THE PRODUCTION AND PERCEPTION OF CODA GLOTTALISATION IN AUSTRALIAN ENGLISH

Joshua James Penney

Bachelor of Arts (Linguistics & German Studies)

Master of Research (Linguistics)

A thesis submitted in fulfilment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Department of Linguistics

Centre for Language Sciences

Faculty of Human Sciences

Macquarie University

Sydney, Australia

Submitted: 14th August, 2019

Table of contents

Thesis abstract

Previous research suggests that glottalisation is a cue to voiceless coda stops in Australian English and may be a recent change to the variety. In this thesis, we examine glottalisation in production and perception by Australian English speaker/listeners. We first consider coda glottalisation in the production of unstressed syllables, with the results confirming previous findings that glottalisation signals coda voicelessness and is more frequent in younger rather than older speakers, supporting suggestions of recent change. We then examine how perception of coda stop voicing is affected by glottalisation, finding that glottalisation facilitates increased perception of coda voicelessness, and that older and younger listeners perceive glottalisation similarly, despite differences in production. In a following study we show that listeners are perceptually sensitive to glottalisation, and do not merely perceive shorter modally voiced vowels when glottalisation is present. We examine glottalisation in production in different phrase positions to disentangle its effects from those of phrase final creaky voice, demonstrating that the effects of glottalisation do indeed occur independently of creaky voice. We also find that glottalisation is more frequent in pre-consonantal environments rather than pre-vocalic environments, and that younger speakers employ glottalisation more often than older speakers, particularly in pre-vocalic environments, indicating a possible progression of change. Finally, we examine links between production and perception at the individual level. Although we find no consistent pattern between perception and production of glottalisation, we observe that a subset of individuals are progressive in both modalities and may be the drivers of change. This thesis thus contributes to our understanding of glottalisation, its use in production and perception, and its spread in

Australian English. More generally, it advances our understanding of the cues that contribute to the phonological coda stop voicing contrast in English.

Declaration and statement of authorship

I hereby declare that the research presented in this thesis is original and has been submitted exclusively to Macquarie University for the degree of Doctor of Philosophy. This work has not been submitted for a degree of any kind at any other institution. All of the research contained in this thesis was undertaken while I was formally enrolled as a higher degree research candidate at Macquarie University. I have made every effort to indicate sources of information and to acknowledge where the work of others has been used.

The research in this thesis was conducted with ethics approval from the Macquarie University Human Sciences Research Ethics Subcommittee (Reference number: 5201819002273 (previously 5201600256)).

Some of the material contained in this thesis has been published, accepted for publication, submitted for publication, or is currently in preparation for submission for publication. These items are listed below:

Chapter 2 has been published as:

Penney, **J**., Cox, F., & Szakay, A. (2019 online). Glottalisation of word-final stops in Australian English unstressed syllables. *Journal of the International Phonetic Association*. https://doi.org/10.1017/S0025100319000045.

Chapter 3 has been revised based on reviewers' comments and resubmitted as:

Penney, **J**., Cox, F., & Szakay, A. (revision submitted). Effects of glottalisation, preceding vowel duration, and coda closure duration on the perception of coda stop voicing. *Journal of Phonetics*.

Chapter 4 has been published as:

Penney, **J**., Cox, F., & Szakay, A. (2019). Perception of coda voicing: Glottalisation, vowel duration, and silence. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS 2019), Melbourne, Australia* (pp. 1863–1867).

Chapter 5 is being prepared for submission to a specialist phonetics journal (such as *Journal of Phonetics* or *Phonetica*) as:

Penney, **J**., Cox, F., & Szakay, A. (in preparation). *Production of coda glottalisation in Australian English*.

Chapter 6 is being prepared for submission to Interspeech 2020 as:

Penney, **J**., Cox, F., & Szakay, A. (in preparation). *Links between production and perception of glottalisation in individual Australian English speaker/listeners*.

I state that I took leadership in conducting this research and was primarily responsible for all parts of this thesis, including the conception, development and design of all experiments, the bulk of data collection (see below), the preparation and annotation of data, stimuli creation, statistical analyses, and the writing of the thesis. My supervisors/co-authors, Prof. Felicity Cox and Dr Anita Szakay, provided advice on the research protocol, analyses, the interpretation of results, and the written component of the thesis. I additionally received detailed feedback from journal editors and reviewers on draft versions of some chapters, which has led to their improvement.

Some of the data for Chapter 2 were provided by Prof. Felicity Cox, and the remaining data for Chapter 2 were extracted from the publicly available AusTalk corpus. Additional annotation of the data for Chapter 2 was provided by Dr Kelly Miles. Additional annotation of the data for Chapter 5 was conducted by Dr Kimiko Tsukada, Linda Buckley, and Sakib Islam.

Joshua Penney

14.08.2019

Additional scientific contributions not included in thesis

Penney, **J**., Cox, F., Miles, K, & Palethorpe, S. (2018). Glottalisation as a cue to coda consonant voicing in Australian English. *Journal of Phonetics, 66*, 161–184. https://doi.org/10.1016/j.wocn.2017.10.001.

Penney, **J**., Cox, F., Szakay, A. (2018). Weighting of coda voicing cues: Glottalisation and vowel duration. In B. Yegnanarayana, C.C. Sekhar, S. Narayanan, S. Umesh, S.R.M. Prasanna, H.A. Murthy, P. Rao, P. Alku & P.K. Ghosh (Eds.), *Proceedings of the 19th Annual Conference of the International Speech Communication Association, INTERSPEECH 2018, Hyderabad, India* (pp. 1422–1426). https://doi.org/10.21437/Interspeech.2018-1677.

Penney, **J**., Cox, F., Szakay, A. (2018). Effects of glottalisation on reaction time in identifying coda voicing. In J. Epps, J. Wolfe, J. Smith, & C. Jones (Eds.), *Proceedings of the 17th Australiasian International Conference on Speech Science and Technology, SST2018, Sydney, Australia* (pp. 133–136).

Acknowledgements

The journey that has led to this thesis has been long and, at times, arduous. For much of the way one person in particular has been a bastion of support and inspiration. I would like to express my deepest gratitude to my primary supervisor (meine Doktormutter!) Felicity Cox, for her patience (even if the face of my naivety and ignorance), for her advice and direction, for challenging me, for her tireless readiness to consider analyses, read (and re-read) drafts, and answer endless questions, and most of all for sowing the seeds that have culminated in this work, and indeed will continue to grow in my future research endeavours.

I am also extremely grateful for the support of my associate supervisor, Anita Szakay, who I would like to thank for her advice and answers in response to my many (quick) questions about experimental design, code, statistics, and interpreting data, and for paying for coffee! Thank you even more for your consistent encouragement and for reminding me to have fun.

I have benefited enormously from the expertise of other members of the MQ phonetics lab, who have always been willing to critique and improve my ideas, analyses, interpretations, and presentations. I would especially like to thank Titia Benders, Mike Proctor, and Ivan Yuen for providing comments and feedback on stimuli, design, statistics, and for generally making time for my persistent enquiries. In addition, I thank Linda Buckley, Rebecca Holt, Shuting Liu, Louise Ratko, and Tünde Szalay for sharing their voices and/or code, helping with piloting experiments, reading drafts, and being generally willing to help out. Thank you as well to Peter Humburg for the patient statistical advice. I also thank Kelly Miles, Chi Lo, and Ping

Tang, who all completed the journey before me and were always willing to provide me with advice, tips, and templates.

I had the very good fortune to spend half a year of my candidature at the Institute of Phonetics and Speech Processing at LMU Munich and I am very grateful to Jonathan Harrington for hosting me, providing feedback on my research, and integrating me into the excellent research environment of the IPS. I am grateful for the advice I received from many IPS members; in particular I thank Eva Reinisch for her thoughtful comments, which substantially improved the chapters on perception, and for sharing her office with me. I also thank Chris Carignan, Felicitas Kleber, Ulrich Reubold, and Raphael Winkelmann, who all provided assistance of one kind or another while I was in Munich. I also gratefully acknowledge the Endeavour Research Fellowship that funded my visit.

To my family, Anik, Luka, and Joseph, thank you for your understanding, support, and patience – particularly when I had to work at night/on weekends/on holidays. Thank you too for forcing me away from my computer at other times and reminding me that there’s a whole world out there outside of phonetics to experience and enjoy.

Chapter 1: Introduction

1.1 Glottalisation

Glottalisation is a term that may be used to refer to a number of different, though related, phonetic phenomena. For example, glottalisation may refer to the inclusion of a glottal stop and/or a period of laryngealised (irregular)[1] phonation at the beginning of a word initial vowel. This occurs in many languages, including English and German (Dilley, Shattuck-Hufnagel, & Ostendorf, 1996; Kohler, 1994; Pierrehumbert & Talkin, 1992; Umeda, 1978), and may function as a form of prosodic strengthening (Garellek, 2014). Glottal stops may also be inserted as a strategy (among other possible strategies such as the insertion of a rhotic or other approximant) to resolve vowel adjacency (hiatus) (Allerton, 2000; Cox, Palethorpe, Buckley, & Bentink, 2014; Uffman, 2007; Yuen, Cox, & Demuth, 2018). In addition, ejective consonants, such as [t’], which are produced with a glottalic egressive as opposed to a pulmonic egressive airstream – increasingly common in Scottish and other northern varieties of British English (BrE) – may sometimes also be included as a form of glottalisation (e.g. Gordeeva & Scobbie, 2013; Henton, Ladefoged, & Maddieson, 1992; Maddieson, 2013; McCarthy & Stuart-Smith, 2013; Simpson, 2014). Phrase final creaky voice is another phenomenon that may be referred to as glottalisation (also sometimes labelled vocal/glottal fry). Phrase final creaky voice is exemplified by laryngealised voicing at the end of a prosodic phrase or utterance, and can extend leftward over multiple (voiced) segments (Garellek, 2015; Garellek & Keating, 2015; Henton & Bladon, 1988; Kreiman, 1982; Redi & Shattuck-Hufnagel, 2001). Finally,

[1] Laryngealisation refers to phonation produced with glottal and laryngeal constriction that perturbs vocal fold vibration resulting in irregular glottal pulses (Blankenship, 2002; Edmondson & Esling, 2006; Esling, Fraser, & Harris, 2005), though we acknowledge that some languages (e.g. Mazatec, Mpi) may display laryngealisation without irregularity (Blankenship, 2002).

glottalisation may refer to the addition of a glottal stop and/or laryngealised voicing to an oral stop in coda position, as occurs in many varieties of English (Docherty & Foulkes, 1999; Esling, Fraser, & Harris, 2005; Gimson, 1962; Huffman, 2005; Pierrehumbert, 1994, 1995; Roach, 1973; Seyfarth & Garellek, 2015), as well as other languages such as Thai (Esling, Fraser, & Harris, 2005; Harris, 2005), Cantonese (Iwata, Sawashima, & Hirose, 1981), Taiwanese, Hakka (Edmondson, Chang, Hsieh, & Huang, 2011), and Siona (Ladefoged & Maddieson, 1996). Coda glottalisation is sometimes referred to as pre-glottalisation or glottal reinforcement, in which the additional glottal stop/gesture precedes[2] and/or reinforces, or strengthens, the oral stop (Docherty & Foulkes, 1999; Esling et al., 2005; Keating, Wymark, & Sharif, 2019; Roach, 1973; Seyfarth & Garellek, 2015), and (more restrictedly) as t-glottalisation, although this suggests that the process only applies to voiceless alveolar stops (Eddington & Channer, 2010; Garellek, 2015; Tollfree, 2001; Wells, 1982). In some cases, generally known as glottal replacement (also sometimes glottalling or t-glottalling), the oral stop gesture may be 'lost', or obscured by the glottal gesture, and the coda consonant is realised solely, or predominately, by the glottal stop/constriction (Andrésen, 1968; Docherty & Foulkes, 1999b; Fabricius, 2002; Roach, 1973, 2004; Schleef, 2013; Seyfarth & Garellek, 2015; Wells, 1982).

This thesis is concerned with the production and perception of coda glottalisation. Therefore, in what follows we will not examine or discuss at length other types of glottalisation such as word initial glottalisation, ejective consonants, or

[2] While the term pre-glottalisation implies that the glottal gesture precedes the oral stop gesture, in the case of glottal reinforcement different researchers refer to either simultaneous constriction of glottal and oral gestures (e.g. [ʔ͡t]), the glottal gesture preceding the oral gesture or release (i.e. pre-glottalisation; e.g. [ˀt]), or the glottal gesture occurring after the oral gesture or release (i.e. post-glottalisation; e.g. [tˀ]) (Docherty & Foulkes, 1999; Esling et al., 2005; Howe & Pulleyblank, 2001; Keating et al., 2019; Seyfarth & Garellek, 2015).

glottal stop insertion in vowel hiatus environments (though note that similarities between coda glottalisation and phrase final creaky voice will be discussed in 1.3 below and Chapter 5 will include an analysis of phrase final creaky voice). Rather, the type of glottalisation examined in this thesis will be limited to glottalisation employed in conjunction with an oral stop in coda position, and any usage of the term glottalisation will be in reference to coda glottalisation, unless otherwise stated.

1.2 Articulation and acoustic effects of glottalisation

Coda glottalisation was traditionally considered to involve the co-production of a canonical glottal stop and an oral coda stop; that is, it was thought that glottalisation involved the complete sustained adduction of the vocal folds, whereby the airflow was blocked by the closed glottis for the period of constriction (analogous to an oral stop closure) (Docherty & Foulkes, 1999, 2005; Gimson, 1962). However, most early studies were based on auditory analysis. Later studies utilising instrumental rather than impressionistic methods have revealed that, while some realisations may involve a complete glottal stop, in many cases of glottalisation there is no complete adduction of the vocal folds. Rather, the impressionistically perceived glottal stop is often produced by a period of laryngealised voicing (sometimes referred to as creaky voice) due to increased glottal constriction but not complete adduction (Docherty & Foulkes, 1999, 2005; Garellek, 2015; Ladefoged & Maddieson, 1996; Pierrehumbert & Talkin, 1992). In such cases, there may be no silent closure period as would be associated with a canonical glottal or oral stop (Docherty & Foulkes, 1999, 2005; Gimson, 1962). Even if complete glottal constriction occurs, a period of laryngealised voicing of the preceding vowel is common (Chong & Garellek, 2018; Garellek, 2015; Seyfarth & Garellek, 2015).

The mechanism for producing various types of glottalisation may involve the use of other articulators, such as the ventricular (false) folds and aryepiglottic folds in addition to (either complete or incomplete) constriction of the vocal folds (Edmondson & Esling, 2006; Esling et al., 2005; Esling, Moisik, Benner, & Crevier-Buchmann, 2019; Moisik & Esling, 2011; Moisik, Esling, Crevier-Buchman, Amelot, & Halimi, 2015). Reports have long been made of glottal stops involving constriction of the ventricular folds (e.g. Catford, 1977; Fujimura & Sawashima, 1971). Esling et al. (2005) identify a continuum of possible glottal stops ranging from weakly constricted, through moderate, to strongly constricted, all of which involve some level of ventricular incursion (i.e. constriction of the ventricular folds which covers and dampens the vibration of the constricted vocal folds) (see also Esling, Zeroual, and Crevier-Buchman, 2007). In addition, the aryepiglottic sphincter, comprising the arytenoid cartilages together with the aryepiglottic folds, may be compressed (along with possible pharyngeal constriction), as occurs with strong glottal (i.e. epiglottal) stops. There is also evidence that ventricular incursion may occur during the laryngealised voicing associated with glottalisation (Esling et al., 2019; Moisik et al., 2015). In some cases it may even be possible that the vocal folds remain abducted and that the acoustic effects of glottalisation are produced solely by ventricular incursion and/or compression of the aryepiglottic sphincter (Edmonson & Esling, 2006). On the other hand, Garellek (2013) found that ventricular incursion occurred commonly in the production of glottal stops, though this was not always present in his data. This may suggest that ventricular incursion is a frequent but perhaps not necessary component of glottalisation, and may hint at the variability with which glottalisation can be achieved. In this thesis we will generally refer to glottal constriction as the articulatory basis for glottalisation, though we acknowledge that a precise articulatory description in many cases likely involves a more complex laryngeal configuration that may not be exclusively due to constriction of the vocal folds.

Auditorily, glottalisation produces the percept of a glottal stop (as discussed above) or creakiness of the vowel preceding a coda stop (Docherty & Foulkes, 1999, 2005; Dilley et al., 1996).[3] From an acoustic perspective, glottalisation is characterised by irregularity in F0 and amplitude, visible most simply as irregularly spaced striations in a wide-band spectrogram and as irregular peaks in an acoustic waveform (Batliner, Burger, Johne, & Kießling, 1994; Docherty & Foulkes, 1999; 2005; Keating, Garellek, & Kreiman, 2015; Pierrehumbert, 1994; Huffman, 2005; Redi & Shattuck-Hufnagel, 2001). It often results in a drop in F0 and amplitude (Batliner et al., 1994; Dilley et al., 1996; Fischer-Jørgensen, 1989; Pierrehumbert & Frisch, 1997; Redi & Shattuck-Hufnagel, 2001; though see Penney, Cox, Miles, & Palethorpe, 2018, for an example of glottalisation co-occurring with a rise in F0). In addition, glottalisation may result in increased energy in the higher spectral frequencies, due to the more rapid closure of the vocal folds, and hence a lower spectral tilt (compared to modal voice) (Fischer-Jørgensen, 1989; Garellek, 2013; Garellek & Seyfarth, 2016; Keating et al., 2015; Pierrehumbert & Frisch, 1997).

1.3 Glottalisation and phrase final creaky voice

As mentioned above, phrase final creaky voice is sometimes referred to as glottalisation in the literature, though we do not adopt this terminology here. Phrase final creaky voice serves as a boundary marker in English (Garellek, 2015; Garellek & Keating, 2015; Henton & Bladon, 1988; Kreiman, 1982) and other languages such

[3] Note that this may also occur in other sonorants preceding coda stops, e.g. nasals, but these are not examined in this thesis.

as Finnish (Ogden, 2001) and Swedish (Carlson, Hirschberg, & Swerts, 2005). Coda glottalisation and phrase final creaky voice may be produced by the same articulatory process, namely glottal (laryngeal) constriction as described in 1.2 above (Garellek & Seyfarth, 2016; Esling et al., 2005; Moisik et al., 2015), though Slifka (2006) and Keating et al. (2015) discuss a type of creaky voice that is not due to glottal constriction, but rather glottal spreading at the end of an utterance where subglottal pressure is low and modal voicing cannot be maintained. There are a number of different possible types of creaky voice, but both phrase final creaky voice and the laryngealised voicing associated with glottalisation are examples of 'prototypical' creaky voice, demonstrating irregular phonation (and a concomitant increase in spectral noise resulting in lower harmonic-to-noise ratios) as well as reduced F0 and spectral tilt (e.g. Garellek & Seyfarth, 2016; Keating et al., 2015). Laryngealised voicing due to coda glottalisation only applies to a single segment (i.e. the vowel preceding the coda stop), whereas phrase final creaky voice may extend across multiple voiced segments within an utterance. Although both glottalisation and phrase final creaky voice may be produced similarly, it has been shown that listeners can differentiate between coda glottalisation and phrase final creaky voice (Garellek, 2015; see 1.8 below). Differences in production have also been found in the timing of the two phenomena, with glottalisation demonstrating a greater change towards the end of the affected segment, consistent with it being localised to a coda stop (Garellek & Seyfarth, 2016).

1.4 Glottalisation in English

Andrésen (1968) identified the earliest known references to glottalisation in BrE in the writings of Bell (1860, 1867), and subsequently Ellis (1875, 1889) and Sweet

(1877, 1880-1881). According to these early references, a glottal stop was produced simultaneously with the oral gesture for /t/ in some Western Scottish dialects.[4] From there glottalisation would appear to have spread to the east of Scotland, before moving south to England. The first reference (as identified by Andrésen, 1968) of glottalisation in England was in Sweet (1908), where it was said to occur in both Scottish and Northern English dialects. Shortly thereafter, Jones (1909) identified glottalisation as a feature of Scottish and London speech, and over the intervening years glottalisation became heavily associated with (working class) London (Andrésen, 1968, and references therein; Mott, 2012; Tollfree, 1999; Wells, 1982). Glottalisation has since spread throughout Britain, and today it is common in many dialects, though its realisation and the environments in which it occurs differ between varieties (Docherty & Foulkes, 1999; Fabricius, 2002; Przedlacka, 2002; Gordeeva & Scobbie, 2013; Mathieson, 1999; Mees, 1987; Mees & Collins, 1999; Milroy, Milroy, Hartley, & Walshaw, 1994; Roach, 2004; Tollfree, 1999; Wells, 1982). Some studies have hypothesised about the influence of London and its role in the rapid spread of glottalisation through the United Kingdom, including into the prestige variety of Received Pronunciation and the more general Estuary English (Fabricius, 2002; Przedlacka, 2002; Mees & Collins, 1999; Milroy, Milroy, Hartley, & Walshaw, 1994; Wells, 1982).

Glottalisation also occurs in American English (AmE), although it is not as easy to trace its development in this variety.[5] Andrésen (1968) identified a single

[4] Bell (1860) and Ellis (1875) both list glottalisation as occurring in the West of Scotland; Sweet (1877) specifies this as occurring in Glasgow (which he attributes to Bell), and Ellis (1889) lists glottalised pronunciations as from the Eastern and Western Mid Lowlands of Scotland.

[5] Wells (1982: 261) for example states, "I know of no systematic investigation of Preglottalization and Glottalling in American speech".

reference to glottalisation in the speech of schoolchildren in New York in 1896 (Babbitt, 1896), where it was stated to occur in place of intervocalic /t/. Kahn (1976) and Selkirk (1982) both discuss a glottalised, unreleased allophone of /t/, and Wells (1982) notes that glottalisation is observable in the speech of *some* Americans, and more so in the speech of those from New York, though not to the same extent as in BrE. Since the 1990s studies on (or including) coda glottalisation in AmE have increased (Byrd, 1993, 1994; Eddington & Channer, 2010; Eddington & Savage, 2012; Eddington & Taylor, 2009; Garellek, 2015; Garellek & Seyfarth, 2016; Huffman, 2005; Pierrehumbert, 1994, 1995; Seyfarth & Garellek 2015).

1.5 Glottalisation in Australian English

As mentioned above, glottalisation appears to have emerged in BrE varieties in the late 1800s or early 1900s.[6] This timeline suggests that glottalisation was not a feature of BrE speech at the time of Britain's colonisation of Australia:

> "From the fact that these accents have little to no T glottalling, we can infer that at the time their accents [AusE, New Zealand English, South African English] were essentially formed this development had not yet taken place in Britain." (Wells, 1982: 592)

Many early commentators provided negative assessments of AusE, especially from the late 1800s (see Mitchell, 1945, for a summary and a sample of such comments from the early to mid-1900s). One of the common criticisms of AusE was its similarity with Cockney (London English). Though glottalisation was associated with London speech, and was stigmatised within Britain (Milroy, Milroy, & Hartley, 1994; Milroy, Milroy, Hartley, & Walshaw, 1994), this feature is never mentioned

[6] It is of course likely that glottalisation was present in speech for a time prior to it being recorded in the literature, although Andrésen (1968) lists a number of reasons to believe that it was not widespread even shortly before its appearance in written accounts.

among those that drew the ire of the commentators, suggesting it was not present in AusE in the latter part of the nineteenth century. Indeed, early descriptions of AusE remarked on the absence of glottalisation: Mitchell (1945: 8) states that “the Cockney use of the glottal stop ... is not heard in Australia.” Baker (1965), too, agrees that glottalisation is “practically unknown in Australia”. Tollfree (2001) additionally points out that there is no evidence from popular Australian writing to suggest that glottalisation was a feature of AusE, or at least that there was no conscious awareness of it among speakers. Glottalisation was not considered to be present in AusE as recently as the 1980s. Both Wells (1982) and Trudgill (1986) commented on the lack of glottalisation in AusE compared to London English:

“Australian lacks the T glottalling ... so typical of contemporary London speech.”

(Wells, 1982: 594);

“Australian English does not have pre-glottalization or glottaling of word final /p/, /t/, /k/. This distinguishes it sharply from London English, where glottalization is the norm.”

(Trudgill, 1986: 141).

