Language Variation and Change, **12** (2000), 51–68. Printed in the U.S.A.

Acoustic evidence for vowel change in New Zealand English

CATHERINE I. WATSON
Macquarie University

MARGARET MACLAGAN
University of Canterbury

JONATHAN HARRINGTON
Macquarie University

ABSTRACT

This study provides acoustic evidence that in the last 50 years New Zealand English (NZE) has undergone a substantial vowel shift. Two sets of data are studied: the Otago corpus, recorded in 1995, and the Mobile Unit corpus, recorded in 1948. Both corpora have male and female speakers. The corpora were labeled, accented vowels were extracted, and formant values were calculated. The results of the formant analysis from the two corpora are contrasted. We provide evidence that in NZE /i/ has centralized, /e/ and /æ/ have raised, and the diphthongs /iə/ and /eə/ have merged. We argue that /i/ changed in quality not only because of crowding in the front vowel space, but also because it would be less likely misperceived as an unaccented vowel (i.e., as ə).

There has been growing evidence from recent experimental and impressionistic data that some vowels in New Zealand English (NZE) have changed in quality in the last 50 years, and it remains an unresolved issue whether these changes were motivated by a drag or a push chain effect. Some of the earliest accounts of NZE can be traced back to the late nineteenth century (e.g., McBurney, 1887). English was the predominant language of the colonizers who began settling in the country in large numbers in the mid-nineteenth century. By 1886 there were more New Zealand–born Europeans than immigrants (McKinnon, 1997: plate 30). Just before the turn of the twentieth century a distinct NZE accent was first noted (Gordon, 1998; McBurney, 1887; Wall, 1938). In 1938 Wall identified several features of the vowels in NZE which differed noticeably from Received Pronunciation; these included retracted first targets for the HIDE and HAY diphthongs, a fronted

We wish to thank the anonymous reviewers and Mary Beckman for their comments and Michelle Dawe for digitizing the tokens from the Mobile Unit database. The project on the study of the origins of New Zealand English was funded by the New Zealand Foundation for Research, Science and Technology. The digitization was also supported by University of Canterbury research grant number U6286.

An earlier version of the article was presented at Laboratory Phonology VI, York, July 1998.

and raised first target for HOW, a fronted first target for HOE, delayed targets for HEED and WHO'D monophthongs, and a fronted HARD.[1] It is likely that these features existed at the turn of the century (Gordon & Abell, 1990; McBurney, 1887), and studies of contemporary NZE (see, e.g., Bauer, 1986; Maclagan, 1982; Watson, Harrington, & Evans, 1998) have shown that these features remain.

The NZE vowel space, however, has not remained static over the last 100 years. Two features of contemporary NZE, the centralized HID vowel and the HEAR/HAIR merger, have only been discussed recently. Although centralized NZE HID has been noted for some time in unstressed positions, centralized NZE HID in stressed positions was first commented on in the mid-1960s (Gordon, 1998; Wall, 1964), and there is no mention of the HEAR/HAIR merger in NZE until the mid-1980s (see Gordon & Maclagan, 1985, 1989). There have also been observations about how NZE vowels have changed relative to the vowels of Australian English and of British English Received Pronunciation. Until the 1940s, two frequent observations were made about NZE: first, that it was similar to, though distinct from, Australian English and, second, that it was more similar to the "King's English" than to Australian English (see, e.g., Bennett, 1943; Gordon, 1991; McBurney, 1887; Wall, 1938). However, modern studies of NZE vowels (e.g., Bauer, 1994; Maclagan, 1982; Watson, Harrington, & Evans, 1998; Wells, 1982) have revealed that, while the similarity to Australian English vowels remains, the differences between NZE and Received Pronunciation vowels are now far greater than those between Australian English and Received Pronunciation vowels.

Recent impressionistic analyses of old and contemporary NZE have confirmed that vowel changes have taken place in NZE. Trudgill, Gordon, and Lewis (1998) showed that the NZE HID vowel was not centralized 50 years ago, and they represented it as the vowel [ɪ] (just above close mid, and nearly front). Furthermore, although they argued that NZE HEAD and HAD were always close in quality, they demonstrated that HEAD and HAD have risen in the last 50 years. Compatibly, Woods (1997) suggested that HEAD has risen much more than HAD. She also showed that the first target of HOW has risen, suggesting that it is linked to the rising of HAD. Maclagan and Gordon (1996) argued that the recent HEAR/HAIR merger in NZE is also evidence of a vowel shift happening in NZE, with the first target of HAIR rising as HEAD rises. This finding is supported by Watson et al. (1998) and Batterham (2000), but see Holmes and Bell (1992) for a different interpretation.

In addition to a raised HEAD and HAD in modern NZE and Australian English, they both have a fronted HUD relative to Received Pronunciation (Watson et al., 1998). We know very little about when the fronting of HUD took place. Although McBurney in 1887 commented on the "peculiarity" of the NZE HUD, Maclagan (in press) found no HUD lowering for ten speakers born in New Zealand between 1865 and 1886. Furthermore, Wall in 1938 said he did not hear any difference between NZE HUD and that spoken by speakers of Received Pronunciation. Nevertheless, modern NZE HUD forms a tense/lax pair with HARD (Watson et al., 1998). For the latter vowel there is strong evidence that it was fronted as early as the late 1880s (McBurney, 1887).

Not all the changes in NZE observed in earlier studies have remained. Turner (1970) observed that the HEAD monophthong was becoming diphthongal. However, in the recent study of NZE by Watson et al. (1998), no evidence for diphthongization was found. Moreover, Maclagan (1998) showed that the diphthongization of HEAD was a change that did not become established. By contrasting NZE speakers aged 45 to 60 years with those aged 20 to 30 years, she showed that the younger group diphthongized HEAD substantially less than did the older group. Maclagan noted that the younger group had closer productions of HEAD compared to the older group. This suggests that there was a time at which the HEAD vowel could have either risen more or become diphthongal, and that although in the older generation of speakers there was a noticeable tendency to diphthongize HEAD "their children reverted to the pattern of [HEAD] rising" (Maclagan, 1998).

