Featural representations of the New Zealand English short front vowels

Mathias Scharinger

Department of Linguistics
University of Konstanz, Germany
mathias.scharinger@uni-konstanz.de

Abstract

The realisation of the short front vowels in NZE differs from British or American English (RP/AE) and is well documented (cf. Bauer, 1986). However, two crucial questions arise: (1) Do the acoustically different vowel realisations also involve different underlying representations as compared to RP/AE? (2) What is the precise nature of these representations? Previous investigations either suggest that most NZE speakers have the same vowel phonemes as British English speakers (cf. Bauer, 1986) or that the NZE vowel change reflects a re-distribution of phonetically detailed exemplars in the mental lexicon (cf. Warren, Hay, & Thomas, to appear). This paper proposes that NZE has in fact different vowel phonemes than RP/AE. In particular, the original three-way height contrast of the short front vowels developed into a two-way height contrast in NZE (cf. Hawkins, 1976). The underlying properties of these vowels is described in the framework of the Featurally Underspecified Lexicon (Lahiri & Reetz, 2002), in which lexical entries consist of phonological feature representations. The differing underlying vowel representations are accessed via a particular feature matching algorithm which allows variation to be resolved in one, possibly underspecified, representation, and not by storing multiple exemplars that interact with pre-lexical biases. Two behavioural experiments provide evidence for the claims that NZE vowels are represented in terms of phonological features and that NZE uses different vowel phonemes than AE. Experiment 1 conducted with NZE subjects using the indirect semantic priming technique showed that both *bet* and *bit* primes could activate the semantic relative of *bat* (*club*). The fact that *bet* and *bit* similarly primed *club* suggests that *bat*, through which the meaning of *club* was retrieved, is not specified for height in NZE such that the vowel in *bet* (with a phonetically high vowel) and the vowel in *bit* (with a mid vowel) could activate *bat* (and thereby, *club*) to the same degree. The results represent a challenge for exemplar-models which would expect priming differences to be based on the acoustic distance of prime and target vowels or on the frequency distributions of the stimuli. However, no such correlations could be found. Experiment 2 used the same material with speakers of AE. The results differed crucially in that the primes with a high vowel did not activate *club* through *bat*. This was predicted on the basis of a mismatch between a high and a low vowel and showed that speakers of AE have a different representation of *bat* words (with a low vowel) as compared to NZE speakers (with a non-low vowel).

1. Introduction

The short front vowels of New Zealand English (NZE[1]) are known for their distinctive properties. Compared to American English (AE)/British English (Received Pronunciation, RP), the non-high short front vowels [æ][2] and [ɛ] raise, while the originally high vowel [ɪ] lowers and centralises. These shifts are well documented, both from an acoustic and an auditory perspective (Allan & Starks, 1999; Bauer, 1986; Easton & Bauer, 2000; Gordon, 1983; Gordon, Maclagan, Hay, Campbell, & Trudgill, 2004; Haggo, 1984; Hawkins, 1973; Maclagan & Hay, 2004; Watson, Maclagan, & Harrington, 2000). The changes of the short front vowels are considered to be the result of a chain shift, which started with the raising of [æ] (Maclagan & Hay, 2004), in turn pushing [ɛ] and then [ɪ] out of their original places. The changes in the NZE vowel system are illustrated in Figure 1.

In some studies on the NZE short front vowel shift, it is assumed that these vowels share the same underlying representations as in AE/RP (e.g. Bauer, 1986).

Proceedings of the 11th Australian International Conference on Speech Science & Technology, ed. Paul Warren & Catherine I. Watson. ISBN 0 9581946 2 9

Accepted after abstract only review

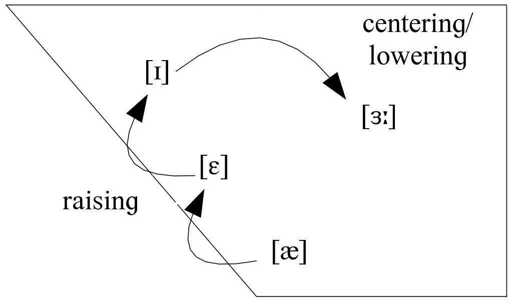

Figure 1: NZE vowel shifts in comparison to AE/RP.

