms that just happen to be produced by a New Zealander or Australian. Thus, (i) stored exemplars of the word *fish* are likely to contain more extreme realizations than other words, and (ii) through the strong association between the word and national stereotypes, the concept of Australianism may boost activation to these fronted/raised versions of the word *fish* more than for any other single word. We would argue that the effect of the stereotype here may be simply that the indexing between the regional labels and the phonetic forms is highly robust and salient, and not as prone to decay as associations which receive less reinforcement and are perceived to be less relevant. This effect, then, can be predicted without positing explicit filters which are put in place during speech perception.

Indeed, because our participants in the 'Australian' condition did not believe that they were listening to an Australian, it seems that expectations about the speaker are not driving this effect. Rather it appears to be mere exposure to the *concept* of Australianism that shifts their behavior in this task. This interpretation makes some rather strong predictions. For example, we predict that we could rerun this experiment in a way in which participants are exposed to the concept, without any implication that this is related to the speaker at all. For example we could run one group with a big stuffed Kiwi on the desk (or even a poster in the background), and a second group with a Kangaroo. If the result was replicated, this would be truly remarkable. Our current interpretation of the results suggests that it should be.

There is some indication in our results that the dialect of the experimenter plays a role in people's perception too. This is hard to conclude from these results alone, as the change in experimenter co-occurred with minor changes in methodology. However in other work we have also seen effects in which the dialect of the experimenter shifts participants' behavior in a speech perception task (Hay et al. to appear; Hay et al. 2006), analogous to the potential experimenter effects we see here. However the identity of the experimenter certainly doesn't shift participants' overt expectations about the voice in the perception experiment. Indeed in the Hay et al. (to appear) experiment, participants were explicitly told that they would be listening to a New Zealander, and the accent of the stimulus voice bore little resemblance to that of the experimenter. It seems to us desirable to be able to account for effects such as the experimenter effect and the regional label effect using a single model. A model that relies on filters expressing listeners' overt expectations about the speaker does not make any predictions about the effect of recent interlocutors. However, an exemplar theoretic account seems more promising in offering a unified explanation of the range of effects of social factors on speech perception. A recent interlocutor will have the effect of raising the activation level of exemplars associated with that speaker, and perhaps (through resonance – see Johnson to appear) with social characteristics associated with that speaker.

The results for /æ/ and /ɛ/ are based on fewer tokens, and so are more speculative. Our interpretation suggests that overt stereotyping is not a necessary precondition for the effect (although it predicts that the effect should be strongest where strong stereotypes exist). The fact that the basic result is also present for /æ/ indicates that the effect may, indeed, surface in the absence of any apparent overt stereotype. What we cannot yet explain is why the effect should be present for /æ/ but not /ɛ/. Potential answers may lie in investigating whether there is a reliable difference in magnitude between the /æ/ and /ɛ/ differences between the two dialects, and/or whether there is a difference in salience or awareness between the two variables. The acoustic comparison provided by Evans and Watson (2004) certainly does show a greater separation between the two dialects for both /ɪ/ and /æ/ than for /ɛ/.

A final mystery relates to the gender differences in this experiment. Much of what we say above applies to the female participants only. It should be noted that we have more female participants than male participants in the experiment. However there is *no sign* in the male data that the manipulation is having the hypothesized effect. This also somewhat reminiscent of effects we have seen in previous work. Drager (2005) shows that females' perceived boundary between /æ/ and /ɛ/ is higher for female voices than male voices (which is consistent with reports that females are ahead in the ongoing sound change; see, Maclagan et al. 1999; Langstrof 2003, 2006; Hay et al. to appear) show that females were more accurate at hearing the difference between the vowels in *air* and *ear* (undergoing a merger in NZE) when listening to a female voice than a male voice. Male participants, on the other hand, were not so sensitive to the gender of the voice.

Eckert (2000) studied a group of high school adolescents, and found that the girls in her study tended to exploit variation to display group membership through language more than boys did. Drager (2005) suggests that, just as females appear to use language more than males to assert their identity, they may also be more aware of the relationship between variability in speech and social characteristics. Drager argues that females may therefore index their exemplars with a larger amount of social detail, and/or place more weight on this social detail. More extensive work is certainly required on this topic. Given that there is a vast literature on gender differences in the production of sociolinguistic variability (see, e.g., Cheshire 2002 for a review); it would perhaps not be surprising if there were pervasive gender effects in sensitivity to social factors in speech perception. As the body of work examining the effect of social factors on speech perception grows (which we certainly hope it does), this will certainly be something to watch out for.