In this study we compare NZE speech recorded 50 years ago (old NZE) with modern NZE on various acoustic measures. Part of the motivation for the present investigation is that the earlier impressionistic studies of vowel change in NZE over the last 50 years were restricted to HID, HEAD, HAD, and HOW, and it is not clear what is happening to the other vowels in the front area of the vowel space. Modern NZE has a raised HEED, a fronted HUD (compared to Received Pronunciation), and a raised and fronted HERD (Maclagan, 1982; Watson et al., 1998), but it is not known when these features became part of NZE. We are interested to see if these features existed 50 years ago. Furthermore, since it has been suggested that the HEAR/HAIR merger is related to a HEAD rising, we also consider the position of the first targets of these diphthongs to see how much they have changed over the last 50 years.

METHOD

Speakers

The data for old NZE was selected from a corpus of recordings gathered in the late 1940s from over 300 speakers from mainly the Waikato and Otago provinces of New Zealand (the Mobile Unit corpus; see Lewis, 1996). Because the aim of the Unit was to record pioneer reminiscences, many of the recorded speakers were elderly, although some middle-aged speakers were recorded too. To avoid any possibility of formant structure changes associated with aging (Jaberi & Stone, 1997), we selected from the youngest speakers in the database. At the time of this study, four speakers were available to us who fit this criterion, two men (referred to as TM and JM) and two women (referred to as EB and MS). They were between 51 and 54 years of age and were all born in New Zealand between 1894 and 1899. The speakers can be considered second generation New Zealanders. Three of the speakers had at least one parent born in New Zealand. The mother of the fourth speaker was born en route to New Zealand. All four speakers were born and lived in the South Island province of Otago and had farming backgrounds. The recorded occupations for the four speakers were spinster for the two women, farmer for one of the male speakers, and provincial newspaper editor for the other.

The speakers would have been school-aged at the time when a distinct New Zealand accent was becoming apparent. The school inspectors of the time were noticing an "objectionable colonial dialect" (Gordon, 1998:66). It is difficult to say how typical these speakers were of their time, but in comparison to the other 300 speakers in the corpus they were not considered "unusual" (Elizabeth Gordon, personal communication).

The data for modern NZE were taken from the Otago speech database (Sinclair & Watson, 1995; Watson et al., 1998), which includes speech from 11 men and 10 women, all between ages of 16 and 33. All speakers were university-educated, with the exception of two of the female speakers who were senior high school students. Most of the speakers were students at the University of Otago. The speakers came from all over New Zealand. Although several speakers came from Southland, where postvocalic /r/ is still used after some vowels, care was taken to factor out this influence (see Watson et al., 1998, for further details).

Materials

We extracted accented vowels—that is, vowels with sentence stress—from the continuous speech of the Mobile Unit corpus. We analyzed only accented vowels because their vowel targets are less likely to be undershot than are those of unaccented vowels (Harrington, Beckman, & Fletcher, 2000; de Jong, 1995; Palethorpe, Beckman, Fletcher, & Harrington, 1999; Summers, 1987; see also Moon & Lindblom, 1994). The phonetic context was variable. Most of the initial consonants that preceded the vowel were either bilabial or dental-alveolar and were either stops or fricatives. Most of the consonants that followed the vowel were dental-alveolar and were either stops or nasals. In addition, any tokens which followed /r/ and /w/ or preceded a dark /l/ were excluded because these consonants cause strong undershoot effects on the formants of neighboring vowels (Harrington & Cassidy, 1999; Lehiste, 1964). We studied 6 monophthongs and 2 diphthongs: these were the vowels in HEED, HID, HEAD, HAD, HUD, HERD, HEAR, and HAIR. Table 1 lists the number of tokens for each vowel analyzed for each speaker.

As described in Watson et al. (1998), the Otago speech database talkers read citation form productions of 129 different words. For this study, the monophthongs and diphthongs of interest were extracted from 20 of the possible 129 words. The phonetic context of the vowels varied. The criterion that determined the choice of consonantal context was to produce a list of predominantly monosyllabic meaningful words in which the context varied minimally from word to word and in which the context minimally affected the vowel. The consonantal context that conformed as far as possible to this criterion was as follows: for the initial consonant, voiced bilabial oral stop; for the final consonant, either voiced or voiceless alveolar oral stop. As with the old NZE corpus, none of the vowel tokens were extracted from words where they followed /r/ or /w/ or preceded a

TABLE 1. *Number of tokens for each vowel, per speaker, from the old NZE data*

	Number of Tokens			
Vowel	EB	MS	TS	JM
HEED	13	14	20	18
HID	16	17	20	18
HEAD	21	19	18	26
HAD	18	20	25	22
HUD	15	24	18	15
HERD	4	19	22	13
HEAR	4	7	14	5
HAIR	3	11	3	6

TABLE 2. *Number of tokens for each vowel studied and words from which they were extracted from the modern NZE data*

		Number of Tokens	
Word	Vowel	Female	Male
bead, eat, pea, peat	HEED	30,30,30,30	33,33,33,33
if, pit, did	HID	29,30,30	33,33,33
egg, beg, pet	HEAD	30,30,30	33,33,33
at, sat, dad	HAD	30,30,30	33,33,33
utter, but, pub	HUD	30,30,28	33,33,33
irk, pert, bird, fur	HERD	27,26,30,30	28,24,26,27
ear, hear	HEAR	23,27	22,25
air, there	HAIR	28,26	30,28

dark /l/. Table 2 lists the vowels used in this study and the words from which the vowels were extracted.

Recording and digitizing

The Mobile Unit corpus was collected by the Mobile Disc Recording Unit of the National Broadcasting Corporation of New Zealand. An ex-army van, fitted out with recording equipment, traveled the country collecting conversations. All recordings were done onto vulcanite discs. There is considerable background noise in the recordings. This is due to three factors: the recording equipment, the recording environment (people's homes and community halls), and the deterioration of the vulcanite disc over time.

The Mobile Unit data was transferred to digital audiotape as part of the "Origins of New Zealand English" project (Lewis, 1996). The data were sampled at

22.05 kHz. However, due to the limitation of the earlier recording equipment, all speech information above 5 kHz was destroyed.

