PERCEPTION OF CODA VOICING: GLOTTALISATION, VOWEL DURATION, AND SILENCE

Joshua Penney, Felicity Cox, and Anita Szakay

Centre for Language Sciences, Department of Linguistics, Macquarie University, Australia
joshua.penney@mq.edu.au; felicity.cox@mq.edu.au; anita.szakay@mq.edu.au

ABSTRACT

Recent research from Australian English has shown that glottalisation of vowels preceding coda stops results in increased perception of coda voicelessness. However, the addition of glottalisation results in a shorter portion of the vowel being modally voiced, raising the question of whether listeners may parse glottalisation as belonging to the coda rather than the preceding vowel. If so, listeners would perceive a shorter preceding vowel duration therefore increasing the perception of coda voicelessness.

This study thus compared listeners' coda voicing responses for words containing glottalised vowels with words containing vowels in which glottalisation was replaced with silence. The results suggest that both glottalisation and shorter vowel duration/longer coda closure duration result in increased voiceless percepts, but that listeners respond differently to these two conditions. The findings indicate that listeners are sensitive to glottalisation and utilise this as a cue to coda voicing rather than simply perceiving shorter modally voiced vowels.

Keywords: Glottalisation; coda stop voicing; vowel duration; closure duration.

1. INTRODUCTION

Glottalisation occurs in conjunction with voiceless coda stops in many varieties of English [5, 6, 8, 15, 18, 19], including Australian English (AusE), where it has been shown to be present 55% of the time in voiceless coda stop contexts compared to 6% in voiced coda contexts [12]. Glottalisation may be realised by the insertion of a glottal stop prior to an oral stop, though in AusE it is most often manifested as irregular, laryngealised phonation (i.e. creaky voice) on the end of the vowel preceding a coda stop. Some researchers hypothesise that glottalisation is employed in association with voiceless coda stops in order to enhance the perception of voicelessness [9, 15], as glottalisation appears to be more common when voiceless coda stops precede sonorants, an environment in which anticipatory coarticulatory voicing is likely [8, 16]. The inclusion of glottalisation may minimise the likelihood of anticipatory voicing.

Recent studies have analysed whether listeners are sensitive to glottalisation in perception. In an eye tracking study, [3] found that American English listeners were slower to identify words with voiced coda stops when glottalisation was present, though they reacted similarly for words with voiceless codas regardless of whether glottalisation was present or absent. This suggests that listeners are sensitive to glottalisation and associate it with voiceless, but not voiced, coda contexts.

[13] conducted a perceptual task with AusE listeners in which preceding vowel duration and glottalisation were co-varied. They found that listeners identified a higher proportion of voiceless stops when glottalisation was present, suggesting that glottalisation has the effect of promoting the perception of coda voicelessness. This effect was found to be stronger for inherently short vowels (e.g. /ɐ/ STRUT) than for inherently long vowels (e.g. /ɐː/ START). Inherently short vowels have been shown to display smaller durational differences across coda voicing contexts than inherently long vowels in production [4, 12]. Therefore, the vowel duration cue to coda voicing may not be as strong in the context of inherently short vowels and, accordingly, glottalisation may be a stronger cue in this context [14]. Importantly, the effect of glottalisation promoting the perception of coda voicelessness was found even when listeners were presented with a preceding vowel with an extended duration, which would otherwise produce the percept of coda voicing [2, 11, 17].

As glottalisation necessarily results in non-modal phonation affecting the vowel preceding the coda stop, it may be possible that listeners in [13] did not parse the glottalised portion of the vowel as belonging to the vowel. That is, listeners may have used only the modally voiced vocalic portion in their perception of stimulus voicing. In other words, listeners would have perceived a shorter vowel duration when glottalisation was present compared to when it was absent. As shorter preceding vowels produce the percept of voiceless coda stops [2, 11, 17], an increase in the proportion of voiceless responses would be expected. In addition, if the glottalisation were interpreted as part of the coda stop closure, this too would be consistent with increased perception of coda voicelessness because increased closure duration is associated with voiceless coda stops [7, 10].

