INTERSPEECH 2020
October 25–29, 2020, Shanghai, China

Links between production and perception of glottalisation in individual Australian English speaker/listeners

Joshua Penney[1], *Felicity Cox*[1], *Anita Szakay*[1]

[1]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

Glottalisation of coda stops is a recent change in Australian English. Previous studies have shown that speakers use glottalisation to signal coda stop voicelessness in production, and that listeners interpret glottalisation as cueing coda stop voicelessness in perception. As is to be expected for a recent change, younger speakers glottalise more than older speakers, but in perception both age groups appear to use glottalisation similarly. This study examines whether links between the production and perception of glottalisation exist at the level of the individual. We determined how frequently individuals used glottalisation in production, and analysed this against how heavily the same individuals weighted glottalisation in perception. Although differences have previously been found at the age group level, at the level of the individual we found no correlation between how heavily listeners weighted glottalisation in perception and how frequently they used glottalisation in production for either the younger or the older listeners. Nevertheless, we did find a small number of individuals who exhibited an alignment of their production and perception repertoires, which may suggest that only a small proportion of individuals exhibit a strong production-perception link, and we propose that these individuals may be important for driving the progression of change.

Index Terms: glottalisation, production-perception link, coda voicing, Australian English

1. Introduction

Glottalisation associated with voiceless coda stops (also referred to as glottal reinforcement) is a recent change in Australian English (AusE). Despite the absence of this feature being noted as recently as the 1980s [1, 2], later studies have shown that contemporary AusE speakers utilise glottalisation frequently [3, 4, 5]. Glottalisation as an acoustic correlate to coda voicelessness [4, 5] is common in many varieties of English [6, 7, 8, 9, 10]. As is expected in the case of a recent change [11, 12], younger speakers have consistently been found to produce glottalisation at higher rates than older speakers, suggesting they are more progressive and are perhaps driving the change [4, 5, 13]. In an analysis of monosyllabic words, [4] found that younger speakers glottalised 71% of items in voiceless coda contexts, whereas older speakers showed lower rates of glottalisation at 36%. Similarly, [5] found that younger speakers glottalised 64% of items with voiceless coda stops in the unstressed syllables of their study, yet older speakers only glottalised 41% of items. [13] examined rates of glottalisation in a range of environments, and similarly found that younger speakers produced more glottalisation than older speakers in each of the environments examined, with the greatest difference in pre-vocalic environments.

While clear age group differences have been found in production, differences between the two age groups were found to be smaller in terms of their perception of glottalisation. [14, 15] found that both younger and older listeners interpreted glottalisation as a cue to coda voicelessness, and did so even when glottalisation occurred in conjunction with a relatively long vowel duration that would otherwise signal a voiced coda. In conjunction with the reported age differences in production, this finding may suggest that the change towards glottalisation is more advanced in perception than in production. The younger group remains more progressive with regard to both modalities. Some theories of sound change posit that changes occur in an individual's perception prior to being reproduced in the individual's productions (e.g. [16, 17]). It may be possible that the glottalising change observed in production is being led by perception with the older group lagging behind the younger group with respect to the change, and hence demonstrating lower rates of glottalisation in production yet showing sensitivity to glottalisation in perception. Perhaps glottalisation was initially perceived by younger listeners as commonly occurring in conjunction with voiceless stops (possibly as produced by speakers of other varieties of English in which glottalisation is frequent, for example American English [8] or British English [9]), and this was subsequently replicated in their own productions. As more younger speaker/listeners then produced this feature, it would in turn be perceived by more listeners (including older listeners), and in due course may possibly have entered into their productions too.