For the modern NZE data, all subjects were recorded in a quiet room. Each word was read three times. The speech was sampled at 22.05 kHz and quantized to a 16-bit number (see Sinclair & Watson, 1995, for further details). The material for each speaker was recorded over several days due to its length. To avoid list intonation, the speaker was presented with each word one at a time. The order of the words was random.

Labeling

The first three formants and their bandwidths for both NZE data sets were automatically tracked using ESPS/Waves (the speech was downsampled to a 10 kHz sampling rate: the formants were then calculated using 12th order LPC analysis, cosine window, 49 ms frame size, and 5 ms frame shift). All automatically tracked formants were checked, and hand corrections were made when these were considered necessary. Formant tracking errors were common because of the noise in the data, and they were far more common in the female data than the male data. All the labeling was done in EMU, a hierarchical speech data management system (Harrington & Cassidy, 1999).

The acoustic onset of the vowel was marked at the onset of periodicity, as shown in the spectrogram by vertical striations. Additional cues were the onset of periodicity in the waveform or, if the preceding consonant was a stop, the cessation of the aperiodicity caused by the aspiration.

The acoustic vowel target was marked as a single time point between the onset and offset. We marked a single target in monophthongs and the first target of the diphthongs. The targets were marked at a point where there was the least movement in the formant tracks. For the high vowels this point tended to occur where F2 reached a peak; for open vowels the target was marked where F1 was a maximum; for back vowels the target was marked where F2 reached a trough. If none of the above criteria was satisfied for a given vowel, the target was marked at the point of maximum amplitude in the waveform.

Speaker normalization

It is well known that in vowel identification listeners can somehow disregard acoustic variability among individual speakers. Speaker normalization techniques aim to transform the acoustic vowel space such that phonetically similar vowels are grouped together and phonetically different vowels are separated. Speaker normalizing is a common technique used in acoustic phonetics to factor out speaker differences. To compare the vowel spaces of the old and modern NZE data, the vowel tokens from each of the two corpora were speaker normalized separately. We used Lobanov normalization, a speaker-dependent strategy based on standardizing the formants means and standard deviations for each of the speaker's vowels (Lobanov, 1971).

In this study we present the averaged formant plots for each of the four speakers of old NZE as well as the average speaker normalized plot. We present the normalized vowel space for the modern NZE corpus. For the non-normalized vowel spaces, readers are referred to Watson et al. (1998).

RESULTS

In this section we present acoustic evidence for vowel change in NZE. In the first part, we focus on differences in the vowel space of the front vowels in old and modern NZE, and in the second part, we analyze the differences between old and modern NZE for the HEAR and HAIR diphthongs. The data from the old NZE corpus are presented two ways: as averaged vowel data for each speaker and as a speaker normalized average. The data for the modern NZE corpus are speaker normalized and presented on one plot.

The comparisons between the old and modern NZE vowel spaces are somewhat confounded since the old NZE data are from continuous speech, whereas the modern NZE data are from citation form speech. However, although the vowels from continuous speech are expected to be more centralized than those from citation form speech, the relative distances between the vowels in F1/F2 space are not expected to change. Thus, the overall shape of the vowel quadrilateral should be the same. Van Bergem (1993) found this result in his study of acoustic vowel reduction when comparing the formant values of vowels extracted from citation form words to those of vowels extracted from continuous speech from syllables with sentence stress. A similar conclusion was also reached in Harrington and Cassidy (1999:70). Therefore any evidence we find that the vowel quadrilateral has changed in shape between the old and modern NZE data can be taken as direct evidence that a vowel shift has occurred. Specifically, since this study considers only the front vowels, we are looking for a change of shape in the front portion of the vowel quadrilateral.

Comparison of the front vowels of old and modern NZE

Figure 1 shows the centroids for the HEED, HID, HEAD, HAD, HUD, and HERD monophthong classes for old NZE for each of the four speakers studied. Despite the degree of speaker variation in the four plots, HID is raised relative to HEAD and HERD for three speakers. By contrast, EB's HID was both more retracted and lower than HEAD, suggesting that some of EB's vowels were already moving in the direction of the modern NZE vowel space. This was supported in casual listening by trained phoneticians at Macquarie University, who consistently identified EB as sounding more "New Zealand" than the other three speakers.

The front vowel spaces of old and modern NZE can be contrasted by comparing the speaker normalized plots of the two data sets in Figure 2. (Only the centroids are plotted in Figure 2; however, the F1 and F2 means and standard deviations for each monophthong class for the normalized data are given in Table A.1 in the Appendix.) The most striking differences between old and mod-