Only a few years after these comments, however, glottalisation was recorded by Ingram (1989) who identified glottalisation of /t/ among a range of other connected speech processes in the speech of teenagers from Brisbane. Shortly thereafter, Haslerud (1995) included glottalisation as one of the variants in her sociolinguistic study of /t/ realisations in teenagers from Sydney, and Tollfree (2001) identified glottalisation as a ‘reduced’ realisation of /t/ alongside tapped and spirantised productions. Tollfree (2001) also reported glottalisation of /p/ and /k/ in addition to /t/, though these data were not included in her analysis. These studies suggest that glottalisation may be a new feature that has entered AusE relatively recently.[7] Tollfree (2001), however, urges caution, and suggests that glottalisation has perhaps been long present, but only in certain environments where it may be likely to go unnoticed, such as preceding consonants or in word final position, rather than more salient locations such as preceding syllabic laterals or in word medial intervocalic environments.

In order to explore whether there was evidence to support a recent change, Penney et al. (2018) conducted an apparent time analysis and found that younger speakers (18-36 years) glottalised coda stops significantly more frequently than older speakers (56+ years). The results provide support for glottalisation having entered AusE recently, and suggest that younger speakers are leading the change.

While Ingram (1989) and Haslerud (1995) report the glottalisation in their respective studies to be cases of glottal replacement, these were based on auditory analysis only.[8] Tollfree (2001), who included an acoustic examination in her study, reports no single canonical glottal stop in her data, but rather a range of glottalised variants including those with evidence of a stop closure (but no release burst) and laryngealised voicing. Similarly, Penney et al. (2018) identified glottalised stops through acoustic analysis of laryngealised voicing of the preceding vowel, and did not report any canonical glottal stops.

[7] Glottalisation has also been reported in New Zealand English (Docherty, Hay, & Walker, 2006; Holmes, 1995), although, as with AusE, the variety was established prior to the emergence of glottalisation in BrE (Wells, 1982). This may also suggest a recent change in New Zealand English. We are not aware of reports of glottalisation in South African English.

[8] Haslerud (1995) did not include glottally reinforced stops in her glottal category, but rather considered these to be instances of canonical /t/.

1.6 Stop voicing contrast

As has been well demonstrated in the literature, multiple acoustic correlates can signal a phonological contrast (Al-Tamini & Khattab, 2018; Lisker, 1986; Jongman, Wayland, & Wong, 2000; Repp, 1982). In particular, the multiple correlates of the English stop voicing contrast have been well researched. Voice onset time (VOT), the duration of the interval between the onset of phonation and the release of the stop, is generally considered to be the primary cue to stop voicing in onset position: a long VOT cues a voiceless stop, such as /t/, whereas a short or negative VOT cues a voiced stop, such as /d/ (Cho, Whalen, & Docherty, 2019; Lisker & Abramson, 1964, 1967, 1970; Zlatin, 1979). In addition to VOT, a number of other acoustic correlates may signal the contrast: for example, differences in F0 at the onset of the following vowel, in the amplitude of the release burst, or in the transitions of the first formant (F1) (House & Fairbanks, 1953; Lisker, 1975; Kohler, 1982; Ohde, 1984; Stevens & Klatt, 1974, Repp, 1979).

In coda position, where VOT does not (generally) apply, phonetic voicing through the closure phase of a voiced stop may signal the contrast, but voicing is often weak in coda position in English (Davidson, 2016; Docherty, 1992; Westbury & Keating, 1986). A number of additional acoustic cues may signal the voicing contrast, such as the duration of the preceding vowel, the duration of the coda closure, differences in F0 at the offset of the preceding vowel, the onset and slope of the F1 transition, and the intensity of the release burst (Gruenefelder & Pisoni, 1980; House & Fairbanks, 1953; Liberman, Delattre, & Cooper, 1958; Lisker 1957, 1975, 1986; Port, 1979). One of the primary cues to coda stop voicing is the duration of the preceding vowel: a relatively long preceding vowel cues a voiced coda stop, such as /d/, and a relatively short preceding vowel cues a voiceless stop, such as /t/ (Chen, 1970; Denes, 1955; Fowler, 1992; Klatt, 1976; Lisker, 1974, 1986; Luce & Charles-Luce, 1985; Malécot, 1970; Peterson & Lehiste, 1960; Raphael, 1972).

Glottalisation may also function as a cue to the coda stop voicing contrast. In the varieties of English in which glottalisation occurs, it is generally used in association with voiceless, rather than with voiced stops. Glottalisation of voiceless coda stops occurs in multiple varieties of BrE (e.g. Estuary English: Przedlacka, 2002; London English: Tollfree, 1999; Wells, 1982; Received Pronunciation: Roach, 1973, 2004; Scottish English: Gordeeva & Scobbie, 2013; Tynside English: Docherty & Foulkes, 1999, 2005; Watt & Allen, 2003), in AmE (Huffman, 2005; Pierrehumbert, 1994, 1995), in AusE (Ingram, 1989; Haslerud, 1995; Tollfree, 2001) and in NZE (Docherty, Hay, & Walker, 2006; Holmes, 1995). One the other hand, glottalisation of voiced coda stops rarely occurs in these varieties, if at all. As far as we are aware, the only variety of English that has been shown to commonly produce glottalisation in conjunction with voiced stops is African American English, where voiced coda stops may be devoiced and glottalised (Anderson & Nguyen, 2004; Fasold, 1981; Farrington, 2018; Koops & Niedielski, 2009; Redi & Shattuck-Hufnagel, 2001; Wells, 1982).

Glottalisation may be utilised in conjunction with voiceless stops to enhance voicelessness (Pierrehumbert, 1995; Seyfarth & Garellek, 2015). Voicelessness is typically achieved by glottal abduction. However, in contexts in which a (strongly) voiced segment follows, coarticulatory voicing in anticipation of the following segment may occur during the production of the stop, particularly as coda stops are generally preceded by strongly voiced vowels or other sonorants (Chong & Garellek, 2018; Westbury & Keating, 1986). In such cases, adduction rather than abduction of the vocal folds (potentially with constriction of other laryngeal articulators: see 1.2

above) may be employed to abruptly extinguish any voicing and thereby ensure that voicelessness is perceived (Chong & Garellek, 2018; Keyser & Stevens, 2006; Pierrehumbert, 1995; Seyfarth & Garellek, 2015; Tollfree, 2001). This may be particularly necessary in the case of following nasals, where abduction may be disfavoured as a strategy for achieving voicelessness due to the high chance of perceptual confusion between a nasal and a stop closure that may be breathily voiced due to abduction of the vocal folds at the end of the preceding vowel (Ohala & Ohala, 1993; Pierrehumbert, 1995; Tollfree, 2001). Indeed, glottalisation of voiceless stops occurs frequently before following sonorant consonants, particularly nasals, in a number of varieties (e.g. *kite maker*) (Pierrehumbert, 1994, 1995; Roach, 1973; Seyfarth & Garellek, 2015; Tollfree, 1999; Tollfree, 2001; Watt & Allen, 2003). However, glottalisation is also found in environments that would not require such enhancement, such as preceding voiceless obstruents in onset position (e.g. *beet counter*) (Huffman, 2005; Seyfarth & Garellek, 2015). It may be possible that glottalisation initially emerged in pre-sonorant environments as an enhancement effect before spreading to other environments (Tollfree, 2001).

1.7 Glottalisation and its relation to other voicing cues in Australian English

In their apparent time analysis of glottalisation in AusE, Penney et al. (2018) found that words with voiceless coda stops were glottalised over half of the time (55% of items), whereas relatively few words with voiced codas showed evidence of glottalisation (6% of items) indicating that glottalisation is an acoustic correlate of coda stop voicelessness in AusE. Their younger group produced more glottalisation in the voiceless coda context than the older group: younger: 71%, older: 36%.

In addition to finding increased glottalisation in younger speakers compared to older speakers, Penney et al. (2018) also found that younger speakers produced less robust differences in preceding vowel duration across coda stop voicing categories than the older speakers. That is, it appears that younger speakers may utilise the vowel duration cue to coda stop voicing less robustly than older speakers, while at the same time increasing their use of glottalisation. Penney et al. (2018) suggest that these two cues to coda voicing may operate in a trading relationship, whereby a reduction in the efficacy of one cue (in this case preceding vowel duration) results in an increase in the use of another cue (in this case glottalisation). Additional evidence for a trading relationship was found in that both older and younger age groups glottalised least when a high vowel preceded a voiceless coda stop but they used the vowel duration cue most robustly on the high vowels (Penney et al., 2018).

1.8 Perception of glottalisation

Little work has been carried out on the perception of glottalisation in English. In an early study, Hillenbrand and Houde (1996) found that a drop in F0 or in amplitude creates the percept of a glottal stop even without a silent stop closure (see 1.2 above). They also found that an F0 drop produced a stronger effect than a drop in amplitude (Hillenbrand & Houde, 1996). Although based on informal listening and thus not a perceptual study per se, Pierrehumbert and Frisch (1997) also reported that manipulating a drop in F0 was sufficient to cue glottalisation in resynthesised speech. They also found that reducing amplitude and spectral tilt added to the naturalness of the glottalisation (Pierrehumbert & Frisch, 1997).

Additionally, perceptual studies suggest that AmE listeners associate glottalisation with /t/ or with voiceless coda stops more generally. Garellek (2011) conducted a phoneme monitoring task in which listeners identified instances of glottalised /t/ with faster reaction times and increased accuracy than non-glottalised variants, suggesting that listeners associate glottalisation with /t/. In an eye-tracking task, Chong and Garellek (2018) found that listeners were slower to identify words with voiced coda stops when glottalisation was present, but the presence of glottalisation did not have this effect on words with voiceless coda stops. This suggests that listeners associate glottalisation with voiceless coda stops, rather than voiced coda stops. However, their results also showed that glottalisation did not lead to improved perception of words with voiceless coda stops, which an enhancement account would predict.

The laryngealised voicing associated with glottalisation can be considered a type of creaky voice, and as different types of creaky voice are perceptually distinguishable (Gerratt & Kreiman, 2001), listeners may also be able to distinguish between glottalisation and phrase final creaky voice. Garellek (2015) presented AmE listeners with near minimal pairs such as *button/bun* and *atlas*/*Alice* in which /t/ was realised with glottalisation (i.e. *atlas* realised as [æ̰ʔləs]), and tested whether listeners could distinguish these words when they occurred with phrase final creaky voice extending across the entire word (e.g. [æ̰ʔl̰ə̰s]). Although he found that listeners were somewhat less accurate at identifying words when both glottalisation and phrase final creaky voice were present, likely due to the acoustic similarities of the two (see 1.3 above), listeners made relatively few errors and were able to correctly identify items well above chance level (Garellek, 2015). This suggests that listeners may be able to perceptually distinguish between glottalisation and phrase final creaky voice.

1.9 Aims, research questions, and organisation of thesis

The overarching aim of this thesis is to examine coda glottalisation in AusE in the modalities of both production and perception in order to further our understanding of how this phonetic feature is utilised. The previous studies reviewed above suggest that glottalisation is a recent change to AusE being led by younger speakers. In addition, the increase in the use of glottalisation appears to coincide with a reduction of the utility of preceding vowel duration to signal coda stop voicing, raising questions about cue weighting. Furthermore, perceptual studies of AmE suggest that listeners associate glottalisation with coda stop voicelessness.

In light of this, the general research questions (RQs) are as follows:

1. *Is there evidence to support previous findings that glottalisation occurs in conjunction with voiceless rather than voiced coda stops?*

2. *Are there differences between older and younger speakers in their production of glottalisation suggestive of a recent change?*

3. *Is there evidence of a trading relationship between preceding vowel duration and glottalisation? Do speakers exhibit a reduction in the production of other cues to coda stop voicing alongside an increase in glottalisation?*

4. *Are AusE listeners sensitive to glottalisation and do they interpret it as a cue to coda voicelessness? How do listeners weight glottalisation and preceding vowel duration as cues to coda voicing?*

5. *Are the patterns of glottalisation identified in the literature found independently of phrase final creaky voice?*

6. *Is the occurrence of glottalisation conditioned by phonetic environment?*

7. Are there links between an individual's use of glottalisation in production and in perception? Are speakers who use more glottalisation in production more sensitive to glottalisation in perception?

The remaining chapters of this thesis present studies that explore these questions in detail. In Chapters 2 and 5 the focus is glottalisation in production, Chapters 3 and 4 examine the perception of glottalisation, and Chapter 6 explores links between these two modalities at the level of the individual. Each of these chapters is presented in the format of a journal article or conference paper. In addition to our primary focus on glottalisation, in the chapters on speech production (Chapters 2 and 5) we will also describe the presence of coda stop variants other than canonical, released stops. This will provide an overview of the occurence of non-canonical coda stop realisations, and will help to inform our understanding of variability in coda stop production in AusE.

Chapter 2 presents an analysis of coda glottalisation in unstressed syllables. Unstressed syllables represent an environment in which the preceding vowel duration cue to coda voicing may already be reduced. Therefore, it may be possible that glottalisation occurs at higher rates in unstressed syllables than has been reported for stressed syllables (Penney et al., 2018). In addition, in Chapter 2 we examine glottalisation in stops at all three places of articulation to explore possible indicators of the progression of change in the variety. The results confirm previous findings (Penney et al., 2018) that glottalisation occurs in conjunction with voiceless coda stops in AusE (RQ 1), and that younger speakers glottalise more than older speakers (RQ 2). In addition, we find that female speakers produce glottalisation more frequently than male speakers. Furthermore, this study shows that glottalisation occurs for voiceless stops at all places of articulation, suggesting that it is an acoustic correlate of voicelessness in coda stops in general, not only for /t/. Although we do not find evidence of increased glottalisation in unstressed compared to stressed syllables, we nevertheless find high rates of glottalisation paired with a reduced coda voicing-related vowel duration difference (though a small difference in coda voicing-related vowel duration is maintained). Furthermore, we find no significant F0 difference between voiced and voiced coda contexts, which may point to a reduction of secondary cues to coda voicing alongside high rates of glottalisation (RQ 3)

Chapter 3 comprises a perceptual study designed to examine how older and younger listeners' perception of coda stop voicing is affected by the presence of glottalisation. We explore how listeners interpret vowel duration and coda closure duration (both separately and in combination) as cues to coda stop voicing, and how these interact with glottalisation. The results show that listeners utilise preceding vowel duration as a cue to coda stop voicing, but that coda closure duration is a weaker signal to the stop voicing contrast. Furthermore, we find that listeners are sensitive to glottalisation, and its presence results in increased listener perception of coda voicelessness, even when this occurs in conjunction with extended preceding vowel duration, which would otherwise signal a voiced coda stop (RQ 4). Despite the differences between older and younger speakers observed in production, we find that glottalisation increases the perception of coda voicelessness in both age groups, suggesting that both older and younger listeners use glottalisation similarly in perception.

Chapter 4 represents an extension to the perceptual study in Chapter 3. As glottalisation results in a period of laryngealised voicing of the vowel preceding the coda stop, it may be possible that listeners interpret only the modally voiced portion of the vowel as belonging to the vowel, and the glottalisation component as belonging to the coda closure. The increased perception of coda stop voicelessness could thus be due to the perception of a shorter preceding vowel and a longer coda closure and not due to the presence of glottalisation. Therefore, in addition to analysing listeners' perception of preceding vowel duration and glottalisation, this study compares listeners' perception of glottalisation with their perception of an extended silent period in the place of glottalisation. We find that both the presence of glottalisation and the presence of an extended silent period both result in an increase in voiceless percepts, but that listeners weight the two cues differently, with glottalisation providing a stronger cue to coda voicelessness. This indicates that listeners are sensitive to glottalisation in perception and interpret it as a cue to coda voicelessness (RQ 4), supporting the findings in Chapter 3.

Chapter 5 includes an analysis of the production of glottalisation in phrase medial and phrase final positions in older and younger speakers. Most previous acoustic descriptions of glottalisation in AusE have been based on single word utterances. Therefore, it is possible that (some of) the glottalisation observed may be attributed to phrase final creaky voice, which, as described in 1.3 above, has similar acoustic characteristics to glottalisation. The results demonstrate that patterns of glottalisation found in phrase medial position are similar to those previously reported for single word utterances (Penney et al., 2018; Penney et al., 2019 [Chapter 2]), suggesting that glottalisation of voiceless stops is not (solely) due to the influence of phrase final creaky voice (RQ 1 & 5). Furthermore, we find that younger speakers produce glottalisation more than older speakers amid a reduction of other cues to coda stop voicing, consistent with previous results (RQ 2 & 3). Chapter 5 also includes an examination of glottalisation in a range of phonetic environments. The results show that glottalisation occurs most frequently in pre-consonantal contexts for both age groups, but that younger speakers also utilise glottalisation in pre-vocalic environments (RQ 6). This may provide an indication of the progression of change of glottalisation in AusE.

Chapter 6 explores whether there are links between the production and perception of glottalisation at the level of the individual. For each speaker/listener we calculated how frequently glottalisation was used in production, and analysed this against how heavily glottalisation was weighted as a cue to coda stop voicelessness in perception. At the level of the individual we find no consistent pattern in which speaker/listeners who rely heavily on glottalisation in perception also use glottalisation more frequently in production (RQ 7). Nevertheless, we find evidence of a small number of individuals who appear to exhibit an alignment of production and perception repertoires. It may be possible that these individuals who exhibit a strong production/perception link are those who are important for driving the progression of change.

Finally, Chapter 7 provides a summary and general discussion of the studies that comprise this thesis, and how they contribute to our understanding of glottalisation, its use in production and in perception, its spread in AusE, and to our understanding of the cues that contribute to the phonological coda stop voicing contrast in English more generally. In addition, the limitations of the work presented here as well as recommendations for future research are discussed.

1.10 References

Allerton, D. J. (2000). Articulatory inertia vs ‘Systemzwang’: Changes in liaison phenomena in recent British English. *English Studies, 81*, 574–581.

Anderson, B., & Nguyen, J. (2004). *The social stratification of glottalized variants of /d/ among Detroit African American speakers*. Paper presented at New Ways of Analyzing Variation (NWAV) 33, University of Michigan, Ann Arbor.

Andrésen, B. S. (1968). *Pre-glottalization in English Standard Pronunciation*. Oslo: Norwegian Universities Press.

Babbitt, E. H. (1896). The English pronunciation of the lower classes in New York and vicinity. *Dialect Notes, 9*, 457–464.

Baker, S. J. (1965). *The Australian Language*. Sydney: Currawong Press.

Bailey, G. (2002). Real and apparent time. In J. K. Chambers, P. Trudgill, & N. Schilling-Estes (Eds.), *The handbook of language variation and change* (pp. 312–322). Malden: Blackwell.

Batliner, A., Burger, S., Johne, B., & Kießling, A. (1993). “MÜSLI: A classification scheme for laryngealizations. In D. House & P. Touati (Eds.), *Proceedings of European Speech Communication Association Workshop on Prosody, Lund, Sweden* (pp. 176–179).

Bell, A. M. (1860). *The elocutionary manual* (3rd ed.). London: Hamilton & Adams.

Bell, A. M. (1867). *Visible speech*. London: Simpkin & Marshall.

Blankenship, B. (2002). The timing of nonmodal phonation in vowels. *Journal of Phonetics, 30*, 163–191.

Byrd, D. (1993). 54,000 American stops. *UCLA Working Papers in Phonetics, 83*, 97–116.

Byrd, D. (1994). Relations of sex and dialect to reduction. *Speech Communication, 15*, 39–54.

Carlson, R., Hirschberg, J. & Swerts, M. (2005). Cues to upcoming Swedish prosodic boundaries: Subjective judgment studies and acoustic correlates. *Speech Communication*, *46*, 326–333.

Catford, J. C. (1977). *Fundamental problems in phonetics.* Edinburgh: Edinburgh University Press.

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

Cho, T., Whalen, D. H., & Docherty, G. J. (2019). Voice onset time and beyond: Exploring laryngeal contrast in 19 languages. *Journal of Phonetics, 72*, 52–65.

Chong, A., & Garellek, M. (2018). Online perception of glottalized coda stops in American English. *Laboratory Phonology, 9,* 1–24.

Cox, F., Palethorpe, S., Buckley, L., & Bentink, S. (2014). Hiatus resolution and linking 'r' in Australian English. *Journal of the International Phonetic Association, 44,* 155–178.

Davidson, L. (2016). Variability in the implementation of voicing in American English obstruents. *Journal of Phonetics, 54,* 35–50.

Denes, P. (1955). Effect of Duration on the perception of voicing. *Journal of the Acoustical Society of America, 27,* 761–764.

Dilley, L., Shattuck-Hufnagel, S., & Ostendorf, M. (1996). Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics*, *24*, 423-444.

Docherty, G. J. (1992). *The timing of voicing in British English obstruents*. Berlin: Foris.

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

Docherty, G. J., & Foulkes, P. (1999b). Sociophonetic variation in 'glottals' in Newcastle English. In J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, & A. C. Bailey (Eds.), *Proceedings of the 14th International Congress of Phonetic Sciences (ICPhS 99), San Francisco* (pp. 1037–1040).

Docherty, G. J., & Foulkes, P. (2005). Glottal variants of /t/ in the Tyneside variety of English. In W. J. Hardcastle & J. M. Beck (Eds.), *A figure of speech: A* *festschrift for John Laver* (pp. 173–199). Mahwah: Lawrence Erlbaum Associates.

Docherty, G. J., Hay, J., & Walker, A. (2006). Sociophonetic patterning of phrase-final /t/ in New Zealand English. In P. Warren & C. I. Watson (Eds.), *Proceedings of the 11th Australasian International Conference on Speech Science & Technology, Auckland, New Zealand* (pp. 378–383).

Eddington, D., & Channer, C. (2010). American English has goʔ a loʔ of glottal stops: Social diffusion and linguistic motivation. *American Speech, 85*, 338–351.

Eddington, D., & Savage, M. (2012). Where are the moun[ʔə]ns in Utah? *American Speech, 87,* 336–349.

Eddington, D., & Taylor, M. (2009). T-glottalization in American English. *American Speech, 84*, 298–314.

Edmondson, J. A., & Esling, J. H. (2006). The valves of the throat and their functioning in tone, vocal register and stress: Laryngoscopic case studies. *Phonology, 23*, 157–191.

Edmondson, J. A., Chang, Y.-C., Hsieh, F.-F., & Huang, H.-C. J. (2011). Reinforcing voiceless finals in Taiwanese and Hakka: Laryngoscopic case studies. In W. S. Lee & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhSXVII), Hong Kong* (pp. 627–630).

Ellis, A. J. (1875). *On early English pronunciation* (Part IV). London: Early English Text Society.

Ellis, A. J. (1889). *On early English pronunciation* (Part V). London: Trübner.

Esling, J. H., Fraser, K. E., & Harris, J. G. (2005). Glottal stop, glottalized resonants, and pharyngeals: A reinterpretation with evidence from a laryngoscopic study of Nuuchahnulth (Nootka). *Journal of Phonetics, 33*, 383–410.

Esling, J., Moisik, S., Benner, A., & Crevier-Buchmann, L. (2019). *Voice quality: The laryngeal articulator model*. Cambridge: Cambridge University Press.

Esling, J. H., Zeroual, C., & Crevier-Buchman, L. (2007). A study of muscular synergies at the glottal, ventricular and aryepiglottic levels. In J. Touvain & W.

J. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences, ICPhS XVI, Saarbrücken* (pp. 585–588).

Fasold, R. W. (1981). The relation between black and white speech in the South. *American Speech, 56*, 163–189.

Fabricius, A. (2002). Ongoing change in modern RP: Evidence for the disappearing stigma of t-glottalling. *English World-Wide, 23*, 15-136.

Farrington, C. (2018). Incomplete neutralization in African American English: The case of final consonant voicing. *Language Variation and Change, 30*, 361–383.

Fischer-Jørgensen, E. (1989). Phonetic analysis of the stød in standard Danish. *Phonetica, 46*, 1–59.

Fowler, C. A. (1992). Vowel duration and closure duration in voiced and unvoiced stops: There are no contrast effects here. *Journal of Phonetics, 20,* 143–165.

Fujimura, O., & Sawashima, M. (1971). Consonant sequences and laryngeal control. *Annual Bulletin, Research Institute of Logopedics and Phoniatrics, University of Tokyo* 5, 1–6.

Garellek, M. (2011). The benefits of vowel laryngealization on the perception of coda stops in English. *UCLA Working Papers in Phonetics*, *109*, 31–39.

Garellek, M. (2013). *Production and perception of glottal stops* (Doctoral dissertation, University of California, Los Angeles).

Garellek, M. (2014). Voice quality strengthening and glottalization. *Journal of Phonetics, 45*, 106–113.

Garellek, M. (2015). Perception of glottalization and phrase-final creak. *Journal of the Acoustical Society of America, 137*, 822–831.