While group level differences may suggest younger speakers/listeners are leading the change, it is likely that within both age groups there will be individuals who are progressive with respect to the change and others who are less innovative. It is often assumed that the production and perception repertoires of individual speaker/listeners are linked [16, 17, 18, 19], which would entail that individuals who are progressive in terms of one modality (e.g. perception) will also be progressive in terms of the other (e.g. production) and vice versa [17]. Prior research has shown mixed results in terms of how closely production and perception are linked at the level of the individual: some recent studies have successfully shown evidence of such links (e.g. [18, 20, 21, 22]). However, other studies have failed to find support for a close alignment of individuals' production and perception repertoires (e.g. [23, 24, 25, 26, 27, 28]).

In this paper, we examine to what extent individual speaker/listeners' production and perception of glottalisation as a cue to coda voicelessness are linked. The paper draws on two previous studies [13, 15], in which the same individuals participated in a production study and a perception study. If the production and perception of glottalisation is closely linked at the individual level, we would anticipate that individuals who use glottalisation more frequently in production will also show

http://dx.doi.org/10.21437/Interspeech.2020-1175

high perceptual sensitivity to glottalisation as a cue to coda voicelessness. Likewise, we would expect that those speakers who make less use of glottalisation in production would correspondingly show less sensitivity to glottalisation in perception.

2. Methods

2.1. Participants

80 participants took part in a session in which a production and a perception task were presented. For most participants this was conducted in a sound-attenuated room in the Department of Linguistics at Macquarie University; 11 of the participants took part in a quiet room of an alternative offsite location. Data for three participants were discarded for the perception task and data for a further three participants were discarded for the production task (see [13, 15] for more details). The remaining 74 participants took part in both the perception and production tasks. These were allocated to either an older group (aged 56+, $n = 31$; f: 22; m: 9) or a younger group (aged 18–36, $n = 43$; f: 36; m: 7). All participants were L1 speakers of AusE and were born and schooled exclusively in Australia (apart from one participant who migrated to Australia as an infant).

2.2. Production task

Participants were fitted with an AKG C520 headset condenser microphone recorded to a Marantz PMD661 MK II recorder with 44.1 kHz sampling rate and 16-bit quantisation. They read aloud 396 sentences, orthographically presented randomly on a notebook computer screen. Each sentence contained a target word with the form /bVC/, where V was one of the vowels /iː, ɪ, ɐː, ɐ/ and C was either a voiced or voiceless alveolar stop. The sentences included the target words in both phrase medial and phrase final positions: in phrase medial position the target word was embedded in a carrier of the form: ***say <TARGET> now one more time***. The word immediately following the target word (e.g. ***now*** in the example given) was varied, beginning with either a vowel, a consonantal sonorant, or a voiced or voiceless obstruent. This enabled an examination of the following phonetic environment on the presence of glottalisation. In phrase final position the target word was embedded in a carrier of the form: ***now one more time say <TARGET>***. Only target words produced in phrase medial position are discussed here to ensure no potential influence of phrase final creak. One of the aims of the production task was to examine the frequency with which participants produced glottalisation to signal a voiceless coda stop. Glottalisation was identified spectrally. Additional details and results of the production task are provided in [13]. Only items with a voiceless coda are examined here. Mispronounced items were excluded, as were items in which a phrase break was inserted after the target word, as boundaries of this type may facilitate phase final creaky voice [29, 30] making it difficult to disentangle creak from coda-glottalisation. This left 3542 items with voiceless codas remaining. A GLMM showed a clear effect of younger speakers producing glottalisation more frequently than the older speakers ($\beta = -1.51$; SE = 0.31; $z = -4.84$; $p < 0.0001$). The younger group produced glottalisation in 72% of the items whereas the older group produced glottalisation in 44% of the items.

