

Congruence between 'word age' and 'voice age' facilitates lexical access

This paper verifies a core prediction of Exemplar Theory (ET). In ET, a lexical representation of a word is a distribution of remembered encounters with that word [1]. These memories (or 'exemplars') are complete in phonetic detail and are tagged with social and contextual information. At the very core of the theory, then, is the prediction that individual words should have a unique phonetic distribution shaped by the environments in which they were most encountered. This hypothesis receives some support from work showing that homophones are disambiguated differently depending on tone of voice. For example, if /bɹaɪdəl/ is said in a happy voice, listeners are more likely to guess 'bridal' than 'bridle' [2]. Similarly, there are several documented cases of 'old-timey' words appearing to lag behind in sound changes [3, 4]. However, these studies do not eliminate the possibility that the observed differences are due to topic-based style shifting rather than being truly representational, and situated at the word level.

We pursue the hypothesis directly by exploring the prediction that words should be more easily processed when they contain characteristics that most resemble the listener's accumulated past experience with that word. More particularly, we predict that old-timey words (e.g. *frock*, *wireless*) should be more easily recognised when produced in an older voice (which should more closely resemble the voices from which that word has been experienced) than when produced in a younger voice.

We conducted an auditory lexical decision task using native speakers of New Zealand English (NZE). Stimulus words were sourced by comparing two corpora in the Origins of New Zealand English Project (ONZE): the Intermediate Archive, a body of interviews with speakers born 1890 and 1930 and recorded between 1989 and the late 1990s, and the Canterbury Corpus, containing speakers born between 1930 and 1984, recorded from 1994 to the present day. A list of words that were considerably overrepresented in one corpus compared to the other were automatically generated, and from these lists we selected 10 'old' and 10 'young' words for each of three vowels: /ɪ/ (e.g. old: *picnic*, *shilling* / young: *video*, *chips*) /e/ (e.g. *sherry*, *pencils* / *restaurant*, *stress*) and /ai/ (e.g. *idle*, *kite* / *lifestyle*, *sky*). These vowels have undergone sound change over the last 40 years in NZE. We also included a set of neutral stimulus words for each vowel which occurred with approximately equal frequency in both corpora.

Two female speakers recorded the target stimuli, together with matched non-words. One speaker's recording was created by having her shadow randomly presented words produced by the other speaker. This was done in an attempt to match pace and intonational contours across the two speakers. The speakers were approximately matched in social class, but differed in age (22 vs 50). This age difference was auditorily perceivable, as verified by independent age ratings provided by our listeners. The auditory lexical decision experiment was conducted using Direct RT. The stimuli were blocked by speaker, and randomized within each block. All 25 participants listened to both blocks. The order of presentation of blocks was counterbalanced. Participants were asked to identify words, and responses and response times were recorded.

Accuracy rates were very high. Old words were correctly recognized as words at 96.2% in the old voice and 94.3% in the young voice. Young words were correctly recognized at 95.6% in the old voice and 96.1% in the young voice. This is a significant interaction in a binomial mixed effects model showing increased error rates when voice age and word age do not match ($p < .02$). After removing outliers, mean reaction times in the congruent conditions were 114ms (old word, old voice) and 118ms (young word, young voice). In the incongruent conditions they were significantly higher: 138ms (young word, old voice) and 127ms (old word, young voice). This interaction between word age and voice age is significant in a linear mixed effects regression model with participant and item as random effects ($p < .03$).

This provides robust evidence that words are more easily processed when they contain characteristics that most resemble the listener's accumulated past experience with that word. This provides verification of a key prediction of exemplar models of the lexicon.

References

- [1] Pierrehumbert, Janet (2001). Exemplar dynamics: Word frequency, lenition and contrast. In Joan Bybee and Paul Hopper (eds.), *Frequency and the Emergence of Linguistic Structure*, pp. 137-157. Amsterdam: Benjamins.
- [2] Nygaard, Lynne C. and Erin R. Lunders (2002). Resolution of lexical ambiguity by emotional tone of voice. *Memory & Cognition* 30 (4): 583-593.
- [3] Gordon, E., Campbell, L., Hay, J., Maclagan, M., Sudbury, A., Trudgill, P. (2004) *New Zealand English: Its Origins and Evolution*. (Studies in English Language) Cambridge: Cambridge University Press.
- [4] Yaeger-Droh, Malcah and William Kemp (1992). Lexical classes in Montreal French (E:). *Language and Speech* 35 (3): 251-293.