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

Garellek, M., & Seyfarth, S. (2016). Acoustic differences between English /t/ glottalization and phrasal creak. In N. Morgan (Ed.), *Proceedings of 17th Annual Conference of the International Speech Communication Association (INTERSPEECH 2016), San Francisco* (pp. 1136–1140).

Gerratt, B.R., Kreiman, J. (2001). Toward a taxonomy of nonmodal phonation. *Journal of Phonetics, 29*, 365–381.

Gimson, A. C. (1962). *An introduction to the pronunciation of English*. London: Arnold.

Gordeeva, O. B., & Scobbie, J. M. (2013). A phonetically versatile contrast: Pulmonic and glottalic voicelessness in Scottish English obstruents and voice quality. *Journal of the International Phonetic Association, 43*, 249–271.

Gruenenfelder, T. M., & Pisoni, D. B. (1980). Fundamental frequency as a cue to postvocalic consonantal voicing: Some data from speech perception and production. *Perception & Psychophysics, 28*, 514–520.

Harris, J. G. (2001). States of the glottis of Thai voiceless stops and affricates. In K. Tingsabadh, & A. S. Abramson (Eds.), *Essays in Tai linguistics* (pp. 3–11). Bangkok: Chulalongkorn University Press.

Haslerud, V. C. D. (1995). *The variable (t) in Sydney adolescent speech: A sociolinguistic study of phonological variation* (Cand. Philol dissertation, University of Bergen).

Henton, C., & Bladon, A. (1988). Creak as a sociophonetic marker. In L. M. Hyman & C. N. Li (Eds.), *Language, speech and mind: Studies in honour of Victoria A. Fromkin* (pp. 3-29). London: Routledge.

Henton, C., Ladefoged, P., & Maddieson, I. (1992). Stops in the world’s languages. *Phonetica, 49*, 65–101.

Hillenbrand, J. M., & Houde, R. A. (1996). Role of F0 and amplitude in the perception of intervocalic glottal stops. *Journal of Speech and Hearing Research, 39*, 1182–1190.

Holmes, Janet. 1995. Glottal stops in New Zealand English: An analysis of variants of word-final /t/. *Linguistics*, 33, 433–463.

House, A, & Fairbanks, G. (1953). The influence of consonant environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America, 25*, 105–113

Howe, D., & Pulleyblank, D. (2001). Patterns and timing of glottalisation. *Phonology, 18*, 45- 80.

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

Ingram, J. C. L. (1989). Connected speech processes in Australian English. *Australian Journal of Linguistics, 9*, 21–49.

Iwata, R., Sawashima, M., & Hirose, H. (1981). Laryngeal adjustments for syllable-final stops in Cantonese. *Annual Bulletin of the Research Institute for Logopedics and Phoniatrics, 15*, 45-54.

Jones, D. (1909). *The pronunciation of English*. Cambridge: Cambridge University Press.

Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America, 108*, 1252–1263.

Kahn, D. (1976). *Syllable based generalizations in English phonology* (Doctoral dissertation, Massachusetts Institute of Technology).

Keyser, S. J., & Stevens, K. N. (2006). Enhancement and overlap in the speech chain. *Language, 82*, 33–62.

Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, *59*, 1208–1221.

Keating, P., Garellek, M., & Kreiman, J. (2015) Acoustic properties of different kinds of creaky voice. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow, Scotland.*

Keating, P., Wymark, D., & Sharif, R. (2019). Proposal for superscript diacritics for prenasalization, preglottalization and preaspiration. *Journal of the International Phonetic Association*. https://doi.org/10.1017/ S0025100319000057.

Kohler, K. J. (1982). F0 in the production of lenis and fortis plosives. *Phonetica, 39*, 199–218.

Kohler, K. J. (1994). Glottal stops and glottalization in German: Data and theory of connected speech processes. *Phonetica, 51*, 38–51.

Koops, C., & Niedielski, N. (2009). */t, d/-glottalization and vowel variation in Houston AAE*. Paper presented at New Ways of Analyzing Variation (NWAV) 38, University of Ottawa, Canada.

Kreiman, J. (1982). Perception of sentence and paragraph boundaries in natural conversation. *Journal of Phonetics, 10*, 163–175.

Ladefoged, P. & Maddieson, I. (1996). *The sounds of the world's languages*. Malden: Blackwell.

Liberman, A.M., Delattre, P.C., & Cooper, F. S. (1958). Some cues for the distinction between voiced and voiceless stops in initial position. *Language and Speech, 1,* 153–167.

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

Lisker, L. (1974) On explaining vowel duration variation. *Glossa, 8*, 233–246.

Lisker, L. (1975). Is it VOT or a first-formant transition detector? *Journal of the Acoustical Society of America, 57,* 1547–1551.

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

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word, 20,* 384–422.

Lisker, L., & Abramson, A. S. (1967). Some effects of context on voice onset time in English stops. *Language and Speech, 10,* 1–28.

Lisker, L. & Abramson, A.S. (1970). The voicing dimension: Some experiments in comparative phonetics. In B. Halá, M. Romportl & P. Janota (Eds.), *Proceedings of the 6th International Conference of Phonetic Sciences* (pp. 563–567). Prague: Academia.

Luce, P. A., & Charles-Luce, J. (1985). Contextual effects on vowel duration, closure duration, and the consonant/vowel ratio in speech production. *Journal of the Acoustical Society of America, 78*, 1949–1957.

Maddieson, I. (2013). Glottalized Consonants. In M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/7.

Malécot, A. (1970). The lenis–fortis opposition: Its physiological parameters. *Journal of the Acoustical Society of America*, *47*, 1588–1592.

Mathisen, A. G. (1999). Sandwell, West Midlands: Ambiguous perspectives on gender patterns and models of change. In P. Foulkes & G. J. Docherty (Eds.), *Urban voices: Accent studies in the British Isles* (pp.107–123). London: Arnold

McCarthy, O., & Stuart-Smith, J. (2013). Ejectives in Scottish English: A social perspective. *Journal of the International Phonetic Association, 43*, 273–298.

Mees, I. M. (1987). Glottal stop as a prestigious feature in Cardiff English. *English World-Wide, 8*, 25–39.

Mees, I. M., & Collins B. (1999). Cardiff: A real-time study of glottalisation. In P. Foulkes & G. J. Docherty (Eds.), *Urban voices: Accent studies in the British Isles* (pp. 85–202). London: Arnold.

Milroy, J., Milroy, L., & Hartley, S. (1994). Local and supra-local change in British English: The case of glottalisation. *English World-Wide*, *15*, 1–33.

Milroy, J., Milroy, L., Hartley, S., & Walshaw, D. (1994). Glottal stops and Tyneside glottalization: Competing patterns of variation and change in British English. *Language Variation and Change, 6*, 327–357.

Mitchell, A. G. (1945). *The pronunciation of English in Australia*. Sydney: Angus & Robertson.

Moisik, S. R., & Esling J.H. (2011). The ‘whole larynx’ approach to laryngeal features. In W. S. Lee & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhSXVII), Hong Kong* (pp. 1406–1409).

Moisik, S. R., Esling, J. H., Crevier-Buchman, L., Amelot, A., & Halimi, P. (2015). Multimodal imaging of glottal stop and creaky voice: Evaluating the role of epilaryngeal constriction. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow, Scotland*.

Mott, B. (2012). Traditional Cockney and popular London speech. *Dialectologia, 9*, 69-94.

Ogden, R. (2001). Turn transition, creak and glottal stop in Finnish talk-in-interaction. *Journal of the International Phonetic Association, 31*, 139–152.

Ohala, J. J., & Ohala, M. (1993). The phonetics of nasal phonology. In M. K. Huffman & R. A. Krakow (Eds.), *Phonetics and phonology 5: Nasals, nazalisation and the velum* (pp. 225–250). New York: Academic Press.

Ohde, R. N. (1984). Fundamental frequency as an acoustic correlate of stop consonant voicing. *Journal of the Acoustical Society of America, 75*, 224–230.

Penney, J., Cox, F., Miles, K., & Palethorpe, S. (2018). Glottalisation as a cue to coda consonant voicing in Australian English. *Journal of Phonetics, 66*, 161–184.

Peterson, G., & Lehiste, I. (1960). Duration of syllabic nuclei in English. *Journal of the Acoustical Society of America, 32*, 693–703.

Pierrehumbert, J. B. (1994). Knowledge of variation. In K. Beals, J. Denton, R. Knippen, L. Melnar, H. Suzuki, & E. Zeinfeld (Eds.), *Papers from the Parasession on Variation, 30th meeting of the Chicago Linguistic Society, Chicago* (pp. 232–256).

Pierrehumbert, J. B. (1995). Prosodic effects on glottal allophones. In O. Fujimura & M. Hirano (Eds.), *Vocal fold physiology: Voice quality control* (pp. 39–60). San Diego: Singular.

Pierrehumbert, J. B., & Frisch, S. (1997). Synthesizing allophonic glottalization. In J. H. van Santen, R. W. Sproat, J. Olive, & J. Hirschberg (Eds.), *Progress in speech synthesis* (pp. 8–26). New York: Springer.

Pierrehumbert, J. B., & Talkin, D. (1992). Lenition of /h/ and glottal stop. In G. Doherty & D. R. Ladd (Eds.), *Papers in laboratory phonology II: Gesture segment prosody* (pp. 90–117). Cambridge: Cambridge University Press.

Port, R. F. (1979). The influence of tempo on stop closure duration as a cue for voicing and place. *Journal of Phonetics, 7,* 45–56.

Przedlacka, J. (2002). *Estuary English: A sociophonetic study of teenage speech in the home counties*. Frankfurt am Main: Peter Lang.

Raphael, L. J. (1972). Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English. *Journal of the Acoustical Society of America, 51*, 1296–1303.

Redi, L., & Shattuck-Hufnagel, S. (2001). Variation in the realization of glottalization in normal speakers. *Journal of Phonetics, 29*, 407–429.

Repp, B. H. (1979). Relative amplitude of aspiration noise as a voicing cue for syllable-initial stop consonants. *Language and Speech*, *22*, 173–189.

Repp, B. H. 1982. Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception. *Psychological Bulletin, 92*, 81–110.

Roach, P. J. (1973). Glottalization of English /p/, /t/, /k/ and /ʧ/ – a reexamination. *Journal of the International Phonetic Association, 3*, 10–21.

Roach, P. (2004). Illustrations of the IPA: British English: Received Pronunciation. *Journal of the International Phonetic Association, 34*, 239–245.

Schleef, E. (2013). Glottal replacement of /t/ in two British capitals: Effects of word frequency and morphological compositionality. *Language Variation and Change*, *25*, 201–223.

Selkirk, E. (1982). The syllable. In H. van der Hulst & N. Smith (Eds.), *The structure of phonological representations* (Part II, pp. 337–383). Dordrecht: Foris.

Seyfarth, S., & Garellek, M. (2015). Coda glottalization in American English. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow, Scotland.*

Simpson, A. (2014). Ejectives in English and German: Linguistic, sociophonetic, interactional, epiphenomenal? In C. Celata & S. Calamai (Eds.), *Advances in Sociophonetics* (pp. 189–204). Amsterdam: Benjamins.

Slifka, J. (2006). Some physiological correlates to regular and irregular phonation at the end of an utterance. *Journal of Voice*, *20*, 171–186.

Stevens, K. N., & Klatt, D. H. (1974). Role of formant transitions in the voiced-voiceless distinction for stops. *Journal of the Acoustical Society of America*, *55*, 653–659.

Sweet, H. (1877). A handbook of phonetics. Oxford: Clarendon Press.

Sweet, H. (1881). Sound notation. *Transactions of the Philological Society*. 177–235.

Sweet, H. (1908). The sounds of English: An introduction to phonetics. Oxford: Clarendon Press.

Tollfree, L. (1999). South East London English: Discrete versus continuous modelling of consonantal reduction. In P. Foulkes & G. J. Docherty (Eds.), *Urban voices: Accent studies in the British Isles* (pp.163–184). London: Arnold.

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

Trudgill, P. (1986). *Dialects in contact.* Oxford: Blackwell.

Uffman, C. (2007). Intrusive [r] and optimal epenthetic consonants. *Language Sciences, 29,* 451–476.

Umeda, N. (1978). Occurrence of glottal stops in fluent speech. *Journal of the Acoustical Society of America, 64*, 88–94.

Watt, D., & Allen, W. (2003). Illustrations of the IPA: Tyneside English. *Journal of the International Phonetic Association, 33*, 267–271.

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

Westbury, J. R., & Keating, P.A. (1986) On the naturalness of stop consonant voicing. *Journal of Linguistics, 22*, 145–166.

Yuen, I., Cox, F., & Demuth, K. (2018). Prosodic effects on the planning of inserted /ɹ/ in Australian English. *Journal of Phonetics, 69*, 29–42.

Zlatin, M. (1974). Voicing contrast: perceptual and productive voice onset time characteristics of adults. *Journal of the Acoustical Society of America, 56,* 981–994.

Chapter 2: Glottalisation of word-final stops in Australian English unstressed syllables

This chapter is based on the following published paper:

Penney, J., Cox, F., & Szakay, A. (2019 online). Glottalisation of word-final stops in Australian English unstressed syllables. *Journal of the International Phonetic Association*. https://doi.org/10.1017/S0025100319000045.

I certify that I was responsible for the development of the concept of this paper, in discussion with my supervisors/co-authors. The data for the first analysis were collected by Prof. Felicity Cox for an unrelated study, and the data for the second analysis were extracted from the publicly available AusTalk Corpus. I took leadership in conducting the research, including the preparation and annotation of the bulk of the data, the analyses, and the writing of all parts of the paper. My co-authors provided advice to improve the research protocol, the analyses, and the interpretation of the data, as well as the presentation of the written component. Additional annotation of the data for reliability was provided by Dr Kelly Miles.

2.0 Abstract

Glottalisation functions as a cue to coda stop voicelessness in many varieties of English, occurring most commonly for alveolar stops, although varieties differ according to the context and frequency with which glottalisation is used. In Australian English, younger speakers glottalise voiceless coda stops at much higher rates than older speakers suggesting a recent change to the variety, yet this change has only been examined in stressed syllables for stops with alveolar place of articulation. In addition, research has found that glottalisation occurs in a trading relationship with preceding vowel duration to cue coda stop voicing: younger speakers make less use of vowel duration and more use of glottalisation. This study investigates glottalisation as a cue to coda voicing in unstressed syllables, an environment in which coda voicing-related vowel durational differences are already reduced. We examine this phenomenon in two separate datasets of Australian English with reference to stops at three places of articulation to explore dialect-specific distributional patterns and to track the potential progression of change. The results suggest that glottalisation occurs in conjunction with voiceless stops at all places of articulation in the unstressed Australian English contexts examined here. The results also confirm that younger speakers employ glottalisation more than older speakers, and show that females glottalise more than males, both results supporting previous suggestions of a recent change to the variety.

2.1 Background

The phenomenon under examination in this paper is the glottalisation of coda stops. Glottalisation as defined here is the addition of glottal constriction to a coda oral stop, resulting in irregular, laryngealised phonation towards the end of the preceding voiced segment (Garellek, 2015; Huffman, 2005).[1] This phenomenon is most commonly associated with alveolar stops (Pierrehumbert, 1994, 1995; Seyfarth & Garellek, 2015; Tollfree, 1999) although varieties of English differ according to the context and frequency of glottalisation used by speakers of the community. Previous studies have found glottalisation to be an important cue to coda stop voicelessness in many varieties of English, such as British English (BrE) (Roach, 1973), American English (AmE) (Huffman, 2005; Pierrehumbert, 1995; Redi & Shattuck-Hufnagel, 2001), and Scottish English (Gordeeva & Scobbie, 2013). Some varieties exhibit place of articulation (POA) asymmetries in the degree of glottalisation, with alveolar stops more likely to be glottalised than bilabial or velar stops (Keyser & Stevens, 2006). Keyser and Stevens (2006) hypothesise, based on an EMA/intra-oral pressure study by Svirsky et al. (1997), that voiceless bilabial and velar stops may be accompanied by an increase in tongue stiffness, which results in rapid inhibition of voicing. They suggest that alveolar stops, on the other hand, require increased flexibility of the tongue for their articulation allowing for greater expansion of the vocal tract, which in

[1] This is sometimes referred to in the literature as glottal reinforcement (e.g. Docherty Docherty, Foulkes, Milroy, Milroy, & Walshaw, 1997; Gimson, 1962; Roach, 1973; Wells, 1982), whereby the glottal constriction/glottal stop 'reinforces' or strengthens the coda stop. We reserve the term glottalisation for this process. We are aware that the term glottalisation is used to refer to various phonetic phenomena, such as glottal replacement, non-pulmonic stops produced with a glottalic airstream mechanism, and creaky voice. We also note that glottalisation does not only occur in conjunction with coda stops; glottalisation of vowels in onset position is well documented in the literature (e.g. Dilley, Shattuck-Hufnagel, & Ostendorf, 1996; Garellek, 2014; Pierrehumbert & Talkin, 1992; Umeda, 1978). However, our use of the term in this paper does not include these phenomena (unless noted) and refers specifically to the addition of glottal/laryngeal constriction to a coda oral stop.

turn promotes maintenance of the transglottal pressure difference and thereby voicing. In other words, the suggestion is that the pressure above and below the glottis does not equalise as quickly for alveolar stops as it does for stops at other places of articulation, and hence voiceless alveolar stops may be prone to prolongation of voicing. Their hypothesis is that speakers employ glottalisation of alveolar stops as a mechanism to successfully extinguish voicing in the face of an undesirably expanding vocal tract (Keyser & Stevens, 2006) and they explain that this expansion (and hence maintenance of the transglottal pressure difference and resulting voicing prolongation) does not occur in labial and velar stop production, making glottalisation less necessary in these contexts. To the best of our knowledge this hypothesis is yet to be empirically tested. Keyser and Stevens (2006) concede that glottalisation of all three POAs may be necessary to enhance voicing contrasts compromised by reduced subglottal pressure in utterance final position. It should also be noted that bilabial and velar stops are nevertheless often glottalised in multiple varieties of English (see below) in both utterance final and non-utterance final position.

Although, as noted above, glottalisation as a cue to coda voicelessness is reported for many varieties of English, varieties differ according to which stops are glottalised and the frequency with which they are glottalised. In BrE, for example, glottalisation in conjunction with /t/ is well known (Docherty, Foulkes, Milroy, Milroy, & Walshaw, 1997; Milroy, Milroy, Hartley, & Walshaw, 1994). Glottal replacement of /t/, a similar and potentially related phenomenon[2] in which the oral

[2] While some researchers take glottalisation and glottal replacement to be two points on a continuum of lenition, others note that while glottal replacement is a weakening process – it is a form of debuccalisation in which the oral gesture is lost while a laryngeal residue is retained – this is not necessarily the case for glottalisation, in which an additional gesture is added to the oral stop resulting in increased rather than reduced structural complexity (Czyżak, 2014; O'Brien, 2012). Docherty et al. (1997: 307) suggest that glottalisation and glottal replacement should perhaps be considered as independent phenomena rather than 'manifestations of a

stop is replaced or completely masked by a glottal stop, is also well attested (e.g. Docherty & Foulkes, 1999a; Docherty et al., 1997; Roach, 1973; Wells, 1982) but will not be examined here. Similarly, we will not focus here on ejective realisations of stops (i.e. glottalic egressive stops), a phenomenon that is attested in Scottish English (Gordeeva & Scobbie, 2013; McCarthy & Stuart-Smith, 2013). However, glottalisation in BrE is not limited to alveolar stops; all of the voiceless stops /p, t, k/ as well as the voiceless affricate /ʧ/ can be glottalised (Roach, 1973), though variation exists among regional British dialects. In London and South-eastern English varieties /t/ is most frequently glottalised, with /p/ and /k/ glottalised less often (Schneider, Burridge, Kortmann, Mesthrie, & Upton, 2004; Tollfree, 1999; Wells, 1982). In Tyneside English /p/ is more frequently glottalised than the other stops (Docherty et al., 1997; Watt & Milroy, 1999; note that the timing of glottalisation in this variety differs from other British varieties, with the glottal gesture occurring prior to the oral gesture and masking the oral stop release (Docherty & Foulkes, 1999b; Wells, 1982)), and in Sandwell /t/ and /k/ are often glottalised, whereas /p/ is affected less frequently (Mathisen, 1999). In AmE glottalisation commonly occurs for /t/ and /p/, though it is reported to occur more frequently and in a greater range of contexts for /t/ than for /p/ (Huffman, 2005; Pierrehumbert, 1994; Seyfarth & Garellek, 2015). Nevertheless, in perception AmE listeners have been found to associate glottalisation with both /t/ and /p/ (Chong & Garellek, 2018).

Recent research has shown that glottalisation is also used in Australian English (AusE) voiceless stop codas. Penney, Cox, Miles, and Palethorpe (2018), in an analysis of monosyllabic words containing alveolar voiced and voiceless stop

---

single process or as points on a single continuum' due to their varied sociolinguistic distribution in Tyneside English speakers.

codas produced by 67 speakers from Sydney, found that 55% of the words containing voiceless codas exhibited glottalisation, whereas only 6% of the words ending in voiced codas contained glottalisation. Whether glottalisation functions in the same way in this variety for stops at other POAs remains an open question. In her paper on /t/ variation in AusE, Tollfree (2001) notes that in addition to /t/, glottalisation is also present for /k/ in some lexical items (such as the almost categorical presence for certain items such as *like*) and that it can also occasionally be found for /p/. However, no empirical evidence is available to support this observation.

A number of studies have highlighted links between glottalisation and social factors in different varieties of English. Glottalisation has been linked to different class affiliations. For example, in Tyneside English, Milroy et al. (1994) found voiceless coda stop glottalisation to be associated with working class speakers, whereas middle class speakers tended to prefer glottal replacement. Similarly, word final glottalisation[3] has been found to serve as a marker of prestige in Cardiff English, where it was present in the speech of the middle class and those with middle class aspirations (Mees, 1987; Mees & Collins, 1999). On the other hand, intervocalic /t/ glottalisation was rarely used by middle class Cardiff English speakers, but commonly used by working class speakers (Mees, 1990). Tollfree (2001) found AusE speaking teenagers from lower socioeconomic backgrounds used glottalised variants of /t/ more frequently in pre-pausal contexts in conversational speaking style than teenagers from higher socioeconomic backgrounds (though this difference was reduced in more formal reading list style).

Some studies have suggested that in some varieties glottalisation is more

[3] Here [ʔ] and [ʔt], that is, both glottal replacement and glottalisation, are classified together as word final glottalisation.

common in male speech (e.g. Docherty, Hay, & Walker, 2006; Milroy et al., 1994[4]). Others have found it to be more common in the speech of females (e.g. Holmes, 1995;[5] Mees, 1987). In a study of primary school aged AusE speaking children, Tait and Tabain (2016) found that girls produced more glottalised variants of /t/ than boys,[6] though they did not report any glottalised variants of stops at other POAs. Interestingly, in their study of glottalisation in AusE, Penney et al. (2018) did not find evidence of gender differences in /t/ glottalisation. They did, however, find age differences; younger speakers employed glottalisation much more so than older speakers (71% in younger compared to 36% in older speakers). Paired with the fact that glottalisation has only been noted in the literature as being present in AusE since the late 1980s (Ingram, 1989; Haslerud, 1995; Tollfree, 2001), this suggests a recent change to the variety. In the light of Tollfree's (2001) findings discussed above, that glottalisation is more common for /t/ than for stops at other POAs, it is conceivable that the change may have originally affected alveolar stops, before progressing to the other POAs. If this is the case, we may expect to see differences according to gender for the POAs other than alveolar, as females are often the leaders of linguistic change (Cameron & Coates, 1989; Eckert, 1989; Labov, 1990, 2001).

---

[4] This study found male Tyneside English speakers used glottalisation more frequently than female speakers, although female Tyneside English speakers used glottal replacement more frequently than the male speakers.

[5] Note that Holmes' study focused on the realisation of /t/ as glottal stop, i.e. glottal replacement rather than glottalisation as defined here. However, she identified the glottal stop variants of /t/ auditorily and notes the difficulty in distinguishing between full glottal replacement and glottalised stops in which the alveolar gesture is obscured, and hence her glottal stop category can be assumed to include both glottal replacement as well as glottalised realisations of /t/.

[6] Tait and Tabain's (2016) glottalised category included two types: *glottal /t/*, in which glottalisation but no formant transitions were visible (i.e. glottal replacement), and *laryngealised /t/*, which they 'identified by a lack of stop closure ***or*** release and the presence of fully laryngealised voicing' (p. 66). Therefore, it is reasonable to assume that their second type included at least some examples of glottalisation as it is defined here.