2.3. Perception task

Directly following the completion of the production task, participants took part in a two-alternative forced-choice word identification task. Participants were presented with single-word audio stimulus items in which the voicing status of the final coda had been manipulated to be ambiguous: the coda stop burst was replaced with a low intensity burst that could not be reliably identified as /t/ or /d/; F1 transitions were removed, and F0 and intensity were standardised (see [15] for full details of stimuli manipulation). The stimuli were presented through Sennheiser HD 380 pro headphones, and for each stimulus a minimal pair differing in coda stop voicing (e.g. *bud/but*) was orthographically displayed on the notebook screen, with participants required to select the word they heard. The stimuli came from three separate continua, which were manipulated such that vowel duration, coda closure duration, and the relative proportions of vowel and coda closure duration were varied respectively in equally spaced steps, to cue perception of coda voicing/voicelessness. For each step of each continuum, participants were randomly presented with a non-glottalised and a glottalised item. The glottalised stimuli were created by splicing natural glottalisation into the end of non-glottalised vowels. We here focus only on participants' responses to the vowel duration continua, in which the duration of the vowel preceding the coda stop increased across nine equally spaced steps (7992 items). Additional details and results of the perception task are available in [14, 15]

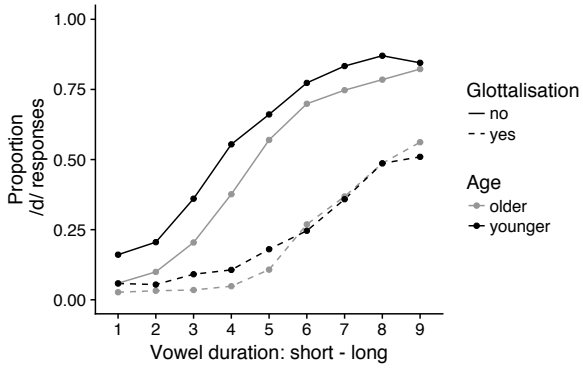

Figure 1: *Proportion of items perceived as voiced in glottalised and non-glottalised conditions by older and younger listeners.*

Figure 1 illustrates the overall response pattern to the vowel duration manipulation stimuli across all vowels. It can be seen that as vowel duration increases, so too does the proportion of listener responses for a voiced coda. In the glottalised condition, shown by dashed lines, this effect is reduced relative to the non-glottalised condition, shown by the solid lines, suggesting that the presence of glottalisation strengthens listeners' perception of coda voicelessness. As discussed above, the effect of glottalisation is visible in both of the age groups. Though it may appear that the effect is stronger in the younger group, as evidenced by the greater difference between the conditions, this is due to the older listeners producing less voiced responses for low vowels in the non-glottalised condition (see [15] for details).

2.4. Calculating individual links between production and perception

In order to examine possible links between participants' production and perception repertoires, we calculated a single production score and a single perception score for each participant. The production score was determined by dividing

the number of voiceless coda target words produced with glottalisation by the overall number of voiceless coda target words that this speaker produced thereby identifying how frequently each participant utilised glottalisation. A score of 1 would demonstrate that a participant employed glottalisation categorically in the voiceless coda context, whereas a score of 0 would show that a participant did not use glottalisation at all.

To identify how heavily each participant weighted glottalisation in perception, we fitted a separate simple regression model (GLM) to each individual participant's responses to the vowel duration stimuli of the perception task, with the predictors *Vowel duration* (nine steps from short to long) and *Condition* (non-glottalised vs glottalised). The resulting models returned coefficients for *Vowel duration* and *Condition*. The coefficient for *Condition* represents the size of the effect glottalisation had on a participant's response, after the effect of *Vowel duration* had been accounted for. For each participant, the coefficient for *Condition* was thus taken to represent a score for how heavily the participant weighted glottalised in perception, with high scores representing increased weighting of glottalisation as a cue to coda voicelessness.

Following [18] and [21], we then plotted each individual participant's production score (on the x-axis) against their perception score (on the y-axis), to visualise possible links between the two modalities (seen in Figure 2 below). If individual speaker/listeners' production and perception repertoires are closely linked, we would expect to see values increasing diagonally from the lower left to the upper right of the plot, such that participants who use glottalisation rarely in production would also exhibit low perceptual weightings for glottalisation, and those who glottalise frequently in production would exhibit high perceptual scores, showing that they are more sensitive to this cue in perception. A linear regression was then performed on the production-perception data for each age group separately to identify potential correlations between the two modalities.

