Word recognition and sound merger: the case of the front-centering diphthongs in NZ English

Paul Warren†, Megan Rae†, & Jen Hay‡

† School of Linguistics and Applied Language Studies, Victoria University of Wellington

‡ Linguistics Department, University of Canterbury

ABSTRACT

Sociolinguistic and phonetic studies of New Zealand English have indicated an ongoing merger of NEAR and SQUARE vowels on the closer form. This paper explores the consequences of this change for spoken word recognition, using an auditory lexical decision experiment with semantic priming. Subjects were young NZE speakers, a group who predominantly merge the vowels in their own speech. These subjects responded to NEAR forms as though they were homophones giving access to the meanings of both NEAR and SQUARE words. Their responses to SQUARE forms, still encountered amongst more conservative speakers, appear to reflect access only to SQUARE meanings. This asymmetry in the pattern of lexical decision responses is compatible with the change-in-progress towards the NEAR vowel, and provides psycholinguistic support for earlier studies of the merger.

1. INTRODUCTION

The research project[1] that provides the data reported in this paper considers how the speech recognition and comprehension system copes with ongoing sound change. In addition, it provides some psycholinguistic evidence that relates to the progress of one particular change that has been studied extensively in recent sociolinguistic and phonetic research, namely the merger of the NEAR and SQUARE diphthongs in New Zealand English (NZE).[2] Popular opinion is that standards are slipping because words like *beer* and *bear* can no longer be distinguished. In reality we suspect that there will be few situations which result in miscommunication, because contextual information will make the intended meaning clear. However, the fact that the NEAR/SQUARE merger is a change in progress, meaning that both NEAR and SQUARE forms are likely to be encountered, suggests that we should perhaps not assume that NEAR/SQUARE words are treated just the same as more established homophones such as *bank* or *write/right*.

A series of production studies conducted with 14-15 year olds in four Christchurch schools in 1983, 1988, 1993 and 1998 ([5, 6]) initially showed a high degree of variability, which by 1993 had settled on a clear merger to NEAR. Other studies confirm that the NEAR and SQUARE vowels are undergoing merger elsewhere in New Zealand [1, 8, 21], though they do not agree on the form that the merger is settling on.

If words with NEAR and SQUARE vowels (e.g. *cheer*, *chair*) have become indistinct for some speaker groups then we are likely to be dealing, for those speakers, with homophony. Studies of the processing of homophones have shown that compared with unambiguous words, their recognition times are longer, while those for polysemous words are shorter. This reflects competition between separate lexical items in the case of homophones, but accumulation of activation from a range of word senses in the case of polysemes [16]. Despite such initial delay, recognition of homophones is nevertheless rapid. Additionally, in sentence contexts, the multiple meanings of homophones are available early on, even when one meaning may be infelicitous to the context [17]. Contextual information is then quickly used to select the appropriate meaning [22], though the relative frequencies of the meanings may influence this process [20].

In the study reported in this paper, we wished to determine whether words with NEAR or SQUARE vowels are treated in NZE as homophones, giving access to the meanings of both words in a NEAR/SQUARE pair. If there is evidence of homophony, then is this true of both pronunciations of the pair, and if so, is it *equally* true of both pronunciations? For instance, when listeners hear [tʃiə], do they access meanings associated both with *cheer* and with *chair*? Is the same true when they hear [tʃeə]? Access to these meanings may be reflected in the activation of words associated with the meanings, such as *shout* for *cheer* and *sit* for *chair*. If the two pronunciations map on to distinct lexical forms, then we would expect *shout* to be active on hearing [tʃiə] and *sit* on hearing [tʃeə], as in (a) in Figure 1. If merger is complete on the [tʃiə] form (reflecting the dominant trend in the auditory and acoustic research), then this form should result in access of both *shout* and *sit*, as in (b). Though not shown in the figure, we imagine in the (b)-scenario that [tʃeə], as a phonetic variant of /tʃiə/, might also lead to access of both *shout* and *sit* meanings, but less strongly than [tʃiə]. A third possibility is given in (c) in Figure 1. This is the situation which we believe is compatible with the fact that the merger is currently still in progress. It presents an asymmetry in the availability of the different meanings, so that [tʃiə] gives access to both meanings (as a

1 "Sound change and spoken word recognition in New Zealand English", funded in part by Victoria University Research Fund (URF) grant 1/11.