Thus, the aim of this study is to investigate whether listeners are perceptually sensitive to glottalisation itself or whether they parse glottalisation as part of the closure and, consequently, respond to the shorter (modally voiced) vowels when perceiving coda voicing. We will examine whether listeners react in the same way to glottalised stimuli as they do to stimuli in which the glottalised portion is replaced with silence, thereby reducing vowel duration and increasing closure duration. If listeners do parse glottalisation as belonging to the coda and hence are sensitive to shorter preceding vowel duration rather than glottalisation, we would hypothesise that the inclusion of both glottalisation and of silence would provide the same results. On the other hand, if listeners are perceptually sensitive to the glottalised portion of the vowel, we would expect increased perception of coda voicelessness when glottalisation is present compared to silence.

2. METHOD

2.1. Participants

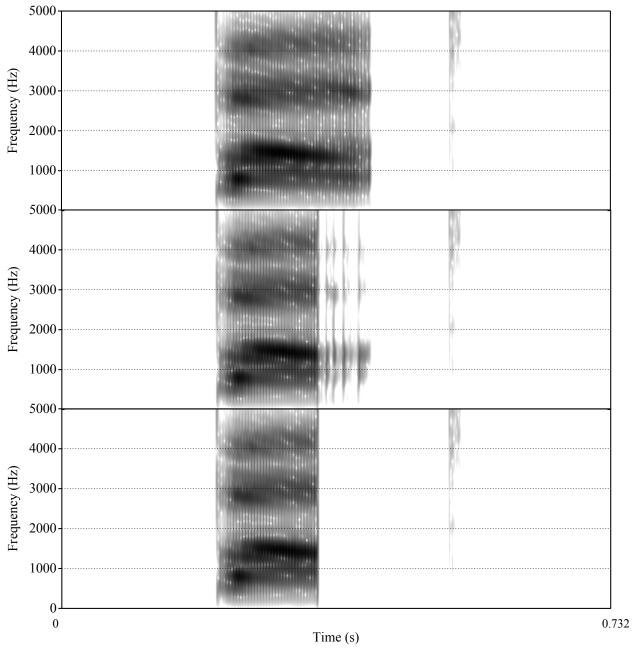
We enlisted 51 participants (female: 46; male: 5) to take part in this task. All were undergraduate students who received course credit for participating. Participants were aged between 18 and 23 years (mean: 19.5) and were native AusE speakers, who were born in and completed all of their schooling in Australia. No participant reported any speech or hearing issues.

2.2. Stimuli

The stimuli were created from natural speech tokens of the words *bead*, *bid*, *bard,* and *bud* produced by a non-rhotic female native AusE speaker aged 25. Recordings were made in a sound treated studio at Macquarie University with an AKG C535 EB microphone and a Presonus StudioLive 16.2.4 AT mixer recorded to an iMac at a sampling rate of 44.1kHz. The speaker produced the target words with modal phonation, and also produced sustained realisations of each of the relevant vowels with laryngealised phonation from which glottalised stimuli were created.

Cues to coda voicing were removed: release bursts were replaced with a perceptually ambiguous burst taken from a low amplitude voiced coda stop from an unstressed syllable; F1 formant transitions at the end of the vowels were removed; intensity contours were manipulated to ensure uniformity; closure periods were replaced with silence; F0 was manipulated to fall from 265 to 203Hz across each of the vowels. We then manipulated the duration of the vowels. For each of the four tokens we created a continuum of nine equally spaced vowel duration steps. The minimum and maximum durations were determined by production values for the relevant vowels reported in [12] for young female AusE speakers. The mean minus two standard deviations preceding a voiceless coda was the minimum duration, and the mean plus two standard deviations preceding a voiced coda was the maximum duration.

Figure 1: S*pectrograms of ninth-step (i.e. longest vowel duration) 'bud' stimuli. Upper panel shows control condition; middle panel shows glottalised condition; lower panel shows silence condition.*



A second set of stimuli was then created for the glottalised condition. This set was identical to the first set of stimuli, with the exception that the final portion of the vowel in each token was replaced with glottalisation taken from the sustained vowels produced with laryngealised phonation, resulting in a drop in F0. The final 35% of the vowel was replaced with glottalisation for inherently short vowels, and the final 25% was replaced for inherently long vowels, as reported for production in [12].