3. Results

Figure 2 illustrates the relationship between each participant's production and perception scores. As differences at the group level have been found between the older and younger speakers in production [4, 5, 13], we examined the individual production-perception scores separately for each age group. Note that scores for three outliers were removed from this plot and the following regression analyses: two of these were from the older age group; the other was from the younger age group. All three outliers had extremely high perception scores – these participants selected voiceless responses for all items in the glottalised condition. The removal of these outliers did not change the overall results. Figure 2 shows, consistent with our previous findings, that the younger speakers produce glottalisation more frequently than the older speakers [4, 5, 13]. The number of data points clustered towards the right-hand side of the younger speakers' plot illustrates that many of these speakers used glottalisation at near categorical levels when producing items with coda /t/ (four of the participants indeed used glottalisation categorically as shown by production scores of 1). Some younger speakers produced glottalisation less frequently, though the majority have production scores above 0.5 (i.e. they produced glottalisation more than 50% of the time). In the older group, there are fewer data points on the right side of the plot, with the majority of production scores clustered between 0.25 and 0.75. This shows that the older speakers also used glottalisation in production, but not as frequently as the younger speakers, and, with the exception of a few speakers, not near categorical level. Figure 2 also illustrates that in addition to producing more glottalisation than the older group, the younger group is also responsible for the highest perception scores. Nevertheless, the younger group as a whole does not appear to be more overly sensitive to glottalisation in perception.

Perhaps more importantly, Figure 2 shows no clear link between individuals' perception and production scores. There are some participants in the younger group with high production scores and high perception scores, visible in the upper right of the plot, who could be considered as demonstrating an alignment of production and perception repertoires. However, there are also a number of participants with high production scores whose perception scores are quite low, visible in the lower right of the plot. These speakers appear to use glottalisation at near categorical levels in production, but nevertheless appear to have very low sensitivity to this cue in perception. Similarly, there are a number of participants (particularly in the older group) who have relatively low production scores (<0.5), but nevertheless show high sensitivity to glottalisation as a perceptual cue. Simple linear regressions performed on the data from each of the age groups confirm that there is no significant correlation between participants' production and perception scores within either of the groups (older: $R^2 = 0.003$; $p = 0.77$; younger: $R^2 = 0.006$; $p = 0.62$).[1]

4. Discussion

We predicted that if individuals' production and perception repertoires were linked we would see a correlation between their perception and production scores, such that individuals who utilised glottalisation heavily in production would also utilise this feature heavily in perception. We found no significant correlation between the two modalities in either age group. In addition, we were able to identify individuals with an apparent mismatch between production and perception, such that one was high where the other was low. Thus, the results presented here do not support a direct correspondence between production and perception at the level of the individual with regard to glottalisation. Although some recent studies have successfully identified individual production-perception links [18, 20, 21, 22], others have failed to find evidence of a correspondence between the two modalities, in accordance with this study [23, 24, 25, 26, 27, 28]. We note that the method employed here analysed participants' (final) lexical decisions.

[1] In order to be confident that the lack of a strong correlation was not due to the method used to calculate the perception score we tested a number of alternative methods: for each participant we calculated the increase in the proportion of voiceless responses in the glottalised compared to the non-glottalised condition; for each participant we identified the 50% crossover point from voiceless to voiced responses in each condition using sigmoid curves fitted to the perception data, then subtracted the crossover point in the glottalised condition from that in the non-glottalised condition; we also extracted the intercept for each participant from a GLMM on the perception data with participant included as a random factor, which we included as a factor in a cascading model to identify predictors of glottalisation in production. In all cases the results proved similar to those presented in this analysis.