There are a number of acoustic cues to the phonological voicing status of coda stops in English other than glottalisation: F0 is lower at the offset of vowels preceding voiced stops compared to voiceless stops; and there is often a voice bar present in the stop closure of voiced stop codas (Gruenefelder & Pisoni, 1980; House & Fairbanks, 1953; Kingston & Diehl, 1994; Kohler, 1982; Lisker, 1986; Ohde, 1984; Song, Demuth, & Shattuck-Hufnagel, 2012; Wright, 2004). Preceding vowel duration is a major cue to coda voicing (e.g. Fowler, 1992; Klatt, 1976; Lisker, 1978; Port & Dalby, 1982). Recent research on AusE suggests that a trading relationship may exist between glottalisation and vowel duration as cues to coda stop voicing. Penney et al. (2018) found that in voiceless coda contexts younger speakers did not only utilise glottalisation more than older speakers, but they also made less use of preceding vowel duration. In addition, glottalisation affected high vowels less than non-high vowels (in both older and younger groups), with high vowels exhibiting greater coda voicing-related preceding vowel duration differences than non-high vowels. Furthermore, coda stop voicing affected inherently long vowels (e.g. /iː/ *beat* vs *bead*) more than inherently short vowels (e.g. /ɪ/ *bit* vs *bid*). These findings were based on observations of stressed syllables only; questions remain about whether the same effects would be present in unstressed syllables, as coda voicing-related durational differences are expected to be already reduced in such unstressed environments (Crystal & House, 1988b; Davis & van Summers, 1989; Klatt, 1975). In cases where the vowel duration cue to coda voicing has been minimised, the trading relationship would predict higher rates of glottalisation in voiceless stop contexts to offset the reduced effectiveness of the vowel duration cue and help to preserve the voicing contrast. Although word initial glottalisation, that is, glottal marking at the onset of a vowel initial word, has been found to be more common in stressed rather than unstressed syllables (Kohler, 1994; Malisz, Żygis, & Pompino-Marschall, 2013), we are not aware of similar findings for coda glottalisation.

The present study investigates glottalisation as a cue to coda voicing in unstressed syllables with reference to all three POAs in AusE to investigate dialect-specific distributional patterns and the potential progression of change. In this study we examine two separate datasets in two analyses. The first dataset contains productions from a cohort of young female university students from Sydney. These data enable an analysis of coda stops in unstressed contexts at all three English stop POAs and a comparison of voiced and voiceless stop codas. The second dataset contains productions extracted from the AusTalk corpus (Burnham et al., 2011). These data allow for an analysis of age and gender with respect to the implementation of glottalisation.

Based on the previous literature, our broad expectations are as follows:

- **Glottalisation will occur at all three POAs.** The general patterns of glottalisation in AusE may be similar to those reported for London and South-eastern BrE; namely, that glottalisation will occur for voiceless stops at all three POAs to cue coda stop voicelessness, but will be more frequently associated with alveolar coda stops than stops at other POAs. This hypothesis is based on the close historical connection between London English and AusE (Cox & Palethorpe, 2012) and the fact that there are a number of similarities between the two dialects (Cochrane, 1989; Wells, 1982; Yallop, 2001). This would also be in accord with comments in Tollfree (2001).
- **Reduced coda voicing-related vowel duration differences will lead to increased glottalisation.** Coda voicing-related preceding vowel duration effects will be reduced in unstressed syllable contexts, and therefore

glottalisation, which is suggested to occur in a trading relationship with preceding vowel duration in the implementation of coda voicelessness (Penney et al., 2018), may occur at increased rates than has previously been reported for stressed syllables.

- **Younger speakers will produce glottalisation at higher rates than older speakers.** Penney et al. (2018) showed that younger speakers were more likely to glottalise voiceless alveolar coda stops than older speakers and interpreted this as support for glottalisation being a recent change to AusE. If it is indeed a recent change and younger speakers are leading the change, then we may expect to see evidence of this not just in the alveolar POA, but across the entire range of voiceless stops. We also expect alveolars to exhibit increased rates of glottalisation relative to labials and velars indicating a progression of change.
- **Female speakers will glottalise at higher rates than male speakers.** We may expect to find increased glottalisation in female speakers consistent with the idea that women are the leaders in linguistic changes (Cameron & Coates, 1989; Eckert, 1989; Labov, 1990, 2001). Penney et al. (2018) did not find support for a gender effect in their study of glottalisation in alveolar POA in AusE; however, if there is evidence for a progression of change across POA we may expect to see gender effects for the other POAs.

2.2 Analysis 1: Young Sydney Females

2.2.1 Method

2.2.1.1 Data

Recordings were made in a sound treated recording studio in the Department of Linguistics at Macquarie University, using an AKG C535 EB microphone, Cooledit 2000 audio capture software via an M-Audio delta66 soundcard to a Pentium 4 PC at 44.1kHz sampling rate. Speakers read words from a computer monitor in a task containing 100 words in total. All 100 words were presented three times in random order. These data were originally collected for an experiment designed to examine the production of unstressed vowels. The word list included, among other forms, trochees with the form /(C)VɹəC/, where the final consonant was either a voiced or voiceless stop at bilabial, alveolar or velar POA. Analyses from words of this type only will be reported here. Note that in English there are no instances of words in the /(C)VɹəC/ form containing a voiced velar coda in word final position so this context is excluded from the analysis. As alveolar stops show a wide range of realisations in different contexts (e.g. /t, d/ may be flapped intervocalically; /t/ may be realised as a glottal stop before nasals and laterals; /t, d/ may be realised as unreleased stops[7] preceding obstruents), we chose the prepausal context to maximise the occurrence of canonical, released realisations. Note, however, that examining this context does not preclude the presence of non-canonical stop variants in the data. We therefore also describe the non-canonically produced stop variants in the results section below. The words

[7] In this paper we use the term unreleased to describe stops that have no audible or acoustically visible release, though we acknowledge that all stops are eventually released.

analysed here are *Arab*, *carob*, *Europe*, *syrup, arid*, *Jarrod*, *ferret*, *parrot, barrack.*[8] Two tokens were discarded due to mispronunciations. A total of 754 items (bilabial: 335; alveolar: 335; velar: 84) were examined (see Table A1 in the Appendix for a list of tokens by word for this and subsequent analyses).

The selected items allow for the comparison between words containing final coda stops in each of the three POAs (e.g. *syrup, parrot*, *barrack*) as well as between voiced and voiceless final coda stops for the bilabial (e.g. *carob, syrup*) and alveolar (e.g. *Jarrod, parrot*) POA. The use of trochees enabled us to explore whether glottalisation may be maximally exploited in an environment in which the coda voicing-related preceding vowel duration differences may already be reduced. Note that for all of the words in this analysis a schwa vowel is used in the unstressed syllable as is common for words of this type in AusE; however, unstressed vowels preceding velar codas may be realised as phonetically higher than unstressed vowels in other contexts in AusE (Cox & Fletcher, 2017) and vowel height has been linked to rate of glottalisation, with high vowels showing less glottalisation than non-high vowels (Penney et al., 2018).

2.2.1.2 Speakers

Twenty-eight female AusE speakers (aged 18-38; mean age: 24; SD: 7) took part in this study. All were Macquarie University undergraduate students who received course credit for their participation. All were born in Australia and had at least one parent born in Australia. All had completed their high school education in Australia,

---

[8] Ideally the data would have contained an equal number of words ending in stops at each POA, but as the data collection protocol was not originally created with this analysis in mind this was unfortunately not possible and hence our data only contains one word ending in a voiceless velar stop.

with all but three having done so in Sydney. All non Australian-born parents were born in an English speaking country and had English as an L1.

In this study we have opted for a design which will allow for a broad-based population-level analysis as opposed to a more focussed individual-level analysis (i.e. we have selected a relatively large sample of speakers from a homogeneous population who produce a restricted set of words repeated only three times). Our rationale for this design structure is that our focus is on patterns pertaining to this population of speakers. However, we acknowledge that individuals are likely to vary in their use of glottalisation so cross speaker variability will be reported below where appropriate.

2.2.1.3 Acoustic analysis

All /(C)VɹəC/ tokens were first processed by the WebMAUS automatic aligner (Kisler, Reichel, & Schiel, 2017) utilising an AusE model. MAUS automatically returns Praat (Boersma & Weenik, 2015) textgrids with phonemic boundaries labelled. These were then checked and hand corrected where necessary. In addition, all textgrids were hand labelled for subsegmental components including:

- the onset of a high energy periodic F2 signalling the onset of the initial (stressed) vowel;
- the cessation of high energy F2 signalling the end of the second (unstressed) vowel;
- the presence and duration of a voice bar in the stop closure;
- the release burst of the coda stop;
- the F3 trough signalling the target of the intervocalic rhotic (Cox, Palethorpe, Buckley, & Bentink, 2014; Espy-Wilson, Boyce, Jackson, Narayanan, &

Alwan, 2000; Foulkes & Docherty, 2000; Hay & Maclagan, 2012; Yuen, Cox, & Demuth, 2018);

- the presence and duration of glottalisation.

The presence and duration of glottalisation was identified through irregular pitch periods in the second half of the voiced /Vɹə/ sequence, visible as irregularity in the waveform and a sudden increase in the duration between periods in conjunction with irregularity in amplitude in the wideband spectrogram. We acknowledge that isolated words each represent a separate intonational phrase and therefore phrase final creak may be evident. We therefore established conservative criteria to ascertain whether irregularity should be considered due to glottalisation associated with the coda. Tokens in which irregularity extended throughout the voiced /Vɹə/ sequence were considered to be examples of phrase final creak or speaker specific creaky voice and were thus not included as examples of coda glottalisation for this study in line with Penney et al. (2018). Tokens in which irregularity began in the second half of the /Vɹə/ sequence (i.e. irregularity began after the F3 trough representing the target of the rhotic segment) but within 15 milliseconds of the F3 trough were also considered examples of creaky voice rather than labelled as glottalised codas. In classifying the tokens in this manner we may have discarded some examples of glottalisation associated with final coda stops in which glottalisation began shortly after the target of the rhotic. It is also plausible that very short instances of phrase final creak occurring only at the end of the unstressed vowel may have been labelled as glottalised, though note that this would be the case for both voiced and voiceless coda contexts. Figure 1 below illustrates an example of a labelled token containing coda glottalisation showing the waveform, spectrogram and textgrid tiers.

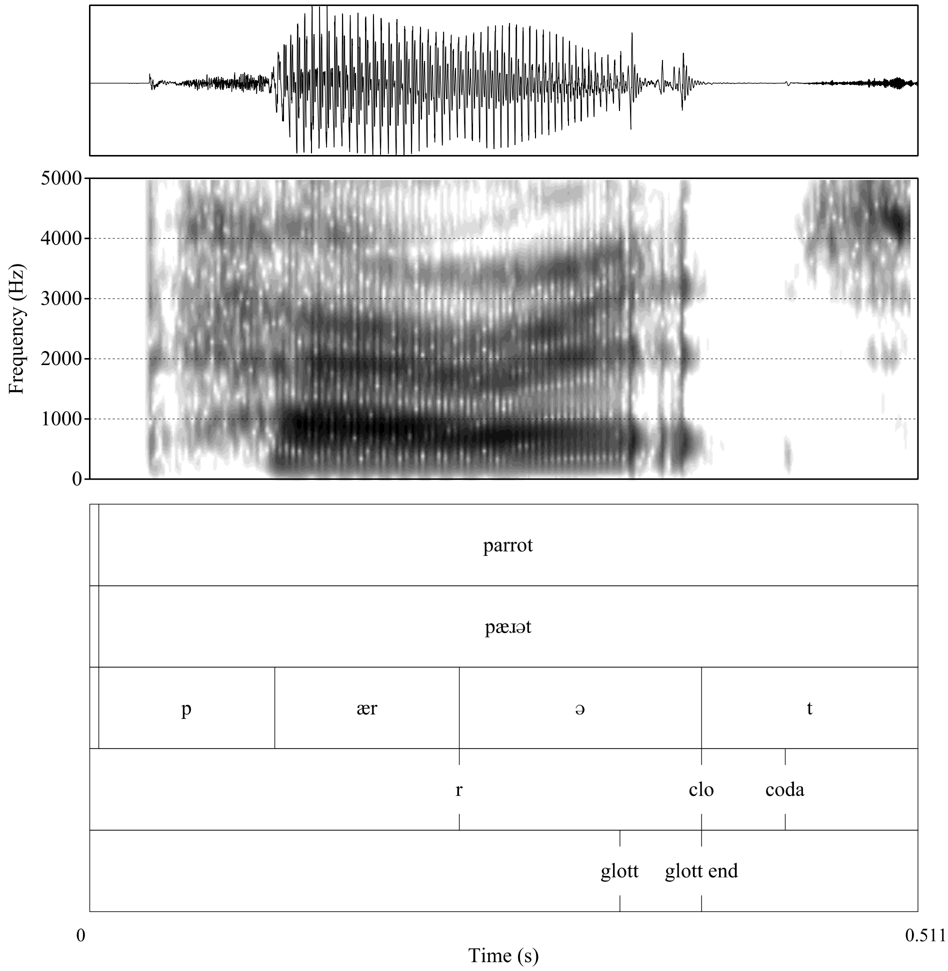

Figure 1. Example of a labelled file of the word *parrot* with coda glottalisation at the end of the voiced sequence. The glottalised portion is shown by labels *glott – glott end*. The closure period of the coda stop is shown by labels *clo – coda*.

We then calculated the duration of the final unstressed vowel. As it is difficult to effectively segment an intervocalic rhotic approximant from a neighbouring vowel, we measured duration from the F3 trough of the rhotic to the cessation of periodic F2 indicating offset of the vowel. We then measured F1 and F2 at the point equivalent to 75% of the duration of this segment, in order to obtain measurements of vowel height and backness. The point equating to 75% of the segment's duration was selected in order to estimate the position where the formants had stabilised from the influence of the rhotic, though we note that unstressed schwa displays highly coarticulated characteristics (Fleming, 2009; van Bergem, 1994). High vowels are less affected by glottalisation than low vowels (Brunner & Żygis, 2011; Malisz et al., 2013; Penney et al., 2018); in addition, as discussed above, in AusE unstressed vowels preceding velar and postalveolar codas (e.g. *paddock*, *marriage*) may be realised as phonetically higher than unstressed vowels in other contexts (Cox & Fletcher, 2017). Thus, formant measurements are an important variable in our analysis. Outliers were checked and hand corrected where necessary. Formant measures were then converted to the Bark scale using the *vowels* package (Kendall & Thomas, 2014) in R (R Core Team, 2016).

Using the STRAIGHT pitch tracker (Kawahara, Masuda-Katsuse, & de Cheveigné, 1999) in VoiceSauce (Shue, Keating, Vicenik, & Yu, 2011), F0 was measured with a window size of 25 milliseconds and a frame shift of one millisecond throughout the unstressed vowel (as segmented from the F3 trough associated with the preceding rhotic to the end of the vowel) and subsequently averaged into five equal subsections. The average F0 measure of the fifth subsection of the unstressed vowel was recorded as F0 at vowel offset.

To ensure annotator reliability 10 per cent of the data were randomly selected and re-labelled. A second trained annotator also re-labelled 10 per cent of the data to assess inter-annotator reliability. Intra-class correlations and 95% confidence intervals were calculated for the continuous measure of vowel duration, using the *irr* package (Wolak, Fairbairn, & Paulsen, 2012). For the categorical measure of the presence or absence of glottalisation a Cohen's Kappa score was calculated. Table 1 contains the results for inter-annotator reliability and intra-annotator reliability. As can be seen, reliability was high in all cases.

Table 1. Intra-class correlation results for vowel duration and Cohen's Kappa results for the presence of glottalisation.

	Vowel duration			Presence of glottalisation	
	Inter-annotator	Intra-annotator		Inter-annotator	Intra-annotator
ICC	0.952	0.875	Kappa	0.955	0.969
p value	< .001	< .001	*p* value	< .001	< .001
95% CI	0.922, 0.97	0.809, 0.919	*z* score	7.65	8.4

2.2.1.4 Statistical analysis

We fitted a number of mixed models using the *lme4* (Bates, Maechler, Bolker, & Walker, 2015) and *lmertest* (Kuznetsova, Brockhoff, & Christensen, 2016) packages in R (R Core Team, 2016). The following analyses were conducted:

- Linear mixed effects modelling was used to conduct three separate analyses of the following dependent variables: vowel formants (either F1 or F2), coda voicing-related durational contrasts, and differences in fundamental frequency;
- Logistic mixed effects modelling was used to conduct two separate analyses of the following dependent variables: glottalisation in voiced and voiceless codas, and glottalisation in voiceless codas.

The details of the dependent variables and fixed factors included in each of the models will be reported below. Unless otherwise specified, we included in each model random intercepts for speaker and repetition and random slopes for repetition by speaker. Where relevant (and noted below) random intercepts were included for word and random slopes were included for word by speaker. We initially included all fixed factors and their interactions, then pruned the models after carrying out model comparisons with the *anova()* function to remove non-significant terms that did not significantly improve the model, beginning with non-significant interactions. Using this approach we arrived at the most parsimonious model for each analysis. For significant effects (at an alpha level of $p = 0.05$) we report F-statistics for the linear mixed effects models and Chi-Square statistics the logistic mixed effects models. Post-hoc analyses were carried out with Tukey HSD corrections for multiple comparisons.

2.2.2 Results

2.2.2.1 Analysis of vowel formants

We first analysed F1 and F2 of the unstressed vowels in all 754 tokens to identify whether the POA of the following stop had an effect on the height and fronting of the unstressed vowel. We fitted separate linear mixed effects models for F1 and F2, in which the respective formant measurement was the dependent variable (i.e. either F1 or F2 measured in Bark). In both models POA was a fixed factor. The syntax for each of the models was as follows: lmer(formant ~ poa + (1+repetition|speaker) + (1|repetition)).

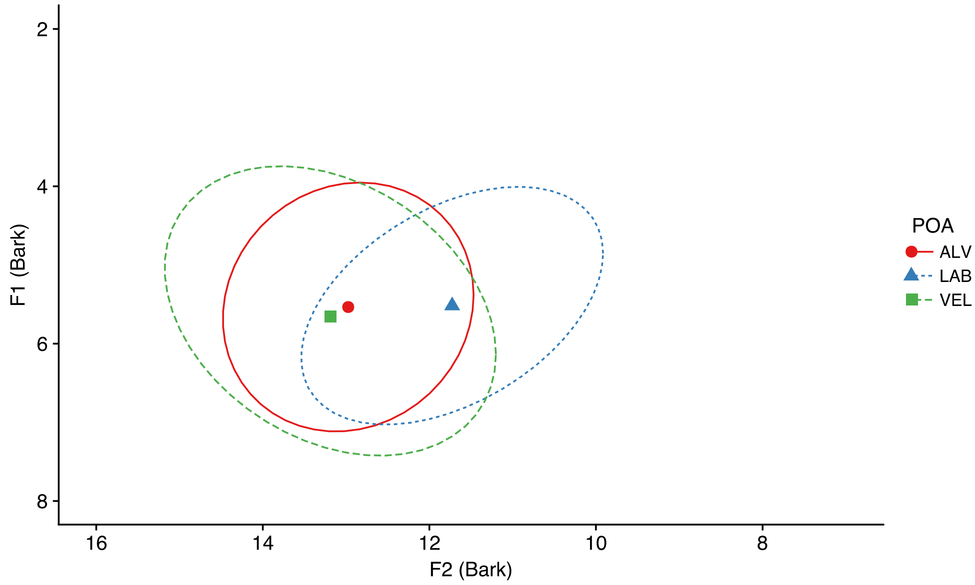

Figure 2. Mean F1 and F2 values (Bark) for unstressed vowels according to POA of following stop. Shapes represent centroids. Ellipses represent 95% confidence intervals.

Figure 2 illustrates the mean F1 and F2 in Bark for each POA for all voiced and voiceless tokens for all speakers. As can be seen, velar POA produced slightly more variation in F1 than the other POAs. However, the linear mixed effects model for F1 showed no significant effect of POA ($F(2,696) = 2.226$; $p = 0.105$). Figure 2 also shows that the bilabial POA had a clear effect of lowering F2, as is to be expected (Modarresi, Sussman, Lindblom, & Burlingame, 2005), and the linear mixed effects model for F2 confirmed POA had a significant effect ($F(2,724) = 561.9$; $p < 0.0001$). Post-hoc tests showed that F2 in each POA context differed significantly from each of the others (alveolar–labial: $p < 0.0001$; alveolar–velar: $p = 0.004$; labial–velar: $p < 0.0001$). Unstressed vowels were more retracted in the bilabial context and most fronted in the velar context, but POA did not have an effect on the height of unstressed vowels.

2.2.2.2 Analysis of coda voicing-related durational contrast

Figure 3 displays the mean durations of the unstressed vowels (which included the transitional component from the trough of F3 of the rhotic) in both the voiced and voiceless coda contexts for alveolar and bilabial POAs. As can be seen, for each POA the unstressed vowels are longer in the voiced coda context (mean: 142 ms; SD: 30 ms) than the voiceless coda context (mean: 124 ms; SD: 33 ms), though the differences between voicing contexts are small.

We fitted a linear mixed effects model in order to identify whether there was a coda voicing-related vowel duration difference between the voiced and voiceless coda contexts. For this model we included only words with bilabial and alveolar POAs to account for the lack of examples of words with a velar POA in the voiced coda set. We also removed two words from the data set that contained no onset consonant in the initial syllable, so that all of the items included in the analysis contained the same number of segments. This was necessary in order to avoid potential durational differences that may result from compression effects, whereby syllables containing a greater number of segments may lead to reduced vowel duration (Fowler, 1983; Katz, 2012; Munhall, Fowler, Hawkins, & Saltzman, 1992). Both of the items removed contained voiced coda stops: one bilabial (*Arab*) and one alveolar (*arid*). The words included in this analysis were therefore *carob, Europe, syrup, Jarrod, ferret, parrot*. A total of 503 tokens (bilabial: 251; alveolar: 252) were included. The duration of the unstressed vowel (represented by the segment identified from the trough of F3 in the /ɹ/ to the offset of F2) was the dependent variable. The voicing of the final coda stop and POA were included as fixed factors. An interaction term was originally included but this did not improve the model and hence it was removed from the final model. Random intercepts were included for speaker, repetition and word, and random slopes were included for repetition and word by speaker. The syntax for the most parsimonious model was as follows: lmer(duration ~ voicing + poa + (1+repetition+word|speaker) + (1|repetition) + (1|word)). The model showed a significant effect for coda voicing ($F(1,4) = 21.613$; $p = 0.012$), confirming that vowels were longer before voiced coda stops. There was also a significant effect for POA ($F(1,4) = 16.136$; $p = 0.013$), with longer vowels preceding alveolar stops (mean: 137 ms; SD: 33 ms) compared to bilabial stops (mean: 123 ms; SD: 32 ms). Unstressed vowels were longer in duration before voiced coda stops, and longer in duration before stops at alveolar POA.

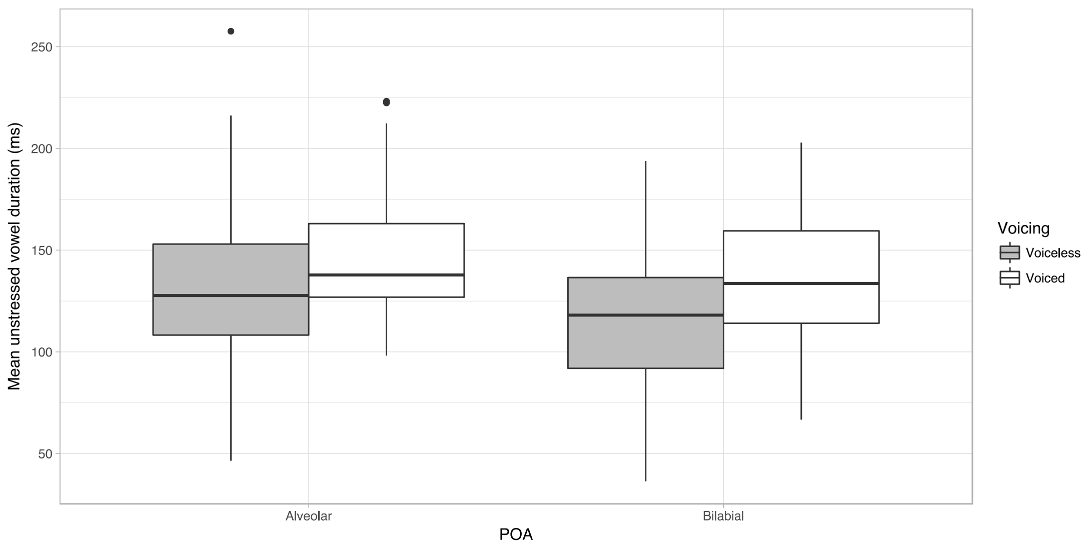

Figure 3. Mean duration (ms) of unstressed vowels (represented by the segment identified from the trough of F3 in the /ɹ/ to the offset of F2) in voiced and voiceless contexts at alveolar and bilabial POA. Boxes represent the middle 50% of unstressed vowel duration values; solid horizontal lines within the boxes represent the median; whiskers represent minimum and maximum values excluding outliers.