A related mystery is why female participants perceived more Australian-like variants than the males did, overall. We have no good explanation for this. Bell (1997) and Maclagan et al. (1999) argue that females lead in the cen-

tralization of /ɪ/ in NZ. Thus, if participants' perception parallels their own perception, we would expect females to perceive more centralized tokens than males, which is the opposite of what we observe here. On the other hand, in Drager's (2005) categorical perception experiment, she finds that females have less innovative boundaries between /æ/ and /ɛ/ when listening to male voices than female voices, but the opposite is true (though less dramatically) for male participants. Untangling the true source of this trend, then, may necessitate re-running the experiment with stimuli produced by a female voice.

8. Conclusion

New Zealanders perceive vowels differently depending on whether their answer-sheet is labeled 'Australian' or 'New Zealander'. This is despite reporting that they recognize the accent as belonging to a New Zealander. We argue that this can be well accounted for by exemplar models of speech perception, in which lexical representations are distributions of remembered exemplars, complete with phonetic detail. These exemplars are indexed to social information such as gender, age and nationality.

The results raise three major challenges for future work. The first is the degree to which comparable effects may be attainable by exposing participants to objects or ideas associated with different groups. Is activating a concept sufficient, or does this manipulation have to be closely linked to the perception task and/or the speaker? The second is the issue of perceptual salience, and the degree to which variables are above or below the level of consciousness. Do listeners need to be overtly conscious of a distinction for it to be used in speech perception? The results with the /æ/ vowel suggest that the answer may be no, but we don't yet possess concrete evidence that individuals are not aware of this difference between NZ and Australia. In order to untangle the role of awareness, this type of work will need to be systematically conducted on variables which are variably salient, while independently controlling (or manipulating) degree of acoustic similarity. Finally, this experiment hints at some intriguing differences between our male and our female participants. It would certainly be premature to conclude that there are gender differences in the storage and/or exploitation of social information. But this will be something to consider as work on this topic progresses.

University of Canterbury

References

Bauer, Laurie and Paul Warren (2004). New Zealand English phonology. In *A Handbook of Varieties of English: A Multimedia Reference Tool*, vol. 1, B. Kortmann, W. Schneider, K. Burridge, R. Mesthrie and C. Upton (eds.), 580–602. Berlin: Mouton de Gruyter.

Bayard, Donn (1990). 'God help us if we all sound like this': Attitudes to New Zealand and other English accents. In *New Zealand Ways of Speaking English*, Janet Holmes and Allan Bell (eds.), 67–96. Clevedon: Multilangual Matters.

— (1995). *Kiwitalk: Sociolinguistics and New Zealand Society*. Palmerston North: Dunmore Printing Company Ltd.

— (2000). New Zealand English: Origins, relationships, and prospects. *Moderna Språk* 94 (1): 8–14.

Bell, Allan (1997). The phonetics of fish and chips in New Zealand: marking national and ethnic identities. *English Worldwide* 18: 243–270.

Bybee, Joan (2001). *Phonology and Language Use.* Cambridge: Cambridge University Press.

Chesire, Jenny (2002). Sex and gender in variationist research. In *The Handbook of Language Variation and Change*, J.K. Chambers, Peter Trudgill and Natalie Schilling-Estes (eds.), 423–443. Massachusetts: Blackwell Publishers Ltd.

Cox, Felicity M. (2004). A question of broadness. Paper presented at the Australian Linguistic Society Conference, Sydney, 2004.

Davis, P., G. Jenkin, and P. Coope (2003). *NZSEI-96: An Update and Revision of the New Zealand Socioeconomic Index of Occupational Status*, Statistics New Zealand, Wellington.

Davis, P., K. McLeod, M. Ransom and P. Ongley (1997). *The New Zealand Socioeconomic Index of Occupational Status (NZSEI)*: Research Report #2, Statistics New Zealand, Wellington.

Docherty, Gerry and Paul Foulkes (2000). Speaker, speech and knowledge of sounds. In *Phonological Knowledge: Conceptual and Empirical Issues*, P. C. N. Burton-Roberts and G. Docherty (eds.), 105–129. Oxford: Oxford University Press.

Drager, Katie (2005). The influence of social characteristics on speech perception. MA Thesis, University of Canterbury, New Zealand.

— (2006). From *bad* to *bed*: the relationship between perceived age and vowel perception in New Zealand English. *Te Reo* 48: 55–68.

Eckert, Penny (2000). *Linguistic Variation as Social Practice*. Massachusetts: Blackwell Publishers Ltd.

Evans, Zoe and Catherine Watson (2004). An acoustic comparison of Australian and New Zealand English vowel change. In *Proceedings of the 10th Australian International Conference on Speech Science and Technology*, Macquarie University, Sydney, December 2004, 195–200.

Foulkes, Paul and Gerry Docherty (to appear). The social life of phonetics and phonology. *Journal of Phonetics*.

Giles, Howard, Nikolas Coupland and Justine Coupland (1991). Accommodation theory: Communication, context, and consequence. In *Contexts of Accommodation: Developments in Applied Sociolinguistics*, Howard Giles, Justine Coupland and Nikolas Coupland (eds.), 1–68. Cambridge: Cambridge University Press.