With respect to the NEAR/SQUARE-diphthong merger (which is effectively a raising of [eə] to [iə], cf. Gordon & Maclagan, 1999, 2001), Warren et al. account for the change by different types of exemplar sets (Warren et al., to appear). They propose a model in which words are represented as distributions of phonetically very detailed exemplars. Words with the raised diphthong (e.g. *hair*) involve sets of exemplars consisting of both [heə] and [hiə] tokens, while words with the originally high diphthong (e.g. *here*) only comprise [hiə] tokens. Lexical access is guided through a pre-lexical processor which filters the incoming speech signal according to particular biases. In the *hair/here* case, the processor is claimed to be biased towards [hiə], accounting for the merging on *here*.

Thomas, 2004, proposes an interesting phonemic pre-lexical representation for the perception of words containing [el] versus [æl] (e.g. in *mellow* versus *mallow*), denoted by [Æl]. This representation enables the access of both [el] and [æl] exemplars and is in a way reminiscent of the underspecification approach to which I will turn in the next section.

In contrast to the proposals briefly illustrated above, this paper approaches the NZE short front vowel shift from a different perspective and pursues the following claims:

- The NZE short front vowels are not only phonetically but also phonologically different to their counterparts in British or American English. The originally three-way height contrast of the short front vowels in AE/RP is restructured into a two-way height contrast in NZE (cf. Hawkins, 1976; Wells, 1982b).
- Words containing these short front vowels in their stems are single and abstract phonological representations.
- The vowel in *bat* is low in AE, but changed to a non-low (mid) vowel in its underlying representation in NZE.

2. Feature-based approach to the NZE vowel change

The claim of this paper is that the NZE short front vowels have underlyingly a two-height contrast. The lexical representation of the relevant vowels is described within the framework of the Featurally Underspecified Lexicon (FUL, Lahiri & Reetz, 2002). In this model, lexical entries do not consist of listed exemplars, but are abstract representations of sets of features. These phonological features are hierarchically organised, but not all sounds are necessarily specified for every conceivable feature (*underspecification*).

For the description of the NZE vowel system, the most crucial of the underlying features are *articulator* and *tongue height* features, which are represented as monovalent specifications. The synchronic representation of the short front vowels in NZE is illustrated in Table 1.

Table 1: Synchronic feature-based representation of the NZE front vowels compared to their counterparts in AE/RP (ORTH.=orthography, PRON.=pronunciation).

ORTH.	LEXICAL FEATURES NZE	PRON. NZE	LEXICAL FEATURES AE/RP	PRON. AE/RP
bit	[LAB] [ ]	[bɜt]	[HIGH]	[bɪt]
bet	[HIGH]	[bɪt]	[ ]	[bɛt]
bat	[ ]	[bɛt]	[LOW]	[bæt]

The crucial differences between NZE and AE/RP are based on the height specifications of the respective vowels. In particular, it is assumed here that NZE *bit* comprises a non-high labial vowel, while the vowel in *bet* is high. At last, the vowel in *bat* is non-low in NZE, thereby contrasting with its low AE equivalent. The experimental focus will be on that vowel (cf. next section).