Finally, a third set of stimuli was produced for the silence condition. This set was identical to the glottalised set of stimuli, except that the glottalised portion of the vowel (final 35% for inherently short vowels; final 25% for inherently long vowels) was replaced with silence. This also resulted in a truncated F0 contour.

Figure 1 illustrates the ninth step of the *bud* continuum in each of the three conditions (control, glottalised, silence). The modally voiced portion of the vowel in the glottalised condition is the same duration as the vowel in the silence conditions. Note that the overall stimulus duration remains the same in all three conditions.

2.3. Procedure

Participants undertook the perception task in a sound attenuated room, using an Apple MacBook Air notebook computer and Sennheiser HD 380 Pro headphones. This was a two-alternative forced choice identification task, in which participants were presented with orthographic representations of minimal pairs differing only in coda stop voicing (e.g. *bard*/*bart*) on the computer screen. The task was to select the word corresponding to the single word audio stimulus. For each stimulus cycle a fixation cross was displayed on the screen for 600 milliseconds, followed by the orthographic presentation of a minimal pair, with one word displayed on the left hand side of the screen and the other on the right hand side (counterbalanced by block and participant). An audio stimulus item was then presented through the headphones after 500 milliseconds and the participant selected by key press the word that they heard. The next cycle then began after the participant's response.

Participants were presented with three repetitions of a nine-step continuum in each of the three conditions for four vowels (/iː, ɪ, ɐː, ɐ/) with each vowel continuum presented in a separate block resulting in 324 responses per participant. We here report on a subset of the data comprising the low vowel contexts (/ɐː, ɐ/; 162 responses per participant).

2.4. Analysis

Using *lme4* [1], we fitted a generalised linear mixed effects model (GLMER) with the dependent variable *listener response*, and fixed factors *continuum step* (short to long vowel duration), *condition* (control/glottalised/silence), *inherent vowel length* (short/long), as well as all two- and three-way interactions between the fixed factors. Random intercepts were included for participant and random slopes were included for all factors by participant.

3. RESULTS

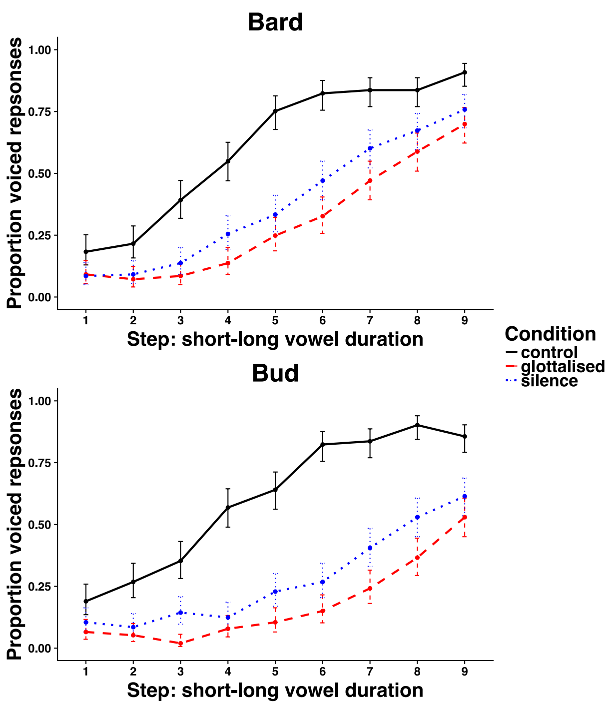
Figure 2 illustrates the proportion of participants' responses for voiced coda stops at each step of the continua across the three conditions. As can be seen, the results of the control condition confirm previous findings that AusE listeners utilise vowel duration as a cue to coda stop voicing: when vowel duration is short (i.e. in the lower steps of the continuum), listeners produce a high proportion of voiceless coda stop responses. As vowel duration increases, so too does the proportion of responses for voiced coda stops [13, 14]. In the glottalised condition the results also mirror what has been previously reported: namely, that the presence of glottalisation at the end of the vowel results in an increase in the proportion of voiceless responses [13, 14]. It can be seen in Figure 2 that in the glottalised condition the proportion of voiced responses is lower than in the control condition. Furthermore, it is evident from Figure 2 that in the silence condition, in which the glottalised portions have been set to silence, there is also an increase in voiceless coda percepts, consistent with our prediction that shorter preceding vowel/longer coda closure duration would increase voiceless coda responses. Importantly, although both experimental conditions show an increase in voiceless coda responses, the proportion of responses differs between the glottalised and the silence conditions, with more voiceless responses in the glottalised condition than in the silence condition.