2.2.2.3 Analysis of glottalisation in voiced and voiceless codas

83% of voiceless coda tokens and 10% of voiced coda tokens exhibited glottalisation as determined by the presence of irregular phonation at the end of the final vowel and identified during the section of the word from the trough of F3 to the coda stop closure. 83% of the voiceless alveolar stops, 83% of the voiceless bilabial stops, and 82% of the voiceless velar stops were glottalised. In contrast, glottalisation was only present for 8% of the voiced alveolar stops, and 11% of the voiced bilabial stops. Recall that there were no items containing final voiced velar stops in the data set. Figure 4 illustrates the proportions of glottalised tokens in voiced and voiceless coda contexts for each of the POAs. In the voiceless context, nine of the 28 speakers produced glottalisation categorically, with the majority of the other speakers glottalising at least 80% of tokens. Only one speaker produced glottalisation in less than half of the tokens in the voiceless context. In the voiced context, 13 speakers produced glottalisation in at least one token, and one speaker glottalised more than 50% of tokens. Table A2 in the Appendix provides details of the individual speakers' rates of glottalisation in each voicing context.

In order to investigate whether the voicing of the coda had an effect on the presence of glottalisation we fitted a logistic mixed effects model to a subset of the data comprising all words (voiced and voiceless contexts) with bilabial and alveolar codas. The presence of glottalisation was the dependent variable. POA and the phonological voicing status of the final coda stop were included as fixed factors, as was an interaction term between the factors, though this did not improve the model fit and was removed from the final model. Random intercepts were included for speaker, repetition, and word, and random slopes were included for repetition and word by speaker. The syntax for the most parsimonious model was as follows: glmer(glottalisation ~ poa + voicing + (1+repetition+word|speaker) + (1|repetition) +(1|word)). The words included in this analysis were *Arab, carob*, *Europe*, *syrup, arid, Jarrod*, *ferret*, *parrot.* A total of 670 tokens (bilabial: 335; alveolar: 335) were included.

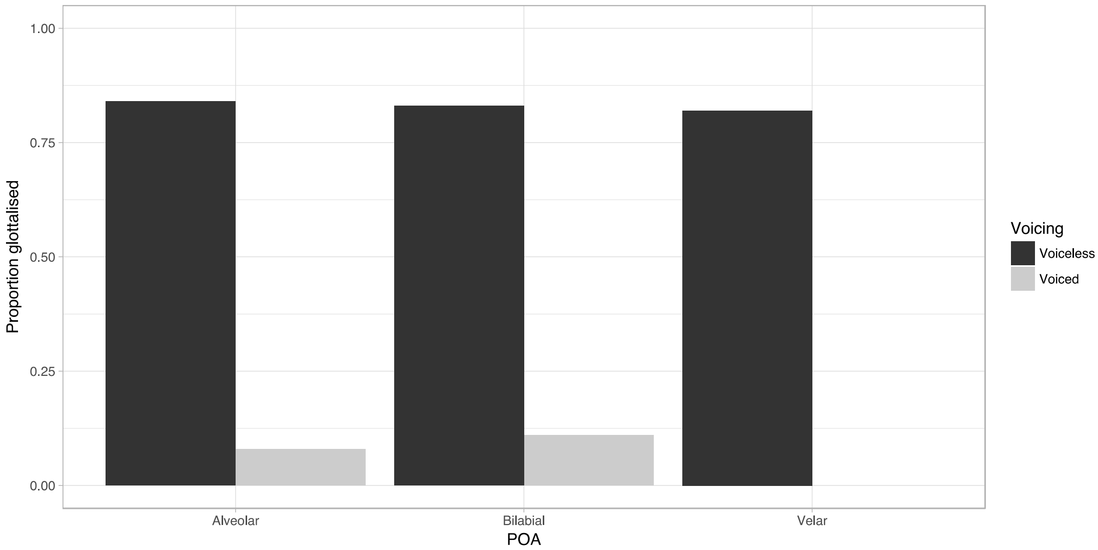

Figure 4. Proportion of items glottalised in each POA and coda voicing context. Note that there were no items containing voiced velar coda stops.

The results showed a significant effect for coda voicing ($\chi^2 = 41.226$; $p < 0.0001$), confirming that glottalisation was less likely to occur in the voiced coda context for words ending in stops at both of these POAs. There was no significant effect found for POA ($\chi^2 = 1.780$; $p = 0.182$). Glottalisation occurred at high rates in the voiceless coda context in both bilabial and alveolar contexts, but occurred only rarely in the voiced coda context.

2.2.2.4 Analysis of glottalisation in voiceless codas

We then fitted a logistic mixed effects model to another subset of the data containing all items in the voiceless coda context only. This enabled an analysis of potential differences in rates of glottalisation for words ending in stops at all three POAs. The presence of glottalisation was the dependent variable. POA of the final coda stop (bilabial, alveolar, velar) was the fixed factor.[9] The syntax for this model was as follows: glmer(glottalisation ~ poa + (1+repetition|speaker)+ (1|repetition)). The words included in this analysis were *Europe*, *syrup*, *ferret*, *parrot*, *barrack*. 419 tokens (bilabial: 167; alveolar: 168; velar: 84) were included in this analysis.

Unsurprisingly given the similar rates of glottalisation in the voiceless coda context, the analysis revealed no significant effect for POA ($\chi^2 = 0.204$; $p = 0.903$), demonstrating that the presence of glottalisation was not dependent on the POA of the final coda stop. Glottalisation occurred at high rates in the voiceless coda context in all three POA contexts.

2.2.2.5 Comparison with stressed syllables

As the rates of glottalisation found in the voiceless coda context here (83%) are numerically higher than those previously reported for stressed syllables produced by a comparable cohort of speakers (i.e. the young speakers reported in Penney et al., 2018, who glottalised at a rate of 71%), we fitted a logistic mixed effects model to analyse whether these differences were statistically significant. We compared the data for unstressed syllables in the alveolar coda context with the data for stressed syllables produced by young female speakers ($n = 17$) included in Penney et al. (2018), which contained stressed CVC monosyllables with an alveolar coda. We examined only the young female speakers as they were analogous in age and gender

[9] As our analysis of F1 showed no POA effect, unstressed vowel height was not included in this model. We did initially include vowel height, but it was not found to be significant and did not improve the model fit. We also originally calculated word frequency scores based on ICE-AUS, the Australian component of the International Corpus of English (Greenbaum, 1991). Two of the words with voiced codas were not included in ICE-AUS, so this was only investigated for words with voiceless codas. Unsurprisingly, we found no significant effect of word frequency on the presence or absence of glottalisation and this factor did not improve model fit, so we do not report the results here.

to the speakers examined here. The young females in Penney et al. (2018) produced glottalisation in 79% of the stressed tokens, a higher rate than the reported rate for the combined younger male and female speakers. The presence of glottalisation was the dependent variable, and whether the item was produced in a stressed or unstressed syllable was included as a fixed factor. The syntax for this model was as follows: glmer(glottalisation ~ stress + (1+repetition|speaker)+ (1|repetition)). The words included in this analysis were *ferret*, *parrot* for the unstressed syllables, and *heat*, *hit*, *heart*, *hut*, *hort*, *hot*, *hoot* for the stressed syllables. 476 tokens (stressed: 308; unstressed: 168) were included in this analysis. The model showed that the difference in rates of glottalisation between stressed and unstressed contexts was not significant (χ^2 =1.492; p = 0.222). Young female speakers glottalised at comparable rates in unstressed and stressed syllable contexts.

2.2.2.6 Analysis of differences in fundamental frequency

Table 2 below lists the mean F0 for each POA in both voiced and voiceless contexts. The analysis of F0 at the offset of the unstressed vowel showed small differences in the expected direction between the voiced and voiceless coda contexts. F0 was slightly lower before voiced coda stops (mean = 200Hz) than before voiceless coda stops (mean = 204Hz). There was also a small difference in F0 between POAs, with F0 marginally lower before alveolar stops (mean = 201Hz) than bilabial stops (mean = 203Hz). In order to examine the relationship between F0 and coda voicing we fitted a linear mixed effects model to all of the items in the bilabial and alveolar POA contexts. F0 at the offset of the unstressed vowel was included as the dependent variable. POA of the final coda stop (bilabial, alveolar) and phonological coda voicing were included as fixed factors. An interaction term was also included but this did not improve the model fit and as such it was not included in the final model. We included random intercepts for speaker, repetition, and word. This was the maximal random effects structure to converge. The syntax for this model was as follows: lmer(F0 ~ voicing + POA + (1|speaker) + (1|repetition) + (1|word)). The words included in this analysis were *Arab*, *carob*, *Europe*, *syrup*, *arid*, *Jarrod*, *ferret*, *parrot*. Eight tokens were excluded from the analysis due to pitch tracking errors. 660 tokens (bilabial: 330; alveolar: 330) were included. We found no significant effect of either coda voicing ($F(1,628) = 0.1079$; $p = 0.743$) or POA ($F(1,628) = 0.3527$; $p = 0.553$) on F0. F0 at the offset of unstressed vowels was comparable preceding voiced and voiceless coda stops.

Table 2. Mean F0 (Hz) at vowel offset in voiced and voiceless coda contexts according to place of articulation of following coda stop.

Voicing context	POA	N	F0 (Hz)	SD	SE
Voiceless	Alveolar	164	200	70.17	5.48
Voiceless	Bilabial	163	203	63.51	4.98
Voiced	Alveolar	166	200	46.44	3.61
Voiced	Bilabial	167	201	43.90	3.40

2.2.2.7 Voice bar

The majority of tokens containing a voiced coda stop exhibited some prolonged voicing into the stop closure. In total, 321 tokens showed evidence of a voice bar: 312/335 (93%) of these were in the voiced coda context (155 alveolar, 157 bilabial with the remaining nine in the voiceless coda context (3 alveolar, 5 bilabial, 1 velar). Of the tokens with a voice bar, 27 (9%) of these exhibited voicing throughout the entire closure.

2.2.2.8 Non-canonical stop realisations

In addition to the glottalised productions, there were a number of other non-canonical realisations of coda stops present in the data (see Table 3 for a summary). First, 5% of the tokens (*n* = 40/754 produced by 12 speakers) contained coda stops which were unreleased. Glottalisation is often associated with unreleased stops (e.g. Blevins, 2006; Kahn, 1976; Selkirk, 1982), and in 31 of the 40 tokens with unreleased stops, glottalisation was also present. In all but three of these glottalised unreleased stops, formant transitions were visible at the end of the unstressed vowel (i.e. a rising F2 preceding alveolar coda stops and a lowering F2 preceding bilabial coda stops), suggesting that an oral articulation was made for the stop even if there was no acoustic evidence for its release. The majority of unreleased stops had an alveolar POA (25/40), though there were also bilabial stops that were unreleased (15/40). Interestingly, none of the velar tokens contained unreleased stops. Velar stops have been reported to display unreleased variants less frequently than the other POAs (Byrd, 1993; Crystal & House, 1988a), and, although bilabial stops have been suggested to occur more frequently than unreleased alveolars (Byrd, 1993), Crystal & House (1988b) note that in unstressed syllables they tend to be released more frequently than the other POAs, which is consistent with the data examined here. There were also tokens in which the coda stop was spirantised; that is, produced with an incomplete occlusion resulting in turbulent (fricative) airflow through the closure period. This was the case in 4% of the tokens (*n* = 28/754 produced by 11 speakers). All but two of these were alveolar stops (with the other two bilabial stops), and spirantisation occurred in conjunction with glottalisation in 14 of the 28 tokens. Spirantisation of /t/ has been reported to be associated with female speakers with high socioeconomic status in AusE (Jones & McDougall, 2009; Tollfree, 2001). In the analysis below, we compare whether differences are present between males and females, or between this cohort of speakers and the speakers in the AusTalk corpus more generally. A further 2% (n = 17/754 produced by six speakers) of the tokens exhibited preaspiration at vowel offset. Preaspiration is a period of voiceless aspiration that may occur at the end of a vowel preceding a voiceless obstruent (Helgason, 2002; Nance & Stuart-Smith, 2013). Occurrences of preaspiration were found at all POAs, but only one token preceded a bilabial, with eight tokens before both velar and alveolar coda stops. One speaker in particular was responsible for seven of the 17 examples. Glottalisation occurred in conjunction with preaspiration in over half of the tokens. Figure 5 shows an example of a token containing both preaspiration and glottalisation.

Table 3. Number of non-canonical stop realisations according to coda voicing context. Brackets indicate number of tokens occurring in conjunction with glottalisation.

	Unreleased	Spirantised	Preaspirated	Squeak
Voiced	6/335 (2)	13/335 (1)	0/335 (0)	0/335 (0)
Voiceless	34/419 (29)	15/419 (13)	17/419 (9)	27/419 (23)
Total	**40/754 (31)**	**28/754 (14)**	**17/754 (9)**	**27/754 (23)**

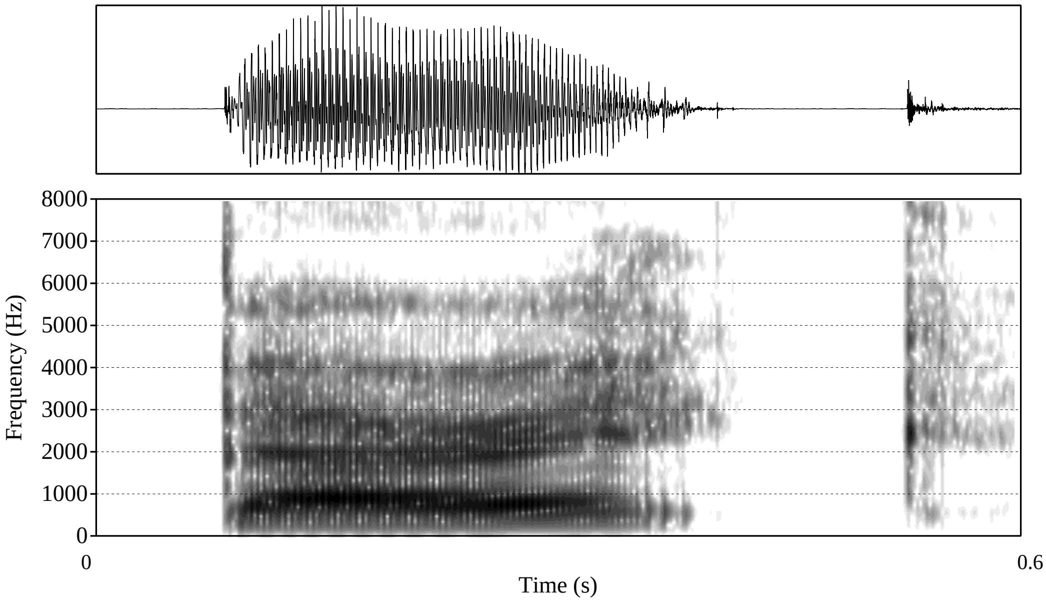

Figure 5. Spectrogram and waveform of the word *barrack* containing preaspiration and glottalisation at the end of the vowel.

Finally, 4% (*n* = 27/754 produced by 11 speakers) of the tokens contained glottal squeaks, which Redi and Shattuck-Hufnagel (2001: 414) describe as ‘a sudden shift to relatively high sustained F0, which [is] usually very low amplitude.’ These squeaks occurred for words at each POA (9 alveolar; 7 bilabial; 7 velar), and, as has previously been suggested to be the case, they occurred almost exclusively in conjunction with glottalisation (Redi & Shattuck-Hufnagel, 2001; Hejná, Palo, & Moisik, 2016). There were four examples of squeaks that were not produced in conjunction with glottalisation; however, each of these tokens was produced in contexts with phrase final creak.[10] The squeaks were relatively short in duration (mean: 26 ms) and generally occurred closer to the left edge of the stop closure period than to the right edge (i.e. the release) (mean duration from closure: 20 ms; mean duration from release: 92 ms), with a mean F0 of 280 Hz. Figure 6 shows an example

[10] Note that it is possible that the squeaks occurring with phrase final creak were also glottalised, though due to our labelling criteria these would not have been labelled as glottalised.

of a token containing a glottal squeak.

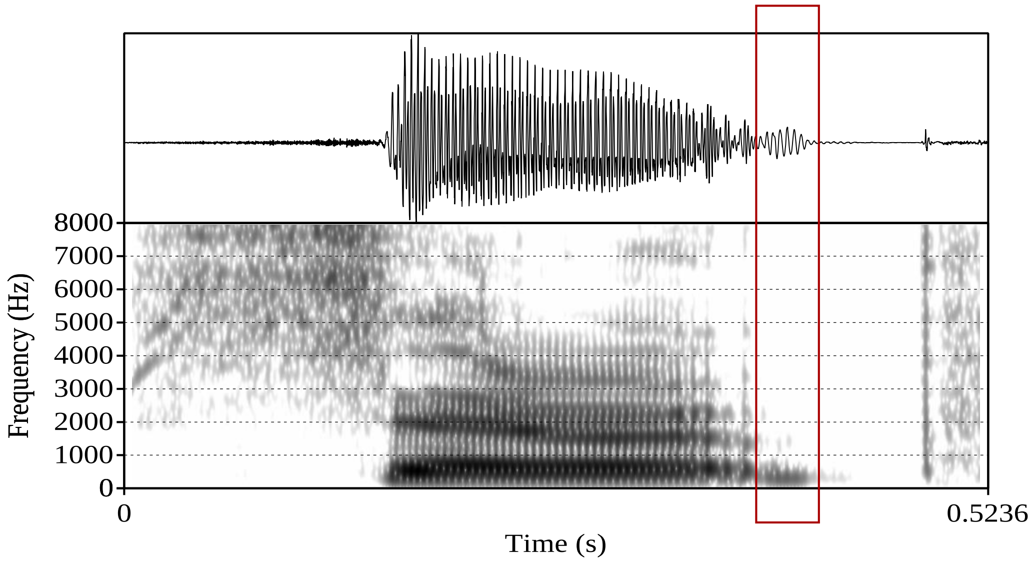

Figure 6. Spectrogram and waveform of the word *syrup* containing glottalisation and a glottal squeak at the end of the vowel. Red box shows the region of the glottal squeak, which has a mean f0 of 241 Hz. The preceding unstressed vowel has a mean f0 of 218 Hz prior to the onset of glottalisation.

2.2.3 Analysis 1 Discussion

The analysis of glottalisation according to coda stop voicing for alveolar and bilabial POAs suggests that glottalisation is employed as a cue to coda stop voicelessness in word final unstressed syllables containing schwa. This supports the hypothesis that, in line with findings for other varieties of English such as BrE and AmE (Huffman, 2005; Pierrehumbert, 1994; Roach, 1973; Seyfarth & Garellek, 2015; Wells, 1982), glottalisation cues coda stop voicelessness generally in AusE, rather than specifically for alveolar stops. Future work comparing voiced and voiceless stops in an extended set of environments including both voiced and voiceless velar coda stops will be required to confirm the suggestion that glottalisation as a cue to voicelessness holds for all three POAs.

We hypothesised that /t/ would exhibit more glottalisation than the other voiceless stops; however, all three POAs demonstrated similar proportions of glottalised tokens and POA did not affect the likelihood of glottalisation occurring. This is in contrast to London English and to AmE, as in both varieties /t/ is more frequently glottalised than the other voiceless stops (Huffman, 2005; Pierrehumbert, 1994; Tollfree, 1999; Wells, 1982). Although glottalisation of velar and bilabial stops has previously been reported for AusE (Tollfree, 2001), it was suggested to occur less frequently for /k/ and only occasionally for /p/. This differs from the pattern shown in our results; however, it must be remembered that the environment examined here was different from the other studies, as we have exclusively examined unstressed syllables in isolated words.

We hypothesised that coda voicing-related vowel duration differences would be reduced in the unstressed environment and that, as a consequence, higher rates of glottalisation may be present compared to what has previously been reported for stressed contexts. We found a mean unstressed vowel duration difference of 18 milliseconds, which was measured from the F3 trough of the preceding rhotic and as such incorporates part of the consonant. Although the difference was smaller than what has been previously reported for young female AusE speakers in stressed contexts, our analysis found that unstressed vowels were nevertheless significantly longer preceding voiced coda stops than before voiceless coda stops. Penney et al. (2018) found a mean vowel duration difference of 46 milliseconds (24 milliseconds when only short vowels were considered) between voiced and voiceless coda contexts in stressed syllables. Recall also that Penney et al. (2018) showed that coda voicing-related vowel duration differences were reduced in their young speakers compared to older speakers. So while coda voicing-related durational differences may be reduced in the unstressed environment, they do not appear to be markedly more so than in the stressed environment. In addition, although we found numerically higher rates of glottalisation than those reported in Penney et al. (2018) for stressed monosyllables, this difference was not found to be significant when compared to young female speakers in the stressed syllable context. In both stressed and unstressed contexts young female speakers exhibited high rates of glottalisation preceding voiceless codas.

Additional support for the idea that glottalisation may help to preserve the voicing contrast comes from the analysis of F0 above. Although we found small differences in F0 in the expected direction for voiced and voiceless stops, these differences were not found to be significant in our model, lending further support to the idea that glottalisation is taking on more of the 'heavy lifting' to maintain the coda voicing contrast in this environment. Note, however, that we observed at least some prolonged voicing in the closure period for the majority of voiced stops (though only relatively few tokens exhibited full voicing throughout closure). Penney et al. (2018) also found a voice bar to be frequently present in voiced coda contexts in their data and suggest that glottalisation and voice bar may serve as complementary cues to coda stop voicing, at least in laboratory speech. The results presented here support this suggestion of complementarity.

Previous research has suggested that glottalisation is less likely to occur in high vowel contexts (Brunner & Żygis, 2011; Malisz et al., 2013; Penney et al., 2018). Although there were examples of raised unstressed vowel realisations in the data, particularly in the velar coda context, we found no relationship between F1 and the presence of glottalisation. This is not surprising, given that the F1 variance in the unstressed vowels was minimal, as shown by the lack of significant effect of POA on F1 in the analysis of formants above. It is possible that examination of unstressed vowels that display more extreme height differences may produce different results.

Analysis 1 has shown that glottalisation occurs in conjunction with voiceless stops in unstressed syllables at all three POAs in AusE, at a level that is comparable to rates of glottalisation in stressed syllables in an alveolar coda context. We also found that the coda voicing-related vowel duration differences are reduced in unstressed contexts, although a significant difference in vowel duration between voiced and voiceless coda contexts remains. It should of course be borne in mind that this analysis is based on analysis of data in a very narrow, controlled environment; accordingly, it is not clear to what extent these results may be generalisable to other contexts. In addition, this analysis examined only young, female speakers. The following analysis is an extension to a different cohort, made up of both male and female speakers from different age groups, in order to examine more closely whether and to what extent glottalisation in unstressed syllables is affected by the factors of age and gender.

2.3 Analysis 2: Older and Younger Sydney Speakers from the AusTalk Corpus

2.3.1 Method

2.3.1.1 Data

Data for Analysis 2 were extracted from AusTalk (Burnham et al., 2011), which is a large corpus of AusE speech that was collected at multiple locations using standardised equipment and procedures between 2011 and 2015. It comprises audio and visual data from 973 speakers aged from 18 to 83 from a range of regional and social backgrounds. The audio data used for the current analysis were sampled at 44.1 kHz via an AudioTechnica headworn AT892c microphone through an M-Audio

FastTrackUltra8R digital recording interface. Specific details about the recording process, setup, equipment and hardware can be found in Burnham et al. (2011). Each speaker was recorded in both standardised and spontaneous speech tasks (Burnham et al., 2011; Cassidy, Estival, & Cox, 2017). One of the standardised tasks was a word list task, in which 322 isolated words were read in random order from a computer screen. For the present analysis trochees with the form /CVɹəC/ with final voiceless coda stops at bilabial, alveolar or velar POA were extracted from the word list recordings. The words analysed here are *syrup, parrot, barrack.* Most speakers attended three sessions, and so produced three repetitions of each word; however, for some speakers only two tokens could be extracted. As in Analysis 1 above, the selected items will enable us to analyse potential differences in rates of glottalisation for words ending in stops at all three POAs. A total of 512 items (bilabial: 174; alveolar: 170; velar: 168) were examined.