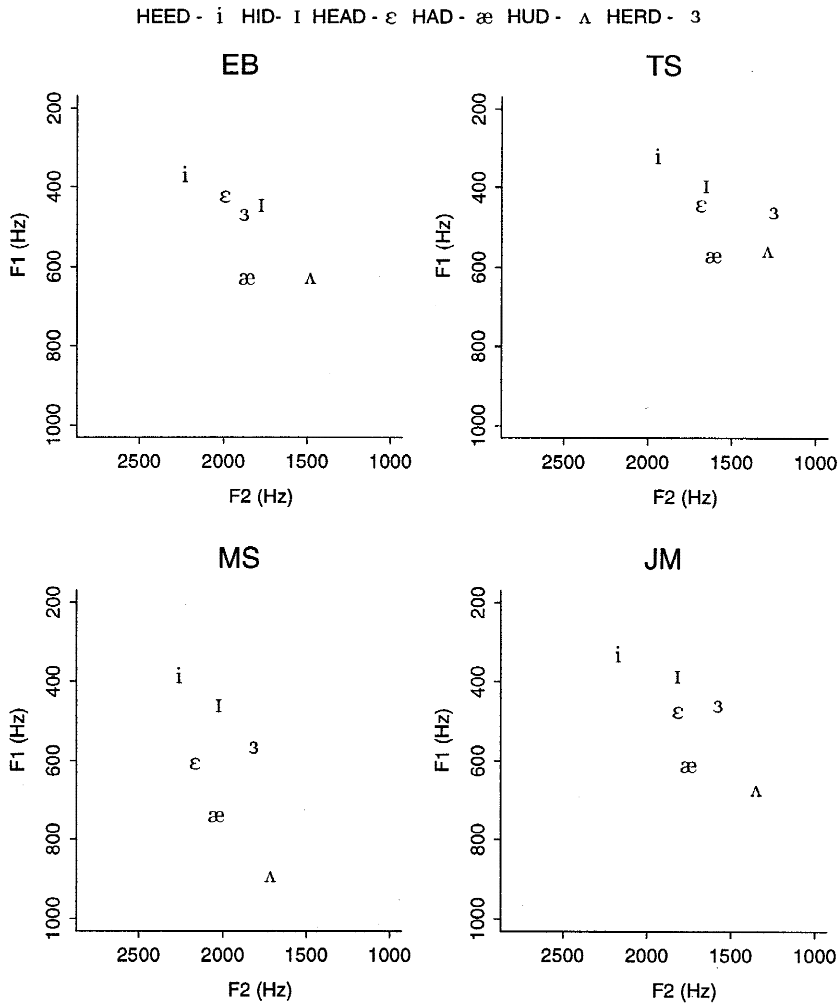

FIGURE 1. The centroids of the monophthong classes HEED, HID, HEAD, HAD, HUD, HERD for each of the four speakers of old NZE.

ern NZE are the relative positions of HEAD, HID, and HERD. In modern NZE, HID is always more lowered than both HEAD and HERD, as central as HERD, and more central than HAD. In old NZE, HID is more raised than HEAD and HERD and no more retracted than HAD.

Figure 2 also shows that HEAD is equidistant between HAD and HEED in old NZE, whereas in modern NZE it is a good deal closer to the HEED vowel. The same plot shows, using HUD as a starting point, that HAD is closer to HEED in modern NZE than in old NZE. Finally, it does not appear that HUD has fronted in

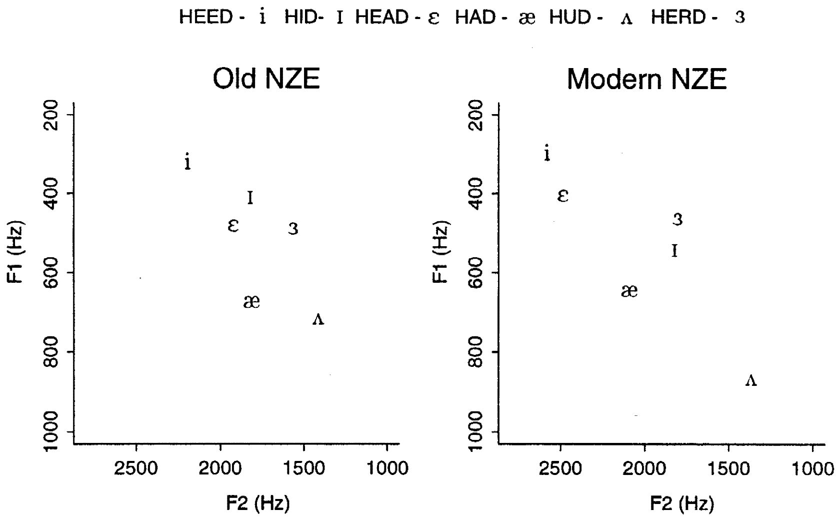

FIGURE 2. The centroids of the monophthong classes HEED, HID, HEAD, HAD, HUD, HERD from the speaker normalized data for old NZE (left) and modern NZE (right).

the last 50 years. The observable height differences in HUD between the old and modern data may be an artifact of continuous speech effects in the former. That is, continuous speech is likely to result in a certain amount of target undershoot and therefore in a less open vocal tract in HUD, resulting in an F1 raising relative to citation form data.

For each vowel class *t* tests were performed on the F1 and F2 values, contrasting the old and modern normalized data. All significance levels were Bonferroni corrected to account for multiple testing. Significant differences in F1 and F2 were found for all vowels except for F2 of HID and F1 of HAD ($\alpha \ll 0.05$). Unfortunately, the significant differences for many of the vowels may not be attributable to vowel shift alone but rather may be due to a conflation of vowel shift and context (continuous old vs. isolated modern data). However, one significant difference is noteworthy. In modern NZE the F1 of HID was greater than that in old NZE. This movement is opposite to the expected influence of the continuous context and is further evidence that the HID vowel has fallen in modern NZE.

Comparison of HEAR *and* HAIR *in old and modern NZE*

Figure 3 shows both the centroids of the first targets of HEAR and HAIR and the centroids from the monophthongs HEED, HID, HEAD, HAD, and HUD for each of the four speakers of old NZE. It can be seen that the HEAR and HAIR first targets are well separated for all four speakers. The HEAR and HAIR centroids are very near the HEED and HEAD centroids, respectively. These results suggest that HEAR and