2 We use NEAR and SQUARE [24] to represent the lexical sets involved in this merger.

ISBN 1-876346-48-5

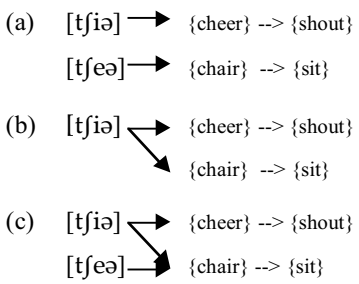

Figure 1: possible lexical mappings from phonetic forms undergoing merger

homophone), but [tʃeə], still heard in the environment as the form for *chair* amongst more conservative speakers, gives access to the *sit* meaning.

2. EXPERIMENT

To investigate these possibilities, we used a semantic priming paradigm that explores access to the meanings of ambiguous words. In this paradigm subjects respond to sequences of items which bear some meaning relationship to one another, as well as to unrelated items. Semantic priming obtains when a word is responded to more rapidly or more accurately when it has been preceded by a related word [12]. Since research has shown that semantic priming is more robust for semantic associates than for words related simply through common semantic features [7, 14, 15], our experiment uses words related by association, as determined from published norms. Semantic priming tasks tend to use either paired presentation (items are presented in pairs, each pair containing either related or unrelated items) or list presentation (a long list of items containing pairs of (un)related items). Since detailed studies of priming paradigms have argued that list presentation is less likely than paired presentation to result in strategic approaches such as post-lexical relatedness checking [11, 18], we elected to use list presentation in our experiment. In consequence this entails an auditory-auditory rather than cross-modal presentation of prime and probe items.

2.1 Materials and design

Test materials comprised 12 pairs of NEAR/SQUARE items (e.g. [tʃiə] and [tʃeə]) serving as *test primes*, *probes* semantically associated to each of these 24 items, and 24 *control primes*, which were words semantically unrelated to the probes [9]. All items were monosyllabic and approximately matched for frequency. Some test pairs have NEAR or SQUARE forms which are themselves homophones (e.g., *deer* and *dear*; *bear* and *bare*; *sheer* and *shear*). As the stimuli are presented auditorily, choosing between the meanings of these homophonous pair members does not affect the form of the prime, but it does affect the selection of the associated probe word. Given that we needed, in the list presentation format, to allow sufficient time after each item for a lexical decision to be made, our preference was to use associates of the dominant meaning of the form (with dominance determined using frequency information from [23]), since the dominant meaning has been shown to persevere over time [7, 19]. However, in some cases the most frequent meaning has less clear associates than another meaning. For example, in the corpora *sheer* has a higher frequency than *shear*, but intuitively, *shear* followed by *sheep* would provide greater priming effects than *sheer* followed by *cliff* or *tights*. In such cases, a lower frequency meaning of the prime was chosen, to ensure that there was a clear associate. It is worth stressing however that the frequency difference between homophonous members such as *sheer* and *shear* was never great.

Probe words were selected on the basis of three sets of published association norms [3, 4, 13], using where possible concrete and natural category words [15]. In some instances the constraints on probe selection described above meant that we did not use the most frequent response from the association data. For prime words not found in the association norms, probe words were chosen based on intuition. Control and filler words were selected from a computerised list of monosyllables. Control words were unrelated to probes and were open syllables, phonologically matched with the NEAR/SQUARE prime words. Twenty real-word fillers helped disguise the test materials so as to reduce the risk of strategies such as relatedness checking. Finally, 44 non-words were included, which were constructed by altering one phoneme in existing words – e.g., *jatt* from *cat*.

