Covarying New Zealand vowels interact with speech rate to create social meaning for NZ listeners

Elena Sheard[1], Jen Hay[12], Robert Fromont[1], Joshua Wilson Black[1], Lynn Clark[12]

[1]*New Zealand Institute of Language, Brain and Behaviour, University of Canterbury*

[2]*Department of Linguistics, University of Canterbury*

Brand et al. [1] and Hurring et al. [2] identify sets of vowels that covary across speakers in New Zealand. One of these is a set of mostly front vowels undergoing sound change, in which speakers tend to be either leaders or laggers (see Fig. 1). The second is a set of back vowels, for which speakers have systematically different configurations, characterized by either an innovative STRUT/START and a conservative THOUGHT, or vice-versa. We refer to these subsets as covariation in the *leader-lagger* space, and in *back-vowel configuration,* respectively. This study asks whether these patterns of co-varying vowels identified in production also surface as perceptually salient to NZ listeners.

We presented 117 New Zealand English speakers with audio stimuli from 38 women between the age of 46 and 55 in the QuakeBox corpus [3], and asked them to place them into groups, using a modified version of the browser-based Audio-Tokens toolbox [4] (see Fig. 2). Participants were asked to make groups of stimuli that they perceived to sound similar and label them with free text [5]. We manually recoded the free text responses into categories for analysis. When we inspect all the labels used, 49% of labels named a speech characteristic (most frequently, speed or pitch). Analysis shows that listeners are highly accurate in the use of these labels. Our focus here is on the 37% of labels specifically related to social evaluations. Examples include 'Rural', 'Old, or 'Strong NZ Accent'.

We applied a two-dimensional non-metric multidimensional scaling analysis to this subset of data [6]. This process gave us two major perceptual Dimensions. We then used Principal Components analysis to explore the labels typically associated with voices at the extremes of these dimensions. As shown in Figure 3, speakers low in Dimension 1 are labelled with characteristics associated with broad New Zealand accents, such as 'Rural', 'Strong NZ accent', or 'low SES'. We therefore refer to D1 as *'perceptual broadness'*. Dimension 2 is more robustly associated with age, with speakers low in D2 identified as old sounding. We refer to D2 as '*perceptual age*'.

To investigate the links with acoustics, we fit two linear regressions modelling the measures of *perceptual broadness* and *perceptual age* as dependent variables, using speakers' position in 'leader-lagger' and 'back vowel configuration' continuums (Principal Components taken from the production study in Hurring et al. [2], see Fig. 1 for leader-lagger PC) in interaction with speaker speech rate and mean pitch as fixed effects. Position in the 'back vowel configuration' did not predict either perceptual dimension. However, the models revealed a significant interaction between speech rate and 'leader-lagger' position in predicting both *'broadness*' (Fig. 4) and '*age'*. Leaders in sound change are heard as 'broad', for example, but only if they are also fairly slow speakers. Laggers in sound change are heard as 'old' if they are slow speakers, but relatively young if they speak more quickly. The perceived age of leaders in sound change, on the other hand, is much less affected by speech rate.

Using a bottom-up approach to speaker grouping, we have revealed perceptual dimensions that are predicted by one of the covarying subsystems of vowels identified by Brand et al. [1] and Hurring et al. [2]. This suggests that covarying vowel patterns can carry social meaning. However, the meanings carried can depend on listener categorisation and perceived social meanings do not operate in isolation of other acoustic cues. Specifically, fast and slow speakers with similar front vowel patterns are not only accurately labelled in terms of speed, but they are also more likely to be labelled with different social characteristics, and their vowel productions are likely to be interpreted differently from each other. The effect of vocalic production is therefore mediated by speech rate in the creation of social meaning. This supports previous work by Campbell-Kibler [7] which suggests that individual variables do not carry social meaning in isolation, but rather that social meaning can be associated with a complex set of characteristics, which work together and interact to affect how a voice is evaluated.

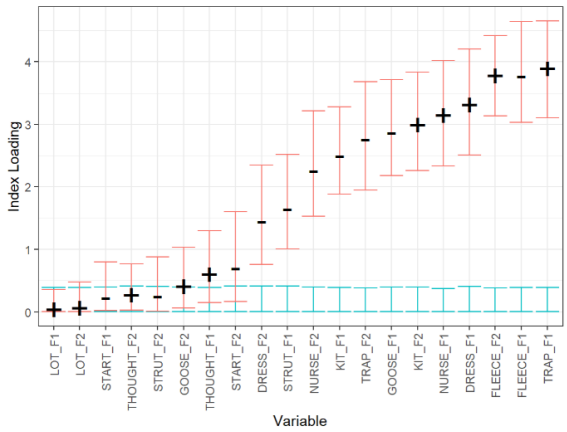

Figure 1 PC1 ('leader-lagger continuum') from Hurring et al. (Under review) showing acoustic characteristics associated with the 'leader-lagger' continuum. Speakers high in PC1 are 'laggers' - they have high TRAP F1, low FLEECE F1, etc. See Wilson Black et al. (2023) for information about how to read the plot.