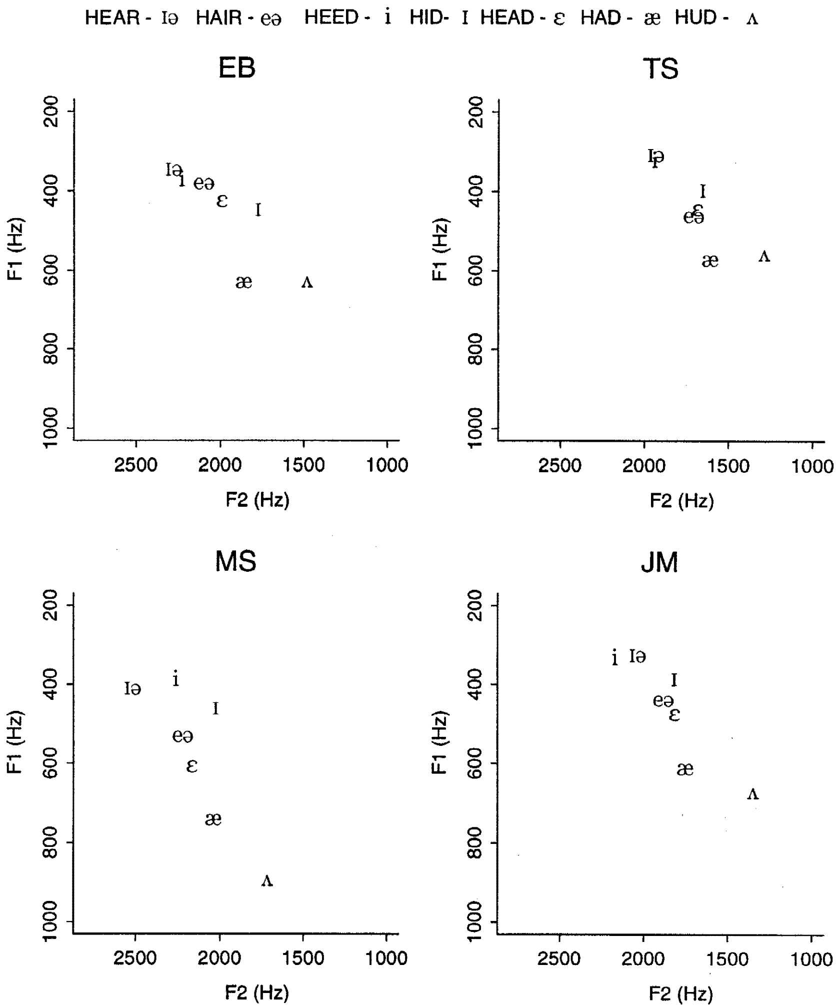

FIGURE 3. The centroids of the first targets of HEAR and HAIR and the centroids from the monophthongs HEED, HID, HEAD, HAD, and HUD for each of the four speakers of old NZE.

HAIR productions were distinct for the four speakers of old NZE, and that the merger of HEAR/HAIR currently in progress in modern NZE had not yet begun (see, e.g., Maclagan & Gordon, 1996; Watson et al., 1998).

Figure 4 shows both the centroids of the first targets of HEAR and HAIR and the centroids from the monophthongs HEED, HID, HEAD, HAD, and HUD for the speaker normalized data of old and modern NZE (see Table A.1 in the Appendix for F1 and F2 means and standard deviations for each vowel class of the normalized data). There are three outstanding differences between old and modern NZE.

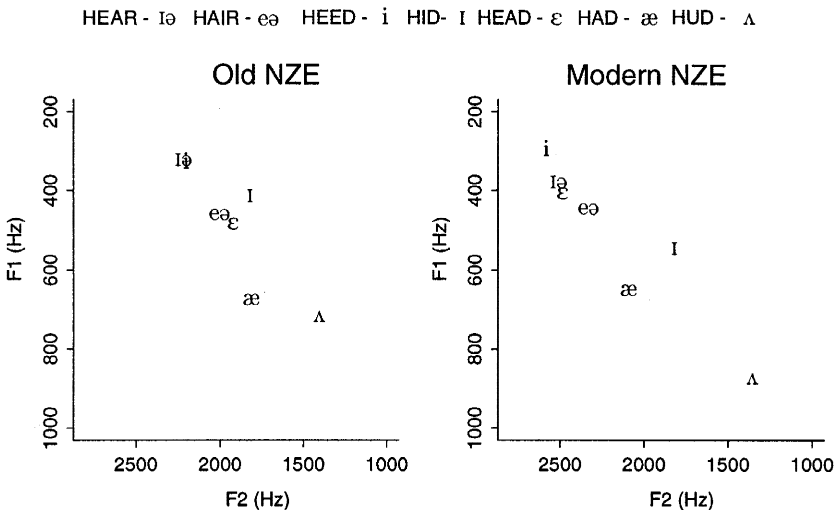

FIGURE 4. The centroids of the first targets of HEAR and HAIR and the centroids from the monophthongs HEED, HID, HEAD, HAD, and HUD for speaker normalized data for old NZE (left) and modern NZE (right).

First, whereas the first target of HEAR is very close to HEED in old NZE, it is close to HEAD in modern NZE. Second, in old NZE the first target of HAIR is very close to HEAD; in modern NZE HAIR is still close to HEAD, but both are raised and more front. Finally, in old NZE the first targets of HEAR and HAIR are well separated, whereas in modern NZE they are not. These effects are all the more striking since the old NZE data is from continuous speech: that is, if HEAR and HAIR in old and modern NZE were phonetically equivalent, our plots should show the first targets of HEAR and HAIR of old NZE as being closer than they are in the modern data. However, since it is in contemporary NZE that the first targets of HAIR and HEAR are closer, in spite of the centralizing effects of continuous speech on the old NZE vowel space, these data are very clear support for the idea that the HEAR and HAIR vowels have merged in the last 50 years.

While our data gives evidence that the first target of HAIR is raising with HEAD, there is also evidence that the first target of HEAR has fallen. The data showed that the first target of HEAR is near the centroid of HEED in old NZE, but is lower in modern NZE. While this may suggest that HEED has risen slightly in the last 50 years, there are limits on the extent to which it can be further raised or fronted because HEED is a phonetically peripheral vowel. Alternatively, it may suggest that the first target of HEAR has fallen slightly. This finding is supported by a comparison of the F1 values. The differences in F1 between the old and modern NZE data were tested for significance in the same manner as for the monophthongs. Significant differences were found, but, as for the monophthongs, the degree of significance was compromised by the possibility of conflation between the direction of vowel shift and by the fact that the old data is from continuous speech, whereas the modern data is from isolated speech. However, the fact that F1 for the first target of HEAR was higher in the modern data than in the old data suggests that it has fallen in the last 50 years.

DISCUSSION AND CONCLUSION

This study has provided evidence that NZE has undergone a vowel shift in the last 50 years, with a rearrangement of the order and position of the front vowels. In agreement with earlier studies (Trudgill et al., 1998; Woods, 1997), this study suggests that HEAD has risen in the last 50 years to become higher than HID in the vowel space. Our study has also shown that HAD has risen to a lesser degree and that HID has centralized. The relatively greater fronting of HERD with respect to HUD in the modern NZE data suggests that HERD also may have moved in the last 50 years. In addition, results from this study suggest that the HAIR/HEAR merger in modern NZE is due to both the first target of HAIR raising and the first target of HEAR falling.