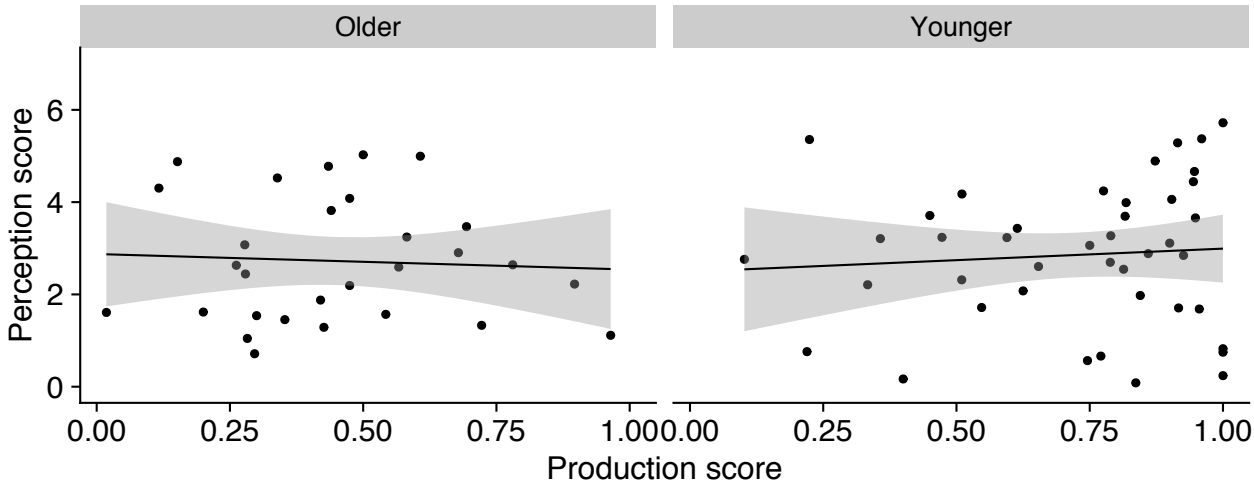

Figure 2: *Scatterplot illustrating relationship between production and perception of glottalisation for each individual speaker/listener. Left panel shows older age group; right panel shows younger age group. Solid lines represent linear regression between production and perception scores within each age group. Shaded areas represent 95% confidence intervals.*

[20] suggest that tracking an individual's perception as the relevant cue unfolds, for example using techniques such as eye-tracking, may provide a better understanding of how efficient listeners are in their perception.

We identified some participants who were perceptually sensitive to glottalisation despite having low production scores. According to [18], it is to be expected that some individuals will be sensitive to a feature in perception despite not using (at least not to the same extent) the feature in production. Perception needs to be flexible so that listeners can adapt to and make sense of productions that differ from their own (e.g. foreign-accented speech), whereas this same level of flexibility is not necessary for production. Hence, it should not be surprising that we found participants who had low production scores for glottalisation but were nevertheless sensitive to glottalisation in perception. We may also expect listeners in the unnatural environment of a perceptual task to make use of any cues that are perceptually available in the signal [18], regardless of whether they produce these themselves, particularly if these are present in the speech of some members of their community [31, 32].

On the other hand, we also found a number of individuals whose behaviour appeared to be what [18] suggests should be exceptional cases. These participants glottalised consistently in production but appeared to be rather insensitive to glottalisation as a perceptual cue. Such individuals in particular are problematic for the idea of an alignment of production-perception repertoires. One possible explanation for the exceptional cases may be that some participants found the demands of the task to be quite onerous. We note that the perception task was conducted after the production task, and the duration of the entire session was quite long. Thus, it may be that fatigue played a role in some participants' responses, though why this would apply to these participants but not be the case for the others remains to be explained.