2.3.1.2 Speakers

Data for this analysis were extracted for 61 speakers in two age groups: an older group aged above 55 years (*n* = 27, 12 female, 15 male; mean age: 65; SD: 8) and a younger group aged between 18 and 35 years (*n* = 34, 16 female, 18 male; mean age: 25; SD: 5). All of the speakers were born in and had completed their entire school education in Sydney. The inclusion of an older and a younger group will allow us to compare potential differences between the age groups in the implementation of voicing, particularly as glottalisation has been suggested to be a recent development in AusE (Penney et al., 2018). Hence, we may expect to find differences between older and younger speakers, particularly if there has been a progression of change across POA. Furthermore, the inclusion of male and female speakers in both age groups will enable us to examine whether differences related to gender play a role in the how glottalisation is employed.

As in Analysis 1, this design allows for a broad-based population-level analysis rather than a more focussed individual-level analysis. However, as individuals are likely to vary in their use of glottalisation, cross speaker variability will be reported below as appropriate.

2.3.1.3 Acoustic analysis

All tokens in the data set were acoustically analysed using the same methods as described in the acoustic analysis section from Analysis 1 above. Vowel formants were also measured as outlined above; however, as this data set contains productions from both male and female speakers, normalisation of the formant measurements was also necessary to account for physiological differences between genders. The data were Lobanov normalised using the *vowels* package (Kendall & Thomas, 2014) in R (R Core Team, 2016).

As in Analysis 1 above, 10 per cent of the data were randomly selected and relabelled by the first author to assess intra-annotator reliability. A second trained annotator also re-labelled 10 per cent of the data to assess inter-annotator reliability. Using the *irr* package (Wolak et al., 2012), Intra-class correlations and 95% confidence intervals were calculated for vowel duration. For the categorical measure of the presence or absence of glottalisation a Cohen's Kappa score was calculated. Table 4 below contains the results for inter-annotator reliability and intra-annotator reliability and shows that reliability was high in all cases.

Table 4. Intra-class correlation results for vowel duration and Cohen's Kappa results for the presence of glottalisation.

	Vowel duration			Presence of glottalisation	
	Inter-annotator	Intra-annotator		Inter-annotator	Intra-annotator
ICC	0.889	0.935	Kappa	0.964	1
p value	< .001	< .001	*p* value	< .001	< .001
95% CI	0.819, 0.933	0.889, 0.962	*z* score	7.28	7.14

2.3.1.4 Statistical analysis

As in Analysis 1 above, we fitted mixed models to the data using the *lme4* package (Bates et al., 2015) in R (R Core Team, 2016). The following analyses were conducted:

- Linear mixed effects modelling was used to conduct an analysis of vowel formants, with the relevant formant (i.e. either F1 or F2) as the dependent variable;
- Logistic mixed effects modelling was used to conduct an analysis of glottalisation in voiceless codas, with the presence of glottalisation as the dependent variable.

As above, we included random intercepts for speaker and repetition and random slopes for repetition by speaker in each model. We initially included all fixed factors and their interactions, then pruned the models after carrying out model comparisons with the *anova()* function to remove terms that did not significantly improve the model, beginning with non-significant interactions. Post-hoc analyses were carried out with Tukey HSD corrections for multiple comparisons. Full details and the syntax of the most parsimonious model for each analysis will be reported below.

2.3.2 Results

2.3.2.1 Analysis of vowel formants

As in Analysis 1, we first examined the F1 and F2 measurements of the unstressed vowels in all 512 tokens to identify whether the POA of the following stop had an effect on the vowels' phonetic height and fronting. We were also interested in whether this differed according to age group and/or gender. Figure 7 illustrates the mean F1 and F2 in Bark (converted from the Lobanov normalised measurements) for each POA for all tokens produced by the speakers according to age group and gender. We fitted linear mixed effects models separately for F1 and F2. In each model the relevant formant measurement (i.e. F1 or F2) was the dependent variable, with POA, age group, gender and their interactions included as fixed factors. The syntax for the most parsimonious F1 model was: lmer(F1 ~ (poa + age group + gender)^3 + (1+repetition|speaker) + (1|repetition)). The syntax for the most parsimonious F2 model was: lmer(F2 ~ (poa + age group + gender)^2 + (1+repetition|speaker) + (1|repetition)).

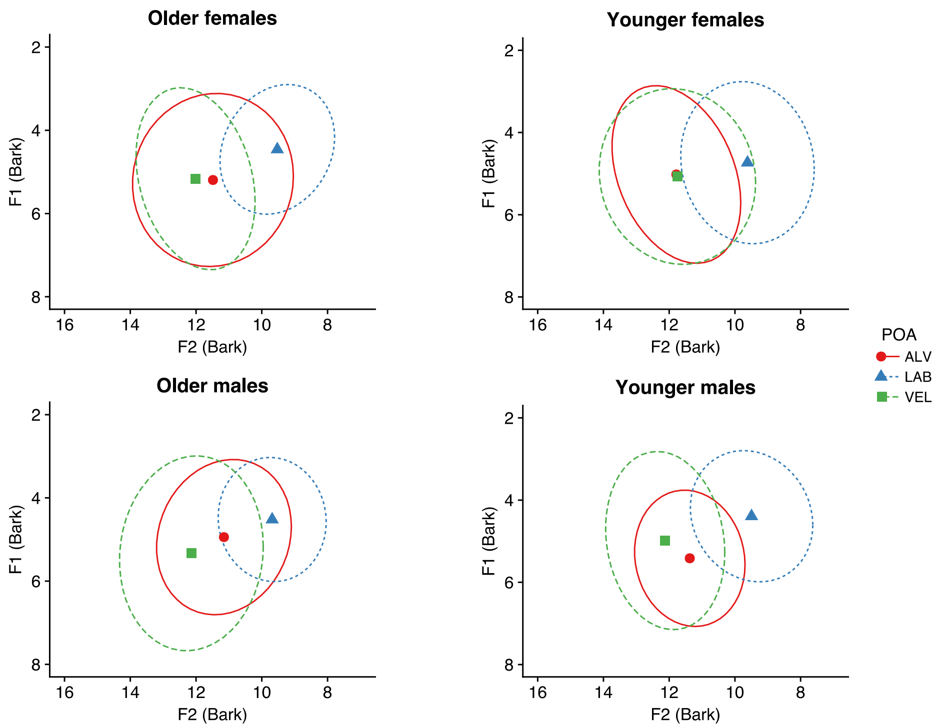

Figure 7. Mean F1 and F2 values (Bark) for unstressed vowels according to POA of following stop. Upper panels represent female speakers; lower panels represent male speakers. Left panels represent older speakers; right panels represent younger speakers. Shapes represent centroids. Ellipses represent 95% confidence intervals.

In the analysis of F1 data we a found significant effect for POA ($F(2,444) = 37.614$; $p < 0.0001$); in general, unstressed vowels before bilabial stops were produced with lower F1 (i.e. phonetically higher realisations) than those before stops at the other POAs. There was also a significant three way interaction between POA, age group, and gender ($F(2,444) = 5.936$; $p = 0.003$). Post hoc analyses showed that within each age and gender group all of the differences were related to the bilabial POA, which was significantly phonetically higher than alveolar and velar POAs for the older females (alveolar–bilabial: $p = 0.005$; bilabial–velar: $p = 0.006$) and the younger males (alveolar–bilabial: $p < 0.0001$; bilabial–velar: $p = 0.018$), and than velar POA for the older males ($p = 0.001$). For the younger females there were no significant differences between the POAs.

The F2 analysis also showed a significant effect of POA ($F(2,502) = 419.612$; $p < 0.0001$), as well as a significant two way interaction between POA and gender ($F(2,502) = 8.373$; $p = 0.0003$). In general, the bilabial context showed lower F2 values than the other POAs, as expected (Modarresi et al., 2005). Post hoc analyses showed significant differences in F2 between the male and female speakers before the alveolar POA ($p = 0.011$), with the male speakers showing lower F2 (i.e. more retracted productions) than the female speakers in this context. The male speakers showed significant differences between all three POAs (all $p < 0.0001$). In contrast, the female speakers showed significant differences between the bilabial and alveolar and between the bilabial and velar contexts (both $p < 0.0001$), but not between alveolar and velar.

To summarise, unstressed vowels were more retracted in the bilabial contexts for all groups, and higher in the bilabial context for all but the younger females. In addition, the male speakers produced more retracted unstressed vowels in the alveolar context compared to the female speakers.

2.3.2.2 Analysis of glottalisation in voiceless codas

In total, 60% of the alveolar stops, 51% of the bilabial stops, and 51% of the velar coda stops were glottalised. Twenty of the 28 female speakers glottalised at least 50% of tokens, with five of the female speakers categorically producing glottalisation (three from the younger group and two from the older group). Three of the female speakers (all from the older age group) produced no glottalisation. 13 of the 33 male speakers produced glottalisation in at least 50% of tokens, none produced glottalisation categorically, and two of the male speakers (both from the older age group) produced no glottalisation. Table A3 in the Appendix provides details of the individual speakers' rates of glottalisation according to age group and gender. We fitted a logistic mixed effects model to all of the data (recall that this data set contained only items in the voiceless coda context) to analyse potential differences in rates of glottalisation for words ending in voiceless stops at all three POAs, as well as differences related to gender and to age group. The presence of glottalisation was the dependent variable. The POA of the final coda stop (bilabial, alveolar, velar), age group, gender and their interactions were included as fixed factors.[11] The syntax for the most parsimonious model was as follows: glmer(glottalisation ~ (poa + age group + gender)^2 + (1+repetition|speaker) + (1|repetition)). The model showed significant effects for age group ($\chi^2 = 12.254$; $p = 0.001$) and for gender ($\chi^2 = 10.502$; $p = 0.001$), as well as a significant interaction between POA and age group ($\chi^2 = 6.174$; $p = 0.047$).

The significant effect for age group reveals that glottalisation was employed more frequently by the younger speakers (younger speakers 64%, older speakers 41% glottalised tokens). The gender effect shows that the male and female speakers utilised glottalisation differently (females 65%, males 46% glottalised tokens). Both of these effects support the suggestion that glottalisation is a recent change to AusE (Penney et al., 2018) and may indicate a change being led by young women.

[11] Note than we initially included phonetic vowel height (measured by F1) here as we found differences in F1 according to POA in our formant analysis. However, vowel height of the preceding vowel was not significant and its inclusion did not improve the model fit ($\chi^2 = 2.3392$; $p = 0.801$), and as such was not included in the final model.

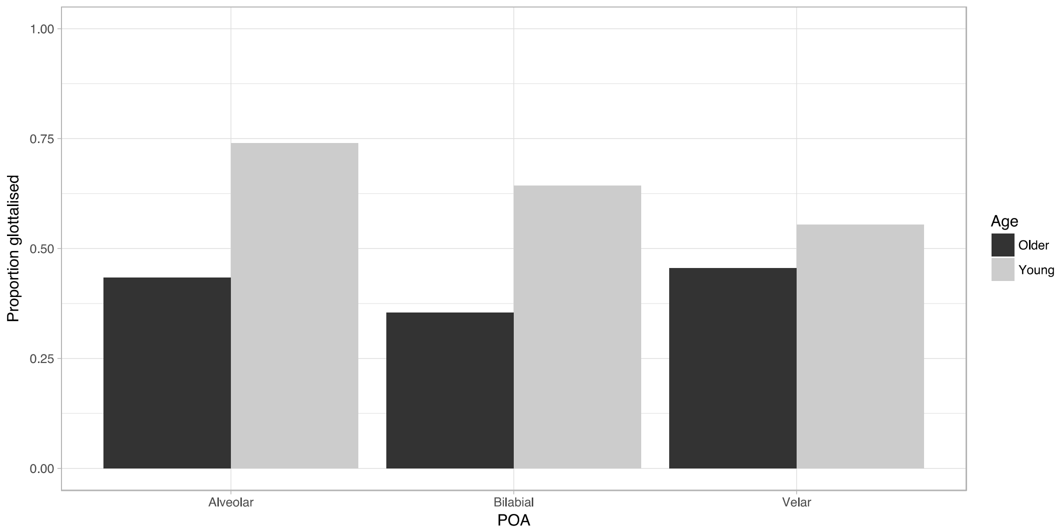

Figure 8. Proportion of tokens glottalised in each POA according to age group.

The significant interaction between POA and age group indicates that the two age groups vary in the incidence of glottalisation at different stop POAs. Figure 8 illustrates the proportion of glottalised tokens within each POA for older and younger speakers. As can be seen, in each POA the younger speakers produced more glottalised tokens than the older speakers. However, the younger speakers show a preference for glottalisation in the alveolar context (74% glottalised), followed by bilabial (64% glottalised), with the velar context showing the least amount of glottalisation (55% glottalised). Older speakers do not exhibit this same preference for POA; instead, they glottalised at similar rates in both the velar (46% glottalised) and alveolar (43% glottalised) contexts, and least in the bilabial context (35% glottalised). Post hoc analyses showed that within each age group there were no significant differences between POAs. However, they also showed that the older and younger speakers differed significantly from each other in the alveolar (p = 0.007) and bilabial (p = 0.008) contexts, but did not differ in the velar context.

To summarise, female speakers produced glottalisation at higher rates than male speakers, and younger speakers produced glottalisation at higher rates than older speakers, particularly in the alveolar and bilabial coda contexts.

2.3.2.3 Comparison with stressed syllables

As in Analysis 1, we then conducted a comparison between these data and the data from the stressed syllable context reported in Penney et al. (2018), which used the same cohort of speakers (n = 67; younger female: 17; younger male: 19; older female: 14; older male: 17). As only alveolar contexts were included in the stressed context data, we included only the unstressed syllables in the alveolar context in this comparison. We fitted a logistic mixed effects model with the dependent variable presence of glottalisation, and fixed factors stress (i.e. whether the item was produced in a stressed or unstressed syllable), age group, and gender. Interactions between the fixed factors were also included but these did not improve the model fit and they were not included in the final model. The syntax for the most parsimonious model was as follows: glmer(glottalisation ~ stress + age + gender + (1+repetition|speaker) + (1|repetition)). The words included in this analysis were *parrot* for the unstressed syllables, and *heat*, *hit*, *heart*, *hut*, *hort*, *hot*, *hoot* for the stressed syllables. 1418 tokens (stressed: 1248; unstressed: 170) were included in this analysis. We found significant effects for age group ($\chi^2 = 41.422$; $p < 0.0001$), showing that younger speakers glottalised more than older speakers in both contexts, and for gender ($\chi^2 = 6.814$; $p = 0.009$), demonstrating that the female speakers glottalised more than the male speakers. As in Analysis 1, we found no effect of stress ($\chi^2 = 0.725$; $p = 0.394$), suggesting that rates of glottalisation were comparable across age and gender groups in both stressed and unstressed environments.

2.3.2.4 Comparison with Analysis 1

As the rates of glottalisation found in Analysis 2 were numerically smaller than those found in Analysis 1 above, we carried out a further comparison to test whether there was a statistical difference between the two cohorts of speakers. To enable a comparison with the speakers in Analysis 1 (n = 28), we examined only the productions of the young female speakers included in Analysis 2 (n = 16). Figure 9 shows the proportion of glottalised tokens containing a voiceless coda stop at each POA according to the respective data set. As can be seen, the speakers in Analysis 1 glottalised at essentially the same rate in each POA, while the speakers from Analysis 2 glottalised slightly less in the bilabial and velar contexts than they did in the alveolar context. We fitted a logistic mixed effects model to the data with presence of glottalisation as the dependent variable and fixed factors POA and data set (i.e. Analysis 1 or 2). An interaction between the fixed factors was also included but this did not improve the model fit and was removed from the final model. The syntax for the most parsimonious model was as follows: glmer(glottalisation ~ poa + corpus + (1+repetition|speaker) + (1|repetition)).

The results showed no significant effect for POA (χ^2 = 0.172; p = 0.918). Overall, the speakers from Analysis 1 glottalised more frequently than the speakers from Analysis 2 (83% compared to 71%), with a trend towards significance for data set (χ^2 = 3.223; p = 0.073). Despite the differences, in both sets glottalisation was produced at each POA at very high rates by the young female speakers.

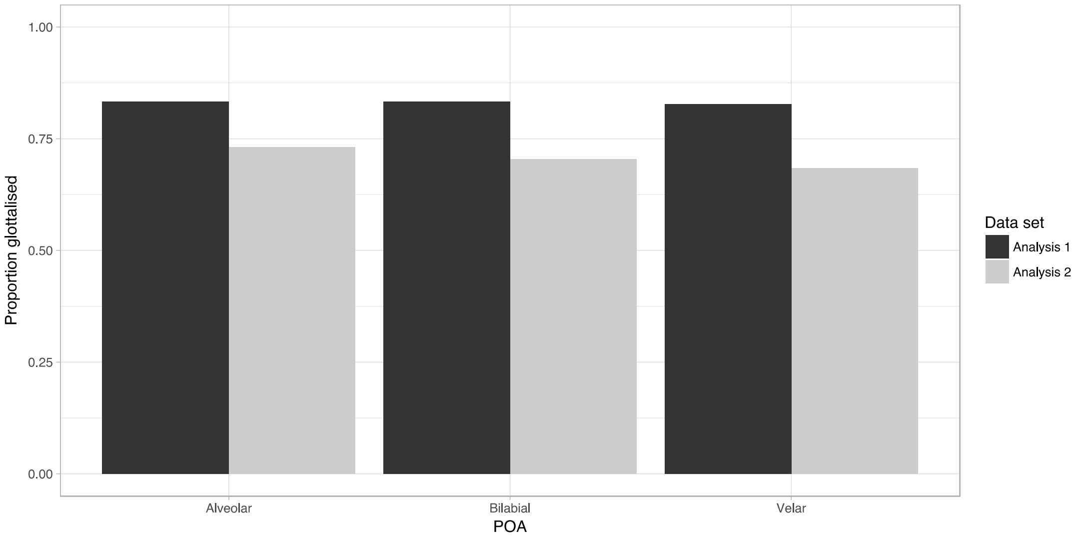

Figure 9. Proportion of glottalised tokens with voiceless coda stops according to POA for young female speakers in two data sets.

2.3.2.5 Non-canonical stop realisations

As in Analysis 1, there were a number of non-canonical stop realisations represented in the data; these are summarised in Table 5 below. 12% of the tokens (*n* = 61/512 produced by 24 speakers) contained unreleased coda stops (i.e. they showed no evidence of a release burst). Of these unreleased stops, more than half were also glottalised. Formant transitions at the end of the unstressed vowel were observable for all but three of these glottalised unreleased stops, indicating that although no release was visible the vast majority of cases nevertheless involved a supralaryngeal articulation. In the younger speakers, the majority of the unreleased stops had alveolar POA, though unreleased bilabial tokens also occurred; unreleased velar tokens were rare (alveolar: 29; bilabial: 18; velar: 2). In the older speakers, bilabials were more often unreleased than alveolars, and there were no unreleased velar tokens (alveolar: 3; bilabial: 9). The small number of unreleased velar stops is consistent with Analysis 1 above and with previous literature which suggests velar stops are released more frequently than stops at the other POAs (Byrd, 1993; Crystal & House, 1988a). There were also some examples of spirantised stops, though these accounted for only 3% of all tokens (*n* =17/512 produced by 11 speakers). As was found in Analysis 1, spirantisation occurred almost exclusively in the alveolar context though there were two occurrences in conjunction with velar stops. Glottalisation was also present in four of the spirantised tokens. The majority of the spirantised stops were produced by males (11/17), and those that were produced by females were produced by older females. Penney et al. (2018) also found that male speakers used spirantisation more than females. These results contrast with previous research based on speakers from Melbourne, which found spirantisation of /t/ (fricated /t/) to be associated with female speakers (Jones & McDougall, 2009; Loakes & McDougall, 2010), possibly suggesting a regional difference.

Another 3% of the tokens (*n* = 15/512 produced by eight speakers) contained preaspiration. The majority of these were produced by older speakers, and they occurred in alveolar contexts, though there were two examples for velar stops. In contrast to Analysis 1, where half of the 17 preaspirated stops were also glottalised, only one preaspirated token occurred in conjunction with glottalisation. Glottal squeaks were rare in this data set, accounting for only 1% of the data (*n* = 7/512 produced by five speakers). As in Analysis 1, each glottal squeak occurred in conjunction with either glottalisation or phrase final creak. Five of the seven squeaks were found in the alveolar context; there was also one squeak observed in each of the bilabial and velar contexts. A single female produced three of the squeaks, with the remaining squeaks each produced by a different speaker. The squeaks found here exhibited similar characteristics to those described in Analysis 1; they were short in duration (mean: 21 ms) and occurred shortly after the stop closure period began (mean duration from closure: 30 ms; mean duration from release: 66 ms). The squeaks in this Analysis had a mean F0 of 295 Hz (female: 301Hz; male: 281Hz).

Table 5. Number of non-canonical stop realisations according to age group and gender. Brackets indicate number of tokens occurring in conjunction with glottalisation.

	Unreleased	Spirantised	Preaspirated	Squeak
Older female	4/103 (1)	6/103 (0)	5/103 (0)	0/103 (0)
Older male	8/129 (3)	7/129 (2)	7/129 (0)	0/129 (0)
Younger female	26/123 (17)	0/123 (0)	2/123 (0)	5/123 (3)
Younger male	23/157 (14)	4/157 (2)	1/157 (1)	2/157 (1)
Total	**61/512 (35)**	**17/512 (4)**	**15/512 (1)**	**7/512 (4)**

2.3.3 Analysis 2 discussion

One of our hypotheses was that younger speakers would produce glottalisation in unstressed syllables at higher rates than the older speakers (previously found for alveolar stops in stressed syllables), which would be consistent with a recent change to AusE. As stated above, glottalisation has been noted in the AusE literature since the late 1980s (Haslerud, 1995; Ingram, 1989; Tollfree, 2001), whereas prior to this it was considered to be absent from the variety (Trudgill, 1986; Wells, 1982). The results presented support the hypothesis and show that younger speakers are significantly more likely to glottalise than older speakers in the unstressed syllable context examined here. In addition, the results provide additional support for the claim that glottalisation is a recent change to AusE (Penney et al., 2018), with younger speakers leading the change. Of course, real time analysis would be required to discount the possibility that the age-related differences are due to age grading effects (Bailey, 2002); likewise, an examination of historical data may be able to shed more light on the processes of how and when this change entered the variety.

The results also demonstrate that female speakers produced glottalisation more frequently than male speakers. While previous studies have found gender effects related to glottalisation in other varieties of English (Holmes (1995), Mees (1987, 1990), and Redi and Shattuck-Hufnagel (2001) all found females exhibited more glottalisation than males, though Milroy et al. (1994) found the opposite), this has not previously been shown to be the case in AusE. Penney et al. (2018) found no evidence of a gender effect; however, their analysis examined glottalisation associated with alveolar coda stops in stressed syllables only. In the present study, we explored glottalisation in unstressed syllables for all three English stop POAs. We might speculate that the gender difference could be due to the different stress contexts, though we found that rates of glottalisation were similar between stressed and unstressed syllables for alveolar stops. Perhaps the difference is rather due to our examination of a greater number of contexts; though we did not find an interaction between gender and POA, Figure 9 shows that the younger speakers produced numerically less glottalisation in the non-alveolar contexts, which may suggest that the relevant gender differences lie within the non-alveolar POAs. Figure A1 in the Appendix also shows that the differences between males and females are greater in the non-alveolar contexts. The gender effect that we found may be interpreted as supporting the notion that women are the drivers of this change to AusE. This would not be surprising, given that women are often at the forefront of sociophonetic change and have been shown to adopt new features sooner than men (Labov, 1990). Of course, other explanations may also be possible; for example, it may be that females glottalise more frequently than males for physiological reasons. For example, women may make greater use of a raised larynx as a strategy to reduce transglottal pressure differences in order to cease voicing but this suggestion requires empirical examination. Such a strategy may introduce a bias for glottalisation (Moisik, 2013).