With regard to the merger of HEAR and HAIR, two possibilities have been suggested for the production of words that traditionally contained these diphthongs: either that the total vowel space originally available for these two diphthongs is now available for realizations of either diphthong (Holmes & Bell, 1992) or that the total available vowel space has shrunk to a space equivalent to one of the original diphthongs (Maclagan & Gordon, 1996). The results presented here indicate that the space available for realizations of HEAR and HAIR has shrunk so that it is equivalent to the space originally available for one diphthong, but situated between the original two diphthong areas. The HEAR/HAIR vowel space now appears to be centered on HEAD, which could indicate that the two diphthongs have merged on the more open member of the pair. However, because NZE HEAD has raised so much, the starting point for the merged centering diphthongs is relatively close.

As mentioned earlier, the comparisons between the old and modern NZE vowel spaces are somewhat confounded since the old NZE data are from continuous speech, whereas the contemporary data are from citation form speech. However, because the old NZE data are from continuous speech, they are likely to be more central than comparable citation form data. Therefore, the dramatic nature of the vowel shift in NZE is indicated by the fact that for the continuous speech data HID and the first target of HEAR are less central and the separations between HEAD and HEED, HEED, and HAD and between HAIR and HEAR are greater.

According to Bauer (1979) and Woods (1997, 2000), the HAD, HEAD, and HID movements are all related in an extended push chain. From the results of this study it is difficult to comment on the movement of HAD, HEAD, and HID. Our evidence of vowel movement in NZE comes from only two periods in time (1948 and 1995), and although we can establish that there has been movement, we cannot establish the direction (push or drag chain) of the movement. There is clear evidence in this study that HEAD has risen; however, it is still not clear whether it is due to a push below from HAD or to HID retracting, leaving a space for HEAD to rise into. There is so far no experimental data of NZE speech recorded prior to the 1950s; therefore, the positions in vowel space of NZE HAD, HEAD, and HID before this time can only be conjecture.

We offer finally some comments on the reason why some of the NZE and Australian English front vowels could have rearranged themselves in the way that they have. We begin with the suggestion that the position of Received Pronunciation and Southern British English HID may be inherently unstable, both because it is in a crowded part of the vowel space—being close to HEED, HEAD, HAY, HEAR, and HAIR—and because HID is similar in many respects to an unstressed, weak vowel. Like schwa, /ɪ/ is low intensity (being a mid-high vowel) and short in duration (at least in Received Pronunciation), and its formant structure bears a greater resemblance to an unstressed vowel than do most of the other phonologically lax vowels: for example, HAD, HUD, HOD, and to a certain extent HEAD have a characteristically raised F1, which makes them much more sonorous and therefore less like an unstressed vowel than /ɪ/. The fact that /ɪ/ so often occurs in metrically weak positions with a variable realization as schwa in Southern British English (e.g., the second vowel in *animal* and the first vowel in *depend*) suggests that /ɪ/ has a closer association with unstressed vowels than most other vowels. We would suggest that, because HID in Received Pronunciation is quite similar in some respects to an unstressed vowel, the strategies for signaling HID as accented when it is marked for sentence stress are somewhat compromised. As various studies have shown (Fletcher & Harrington, 1999; Harrington et al., 2000; de Jong, 1995), accented vowels tend to be hyperarticulated (e.g., the tongue dorsum in HEED is fronted or raised to a greater extent; the tongue dorsum is backed to a greater extent in back vowels; the jaw is lowered to a greater extent in open vowels), but the possibility of hyperarticulating a Received Pronunciation of HID is diminished precisely because it is inherently a mid-high lax vowel not too distant in quality from an unstressed vowel. For example, the tongue could not be fronted or raised too much without changing the quality of HID beyond what is an "acceptable" production of HID in Received Pronunciation and/or without it straying too far into the HEAD or HEED space. Therefore, we would suggest that there is pressure on Received Pronunciation of HID to change in quality not only because of the crowding in the front vowel space, but also for prosodic reasons. HID is more likely to be misperceived as unaccented when a speaker intended it to be marked for sentence stress precisely because the possibility of hyperarticulating a Received Pronunciation of HID is so restricted (so that listeners are forced to rely on tonal cues of pitch accent placement to identify HID as accented).

We can therefore ask in what ways might HID change either to increase its perceptual contrast with the other front vowels or to enhance the supralaryngeal cues for signaling sentence stress? One possibility that addresses the latter, but perhaps not so much the former, is for HID to become a tenser vowel—that is, more like HEED. We would suggest that this is precisely the change that has taken place in Australian English, which is known to have a more front and raised HID vowel than in Southern British English (Watson et al., 1998). As a result of its greater peripherality, Australian English HID is perceptually much more distant from an unstressed vowel than in Received Pronunciation and Southern British English. Moreover, when Australian English HID is accented, its distance from the center of the vowel space is further exaggerated (Harrington et al., 2000; Palethorpe et al., 1999), and so the perception of HID as accented can more easily be enhanced by hyperarticulation than in Received Pronunciation and Southern British English.

We would suggest that this increased peripherality of Australian English HID compared with its Received Pronunciation counterpart may be related to two other vowel changes in Australian English: the diphthongization of the Australian English HEED vowel (Harrington, Cox, & Evans, 1997) and the loss of the /ɪ/–/ə/ contrast in unstressed position. The first of these would dramatically reduce the confusability between HEED and a HID vowel that was becoming progressively more tense in Australian English, especially since a diphthongal production is perceptually quite salient (Bladon, 1985). We would propose that the loss of the /ɪ/–/ə/ contrast (making *boxes*/*boxers* and *dances*/*dancers* homophonous in Australian English but not in Southern British English) came about because tense vowels tend not to occur in nonfinal metrically weak positions. Consequently, since the second syllable of *animal* could not be produced with a tense front vowel (given the risk that it would then be incorrectly perceived as strong), the metrically weak vowels in words like *animal*, *boxes*, and *dances* are restricted to a schwa in Australian English.