All stimuli were recorded onto digital tape by a native NZ English speaker who was asked to be careful to distinguish the NEAR and SQUARE vowels. Her own spontaneous speech does distinguish the two, but not consistently for all words. Stimuli were then transferred to Macintosh computer for use in the experiment.

The overall design consisted of 4 test conditions (NEAR and SQUARE primes followed by probes related to each) and 2 control conditions (unrelated primes before each of these probes). The four test conditions were rotated across the 24 items to produce four test blocks, each containing six items in each condition, along with filler and nonword items. A fifth block included the control condition. The experiment was run in a repeated measures design as our research goals included a comparison of priming effects for each subject with their own realisation of the NEAR/SQUARE difference. Each subject completed two sessions (a week apart) with two test blocks in each session. Session one included blocks 1 and 2 and session two blocks 3 and 4, but the order of blocks within sessions varied across subjects. For any item the two blocks in a session did not include a repetition of either the prime or the probe word for that item. Filler tasks were completed between the test blocks, in order to reduce long-term priming of probes by related items occurring in the preceding block. However, because of fatigue effects that showed up in the data, we present results here only for responses from the first block in each session.

The 16 subjects were Linguistics undergraduates at Victoria University of Wellington. They were all fluent native NZE speakers, aged 18-25 with no known hearing

impairment, and were naïve as to the purpose of the experiment. They were tested individually in a quiet room. They heard the sequence of real word and non-word stimuli over closed-ear headphones and pressed one of two response buttons on a millisecond timer to indicate a real word or non-word decision. The experiment was controlled by the PsyScope programme [2]. The inter-stimulus interval was set at 1400 milliseconds - long enough to allow subjects to respond, but short enough to maintain a fast response rate.

2.2 Results

Response times (RTs) for each item were recorded in milliseconds. One item set was excluded because of experimenter error. Missing and incorrect responses were also excluded, affecting 3.69% of test and control data. Facilitation times (FTs) were calculated by subtracting a subject's RT for a probe in a test condition from that subject's RT to the same probe in the control condition. Average FTs in each of the four test conditions were calculated for each subject, and the resulting averages were subjected to Analysis of Variance (ANOVA), with Facilitation Time as the dependent variable and Prime (NEAR word vs SQUARE word) and Probe (*appropriate* [e.g. word related to the NEAR word occurring after the NEAR word] or *inappropriate* [e.g. word related to the SQUARE word occurring after the NEAR word]) as independent variables.[3] In addition, planned comparisons of probe types following each kind of prime were used to test the predictions entailed in Figure 1. Overall averages for the four conditions are shown in Figure 2.

The statistical analysis confirmed that *appropriate* probes received more priming than *inappropriate* probes ($F(1,15)=5.85$, $p<0.03$). In addition, there was a significant interaction of prime and probe type ($F(1,15)=4.90$, $p<0.05$). In the planned comparisons, SQUARE primes facilitated *appropriate* probes more than *inappropriate* probes ([tʃeə]-*sit* > [tʃeə]-*shout*; $F(1,30)=6.10$, $p<0.02$), but there was no difference in priming of the two probe types by NEAR primes ([tʃiə]-*shout* ≈ [tʃiə]-*sit*) ($F(1,30)=0.23$, $p>0.60$).

3. DISCUSSION

The result for NEAR forms supports our prediction that they are ambiguous between NEAR and SQUARE words, following the trend in NZE towards merger on NEAR. Thus, forms containing the [iə] vowel are treated as though they are homophones, resulting in semantic priming of both NEAR-related and SQUARE-related associates. The fact that there is facilitation of SQUARE-related probes after SQUARE forms shows that these forms are still recognised. However, the lack of facilitation of NEAR-related words after the [eə] vowel suggests that this form is not ambiguous between SQUARE and NEAR - the SQUARE form is not a phonetic variant of the NEAR form, but continues to be recognised as a separate form.