3. Behavioural evidence for the vowel change in NZE

The crucial assumption of FUL (Lahiri & Reetz, 2002) is that lexical representations can only be activated if their phonological features do not mismatch with the features phonetically extracted from a particular speech input. Thus, if the vowel in NZE *bat* is in fact non-low (i.e. not specified for height, illustrated with empty brackets), a phonetically high vowel (such as in NZE *bet*) or a non-high labial vowel (such as in NZE *bit*) should be a so-called *nomismatch* for *bat.* One experimental technique suitable to test these predictions is the *indirect semantic priming*. In such a design, the reaction time of a particular target (e.g. *club*) is compared across different conditions. In the control condition, the preceding prime has no semantic relation to the target (e.g. *camp*), while in the direct semantic condition, the prime is a semantic relative of the target (e.g. *bat*) and ought to speed up its lexical decision (priming). The crucial test conditions are the ones in which the prime has a phonological relation to the direct semantic prime. Thus, in these indirect semantic priming conditions, it is assumed that the primes *bet* and *bit* activate the meaning of *club* through *bat* if the amount of priming does not differ from the direct semantic condition. This is only possible if the vowels in *bet* and *bit* do not mismatch with the vowel in *bat,*

Proceedings of the 11th Australian International Conference on Speech Science & Technology, ed. Paul Warren & Catherine I. Watson. ISBN 0 9581946 2 9
University of Auckland, New Zealand. December 6-8, 2006.

Accepted after abstract only review

Table 2: Priming predictions for an indirect semantic priming design comparing the vowel in NZE and AE *bat* (TH=tongue height).

CONDITION	**PRIMES (NZE VOWELS)**	**EXTRACTED FEATURES**	**V IN *bat***		**PRIMING OF *club***	
			NZE	AE	NZE	AE
(1) indirect	*bit* [ɜ]	[LAB] $[\ ]_{TH}$	/ɛ/ $/\ /_{TH}$	/æ/ $/LOW/_{TH}$	✓	✗
(2) indirect	*bet* [ɪ]	$[HIGH]_{TH}$			✓	✗
(3) direct	*bat* [ɛ]	$[\ ]_{TH}$			✓	✓

which is true for NZE ([ɛ]), but not for AE ([æ]; illustrated in Table 2.

For NZE listeners, the features extracted from all three prime types are compatible with the vowel specification in *bat* (underspecified for tongue height). For AE listeners, in contrast, the labial coronal vowel [ɜ] in the prime is incompatible with the low vowel /æ/ in the lexicon. Likewise, the high vowel [ɪ] in the signal is a mismatch for the low /æ/ vowel. Therefore, no priming is predicted in conditions (1) and (2) for AE listeners.

3.1. Experimental design and material

16 triplets of English minimal pairs (e.g. *bat–bet–bit*) were selected as prime items. Targets were selected as semantic relatives of the primes with <a> as their stem vowel (e.g. *club* for *bat*). Targets were approximately matched to the frequency of their primes (48 per million [targets] vs. 49 per million [primes], based on CELEX (Baayen, Piepenbrock, & van Rijn, 1993)).

The experimental design involved three test conditions (direct: *bat→club*; indirect: *bet→club; bit→club*) and one control condition (*camp→club*). All four conditions were distributed over four subject groups in a Latin Square design, such that no subject heard a target more than once. Each subject list involved 104 fillers of which 68 were nonwords, derived from existing English monosyllabic nouns by changing one or more segments. Primes were always words. The total number of items per group was 136 and there were as many words as nonwords. Primes and targets were 250 ms apart.

Subjects heard the prime-target pairs over headphones and had to decide as quickly and accurately as possible whether the 2nd word of each pair was an actual English word or not by pressing the appropriate buttons of a mobile reaction time measurement device (Reetz & Kleinmann, 2003).

3.2. Experiment 1: NZE listeners (110)

In the ANOVA for the NZE listeners, there was no VOWEL effect, i.e. reaction times on the targets in the three test conditions were independent of the stem vowels of the three test primes. The factor PRIME TYPE (control, direct *bat*, indirect *bit* and indirect *bet*) was significant (F[3,1437]=4.79, p<0.003); priming occurred in the direct and indirect conditions (*bat* primes [semantically related] versus unrelated primes: t-value=3.62, p<0.001; *bet* primes [indirectly related] versus unrelated primes: t-value=2.59, p<0.01; *bit* primes [indirectly related] versus unrelated primes: t-value=2.61, p<0.01). No differences were found between these three conditions. There was also no correlation between the frequency of the prime and the reaction time on the target (F[1,1561]=0.03, p>0.87) nor were the target latencies dependent on the Euclidian distances between the prime vowels and the vowel in *bat* (F[1,812]=0.19, p>0.66).