The results of the mixed model showed significant effects for *continuum step* (χ^2 (1) = 176.67, $p < 0.0001$), demonstrating that as vowel duration increased listeners produced more voiced responses, and *condition* (χ^2 (2) = 130.72, $p < 0.0001$), confirming the differences between responses across the three conditions. There was also a significant two-way interaction between *condition* and *inherent vowel length* (χ^2 (2) = 32.77, $p < 0.0001$), indicating differences between the conditions for the two vowel contexts (i.e. *bard* and *bud*). Post-hoc tests revealed that the control condition differed significantly from the other two conditions for both vowel contexts (all $p < 0.0001$). The glottalised and silence conditions also differed from one another in both vowel contexts, but the difference was greater in the short vowel context ($p < 0.0001$) than in the long vowel context ($p = 0.0113$).

4. DISCUSSION

The results presented above confirm previous findings that AusE listeners use vowel duration to cue coda stop voicing [13, 14]. Shorter preceding vowels are associated with voiceless coda stops and longer preceding vowels are associated with voiced coda stops. In addition, the presence of glottalisation promotes the perception of coda voicelessness, as seen by the higher proportion of responses for voiceless codas in the glottalised condition. As has previously been found [13, 14], the effect of glottalisation producing increased voiceless coda percepts was visible even in the presence of extended preceding vowel duration, which is otherwise a strong cue to a voiced coda stop. Note that these results differ from findings for American English listeners: [3] found that the presence of glottalisation resulted in slower identification of voiced codas, but they did not find it improved listeners' perception of coda voicelessness. Nevertheless, they concluded that American English listeners are sensitive to glottalisation and associate it with voiceless rather than voiced codas, as appears to be the case for AusE listeners.

Figure 2: *Proportion of voiced coda responses in control (solid lines), glottalised (dashed lines), and silence (dotted lines) conditions.*



Although setting the glottalised portions of the vowels to silence also resulted in an increase in voiceless responses, this was shown to be significantly different from the effect of glottalisation, with glottalisation producing a stronger effect compared to silence (see Figure 2). This provides strong evidence that listeners are sensitive to the glottalisation itself, and do not parse glottalisation as belonging to the coda. If it were the case that listeners heard a shorter vowel (i.e. the modally voiced portion of the glottalised vowel) and perceived the glottalisation as part of a longer coda, then the stimuli in the glottalised condition should have elicited the same responses as the stimuli in the silence condition. The results here suggest that listeners do perceive the glottalisation and that this facilitates the increased perception of coda voicelessness. It is possible that listener responses may also have been affected by the F0 differences in these conditions. This feature will be examined in future work.

We also found that the difference between glottalisation and silence was more pronounced in the inherently short vowel context than in the inherently long vowel context. [14] previously found that the effect of glottalisation in promoting the perception of coda voicelessness was stronger for inherently short vowels, and it appears that this is driving the difference in our data here as well. Figure 2 shows that in the glottalised condition for the inherently short vowel context (*bud*) the proportion of voiced responses remains below 50%, even at the highest steps of the continuum where vowel duration is at its longest and hence should serve as a strong cue to a voiced coda. As discussed above, vowel duration has been shown to differ less between voiced and voiceless coda contexts for inherently short vowels than for inherently long vowels in AusE [4, 12]. That is, coda voicing related vowel duration differences are greater for inherently long vowels, such as /ɐː/, than they are for inherently short vowels, such as /ɐ/. Thus, it may be that glottalisation is a stronger perceptual cue to voicelessness for inherently short vowels as the vowel duration cue is less reliable in this context.