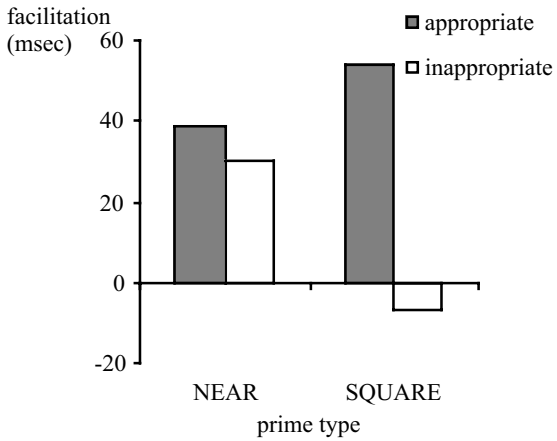

Figure 2. Average facilitation times for appropriate and inappropriate probes after NEAR and SQUARE primes (eg., from left to right, *cheer-shout; cheer-sit; chair-sit; chair-shout*).

A further observation from our results provides some support for the contention that NEAR forms are homophonous but SQUARE forms are not. Recall that recent work has shown that homophones are responded to more slowly than unambiguous words [16], reflecting competition between two different words with the same form. Analysis of lexical decision responses to the NEAR and SQUARE words in our experiment confirms that the former are responded to more slowly (894 vs 868 msec, $t=2.94$, df:15, $p<0.02$). In addition, the data in Figure 2 appear to show a knock-on effect for the associates presented subsequently, with lower levels of facilitation for either probe word following NEAR primes than for the SQUARE probe word following the SQUARE prime.

Overall the results of this experiment provide support from a spoken word recognition study for the view that the merger of NEAR and SQUARE vowels is towards the closer NEAR form. The asymmetry in the data would be difficult to explain if the merger was in the opposite direction, towards SQUARE, since this would predict that the SQUARE form would prime associates of the NEAR word, which is the condition in our experiment that shows no facilitation. It would seem that the presence of the SQUARE form (mainly in the speech of more conservative NZE speakers) is sufficiently strong to result in recognition of this form and significant facilitation of associates of the SQUARE words.

Future research will examine effects of sentence context on the recognition of our NEAR/SQUARE tokens. For instance, will there be immediate access to multiple meanings and rapid integration of contextual information in the process of selecting from those meanings for NEAR words [22]? What will we find for SQUARE words – will there be early access to multiple meanings that is not evidenced in the auditory-auditory priming task with its longer inter-stimulus interval?

[3] An ANOVA was not performed on item means, since the item selection almost exhausted the set of relevant words in the language, so that the items were not selected randomly [10].

ISBN 1-876346-48-5

Finally, we wish in our future research to examine the effects on perception and recognition of demographic information concerning the speaker. If the distinction between NEAR and SQUARE forms is indicative of a more conservative speaker, will participants who believe the speech to have come from an older speaker respond differently to those who believe the same tokens to have come from a young speaker?

REFERENCES

[1] M. Batterham, "The apparent merger of the front centring diphthongs - EAR and AIR - in New Zealand English," in *New Zealand English,* A. Bell, K. Kuiper, Eds., pp. 111-145. Amsterdam: John Benjamins, 2000.

[2] J.D. Cohen, B. MacWhinney, M. Flatt, and J. Provost, "Psyscope: A new graphic interactive environment for designing psychology experiments," *Behavioural Research Methods, Instruments & Computers*, vol. 25(2), pp. 257-271, 1993.

[3] Edinburgh Associative Thesaurus. Accessed: October 2002. http://monkey.cis.rl.ac.uk/Eat/htdocs/eat_old.html.

[4] Florida Free Association Norms. Accessed: October 2002. http://w3.usf.edu/freeassociation.