3.3. Experiment 2: AE listeners (85)

In the ANOVA for the AE listeners, the factor VOWEL was significant (F[2,678]=6.60, p<0.002). Crucially, reaction times on the targets which were preceded by *bat* primes significantly differed from those which were preceded by *bit* or *bet* primes (*bat* vs. *bit*: t-value=-3.14, p<0.002; *bat* vs. *bet*: t-value=-3.15, p<0.002). There was also a PRIME TYPE effect (F[3,985]=3.52, p<0.02). In particular, priming was significant in the direct semantic condition (*bat* primes [semantically related] versus unrelated primes; t-value=2.08, p<0.04), but not in the two indirect priming conditions (*bet* primes [indirectly related] versus unrelated primes: t-value=-0.52, p>0.60, and *bit* primes [indirectly related] versus unrelated primes: t-value=-0.95, p>0.34). Again, the target latencies did not depend on the prime frequencies (F[1,1087]=0.34, p>0.56) or the Euclidian distances between the prime vowels and the vowel in *bat* (F[1,535]=0.23, p>0.63).

3.4. Comparison experiment 1 & 2

The two experiments were compared in a combined ANOVA (F[212,2429]=9.63, p<0.001, cf. Figure 2) with the factors SUBJECT, TARGET, PRIME TYPE, EXPERIMENT (NZE listeners, AE listeners) and the interaction PRIME TYPE X EXPERIMENT. The factor PRIME TYPE was significant (F[3,2429]=6.02, p<0.001). Crucially, PRIME TYPE interacted with EXPERIMENT (F[3,2429]=2.46, p<0.05).

Proceedings of the 11th Australian International Conference on Speech Science & Technology, ed. Paul Warren & Catherine I. Watson. ISBN 0 9581946 2 9
University of Auckland, New Zealand. December 6-8, 2006.