Goldinger, Stephen (1997). Words and voices: perception and production in an episodic lexicon. In *Talker Variability in Speech Processing*, Keith Johnson and John W. Mullenix (eds.), 33–66. London: Academic Press.

Harrell, Frank (2001). Regression Modelling Strategies. New York: Springer.

Harrington, Jonathan (to appear). An acoustic analysis of happy-tensing in the Queen's Christmas broadcasts. *Journal of Phonetics*.

Hawkins, Sarah (2003). Roles and representations of systematic fine phonetic detail in speech understanding. *Journal of Phonetics 31*: 373–405.

Hay, Jennifer, Paul Warren, and Katie Drager (to appear). Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics*.

Hay, Jennifer, Katie Drager and Paul Warren (2006). Cross-dialectal exemplar priming. Poster presented at 10th Conference on Laboratory Phonology, Paris, June 2006.

Hay, Jennifer and Andrea Sudbury (2005). How rhoticity became /r/-sandhi. *Language* 81 (4): 799–823.

Ivory, Helen (2005). Get with the lingo. *The New Zealand Listener* 199.3396: 11–17.

Jansen, Tore and Richard Schulman (1983). Non-distinctive features and their use. *Journal of Linguistics* 19: 321–336.

Johnson, Keith (1997). Speech perception without speaker normalization. In *Talker Variability in Speech Processing*, K. A. Johnson and J. W. Mullennix (eds.), 145–166. San Diego, CA: Academic Press.

— (to appear). Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics*.

Johnson, Keith, Elizabeth A. Strand and Mariapaola D'Imperio (1999). Auditory-visual integration of talker gender in vowel perception. *Journal of Phonetics* 27: 359–384.

Langstrof, Christian (2003). The short front vowels in New Zealand English in the intermediate period. *New Zealand English Journal* 17: 4–16.

— (2006). Acoustic evidence for a push-chain shift in the intermediate period of New Zealand English. *Language Variation and Change* 18 (2): 141–164.

Ludwig, Ilka (2005). "Six dancers" – An accent identification study on New Zealand and Australian English. Unpublished course project, University of Canterbury.

Maclagan, Margaret A., Elizabeth Gordon and Gillian Lewis (1999). Women and sounds change: Conservative and innovative behaviour by the same speakers. *Language Variation and Change* 11: 19–41.

Maclagan, Margaret and Jennifer Hay (forthcoming). Getting fed up with our feet. Contrast maintenance and the New Zealand English 'short' front vowel shift. *Language Variation and Change* 19: 1–25.

McMurray, Bob (in preparation). Klattworks: A [somewhat] new systematic approach to formant-based speech synthesis for empirical research.

Mendoza-Denton, Norma (1999). Style. *Journal of Linguistic Anthropology* 9 (1, 2): 238–240.

Mendoza-Denton, Norma, Jennifer Hay and Stefanie Jannedy (2003). Probabilistic Sociolinguistics: Beyond Variable Rules. In *Probabilistic Linguistics*, Rens Bod, Jennifer Hay and Stefanie Jannedy (eds.), 97–138. Cambridge/MA: MIT Press.

Niedzielski, Nancy (1997). The effect of social information on the phonetic perception of sociolinguistic variables. PhD Dissertation, UC Santa Barbara.

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

Pierrehumbert, Janet (2001). Exemplar dynamics: Word frequency, lenition and contrast. In *Frequency and the Emergence of Linguistic Structure*, Joan Bybee and Paul Hopper (eds.), 137–157. Amsterdam: John Benjamins.

— (2002). Word-specific phonetics. In *Laboratory Phonology*, vol. 7, Carlos Gussenhoven and Natasha Warner (eds.), 101–139. Berlin and New York: Mouton de Gruyter.

R Development Core Team (2004). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org.

Schilling-Estes, Natalie (2002). Investigating stylistic variation. In *The Handbook of Language Variation and Change*, J. K. Chambers, Peter Trudgill and Natalie Schilling-Estes (eds.), 375–401. Massachusetts: Blackwell Publishers Ltd.

Scobbie, James (2005). The phonetics-phonology overlap. QMUC Speech Science Research Centre working Paper WP1.

Smith, Rachel (2003). Influence of talker-specific phonetic detail on word segmentation. *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, 1466–1468. ICPhS.

Strand, Elizabeth A. (1999). Uncovering the role of gender stereotypes in speech perception. *Journal of Language and Social Psychology* 18 (1): 86–99.

Watson, Catherine, Jonathan Harrington and Zoe Evans (1998). An acoustic comparison between New Zealand, and Australian English vowels. *Australian Journal of Linguistics* 18 (2): 185–207.