Despite not finding an overall effect for a link between individuals' production and perception, we did find that a number of participants (particularly in the younger group) demonstrated production and perception repertoires that were aligned. These participants produced high levels of glottalisation and were highly sensitive to glottalisation in perception. Therefore, they appear to be innovative with regard to both production and perception of glottalisation. It has been suggested that, rather than production and perception being aligned throughout all members of a community, the strength of the link between production and perception may vary and may only be aligned in a subset of particular individuals [24, 33]. It is perhaps this aligning subset who may drive a change, particularly in its early stages. The data examined here suggest that some individuals are innovative and show a strong production-perception link, whereas others show a mismatch between production and perception. It may be the case that the innovative individuals with aligned repertoires are those responsible for the spread of glottalisation, whereas the other, non-aligned individuals exhibit the instability of an ongoing shift from one stable alignment of repertoires to another [34].

To conclude, this study did not find strong evidence to support individual links between the production and perception of glottalisation in AusE speaker/listeners. Although we identified some individuals with aligned production-perception repertoires, we found no consistent correlation between these modalities in either age group. It is possible that an examination of participants' real time perception could provide a more fine-grained understanding of listener sensitivity to glottalisation; therefore, future work utilising eye-tracking methods is necessary to further our understanding of the relation between production and perception at the level of the individual.

5. Acknowledgements

This work was supported by an Australian Government Research Training Program scholarship to the first author and by Australian Research Council grant FT180100462 to the second author. We are grateful to Titia Benders and Jonathan Harrington for their insightful comments and suggestions.

6. References

[1] P. Trudgill, *Dialects in contact*. Oxford: Blackwell, 1986.

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

[3] L. Tollfree, "Variation and change in Australian consonants," in *Varieties of English around the world: English in Australia*, D. B. Blair and P. Collins, Eds. Amsterdam: Benjamins, pp. 17–44, 2001.

[4] J. Penney, F. Cox, K. Miles, and S. Palethorpe, "Glottalisation as a cue to coda consonant voicing in Australian English," *Journal of Phonetics*, vol. 66, pp. 161–184, 2018.

[5] J. Penney, F. Cox, and A. Szakay, "Glottalisation of word-final stops in Australian English unstressed syllables," *Journal of the International Phonetic Association,* 2019 (online).

[6] O. B. Gordeeva and J. M. Scobbie, "A phonetically versatile contrast: Pulmonic and glottalic voicelessness in Scottish English obstruents and voice quality," *Journal of the International Phonetic Association*, vol. 43, pp. 249–271, 2013.

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

[8] J. Pierrehumbert, "Prosodic effects on glottal allophones", in *Vocal fold physiology: Voice quality control*, O. Fujimura and M. Hirano, Eds. San Diego: Singular, pp. 39–60, 1995.

[9] P. J. Roach, "Glottalization of English /p/, /t/, /k/ and /ʧ/ – a re-examination", Journal of the International Phonetic Association, vol. 3, pp.10–21, 1973.

[10] S. Seyfarth and M. Garellek, "Coda glottalization in American English," in *Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow, Scotland*, 2015.

[11] P. Eckert, "Adolescent social structure and the spread of linguistic change," *Language in Society*, vol. 17, pp. 183–207, 1988.

[12] W. Labov, *Principles of linguistic change* (Vol. 2): *Social factors*. Oxford: Blackwell, 2001.

[13] J. Penney, *The production and perception of coda glottalisation in Australian English*. Unpublished doctoral dissertation, Macquarie University, 2019.

[14] J. Penney, F. Cox, and A. Szakay, "Weighting of coda voicing cues: Glottalisation and vowel duration," in *INTERSPEECH 2018 – 19th Annual Conference of the International Speech Communication Association, September 2-6, Hyderabad, India, Proceedings*, 2018, pp. 1422–1426.

[15] J. Penney, F. Cox, and A. Szakay, "Effects of glottalisation, preceding vowel duration, and coda closure duration on the perception of coda stop voicing," *Phonetica*, to appear.