The other way for HID to increase its perceptual salience with respect to the crowded part of the mid-high front vowel space is to become more retracted rather than fronted. We propose that this is the vowel change that characterizes NZE and some accents of South African English and that makes these accents so different from Australian English in this respect. It should be noted that some studies of NZE have suggested that the NZE HID has the same quality as a schwa (e.g., Bauer, 1979; Hawkins, 1976; Wells, 1982). Although Watson et al. (1998) proposed a transcription of NZE HID with /ə/, a more recent kinematic and acoustic analysis of NZE HID (Watson, Harrington, & Palethorpe, 1998) showed that, although retracted, the vowel is still a high vowel. In the F1/F2 plots the HID vowel was retracted and lower than HEAD and HERD, as expected. However, the physiological data showed that HID was produced with approximated lips, a feature of high vowels, and the tongue tip was higher than in productions of both HEAD and HERD. Re-examination of the F1/F2 plots revealed that, although HID appeared to be centralized, its F1 frequency value was still very low, and that it was lower than the F1 frequency value of the Australian English HEAD, which is considered to be a high vowel. Considering the high functional load in English of HID and schwa, and keeping in mind the principle of maximal perceptual contrast

(Liljencrants & Lindblom, 1972), it does not make sense for HID to take on the quality of schwa, since this would decrease perceptual contrast and lead to a considerable loss in comprehension.

We would propose that the HID movement was innovative in NZE in the same way that it was innovative in Australian English, motivated in both cases by a lack of perceptual contrast, either paradigmatically to the other front vowels or because of the weakened cues for signaling the prosodic accent contrast. Consequently, we would favor a drag chain account of front vowel raising of HEAD and HAD in Australian and NZE rather than, or in addition to, a rearrangement of the lax vowel space as a result of a push chain initiated by HUD (Bauer, 1979; Woods, 1997, 2000). This suggestion is supported by Trudgill et al. (1998), who demonstrated that HEAD and HAD were already relatively high in NZE during the nineteenth century, and by Maclagan (in press), who found that, although HID was higher than HEAD for the ten speakers studied (all born between 1865 and 1886), some speakers were already starting to centralize HID. HID was therefore starting to centralize while it was still considerably higher than HEAD.

The raising of NZE HEAD and HAD can also be accounted for by the theory that vowel shift is motivated, to a degree, by the principles of maintaining perceptual contrast. This can be seen by looking at the durations of HEAD and HAD compared to the neighboring short vowels. Since HEAD is phonologically a lax vowel in English, it would be expected to have a shorter duration than HID, because vowel duration is inversely correlated to tongue height. However, even though in NZE HEAD is higher than HID and HAD is higher than HUD, there is strong evidence to suggest that NZE HEAD is longer than NZE HID, and that NZE HAD is longer than NZE HUD (see Table A.2 in the Appendix; see also Watson, Harrington, & Palethorpe, 1998). Since HEAD is both more raised and longer than HID, this suggests that it would be perceived as a tense vowel. A similar conclusion can be drawn for HAD when contrasting it to HUD. Interestingly, in both the old and the modern NZE corpora, HID had the shortest vowel duration, followed by HUD, HEAD, and then HAD. This might suggest that as HAD, HEAD, and HID moved they retained the durations associated with their earlier positions in the vowel space. In doing so, HEAD and HAD would then maintain their contrast with the unstressed vowel.

Two additional predictions follow from the idea that the rearrangement of the HID–HEAD–HAD spaces in Australian English and in NZE can be attributed to an innovative movement of HID (in different directions for the two accents). First, we would expect NZE HEED to be less diphthongal than in Australian English if HEED diphthongization were related to the increased tensity and fronting of HID. Second, we would predict two ways that the retraction of NZE HID would favor the conditions for its supralaryngeal enhancement when it is accented: either the tongue dorsum could be further retracted (there is no other unrounded back vowel with which it could be confused), which would make it a more peripheral vowel, and/or the tongue dorsum and jaw might be lowered, which would make it a more open, sonorous vowel resulting in F1 raising. We are currently investigating both of these possibilities using kinematic techniques to track the movement of the tongue, jaw, and lips. We would therefore suggest that HID fronting (Australian English) or retraction (NZE) renders this vowel perceptually more salient, both by increasing its contrast to the other front vowels and by providing greater opportunities to hyperarticulate it when it occurs in accented position.

APPENDIX

TABLE A.1. *Mean and standard deviations of the first and second formants for each vowel studied for both the old NZE and modern NZE normalized data*

		Old NZE		Modern NZE	
Vowel	Formant	Mean (Hz)	*SD*	Mean (Hz)	*SD*
HEED	F1	327	51	296	54
	F2	2,199	170	2,581	102
HID	F1	413	71	543	94
	F2	1,837	204	1,835	99
HEAD	F1	480	78	403	53
	F2	1,921	158	2,485	107
HAD	F1	671	90	648	79
	F2	1,829	135	2,104	125
HUD	F1	714	107	871	70
	F2	1,408	133	1,356	100
HERD	F1	489	51	462	68
	F2	1,532	144	1,770	155
HEAR	F1	322	61	378	62
	F2	2,207	173	2,499	112
HAIR	F1	456	82	442	69
	F2	1,978	112	2,308	154

TABLE A.2. *Mean durations of the vowels HID, HEAD, HAD, and HUD and their standard deviations*

		Duration (msec)			Duration (msec)	
Vowel	Word	Mean	*SD*	Word	Mean	*SD*
HID	pit	103.6	31.9	did	147.3	33.6
HEAD	pet	138.5	36.6	dead	191.7	45.3
HAD	pat	143.5	32.3	dad	289.4	52.2
HUD	but	133.7	35.0	pub	134.5	36.5

Note: The vowels were extracted from the words shown in the table from the modern NZE corpus. The data was taken from both male and female speakers. The duration data could be combined because there was no significant differences between the two data sets.

NOTE

1. In this study the vowels will be indicated not by phonetic symbols but by an HVD/HV lexical set, as used by Watson, Harrington, and Evans (1998).

REFERENCES

Batterham, M. (2000). The apparent merger of the front centering diphthongs—EAR and AIR. In A. Bell & K. Kuiper (eds.), *New Zealand English*. Wellington: Victoria University Press. 111–145.

Bauer, L. (1979). The second great vowel shift. *Journal of the International Phonetic Association* 9:57–66.

______ (1986). Notes of New Zealand phonetics and phonology. *English World-Wide* 7:225–258.

______ (1994). English in New Zealand. In R. Burchfield (ed.), *Cambridge history of the English language* (Vol. 5). Cambridge: Cambridge University Press. 382–429.