5. CONCLUSION

This study confirmed previous findings that AusE listeners utilise vowel duration as a cue to coda stop voicing, and that the presence of glottalisation results in increased perception of coda voicelessness. In addition, we showed that shortening the preceding vowel/lengthening the coda closure by the same duration as the glottalisation also results in an increase in responses for voiceless codas. Although both of these conditions resulted in increased perception of coda voicelessness, glottalisation was found to be a stronger cue than the shortened vowel alone, particularly in the context of an inherently short vowel. These results indicate that listeners are sensitive to glottalisation and do not simply perceive shorter modally voiced vowels when glottalisation is present in the signal.

6. ACKNOWLEDGEMENTS

This work was supported by an Australian Government Research Training Scheme scholarship to the first author. We thank Eva Reinisch, Jonathan Harrington, and Titia Benders for their comments and assistance in the development of the task.

7. REFERENCES

[1] Bates, D., Maechler, M., Bolker, B., Walker, S. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software 67*, 1–48.

[2] Chen, M. 1970. Vowel length variation as a function of the voicing of consonant environment. *Phonetica 22*, 129–159.

[3] Chong, A., Garellek, M. 2018. Online perception of glottalized coda stops in American English. *Laboratory Phonology 9*, 1–24.

[4] Cox, F., Palethorpe, S. 2011. Timing differences in the VC rhyme of Standard Australian English and Lebanese Australian English. *Proc. 17th ICPhS, Hong Kong*, 528–531.

[5] Docherty, G., Foulkes, P. 1999. Derby and Newcastle: Instrumental phonetics and variationist studies. In: Foulkes, P., Docherty, G. J. (eds.), *Urban Voices: Accent studies in the British Isles*. London: Arnold, 47–71.

[6] Gordeeva, O. B., Scobbie, J. M. 2013. A phonetically versatile contrast: Pulmonic and glottalic voicelessness in Scottish English obstruents and voice quality. *J. IPA 43,* 249–271.

[7] Hogan, J. T., Rozsypal, A. J. 1980. Evaluation of vowel duration as a cue for the voicing distinction in the following word-final consonant. *J. Acoust. Soc. Am. 67,* 1764–1771.

[8] Huffman, M. K. 2005. Segmental and prosodic effects on coda glottalization. *Journal of Phonetics 33*, 335–362.

[9] Keyser, S. J., Stevens, K. N. 2006. Enhancement and overlap in the speech chain. *Language 82*, 33–62.

[10] Lisker, L. 1957. Closure duration and the intervocalic voiced-voiceless distinction in English. *Language 33*, 42–49.

[11] Lisker, L. 1986. 'Voicing' in English: A catalogue of acoustic features signaling /b/ versus /p/ in trochees. *Language and Speech 29*, 3–11.

[12] Penney, J., Cox, F., Miles, K., Palethorpe, S. 2018. Glottalisation as a cue to coda consonant voicing. *Journal of Phonetics 66*, 161–184.

[13] Penney, J., Cox, F., Szakay, S. 2018. Weighting of vowel duration and glottalisation in the perception of coda stop voicing. Poster presented at *LabPhon 16*, Lisbon, Portugal.

[14] Penney, J., Cox, F., Szakay, S. 2018. Weighting of coda voicing cues: Glottalisation and vowel duration. *Proc. 19th INTERSPEECH*, Hyderabad, India, 1422-1426.

[15] Pierrehumbert, J. 1994. Knowledge of variation. In: *Papers from the Parasession on Variation, 30th meeting of the Chicago Linguistic Society.* Chicago: Chicago Linguistic Society, 232–256.

[16] Pierrehumbert, J. 1995. Prosodic effects on glottal allophones. In: Fujimura, O., Hirano, M. (eds.), *Vocal fold physiology: Voice quality control.* San Diego: Singular, 39–60.

[17] Raphael, L. J. 1972. Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English. *J. Acoust. Soc. Am. 51*, 1296–1303.

[18] Roach, P. J. 1973. Glottalization of English /p/, /t/, /k/ and /ʧ/ – a reexamination, *J. IPA 3*, 10–21.

[19] Seyfarth, S., Garellek, M. 2015. Coda glottalization in American English. Proc. *18th ICPhS, Glasgow, Scotland.*