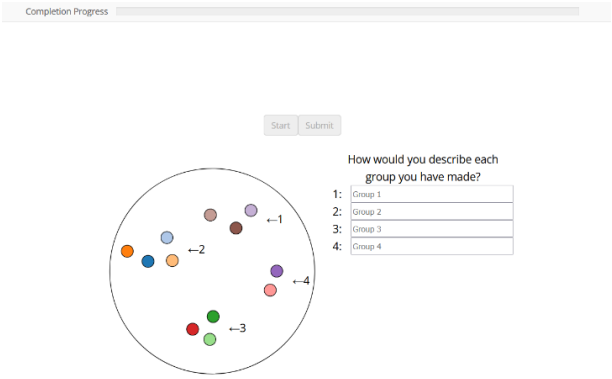

Figure 2 Screenshot of the modified browser-based free classification task. In response to feedback on pilot tasks, audio stimuli were randomly allocated across three iterations of task to reduce cognitive load.

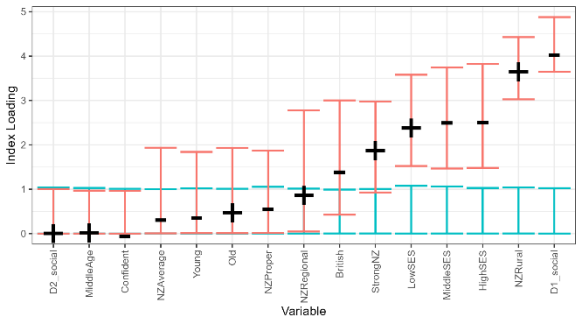

Figure 3 PC from analysis of perceptual Dimension 1 ('Broadness') and its association with label types. Speakers that are low D1 tend to be labelled as Rural, low SES and having a strong accent, whereas speakers that are high D1 tend to be labelled high and middle SES.

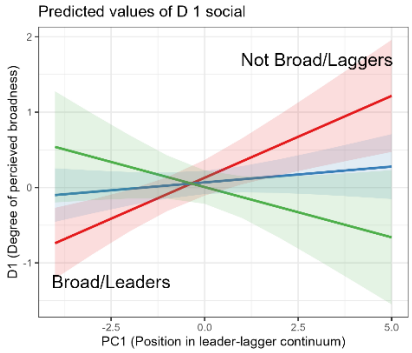

Figure 4 Interaction between position in the leader-lagger continuum (PC1 from Fig 1), and perceived 'broadness' (D1 from perception study, see Fig 3 for social associations) as mediated by speech rate. 'Leaders' (low PC1) are heard as 'Broad' (low D1), and laggers (high PC1) as 'not Broad' (high D1), but only for slow speakers (in red), not for fast speakers (in green).

References

[1] Brand, J., J. Hay, L. Clark, K. Watson & M. Sóskuthy (2021). Systematic co-variation of monophthongs across speakers of New Zealand English. *Journal of Phonetics,* 88, 101096.

[2] Hurring, G., J. Wilson Black, J. Hay, & L. Clark (Under review). How stable are patterns of covariation across time?. *Language Variation and Change,*.

[3] Walsh , L., J. Hay, D. Bent, L. Grant, J. King, P. Millar, V. Papp & K. Watson (2013). The UC QuakeBox Project: Creation of a community-focused research archive. *New Zealand English Journal,* 27, 20-32.

[4] Donhauser, P. W. & D. Klein (2023). Audio-Tokens: a toolbox for rating, sorting and comparing audio samples in the browser. *Behavior research methods,* 55, 508-515.

[5] Clopper, C. G. (2008). Auditory free classification: Methods and analysis. *Behavior Research Methods,* 40, 575-581.

[6] Borg, I. & P. J. F. Groenen (2005). *Modern multidimensional scaling: theory and applications*, 2nd ed. New York: Springer.

[7] Campbell-Kibler, K. (2011). Intersecting variables and perceived sexual orientation in men. *American Speech,* 86, 52-68.