Analysis 2 adds further support to the finding in Analysis 1 that glottalisation is employed for voiceless coda stops at all three POAs in unstressed syllables in this variety of English. Glottalisation was present in at least half of all tokens examined at each POA suggesting that glottalisation is an important cue for voiceless coda stops in unstressed syllables, regardless of POA. We hypothesised that /t/ would exhibit more glottalisation than the other voiceless stops; although in Analysis 2 we found glottalisation to be numerically more frequent for alveolar stops compared to the other POAs in the younger group, we found no significant differences in rates of glottalisation according to POA within either age group. As in Analysis 1 above, this finding contrasts with the pattern of /t/ being most frequently glottalised in other varieties of English, such as London English (Tollfree, 1999; Wells, 1982) and AmE (Huffman, 2005; Pierrehumbert, 1994; Seyfarth & Garellek, 2015). Tollfree (2001) suggests that glottalisation is frequently present for /k/ for particular lexical items, and occasionally present for /p/ in AusE. Based on this, we anticipated evidence of a progression of change from alveolar to the other POAs. We did find that the younger speakers differed significantly from the older speakers in the alveolar and bilabial contexts, but not in the velar context. The alveolar and bilabial were also the contexts that showed the highest proportions of glottalisation for the younger speakers. This may indicate some marginal support for a progression of change. Alternatively, the difference in findings may be due to the unstressed context investigated here.

As in Analysis 1 above, we found no evidence of an effect of unstressed vowel height on the presence of glottalisation, which is not altogether surprising as all of the tokens contained a schwa. We also hypothesised that increased rates of glottalisation may be present in unstressed syllables compared to stressed syllable codas, given vowel duration cues may be reduced in such contexts (Crystal & House, 1988b; Davis

& van Summers, 1989; Klatt, 1975). The younger speakers in Analysis 2 produced glottalisation in unstressed syllables containing coda /t/ at a comparable rate to that reported for stressed syllables in Penney et al. (2018), with a slight increase in the unstressed context (unstressed: 74%; stressed: 71%). The older speakers in this analysis produced glottalisation at higher rates than has been reported for stressed syllables (unstressed: 43%; stressed: 36%). However, as in Analysis 1 above, the comparison of glottalisation of coda /t/ in stressed and unstressed syllables showed no difference for any of the age or gender groups, suggesting that rates of glottalisation, though slightly increased, were statistically comparable between the two contexts.

2.4 General discussion

Taken together, the two analyses here demonstrate that glottalisation occurs for voiceless stops at all three POAs in AusE utterance final unstressed syllable codas. Therefore, it seems that glottalisation serves as a cue to coda voicelessness generally for stops at each POA, and is not specifically related to alveolar stops. In contrast to other varieties of English such as London English and AmE, where glottalisation occurs most frequently with alveolar stops, we found no significant differences in rates of glottalisation between the POAs. It may be possible that glottalisation initially occurred for alveolars in AusE but has now spread to the other stops as well. Glottalisation of /t/ has been noted in the AusE literature since the late 1980s (Haslerud, 1995; Ingram, 1989), whereas the only mention of glottalisation of /p, k/ prior to this paper was in Tollfree (2001), where glottalisation of non-alveolar stops was suggested to be less common. In Analysis 2 we found that the younger speakers differed from the older speakers for /t/ and /p/, and these contexts were also those that were numerically most frequently glottalised in the younger speakers. Taken together with Tollfree's findings, this may provide some limited support for a progression from alveolar to stops at other POAs.

Analysis 2 showed that younger speakers were more likely to glottalise than older speakers for stops at all three POAs. This is in accord with previous findings and supports the idea that glottalisation is a recent change to AusE, though alternative explanations may also be possible (e.g. age grading, physiological differences between older and younger speakers). In addition, glottalisation occurred at high rates for stops at each POA in the younger female speakers in both of the analyses. Combined with the finding from Analysis 2 that females glottalised more than males, these results may support the idea that glottalisation is a recent change led by young female speakers. Tait and Tabain (2016) also found glottalisation to be more common in female than male primary school aged children, perhaps indicating an early onset of a gender difference.

Unsurprisingly given the unstressed context analysed here, we found that preceding unstressed vowel height did not have an effect on the presence of glottalisation. Vowel height has previously been linked to glottalisation, with high vowels less likely to be glottalised (Brunner & Żygis, 2011; Malisz et al., 2013; Penney et al., 2018). In our examination of unstressed syllables we did not find a height effect, although our data were restricted to schwa contexts and, hence, minimal height variance is to be expected.

In both of the analyses we found that glottalisation in unstressed syllables occurred at comparable rates to stressed syllables. Penney et al. (2018) found evidence of a trading relationship between glottalisation and vowel duration, whereby rates of glottalisation increased in combination with decreased coda voicing-related vowel duration differences. Such a trading relationship would predict increased glottalisation in unstressed contexts, if vowel duration differences were reduced compared to stressed contexts. In Analysis 1 we found a reduced vowel duration difference between voiced and voiceless coda contexts compared to stressed contexts, though a small but significant difference was maintained. The high rates of glottalisation observed in conjunction with reduced coda-voicing related vowel duration differences lends supports to the idea that glottalisation may assist in maintaining the phonological coda voicing contrast if the vowel duration cue is minimised. This is further supported by the F0 analysis, which found that differences between the voiced and voiceless coda contexts appear to have been minimised in the data examined here. As all of the tokens examined were produced as single words, and hence all were in utterance final position, this finding is also partially in accord with Keyser and Stevens (2006), who suggest that voiceless stops at all three POAs may show glottalisation in final position to enhance the voicing contrast. Note, however, that they posit that glottalisation may be necessary as subglottal pressure is reduced in final position and hence may not support voicing in the stop closure period thereby undermining the voicing contrast. We found a voice bar in the majority of voiced coda stops, suggesting that this particular cue was not reduced in our data, though, as noted above, the vowel duration and F0 cues were reduced.

We found some evidence of differential rates of glottalisation between the young female speakers in Analysis 1 and 2, despite these speakers being matched for age and gender. Although these differences were not found to be significant in our modelling, it is possible that further variation may exist in the community and may be attributable to social factors that we have not accounted for in our analysis, such as differences in socioeconomic status, regional (area specific) variation, or differing ethno-cultural affiliations. Tollfree (2001) found differences in rates of glottalisation of /t/ related to socioeconomic status in conversational speech of AusE teenagers, and Cox and Palethorpe (1998, 2011) have shown differences between speakers from different regions of Sydney and between those with different ethno-cultural backgrounds. It is also possible that individual differences may play a role.

In both analyses speakers produced a range of non-canonical stop realisations including unreleased stops, which were often associated with glottalisation. Previous research has found that glottalisation often co-occurs with unreleased stops (Blevins, 2006; Kahn, 1976; Selkirk, 1982) and it has been suggested that glottalisation may promote perception of the stop's POA through increasing the amplitude of the formants where transitions occur (Garellek, 2011). Spirantised stops were also present in both analyses: in Analysis 1 there were examples of spirantised stop realisations produced by the young females; in Analysis 2, the young females were not responsible for any of the spirantised stops which were instead produced by males. Previous research on speakers from Melbourne has linked spirantisation to female speakers, particularly females from high socio-economic backgrounds (Jones & McDougall, 2009; Loakes & McDougall, 2010; Tollfree, 2001), though Penney et al. (2018) also found that male speakers produced spirantised realisations more frequently than female speakers in their data from Sydney. Spirantisation is perhaps related to individual speaker differences rather than gender.

Tokens exhibiting preaspiration were also present in each of the analyses. In Analysis 2 these were mainly produced by older speakers and did not occur with glottalisation. Curiously, however, in Analysis 1 over half of the cases preaspiration occurred in conjunction with glottalisation. Although it might seem that preaspiration and glottalisation may be very separate strategies for achieving voicelessness: the former through glottal abduction, and the latter through glottal adduction, it may be possible that in cases where the two strategies are combined, preaspiration occurs when the vocal folds are abducted following the final glottal closure associated with glottalisation (before the oral closure is made). However, Figure 5 seems to suggest that, rather than occurring in sequence, these two phenomena overlap, as can be seen by the aperiodic energy associated with the aspiration occurring at the same time as the irregular phonation due to glottalisation. This may be explained by models of laryngeal activity that consider the entire larynx, rather than focussing solely on the glottis, such as those posited by Edmondson and Esling (2006) and Moisik and Esling (2011). According to this view, glottalisation may be achieved through epilaryngeal constriction (i.e. ventricular incursion), which could take place at the same time as the vocal folds are abducted, resulting in preaspiration.[12]

Finally, both data sets contained examples of glottal squeaks. Interestingly, squeaks were produced exclusively by young speakers in each of the analyses. In Analysis 2, more squeaks were produced by female speakers, though the female speakers in Analysis 1 produced more squeaks than the females in Analysis 2. Redi and Shattuck-Hufnagel (2001) and Hejná et al. (2016) suggest that squeaks are more likely to be produced by female speakers, and are speaker specific; that is, certain (mostly female) speakers tend to produce squeaks in conjunction with glottalisation, but they are not necessary in order to produce glottalisation. 11 of the 28 speakers in Analysis 1 did produce at least one squeak, which to our knowledge is a higher proportion than has previously been reported (and, indeed, is higher than in Analysis 2). Why this might be so is uncertain; however, it does suggest future examination of glottal squeaks in AusE speakers may reveal interesting patterns of individual stop

[12] We are grateful to an anonymous reviewer for bringing this possibility to our attention.

realisation.

It should again be pointed out that this examination was based on small sets of data collected in highly controlled contexts, and therefore the interpretations offered in this paper may not be generalisable to other, less restricted contexts. Accordingly, in our future work we plan to explore glottalisation in a greater variety of contexts and with speakers from a wider range of social backgrounds. In addition, it would be interesting to extend this research to examine when children begin to use glottalisation as a cue to coda voicing. This may help to determine whether glottalisation has a primarily social or physiological basis if prepubescent boys and girls whose anatomical structures are not yet differentiated use glottalisation differently. Future work may also further explore the suggestion that glottalisation is a recent change to AusE, by analysing archival data collected at different time points since the 1980s.

2.5 Conclusion

This study has shown that glottalisation occurs in conjunction with coda voiceless stops at each POA in utterance final unstressed syllables in single words in AusE. Glottalisation was shown to occur at high rates in unstressed contexts, in conjunction a reduction in the use of vowel duration as a cue to coda voicing. In the face of a reduced vowel duration cue to coda stop voicing glottalisation may be utilised to cue voicelessness. Younger speakers used glottalisation more than older speakers, and females were more likely to glottalise than males, both results supporting previous suggestions of a recent change. Further research is needed to examine whether the patterns found here are replicated in unscripted speech. In addition, an analysis of the links between production and perception will further our understanding of potential trading relationships between cues to coda stop voicing.

2.6 Acknowledgements

This research was supported in part by an Australian Government Research Training Program scholarship to the first author. Parts of this research were presented at the 2017 Australian Linguistic Society annual conference at the University of Sydney, and the fifth New Ways of Analyzing Variation – Asia Pacific conference (NWAV-AP5) at the University of Queensland; we thank audience members for their comments and questions. In addition, we would like to thank Steve Cassidy for assistance with AusTalk, Kelly Miles for additional annotation, and members of the MQ phonetics lab for their feedback. We also thank Amalia Arvaniti, Gerry Docherty, and two anonymous reviewers for their insightful comments and suggestions.

2.7 Appendix

Table A1. Number of tokens by word in each analysis.

	Analysis 1	Unstressed vowel duration	Glottalisation in voiced/voiceless codas	Glottalisation in voiceless codas	Fundamental frequency	Analysis 2
Arab	84	–	84	–	83	–
Arid	83	–	83	–	82	–
Barrack	84	–	–	84	–	168
Carob	84	84	84	–	84	–
Europe	84	84	84	84	82	–
Ferret	84	84	84	84	82	–
Jarrod	84	84	84	–	84	–
Parrot	84	84	84	84	82	170
Syrup	83	83	83	83	81	174
Total	**754**	**503**	**670**	**419**	**660**	**512**

Table A2. Rates of glottalisation per speaker in Analysis 1 according to voicing context.

Speaker	Voiced context	Voiceless context	Speaker	Voiced context	Voiceless context
1	8%	80%	15	0%	93%
2	0%	100%	16	0%	87%
3	0%	87%	17	8%	100%
4	0%	93%	18	8%	93%
5	0%	100%	19	17%	100%
6	0%	100%	20	0%	67%
7	58%	7%	21	8%	100%
8	8%	100%	22	0%	80%
9	25%	67%	23	0%	100%
10	0%	53%	24	0%	93%
11	17%	73%	25	0%	100%
12	17%	67%	26	33%	87%
13	0%	53%	27	33%	93%
14	0%	87%	28	25%	67%

Table A3. Rates of glottalisation per speaker in Analysis 2 according to age group and gender.

Speaker	Older female	Older male	Younger female	Younger male
1	11%	22%	67%	67%
2	44%	14%	78%	38%
3	89%	22%	83%	38%
4	78%	56%	75%	89%
5	0%	67%	100%	38%
6	78%	33%	100%	78%
7	100%	33%	78%	89%
8	0%	0%	89%	67%
9	67%	33%	100%	11%
10	100%	33%	50%	44%
11	63%	11%	89%	22%
12	0%	22%	83%	44%
13	–	0%	50%	75%
14	–	56%	38%	78%
15	–	38%	44%	25%
16	–	–	33%	56%
17	–	–	–	89%
18	–	–	–	89%

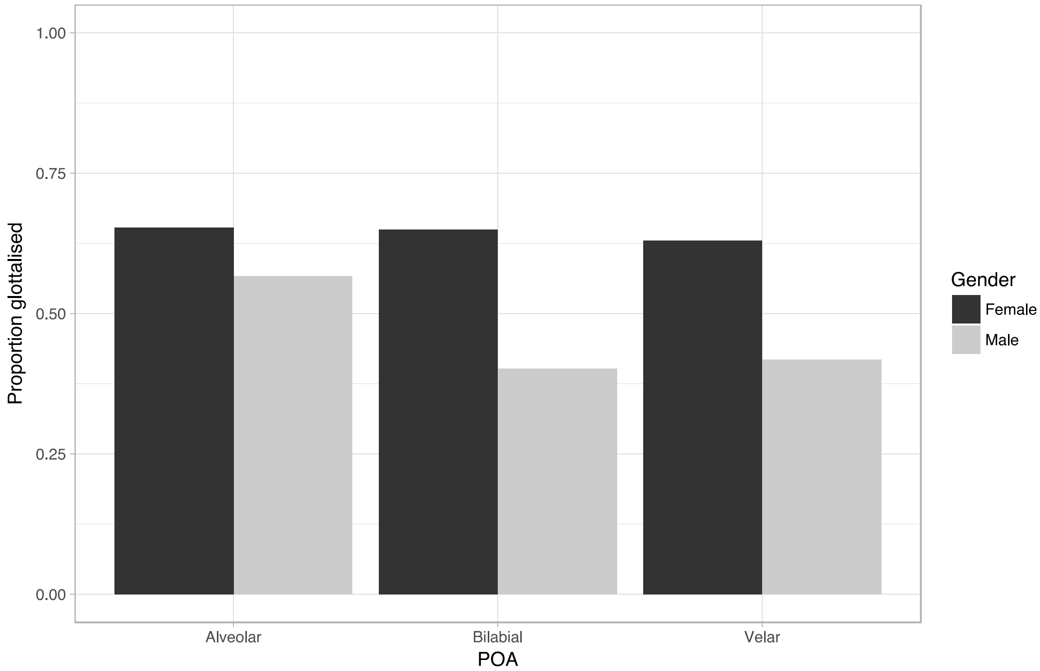

Figure A1. Proportion of tokens glottalised in each POA according to gender.

2.8 References

Bailey, G. (2002). Real and apparent time. In J. K. Chambers, P. Trudgill, & N. Schilling-Estes (Eds.), *The handbook of language variation and change* (pp. 312–322). Malden: Blackwell.

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

Blevins, J. (2006). A theoretical synopsis of evolutionary phonology. *Theoretical Linguistics, 32*, 117–165.

Boersma, P., & Weenink, D. (2015). *Praat: Doing phonetics by computer* (Version 5.4.09). http://www.praat.org/.

Brunner, J., & Żygis, M. (2011). Why do glottal stops and low vowels like each other? In W. Sum Lee & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS XVII), Hong Kong* (pp. 376–379).

Burnham, D., Estival, D., Fazio, S., Viethen, J., Cox, F., Dale, R., Cassidy, S., Epps, J., Togneri, R., Wagner, M., Kinoshita, Y., Göcke, R., Arciuli, J., Onslow, M., Lewis, T., Butcher, A., & Hajek, J. (2011). Building an audio-visual corpus of Australian English: Large corpus collection with an economical portable and replicable black box. In P. Cosi, R. De Mori, G. Di Fabbrizio & R. Pieraccini (Eds.), *Proceedings of 12th Annual Conference of the International Speech Communication Association (INTERSPEECH 2011), Florence, Italy* (pp. 841–844).

Byrd, D. (1993). 54,000 American stops. *UCLA Working Papers in Phonetics, 83*, 97–116.

Cameron, D., & Coates, J. (1989). Some problems in the sociolinguistic explanation of sex differences. In D. Cameron & J. Coates (Eds.), *Women in their speech communities: New perspectives on language and sex* (pp. 13–26). London: Longman.

Cassidy, S., Estival, D., & Cox, F. (2017). Case study: The AusTalk corpus. In N. Ide & J. Pustejovsky (Eds.), *Handbook of linguistic annotation* (pp. 209–227). Berlin: Springer.

Chong, A. J., & Garellek, M. (2018). Online perception of glottalized coda stops in American English. *Laboratory Phonology, 9*, 1–24.

Cochrane, G. R. (1989). Origins and development of the Australian accent. In P. Collins & D. Blair (Eds.), *Australian English: The language of a new society* (pp. 176–186). St Lucia: University of Queensland.

Cox, F., & Fletcher, J. (2017). *Australian English: Pronunciation and Transcription* (2nd ed.). Melbourne: Cambridge University Press.

Cox, F., & Palethorpe, S. (1998). Regional variation in the vowels of female adolescents from Sydney. In R. H. Mannell & J. Robert-Ribes (Eds.), *Proceedings of the 5th International Conference on Spoken Language Processing (ICSLP'98), Sydney, Australia* (pp. 2359–2362).

Cox, F., & Palethorpe, S. (2011). Timing differences in the VC rhyme of standard Australian English and Lebanese Australian English. In W. Sum Lee & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS XVII), Hong Kong* (pp. 528–531).

Cox, F., & Palethorpe, S. (2012). Standard Australian English: The sociostylistic broadness continuum. In R. Hickey (Ed.), *Standards of English: Codified Varieties around the World* (pp. 294–317). Cambridge: Cambridge University Press.

Cox, F., Palethorpe, S., Buckley, L., & Bentink, S. (2014). Hiatus resolution and linking 'r' in Australian English. *Journal of the International Phonetic Association*, *44*, 155–178.

Crystal, T. H., & House, A. S. (1988a). Segmental durations in connected-speech signals: Current results. *Journal of the Acoustical Society of America, 83*, 1553–1573.

Crystal, T. H., & House, A. S. (1988b). Segmental durations in connected-speech signals: Syllabic stress. *Journal of the Acoustical Society of America*, *83*, 1574–1978.

Czyżak, I. (2014). Glottalization and laryngeal node faithfulness in English. In E. Cyran & J. Szpyra-Kozłowska (Eds.), *Crossing phonetics-phonology lines* (pp. 161–189). Newcastle upon Tyne: Cambridge Scholars.

Davis, S. & van Summers, W. (1989). Vowel length and closure duration in word-medial VC sequences. *Journal of Phonetics*, *17*, 339–353.

Dilley, L., Shattuck-Hufnagel, S., & Ostendorf, M. (1996). Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics*, *24*, 423–444.

Docherty, G. J., & Foulkes, P. (1999a). Sociophonetic variation in 'glottals' in Newcastle English. In J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, & A. C. Bailey (Eds.), *Proceedings of the 14th International Congress of Phonetic Sciences* (*ICPhS 99), San Francisco* (pp. 1037–1040).

Docherty, G. J. & Foulkes, P. (1999b). Derby and Newcastle: Instrumental phonetics and variationist studies. In P. Foulkes & G. J. Docherty (Eds.), *Urban voices: Accent studies in the British Isles* (pp. 47–71). London: Arnold.

Docherty, G. J., Foulkes, P., Milroy, J., Milroy, L., & Walshaw, D. (1997). Descriptive adequacy in phonology: A variationist perspective. *Journal of Linguistics*, *33*, 275–310.

Docherty, G. J., Hay, J., & Walker, A. (2006). Sociophonetic patterning of phrase-final /t/ in New Zealand English. In P. Warren & C. I. Watson (Eds.), *Proceedings of the 11th Australasian International Conference on Speech Science & Technology, Auckland, New Zealand* (pp. 378–383).

Eckert, P. (1989). The whole woman: Sex and gender differences in variation. *Language Variation and Change*, *1*, 245–267.

Edmondson, J. A., & Esling, J. H. (2006). The valves of the throat and their functioning in tone, vocal register and stress: Laryngoscopic case studies. *Phonology*, *23*, 157–191.

Espy-Wilson, C. Y., Boyce, S. E., Jackson, M., Narayanan, S., & Alwan, A. (2000). Acoustic modeling of American English /r/. *Journal of the Acoustical Society of America*, *108*, 343–356.

Fleming, E. (2009). The phonetics of schwa vowels. In D. Minkova (Ed.), *Phonological weakness in English*: *From old to present-day English* (pp. 78–95). Basingstoke: Palgrave Macmillan.

Foulkes, P., & Docherty, G. J. (2000). Another chapter in the story of /r/: ‘Labiodental’ variants in British English. *Journal of Sociolinguistics*, *4*, 30–59.

Fowler, C. A. (1983). Converging sources of evidence on spoken and perceived rhythms of speech: Cyclic production of vowels in monosyllabic stress feet. *Journal of Experimental Psychology*, *112*, 386–412.

Fowler, C. A. (1992). Vowel duration and closure duration in voiced and unvoiced stops: There are no contrast effects here. *Journal of Phonetics*, *20*, 143–165.

Garellek, M. (2011). The benefits of vowel laryngealization on the perception of coda stops in English. *UCLA Working Papers in Phonetics*, *109*, 31–39.

Garellek, M. (2014). Voice quality strengthening and glottalization. *Journal of Phonetics*, *45*, 106–113.

Garellek, M. (2015). Perception of glottalization and phrase-final creak. *Journal of the Acoustical Society of America*, *137*, 822–831.

Gimson, A. C. (1962). *An introduction to the pronunciation of English*. London: Arnold.

Gordeeva, O. B., & Scobbie, J. M. (2013). A phonetically versatile contrast: Pulmonic and glottalic voicelessness in Scottish English obstruents and voice quality. *Journal of the International Phonetic Association*, *43*, 249–271.

Greenbaum, S. (1991). ICE: The international corpus of English. *English Today*, *7*, 3–7.

Gruenenfelder, T. M., & Pisoni, D. B. (1980). Fundamental frequency as a cue to postvocalic consonantal voicing: Some data from speech perception and production. *Perception & Psychophysics*, *28*, 514–520.

Haslerud, V. C. D. (1995). *The variable (t) in Sydney adolescent speech: A sociolinguistic study of phonological variation* (Cand. Philol dissertation, University of Bergen).

Hay, J., & Maclagan, M. (2012). /r/-sandhi in early 20th century New Zealand English. *Linguistics*, *50*, 745–763.

Hejná, M., Palo, P., & Moisik, S. (2016). Glottal squeaks in VC sequences. In N. Morgan (Ed.), *Proceedings of 17th Annual Conference of the International*

Speech Communication Association (INTERSPEECH 2016), San Francisco (pp. 1136-1140).

Helgason, P. (2002). *Preaspiration in the Nordic languages: Synchronic and diachronic aspects* (Doctoral dissertation, Stockholm University).

Holmes, J. (1995). Glottal stops in New Zealand English: An analysis of variants of word-final /t/. *Linguistics*, *33*, 433–463.

House, A. S., & Fairbanks, G. (1953). The influence of consonant environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America*, *25*, 105.

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

Ingram, J. C. L. (1989). Connected speech processes in Australian English. *Australian Journal of Linguistics*, *9*, 21–49.

Jones, M. J., & McDougall, K. (2009). The acoustic character of fricated /t/ in Australian English: A comparison with /s/ and /ʃ/. *Journal of the International Phonetic Association*, *39*, 265–289.