Accepted after abstract only review

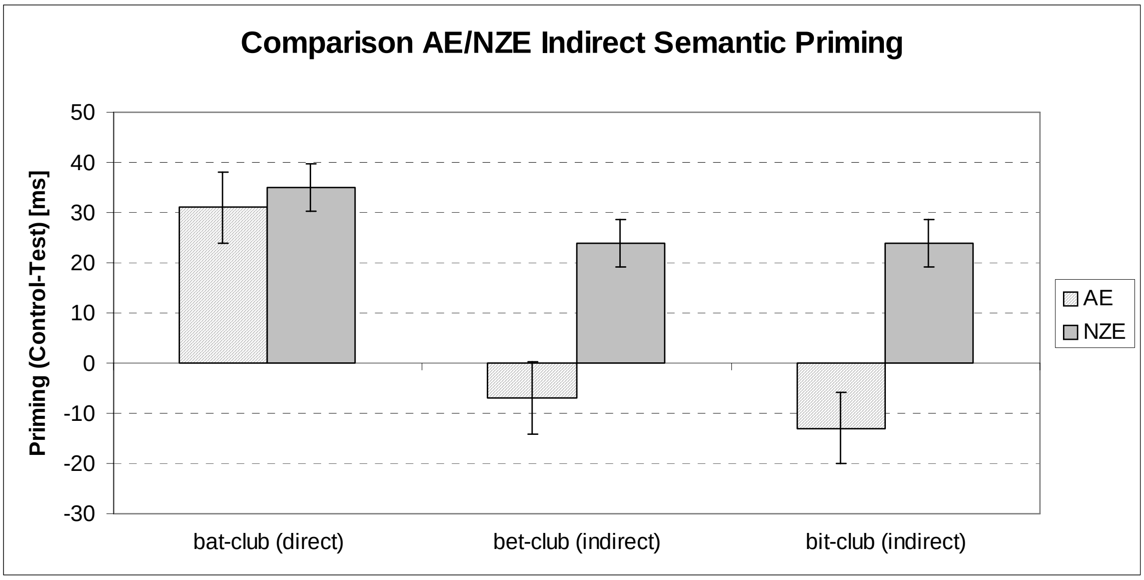

Figure 2: Priming comparison between American English (AE: diagonal stripes) and New Zealand English (NZE: solid grey) listeners in experiment 1 and 2. The amount of priming corresponds to the differences between the control and the corresponding test conditions. Reaction times are given as Least Square Means (LSM, in milliseconds). Standard errors (+/-) are indicated on top of each bar and are based on the t-tests within PRIME TYPE.

4. Discussion

The purpose of experiment 1 and 2 which compared New Zealand English and American English listeners in an indirect semantic priming was to provide evidence that the change in the NZE short front vowels is indeed a phonemic change and not a mere development towards a greater phonetic variability in pronunciation. In addition, each individual experiment provided evidence for featurally organised abstract phonological representations.

The comparison of the priming results between NZE and AE listeners showed that priming depended on (height) matches and nomismatches, respectively. The difference between the specification of the vowel in NZE *bat* and the vowel in AE *bat* was manifested in the priming differences between the two studies (experiment 1 and 2). Experimental evidence for the shifts of [ɛ] and [ɪ] must be provided by future research.

The differences between experiment 1 and 2 depended on the differing representations of the vowel in *bat* in NZE versus AE. The acoustically presented stimulus material was the same in each experiment, and it is assumed that listeners extracted the same phonetic features from the test primes, irrespective of their English variety.

The feature-based model presented here is generally compatible with the view of Hawkins, 1976, who claims that the NZE vowel system has only a two-way height distinction (high versus non-high) and that this pattern is more stable than a pattern involving a three-way height contrast. However, while he considers NZE [ɪ] to be non-front, high and unrounded, we believe that the vowel is front, non-high and rounded, as indicated by some NZE transcriptions such as *fush* or *chups* (Gordon et al., 2004).

On the other hand, the results are a challenge for exemplar-based models which emphasise the importance of distances in the vowel space and frequency distributions among exemplars. Both experiments showed that the amount of priming did not correlate with the acoustic closeness of the prime vowels to the vowel in *bat* or with the frequency of the primes. Yet it is possible that the priming pattern in NZE reflected a restructured exemplar set for *bat* words, containing both *bet* and *bit* tokens. Nevertheless, one would possibly expect a pre-lexical bias towards the token with the vowel closest to the vowel in *bat*. Such a bias was not seen in the priming results, though.

FUL does not deny that indexical properties of speech are important. It is certainly possible that top-down processes exploit the knowledge of social factors, for instance, which are transmitted by accent, pitch or specific phonetic cues. These processes may alter certain expectations in the corresponding communicative situations. However, do indexical properties of speech in fact directly determine whether or not a given word is recognised? We believe not.

Altogether, FUL claims that it is advantageous to encode variability directly in lexical representations, rather than via extensive storage. Exemplar models resolve variation by storing many phonetic details, but it may be that for speech recognition, very fine grained phonetic and extra-linguistic information is subsidiary, but not necessary knowledge.

5. Acknowledgements

The research for this paper was founded in part by the German Research Foundation (SFB471, Leibniz-Prize, and a grant from the ministry of Science, Research, and the Arts of Baden-Württemberg to Aditi Lahiri) and the German Academic Exchange Program. Special thanks to Jen Hay for her support in realising the experiments at Canterbury University, Christchurch. Many thanks also to John Kingston for his invaluable comments on earlier versions of this paper.

6. References

Allan, W. S., & Starks, D. (1999). 'No-one sounds like us?' A comparison of New Zealand and other Southern Hemisphere Englishes. In A. Bell & K. Kuiper (Eds.), *New Zealand English* (pp. 53-83). Amsterdam: John Benjamins.