[5] E. Gordon and M. Maclagan, "'Are we NEAR the SQUARE?' A real time study of the merger of two centering diphthongs in New Zealand English," *Linguistic Society of New Zealand*, Palmerston North, 1999.

[6] E. Gordon and M. Maclagan, "Capturing a sound change: A real time study over 15 years of the NEAR/SQUARE diphthong merger in New Zealand English," *Australian Journal of Linguistics*, vol. 21(2), pp. 215-238, 2001.

[7] Y. Hino, S.J. Lupker, and C.R. Sears, "The effects of word association and meaning frequency in a cross-modal lexical decision task: Is the priming due to "semantic" activation?," *Canadian Journal of Experimental Psychology*, vol. 51(3), pp. 195-210, 1997.

[8] J. Holmes and A. Bell, "On shear markets and sharing sheep: The merger of EAR and AIr diphthongs in New Zealand English," *Language Variation and Change*, vol. 4, pp. 251-273, 1992.

[9] J. Jonides and R. Mack, "On the cost and benefit of cost and benefit," *Psychological Bulletin*, vol. 96(1), pp. 29-44, 1984.

[10] G. Keppel, "Words as random variables," *Journal of Verbal Learning and Verbal Behaviour*, vol. 15, pp. 263-265, 1976.

[11] T.P. McNamara and J. Altarriba, "Depth of spreading activation revisited: semantic mediated priming occurs in lexical decisions," *Journal of Memory and Language*, vol. 27, pp. 545-559, 1988.

[12] D.E. Meyer and R.W. Schvaneveldt, "Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations," *Journal of Experimental Psychology*, vol. 21, pp. 863-883, 1971.

[13] H.E. Moss and L. Older, *Birkbeck Word Association Norms*, Hove, UK: Lawrence Erlbaum, 1996.

[14] H.E. Moss, R.K. Ostrin, L.K. Tyler, and W.D. Marslen-Wilson, "Semantic and associative priming within and across modalities," *25th International Congress of Psychology*, Brussels, Belgium, 1992.

[15] H.E. Moss, R.K. Ostrin, L.K. Tyler, and W.D. Marslen-Wilson, "Accessing different types of lexical semantic information: evidence from priming," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 21(4), pp. 863-883, 1995.

[16] J. Rodd, M.G. Gaskell, and W.D. Marslen-Wilson, "Making sense of semantic ambiguity: semantic competition in lexical access," *Journal of Memory and Language*, vol. 46, pp. 245-266, 2002.

[17] M.S. Seidenberg, M.K. Tanenhaus, J.M. Leiman, and M. Bienkowski, "Automatic access of the meanings of ambiguous words in context: some limitations of knowledge-based processing," *Cognitive Psychology*, vol. 14, pp. 489-537, 1982.

[18] J.R. Shelton and R.C. Martin, "How semantic is automatic semantic priming?," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 18(6), pp. 1191-1210, 1992.

[19] G.B. Simpson, "Meaning dominance and semantic context in the processing of lexical ambiguity," *Journal of Verbal Learning and Verbal Behavior*, vol. 20, pp. 120-136, 1981.

[20] G.B. Simpson, "Context and the processing of ambiguous words," in *Handbook of psycholinguistic research,* M.A. Gernsbacher, Ed., pp. 359-374. San Diego: Academic Press, 1994.

[21] D. Starks, S. Allan, and C. Kitto, "Why vernacular speech? Speech samples from the taped Auckland rapid and anonymous survey," *New Zealand Language and Society Conference*, Victoria University of Wellington, 1998.

[22] D.A. Swinney, "Lexical access during sentence comprehension: (Re) consideration of context effects," *Journal of Verbal Learning and Verbal Behaviour*, vol. 18, pp. 645-659, 1979.

[23] Wellington Corpora of New Zealand English. Accessed: February 2003. http://www2.vuw.ac.nz/lals/corpora.htm.

[24] J.C. Wells, *Accents of English*, Cambridge: Cambridge University Press, 1982.