[16] J. J. Ohala, "The phonetics of sound change," in *Historical linguistics: Problems and perspectives*. C. Jones, Ed. London: Longman, pp. 237–278, 1993.

[17] J. B. Pierrehumbert, "Exemplar dynamics: Word frequency, lenition and contrast," in *Frequency effects and the emergence of linguistic structure*, J. Bybee and P. Hopper, Eds. Amsterdam: Benjamins, pp. 137–57, 2001.

[18] P. S. Beddor, "The relation between language users' perception and production repertoires," in *Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow, Scotland*, 2015.

[19] A. C. L. Yu, "Individual differences in socio-cognitive processing and the actuation of sound change," in *Origins of sound change: Approaches to phonologization*, A. C. L. Yu, Ed., Oxford: Oxford University Press, pp. 201–227, 2013.

[20] P. S. Beddor, A. W. Coetzee, W. Styler, K. B. McGowan, and J. B. Boland, "The timecourse of individuals' perception of coarticulatory information is linked to their production: Implications for sound change," *Language*, vol. 94, pp. 931–968. 2018.

[21] A. W. Coetzee, P. S. Beddor, K. Shedden, W. Styler, and D Wissing, "Plosive voicing in Afrikaans: Differential cue weighting and tonogenesis," *Journal of Phonetics*, vol. 66, pp. 185–216, 2018.

[22] A. C. L Yu, "On the nature of the perception-production link: Individual variability in English sibilant-vowel coarticulation," *Laboratory Phonology*, vol. 10, pp. 1–29, 2019.

[23] M. Grosvald, "Interspeaker variation in the extent and perception of long-distance vowel-to-vowel coarticulation," *Journal of Phonetics*, vol. 37, pp. 173–88, 2009.

[24] M. Grosvald and D. Corina, "The production and perception of sub-phonemic vowel contrasts and the role of the listener in sound change," in *The initiation of sound change: Production, perception, and social factors,* M.-J. Solé and D. Recasens, Eds. Amsterdam: Benjamins, pp. 77–100, 2012.

[25] R. Kataoka, *Phonetic and cognitive bases of sound change*. Unpublished doctoral dissertation, University of California, Berkeley, 2011.

[26] J. Schertz, T. Cho, A. Lotto and N. Warner, "Individual differences in phonetic cue use in production and perception of a non-native sound contrast," *Journal of Phonetics*, vol. 52, pp. 183–204, 2015.

[27] A. A. Shultz, A. L. Francis, and F. Llanos, "Differential cue weighting in perception and production of consonant voicing," *Journal of the Acoustical Society of America,* vol. 132, pp. EL95–EL101, 2012.

[28] M. Stevens and U. Reubold, "Pre-aspiration, quantity, and sound change," *Laboratory Phonology*, vol. 5, pp. 455–488, 2014.

[29] M. Garellek, "Perception of glottalization and phrase-final creak," *Journal of the Acoustical Society of America*, vol. 137, pp. 822–831, 2015.

[30] M. Garellek and P. Keating, *Phrase-final creak: Articulation, acoustics, and distribution*. Paper presented at the 89th Annual Meeting of the Linguistic Society of America, Portland, 2015.

[31] J. Hay, P. Warren, and K. Drager, "Factors influencing speech perception in the context of a merger-in-progress," *Journal of Phonetics*, vol. 34, pp. 458–484, 2006.

[32] S. Jannedy and M. Weirich, "Sound change in an urban setting: Category instability of the palatal fricative in Berlin," *Laboratory Phonology*, vol. 5, pp. 91–122, 2014.

[33] M. Stevens and J. Harrington, "The individual and the actuation of sound change," *Loquens*, vol. 1, e003, 2014.

[34] F. Kleber, J. Harrington, and U. Reubold, "The relationship between the perception and production of coarticulation during a sound change in progress," *Language and Speech*, vol. 55, 383–405, 2012.