Baayen, H., Piepenbrock, R., & van Rijn, H. (1993). The CELEX lexical database (Release 1) [CD-Rom]. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

Bauer, L. (1986). Notes on New Zealand English phonetics and phonology. *English World-Wide: A Journal of Varieties of English, 7*(2), 225-258.

Easton, A., & Bauer, L. (2000). An acoustic study of the vowels of New Zealand English. *Australian Journal of Linguistics: Journal of the Australian Linguistic Society, 20*(2), 93-117.

Gordon, E. (1983). New Zealand English pronunciation: an investigation into some early written records. *Te Reo: Journal of the Linguistic Society of New Zealand, 26*, 29-42.

Gordon, E., & Maclagan, M. (1999, November 1999). "Are we NEAR the SQUARE?" A real time study of the merger of two centering diphthongs in New Zealand English: *13th Conference of the Linguistic Society of New Zealand*, Palmerston North.

Gordon, E., & Maclagan, M. (2001). 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: Journal of the Australian Linguistic Society, 21*(2), 215-238.

Gordon, E., Maclagan, M., Hay, J., Campbell, L., & Trudgill, P. (2004). *New Zealand English: Its Origins and Evolution*. Cambridge, UK: Cambridge University Press.

Haggo, D. (1984). Transcribing New Zealand English vowels. *Te Reo: Journal of the Linguistic Society of New Zealand, 27*, 63-67.

Hawkins, P. (1973). A phonemic transcription system for New Zealand English. *Te Reo: Journal of the Linguistic Society of New Zealand, 16*, 15-21.

Hawkins, P. (1976). The role of NZ English in a binary feature analysis of English short vowels. *Journal of the International Phonetic Association, 6*(2), 50-66.

Lahiri, A., & Reetz, H. (2002). Underspecified recognition. In C. Gussenhoven & N. Warner (Eds.), *Laboratory Phonology VII* (Vol. 7, pp. 637-677). Berlin: Mouton de Gruyter.

Maclagan, M., & Hay, J. (2004). The rise and rise of New Zealand English DRESS. In S. Cassidy, F. Cox, R. Mannell & S. Palethorpe (Eds.): *10th Australian International Conference on Speech Science and Technology*, Macquarie University, Sydney: Australian Speech Science and Technology Association Inc, 183-188.

Reetz, H., & Kleinmann, A. (2003). Multi-subject hardware for experiment control and precise reaction time measurement. In M. J. Solé, D. Recasens & J. Romero (Eds.): *15th International Congress of Phonetic Sciences (ICPHS)*, Barcelona: Universitat Autònoma de Barcelona, 1489-1492.

Thomas, B. (2004). *In Support of an Exemplar-Based Approach to Speech Perception and Production: A Case Study on the Merging of Pre-Lateral DRESS and TRAP in New Zealand English.* Unpublished Master Thesis, University of Canterbury, Christchurch.

Warren, P., Hay, J., & Thomas, B. (to appear). The loci of sound change effects in recognition and perception. In J. Cole & J. I. Hualde (Eds.), *Laboratory Phonology IX*. New York: Mouton de Gruyter.

Watson, C. I., Maclagan, M., & Harrington, J. (2000). Acoustic evidence for vowel change in New Zealand English. *Language Variation and Change, 12*(1), 51-68.

Wells, J. C. (1982a). *Accents of English I: Introduction*. Cambridge, UK: Cambridge University Press.

Wells, J. C. (1982b). *Accents of English III: Beyond the British Isles*. Cambridge, UK: Cambridge University Press.

Notes

[1] New Zealand English refers to the English spoken in New Zealand and is considered to be a homogeneous dialect (Bauer, 1986:227).

[2] The focus is on the [æ]-realisation of orthographic <a> which was not affected by the TRAP-BATH split of 18th century British English (cf. Wells, 1982a:232 pp. and further discussion).

Proceedings of the 11th Australian International Conference on Speech Science & Technology, ed. Paul Warren & Catherine I. Watson. ISBN 0 9581946 2 9
University of Auckland, New Zealand. December 6-8, 2006.

Accepted after abstract only review