Bennett, J. A. W. (1943). English as it is spoken in New Zealand. *American Speech* 18:69–83.

Bladon, R. A. W. (1985). Diphthongs: A case study of dynamic auditory processing. *Speech Communication* 4:145–154.

de Jong, K. (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *Journal of the Acoustical Society of America* 97:491–504.

Fletcher, J., & Harrington, J. (1999). Lip and jaw coarticulation. In W. J. Hardcastle & N. Hewlett (eds.), *Coarticulation*. Cambridge: Cambridge University Press. 164–175.

Gordon, E. (1991). The development of spoken English in New Zealand. In G. MacGregor (ed.), *Dirty silence: Aspects of language and literature in New Zealand*. Auckland: Oxford University Press. 19–29.

______ (1998). The origins of New Zealand English: The limits of recovering historical information from written records. *English World-Wide* 19:61–85.

Gordon, E., & Abell, M. (1990). 'This objectionable colonial dialect': Historical and contemporary attitudes to New Zealand speech. In A. Bell & J. Holmes (eds.), *New Zealand ways of speaking English*. Clevedon: Multilingual Matters. 21–48.

Gordon, E., & Maclagan, M. (1985). A study of the /Iə/~/eə/ contrast in New Zealand English. *New Zealand Speech-Language Therapists' Journal* 40:16–26.

______ (1989). Beer and bear, cheer and chair. *Australian Journal of Linguistics* 9:203–220.

Harrington, J., Beckman, M. E., & Fletcher, J. (2000). Manner and place conflicts in the articulation of accent in Australian English. In M. Broe (ed.), *Papers in laboratory phonology 5*. Cambridge: Cambridge University Press. 40–51.

Harrington, J., & Cassidy, S. (1999). *Techniques in speech acoustics*. Dordrecht: Kluwer.

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

Hawkins, P. (1976). The role of New Zealand English in a binary-feature analysis of English short vowels. *Journal of the International Phonetic Association* 6:50–66.

Holmes, J., & Bell, A. (1992). On shear markets and sharing sheep: The merger of EAR and AIR diphthongs in New Zealand English. *Language Variation and Change* 4:251–273.

Jaberi, M., & Stone, M. (1997). The effects of aging and rate on tongue shape variability and on preservation of position of glides. *Journal of the Acoustical Society of America* 102:3165.

Lehiste, I. (1964). *Acoustical characteristics of selected English consonants*. The Hague: Mouton.

Lewis, G. (1996). The origins of New Zealand English: A report on work in progress. *New Zealand English Journal* 10:25–30.

Liljencrants, J., & Lindblom, B. (1972). Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language* 48:839–862.

Lobanov, B. M. (1971). Classification of Russian vowels spoken by different speakers. *Journal of the Acoustical Society of America* 49:606–608.

Maclagan, M. A. (1982). An acoustic study of New Zealand vowels. *New Zealand Speech Therapist's Journal* 37:20–26.

______ (1998). Diphthongisation of /e/ in NZE: A change that went nowhere? *New Zealand English Journal* 12:43–54.

______ (in press). How long have women been leading language change. In J. Holmes (ed.), *Gendered speech in social context: Perspectives from gown and town*. Wellington: Victoria University Press.

Maclagan, M. A., & Gordon, E. (1996). Out of the AIR and into the EAR: Another view of the New Zealand diphthong merger. *Language Variation and Change* 8:125–147.
McBurney, S. (1887). *Colonial pronunciation*. Christchurch Press. Reproduced in Turner (1967).
McKinnon, M. (ed.). (1997). *New Zealand historical atlas*. Auckland: David Bateman.
Moon, S.-J., & Lindblom, B. (1994). Interaction between duration, context, and speaking style in English stressed words. *Journal of the Acoustical Society of America* 96:40–55.
Palethorpe, S., Beckman, M. E., Fletcher, J., & Harrington, J. (1999). The contribution of schwa vowels to the prosodic accent contrast in Australian English. *Proceedings of the International Congress of Phonetic Sciences, San Francisco*.
Sinclair, S., & Watson, C. (1995). The development of the Otago speech database. In N. Kasabov & G. Cogill (eds.), *Proceedings of the 2nd New Zealand International Conference on Artificial Neural Networks and Expert Systems*. Los Alamitos, CA: IEEE Computer Society Press. 298–301.
Summers, W. V. (1987). Effects of stress and final-consonant voicing on vowel production: Articulatory and acoustic analyses. *Journal of the Acoustical Society of America* 82:847–863.
Trudgill, P., Gordon, E., & Lewis, G. (1998). New-dialect formation and Southern Hemisphere English: The New Zealand short front vowels. *Journal of Sociolinguistics* 2:35–51.
Turner, G. W. (1967). Samuel McBurney's newspaper article on colonial pronunciation. *Australian Universities Modern Language Association* 27:81–85.
_______ (1970). New Zealand English today. In W. Ramson (ed.), *English transported*. Canberra: Australian National University Press. 84–101.
van Bergem, D. R. (1993). Acoustic vowel reduction as a function of sentence accent, word stress and word class. *Speech Communication* 12:1–23.
Wall, A. (1938). *New Zealand English: How it should be spoken*. Auckland: Whitcombe and Tombs.
_______ (1964). *The jeweller's window*. Christchurch: Whitcombe and Tombs.
Watson, C., Harrington, J., & Evans, Z. (1998). An acoustic comparison between New Zealand and Australian English vowels. *Australian Journal of Linguistics* 18:185–207.
Watson, C., Harrington, J., & Palethorpe, S. (1998). A kinematic analysis of New Zealand and Australian English vowel spaces. *Fifth International Conference on Spoken Language Processing*. 2363–2366.
Wells, J. C. (1982). *Accents of English: Beyond the British Isles*. Cambridge: Cambridge University Press.
Woods, N. J. (1997). The formation and development of New Zealand English: Interaction of gender-related variation and linguistic change. *Journal of Sociolinguistics* 1:95–125.
_______ (2000). New Zealand English across the generations: An analysis of selected vowel and consonant variables. In A. Bell & K. Kuiper (eds.), *New Zealand English*. Wellington: Victoria University Press. 84–110.