Kahn, D. (1976). *Syllable based generalizations in English phonology* (Doctoral dissertation, Massachusetts Institute of Technology).

Katz, J. (2012). Compression effects in English. *Journal of Phonetics*, *40*, 390–402.

Kawahara, H., Masuda-Katsuse, I., & de Cheveigné, A. (1999). Restructuring speech representations using a pitch-adaptive time–frequency smoothing and an instantaneous-frequency-based F0 extraction: Possible role of a repetitive structure in sounds. *Speech Communication*, *27*, 187–207.

Kendall, T., & Thomas, E. R. (2014). *vowels: Vowel manipulation, normalization and plotting* (version 1.2-1). https://CRAN.R- project.org/package=vowels

Keyser, S. J., & Stevens, K. N. (2006). Enhancement and overlap in the speech chain. *Language*, *82*, 33–62.

Kingston, J., & Diehl, R. L. (1994). Phonetic Knowledge. *Language*, *70*, 419–454.

Kisler, T., Reichel, U. D., & Schiel, F. (2017). Multilingual processing of speech via web services. *Computer Speech & Language*, *45*, 326–347.

Klatt, D. H. (1975). Vowel lengthening is syntactically determined in connected discourse. *Journal of Phonetics, 3*, 129–140.

Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, *59*, 1208–1221.

Kohler, K. J. (1982). F0 in the production of lenis and fortis plosives. *Phonetica*, *39*, 199–218.

Kohler, K. J. (1994). Glottal stops and glottalization in German: Data and theory of connected speech processes. *Phonetica*, *51*, 38–51.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). *lmerTest: Tests in linear mixed effects models* (Version 2.0-33). https://CRAN.R-project.org/package=lmerTest

Labov, W. (1990). The intersection of sex and social class in the course of linguistic change. *Language Variation and Change*, *2*, 205–254.

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

Lisker, L. (1978). Rapid vs rabid: A catalogue of acoustic features that may cue the distinction. *Haskins Laboratories Status Report on Speech Research SR-54*, 127–132.

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

Loakes, D., & McDougall, K. (2010). Individual variation in the frication of voiceless plosives in Australian English: A study of twins' speech. *Australian Journal of Linguistics*, *30*, 155–181.

McCarthy, O., & Stuart-Smith, J. (2013). Ejectives in Scottish English: A social perspective. *Journal of the International Phonetic Association*, *43*, 273–298.

Malisz, Z., Żygis, M., & Pompino-Marschall, B. (2013). Rhythmic structure effects on glottalisation: A study of different speech styles in Polish and German. *Laboratory Phonology*, *4*, 119–158.

Mathisen, A. G. (1999). Sandwell, West Midlands: Ambiguous perspectives on gender patterns and models of change. In P. Foulkes & G. J. Docherty (Eds.), *Urban voices: Accent studies in the British Isles* (pp. 107–123). London: Arnold.

Mees, I. M. (1987). Glottal stop as a prestigious feature in Cardiff English. *English World-Wide*, *8*, 25–39.

Mees, I. M. (1990). Patterns of sociophonetic variation in the speech of Cardiff schoolchildren. In N. Coupland & A. Richard Thomas (Eds.), *English in Wales: Diversity, conflict and change* (pp. 167–194). Clevedon: Multilingual Matters.

Mees, I. M., & Collins, B. (1999). Cardiff: A real-time study of glottalisation. In P. Foulkes & G. J. Docherty (Eds.), *Urban voices: Accent studies in the British Isles* (pp. 85–202). London: Arnold.

Milroy, J., Milroy, L., Hartley, S., & Walshaw, D. (1994). Glottal stops and Tyneside glottalization: Competing patterns of variation and change in British English. *Language Variation and Change*, *6*, 327–357.

Modarresi, G., Sussman, H. M., Lindblom, B., & Burlingame, E. (2005). Locus equation encoding stop place: revisiting the voicing/VOT issue. *Journal of Phonetics*, *33*, 101–113.

Moisik, S. R. (2013). *The epilarynx in speech.* (Doctoral dissertation, University of Victoria).

Moisik, S. R., & Esling, J. H. (2011). The 'whole larynx' approach to laryngeal features. In W. Sum Lee & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS XVII), Hong Kong* (pp. 1406–1409).

Munhall, K., Fowler, C., Hawkins, S., & Saltzman, E. (1992). "Compensatory shortening" in monosyllables of spoken English. *Journal of Phonetics*, *20*, 225–239.

Nance, C., & Stuart-Smith, J. (2013). Pre-aspiration and post-aspiration in Scottish Gaelic stop consonants. *Journal of the International Phonetic Association*, *43*, 129–152.

O'Brien, J. (2012). *An experimental approach to debuccalization and supplementary gestures* (Doctoral dissertation, University of California, Santa Cruz).

Ohde, R. N. (1984). Fundamental frequency as an acoustic correlate of stop consonant voicing. *Journal of the Acoustical Society of America*, *75*, 224–230.

Penney, J., Cox, F., Miles, K., & Palethorpe, S. (2018). Glottalisation as a cue to coda consonant voicing in Australian English. *Journal of Phonetics*, *66*, 161–184.

Pierrehumbert, J. B. (1994). Knowledge of variation. In K. Beals, J. Denton, R. Knippen, L. Melnar, H. Suzuki, & E. Zeinfeld (Eds.), *Papers from the Parasession on Variation, 30th meeting of the Chicago Linguistic Society* (pp. 232–256). Chicago, Chicago Linguistic Society.

Pierrehumbert, J. B. (1995). Prosodic effects on glottal allophones. In O. Fujimura & M. Hirano (Eds.), *Vocal fold physiology: Voice quality control* (pp. 39–60). San Diego: Singular.

Pierrehumbert, J. B., & Talkin, D. (1992). Lenition of /h/ and glottal stop. In G. J. Doherty & D. R. Ladd (Eds.), *Papers in laboratory phonology II: Gesture segment prosody* (pp. 90–117). Cambridge: Cambridge University Press.

Port, R. F., & Dalby, J. (1982). Consonant/vowel ratio as a cue for voicing in English. *Perception and Psychophysics*, *32*, 141–152.

R Core Team (2016). *R: A language and environment for statistical computing* (Version 3.3.1). https://www.R-project.org/.

Redi, L., & Shattuck-Hufnagel, S. (2001). Variation in the realization of glottalization in normal speakers. *Journal of Phonetics*, *29*, 407–429.

Roach, P. J. (1973). Glottalization of English /p/, /t/, /k/ and /ʧ/ – a reexamination. *Journal of the International Phonetic Association*, *3*, 10–21.

Schneider, E. W., Burridge, K., Kortmann, B., Mesthrie, R., & Upton, C. (2004). *A handbook of varieties of English*, vol. 1: *Phonology*. Berlin: Mouton de Gruyter.

Selkirk, E. (1982). The syllable. In H. van der Hulst & N. Smith (Eds.), *The structure of phonological representations*, (Part II, pp. 337–383). Dordrecht: Foris.

Seyfarth, S., & Garellek, M. (2015). Coda glottalization in American English. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*, *Glasgow, Scotland*.

Shue, Y.-L., Keating, P., Vicenik, C., & Yu, K. (2011). VoiceSauce: A program for voice analysis. In W. Sum Lee & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS XVII), Hong Kong* (pp. 1846–1849).

Song, J. Y., Demuth, K., & Shattuck-Hufnagel, S. (2012). The development of acoustic cues to coda contrasts in young children learning American English. *Journal of the Acoustical Society of America*, *131*, 3036–3050.

Svirsky, M. A., Stevens, K. N., Matthies, M. L., Manzella, J., Perkell, J. S., & Wilhelms-Tricarico, R. (1997). Tongue surface displacement during bilabial stops. *Journal of the Acoustical Society of America*, *102,* 562–571.

Tait, C., & Tabain, M. (2016). Patterns of gender variation in the speech of primary school-aged children in Australian English: the case of /p t k/. In C. Carignan & M. D. Tyler (Eds.), *Proceedings of the 16th Australasian International Conference on Speech Science and Technology, Sydney, Australia* (pp. 65–68).

Tollfree, L. (1999). South East London English: Discrete versus continuous modelling of consonantal reduction. In P. Foulkes & G. J. Docherty (Eds.), *Urban voices: Accent studies in the British Isles* (pp. 163–184). London: Arnold.

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

Trudgill, P. (1986). *Dialects in contact.* Oxford: Blackwell.

Umeda, N. (1978). Occurrence of glottal stops in fluent speech. *Journal of the Acoustical Society of America*, *64,* 88–94.

van Bergem, D. R. (1994). A model of coarticulatory effects on the schwa. *Speech Communication*, *14,* 143–162.

Watt, D., & Milroy, L. (1999). Patterns of variation and change in three Newcastle vowels: Is this dialect levelling? In P. Foulkes & G. J. Docherty (Eds.), *Urban voices: Accent studies in the British Isles* (pp. 25–46). London: Arnold.

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

Wright, R. (2004). A review of perceptual cues and cue robustness. In B. Hayes, R. Kirchner, & D. Steriade (Eds.), *Phonetically-based phonology* (pp. 34–57). New York: Cambridge University Press.

Wolak, M. E., Fairbairn, D. J., & Paulsen, Y. R. (2012). Guidelines for estimating repeatability. *Methods in Ecology and Evolution*, *3*, 129–137.

Yallop, C. (2001). A. G. Mitchell and the development of Australian English pronunciation. In D. B. Blair & P. Collins (Eds.), *Varieties of English around the world: English in Australia* (pp. 287–302). Amsterdam: Benjamins.

Yuen, I., Cox, F., & Demuth, K. (2018). Prosodic effects on the planning of inserted /ɹ/ in Australian English. *Journal of Phonetics*, *69*, 29–42.

Chapter 3: Effects of glottalisation, preceding vowel duration, and coda closure duration on the perception of coda stop voicing

This chapter is based on the following paper, which has been resubmitted after revisions based on reviewer feedback to the original submission:

Penney, J., Cox, F., & Szakay, A. (revision submitted). Effects of glottalisation, preceding vowel duration, and coda closure duration on the perception of coda stop voicing. *Journal of Phonetics*.

I certify that I was responsible for the development of the concept of this paper, in discussion with my supervisors/co-authors. I took leadership in conducting the research, and was responsible for all data collection, the construction of the stimuli, the analyses, and the writing of all parts of the paper. My co-authors provided advice to improve the research protocol, the analyses, the interpretation of the data, as well as the presentation of the written component.

3.0 Abstract

English has multiple potential acoustic cues to coda stop voicing, including the duration of the preceding vowel, the duration of the coda closure, and, in some varieties, glottalisation. Glottalisation appears to be a recent change to Australian English and correspondingly in production younger speakers use glottalisation more than older speakers. In this paper we report on a study designed to examine listeners' perception of cues to coda stop voicing. We presented listeners with audio stimuli in which preceding vowel duration, coda closure duration, and the relative proportions of the rhyme that these occupy were manipulated and co-varied with the presence or absence of glottalisation. The results show that listeners use preceding vowel duration to cue coda stop voicing, and that coda closure duration is a weaker cue to voicing when it is not varied in conjunction with preceding vowel duration. In addition, glottalisation results in increased perception of coda voicelessness, even when paired with extended preceding vowel duration, which otherwise signals coda voicing. Although age-related differences in production have been reported, we found that both older and younger listeners used glottalisation similarly in perception. These results may provide support for a sound change led by a shift in perception.

3.1 Introduction

It is well established that the stop voicing contrast in English is supported by a range of acoustic correlates other than the presence of phonetic voicing (Abramson, 1977; Docherty, 1992; Flege & Brown, 1982; Lisker, 1986; Raphael, 1981). For stops in syllable onset position, voice onset time (VOT) - the interval between the release of a stop closure and the onset of phonation - is a major cue to voicing, with voiced stops exhibiting shorter VOT than voiceless stops (Cho, Whalen, & Docherty, 2019; Lisker & Abramson, 1964, 1967, 1970; Zlatin, 1979). In coda position voicing may extend partially or fully through the stop closure for voiced stops and thereby serve as a cue to the contrast; however, voicing is often weak or absent in this position (Davidson, 2016; Docherty, 1992; Westbury & Keating, 1986). The VOT cue does not generally apply to coda stops, unless there is a following voiced segment. Nevertheless, many other acoustic correlates can cue coda voicing; for example, voiced stops have longer preceding vowel durations, shorter closure durations, a drop in the first formant (F1) frequency prior to the closure, lower fundamental frequency (F0) at the offset of the preceding vowel, and shorter and less intense release bursts when compared to voiceless stops (Cox & Palethorpe, 2011; Gruenefelder & Pisoni, 1980; Hanson, 2009; House & Fairbanks, 1953; Kingston & Diehl, 1994; Kohler, 1982; Liberman, Delattre, & Cooper, 1958; Lisker, 1957, 1975, 1986; Ohde, 1984; Port, 1979; Song, Demuth, & Shattuck-Hufnagel, 2012; Wright, 2004). Preceding vowel duration is a strong perceptual cue to coda stop voicing (Chen, 1970; Denes, 1955; Fowler, 1992; Klatt, 1976; Lisker, 1974, 1986; Luce & Charles-Luce, 1985; Malécot, 1970; Peterson & Lehiste, 1960; Raphael, 1972). However, the duration of the preceding vowel and the duration of the stop closure function in a trading relation, with a longer vowel but shorter closure for voiced stops, and vice versa for voiceless stops (Lisker, 1957;

Hogan & Rozsypal, 1980; Repp, 1982). Accordingly, some researchers have suggested that it may not be so much the duration of the vowel that cues the voicing of the following stop, but rather the interaction between vowel and closure duration, and the relative proportion that each of these occupies in the rhyme (Denes, 1955; Fitch, 1981; Port 1981; Port & Dalby, 1982). A useful metric for measuring the relative proportions of the components of the rhyme is the ratio of the closure duration to the vowel duration (C/V ratio) (Port & Dalby, 1982).

In some varieties of English, the presence of glottalisation co-occurs with coda voicelessness in speech production. Glottalisation, as we use the term here, refers to the addition of glottal/laryngeal constriction to an oral stop, which results in laryngealised phonation of the preceding vowel and auditorily produces a percept of creakiness. This may, but need not, occur together with a full glottal stop, though laryngealisation of the preceding vowel is present in either case (Garellek, 2015). Glottalisation has been shown to occur in conjunction with voiceless stops in several varieties of British English (Docherty & Foulkes, 1999; Foulkes & Docherty, 2006; Gordeeva & Scobbie, 2013; Higginbottom, 1964; Mathisen, 1999; Roach, 1973; Tollfree, 1999), in American English (AmE) (Huffman, 2005; Pierrehumbert, 1995; Seyfarth & Garellek, 2015), and it has recently been documented for Australian English (AusE) (Penney, Cox, Miles, & Palethorpe, 2018). In a study of single words produced by 67 AusE speakers, Penney et al. (2018) found that glottalisation was present in 55% of items containing a voiceless coda, whereas it occurred only in 6% of the items containing a voiced coda. In another study comparing the production of unstressed syllables with voiced and voiceless final coda stops, Penney, Cox, and Szakay (2019 [Chapter 2]) found glottalisation to occur in 83% of items with voiceless codas, compared to 10% of items with voiced codas. These data suggest that, in production at least, glottalisation is strongly associated with coda stop voicelessness in AusE.

Glottalisation appears to be a recent change to AusE. There is no mention of glottalised forms in early phonetic descriptions of the variety (see Tollfree, 2001, for a detailed review of early linguistic descriptions and popular literature) and its absence was noted as late as the 1980s (Trudgill, 1986; Wells, 1982). However, over the past thirty years greater reference has been made to glottalisation as a feature of AusE (Cox & Palethorpe, 2007; Ingram, 1989; Haslerud, 1995; Tollfree, 2001; Penney et al., 2018; Penney et al., 2019 [Chapter 2]). Consistent with glottalisation being a recent change, it is more commonly found in younger compared to older speakers (Penney et al., 2018; Penney et al., 2019 [Chapter 2]). In these production studies, younger speakers not only used glottalisation more frequently, but also made less use of preceding vowel duration than older speakers to cue coda voicing (Penney et al., 2018). Furthermore, for both younger and older speakers, glottalisation occurred least on high vowels (/iː, ʉː/),[1] and these high vowels demonstrated the greatest preceding vowel duration differences across coda voicing categories. This raises the possibility that a trading relationship may exist between preceding vowel duration and glottalisation as cues to coda stop voicing. In this paper, we report on a study designed to test listeners' perception of these cues.

The AusE vowel inventory contains both inherently short (e.g. /ɐ/ as in *cud*) and inherently long vowels (e.g. /ɐː/ as in *card*), with the duration of the inherently short vowels approximately 60% of the duration of the inherently long vowels in voiced coda contexts (Cox, 2006). While both inherently short and long vowels

[1] We use the phonemic symbols recommended by Harrington, Cox, and Evans (1997) for describing Australian English.

demonstrate coda voicing related durational differences (with shorter vowel duration preceding voiceless codas and longer vowel duration preceding voiced codas), inherently short vowels exhibit smaller differences across these contexts than long vowels (Cox & Palethorpe, 2011; Penney et al., 2018). Therefore, preceding vowel duration may be less informative as a cue to coda voicing in short vowel compared to long vowel contexts. In Penney et al. (2018), younger speakers glottalised at higher rates in short vowel contexts than in long vowel contexts, suggesting that the addition of glottalisation may be of benefit in contexts where coda voicing related durational differences are not as robust. In this study we also explore the effect of inherent vowel length on the perception of coda voicing.

As reported above, glottalisation occurs primarily in conjunction with voiceless stops in studies of speech production, possibly as a strategy to enhance voicelessness and thereby support or maintain the voicing contrast (Keyser & Stevens, 2006; Pierrehumbert, 1995; Seyfarth & Garellek, 2015).[2] Glottalisation occurs frequently before sonorants, particularly nasals (Huffman, 2005; Pierrehumbert, 1994, 1995; Seyfarth & Garellek, 2015), an environment in which voicing is likely to spread due to anticipatory co-articulation (Huffman, 2005). As such, constricting the glottis may be employed as a strategy to ensure that voicing is rapidly extinguished. Glottalisation is also reported to occur frequently, possibly as an enhancement strategy (Keyser & Stevens, 2006), in phrase- or utterance-final environments (Huffman, 2005; Redi & Shattuck-Hufnagel, 2001; Seyfarth & Garellek, 2015), where voicing is weak and other cues, such as F0 differences, may not be recoverable.

[2] Note that Huffman (2005) and Seyfarth and Garellek (2015) state that an enhancement account cannot explain glottalisation in all of the environments in which it occurs, for example preceding voiceless obstruents.

However, few studies have explored whether listeners are sensitive to glottalisation and how they interpret it in perception. In a phoneme monitoring task in which participants were required to respond as soon as they heard /t/, Garellek (2011) found that AmE listeners' were faster and more accurate when glottalisation was present, though it should be noted that listeners were not presented with voiced coda stops in that experiment. More recently, Chong and Garellek (2018) conducted an eye-tracking experiment in which AmE participants looked towards an orthographic word from a minimal pair differing in coda stop voicing upon presentation of non-glottalised and glottalised audio stimuli. Listeners were slower at identifying words with voiced codas (/b, d/) when glottalisation was present, suggesting that glottalisation may impede perception of voiced codas. This effect was not found for words with voiceless codas (/p, t/), implying that listeners do indeed associate glottalisation with voiceless coda stops but not with voiced coda stops. However, contra to the assumptions of an enhancement account, Chong and Garellek (2018) found that the addition of glottalisation provided no improvement to listeners' perception of words with voiceless coda stops, despite hindering the perception of voiced coda stops. While these results suggest that AmE listeners are sensitive to glottalisation, they also raise questions about how listeners might deal with multiple voicing cues. It may be possible, for example, that listeners were able to identify other cues to voicelessness in the stimuli and hence were not reliant on glottalisation.

As age differences have been shown in the use of glottalisation in AusE speech production, we are interested in examining whether similar differences may also exist between older and younger listeners in perception. We might expect that younger listeners, who glottalise more than older listeners in production and can therefore be considered to be more progressive with respect to the change, may also show more sensitivity to glottalisation in perception. Correspondingly, we might also expect that older listeners would be less sensitive to glottalisation in perception, consistent with older speakers utilising glottalisation less in production. Such expectations are founded on the assumption of a link between speaker/listeners' production and perception repertoires (Beddor, 2015; Coetzee, Beddor, Shedden, Styler, & Wissing, 2018; Harrington, Kleber, & Reubold, 2008). On the other hand, a mismatch between production and perception repertoires may be indicative of a change in progress. Kleber, Harrington, and Reubold (2012) suggest that sound change represents a shift from a stable alignment of speaker/listeners' production and perception repertoires to an alternative (stable) alignment, and that this may come about due to a misalignment during which the two are out of sync, for instance if listeners cease to compensate for coarticulation in their perception (Ohala, 1983). There is evidence in the literature for changes in which perception follows production (Coetzee et al., 2018; Pinget, 2015) as well as for changes in which perception may lead production (Harrington et al., 2008; Kleber et al., 2012; Kuang, 2018; Ohala, 1981). Younger speakers appear more innovative in their use of glottalisation in production. Therefore, if glottalisation is a change in which production precedes perception, we may expect to observe that younger speaker/listeners do not attend to glottalisation in perception, despite glottalising in production, or that younger listeners show some sensitivity to glottalisation in perception (representing an incipient change in perception) whereas older listeners do not. Alternatively, if perception is leading production, we may expect to see that older speaker/listeners are sensitive to glottalisation in perception despite their reduced use of this feature in their speech production.

3.1.1 Aims and hypotheses

Our main aim is to examine how AusE listeners use glottalisation and preceding vowel duration in the perception of coda stop voicing through a study in which listeners are presented with audio stimuli co-varying in preceding vowel duration and glottalisation. We also explore how listeners use coda closure duration and C/V ratio (Port & Dalby, 1982) in coda stop perception and we are further interested in how inherent vowel length, vowel height, and age interact with the voicing cues. This perceptual study builds on and complements our previous research on glottalisation in production in AusE (Penney et al., 2018; Penney et al., 2019 [Chapter 2]). Based on the results of prior research we hypothesise:

1: Listeners will use preceding vowel duration to cue coda voicing

Longer vowel durations will cue voiced codas and shorter vowel durations will cue voiceless codas.

2: Listeners may use closure duration to cue coda voicing

Longer closure durations will cue voiceless codas and shorter closure durations will cue voiced codas, but we expect listeners not to be as sensitive to this cue as they are to the vowel duration cue (Luce & Charles-Luce, 1985).

3: Listeners will use changes in the proportions of the rhyme components to cue coda voicing

Rhymes in which the vowel proportion dominates will cue more voiced coda responses, whereas rhymes in which the vowel proportion is lower/closure proportion is higher will cue more voiceless coda responses.

In addition, our novel hypotheses are that:

4: The presence of glottalisation will facilitate the perception of coda voicelessness

We expect the presence of glottalisation will result in an increase in voiceless coda responses. This hypothesis is based on findings that glottalisation occurs in voiceless coda contexts in production and that listeners associate glottalisation with voiceless codas, though we note that Chong and Garellek (2018) did not find evidence of a facilitation effect.

5: Preceding vowel duration will have a weaker effect on coda voicing for inherently short vowels

Differences in vowel duration in production are smaller between voicing contexts for inherently short vowels than for inherently long vowels (Cox & Palethorpe, 2011); accordingly, we expect listeners to be more sensitive to duration for inherently long vowels.

Supposing hypothesis 4 is confirmed, we further hypothesise that:

6: The effect of glottalisation in promoting the percept of voicelessness will be stronger for inherently short vowels than for inherently long vowels

Inherently long vowels make greater use of vowel duration differences in production (Cox & Palethorpe, 2011); therefore, vowel duration is a stronger voicing cue for inherently long vowels compared to inherently short vowels. For inherently short vowels, the vowel duration cue to voicing is not so robust and therefore we may expect glottalisation to be given greater weighting.

7: The effect of glottalisation in promoting the percept of voicelessness will be stronger for low vowels than for high vowels

High vowels make greater use of vowel duration differences to cue coda voicing than non-high vowels and less use of glottalisation in production (Penney et al., 2018). Therefore we expect listeners to be more sensitive to glottalisation on low vowels.

8: The effect of glottalisation in promoting the percept of voicelessness will be stronger in younger listeners than in older listeners

As younger speakers use glottalisation more in production whereas older speakers utilise preceding vowel duration more to cue coda voicing, we expect younger speakers to make greater perceptual use of glottalisation. We note that competing hypotheses may be possible: younger speakers may not be sensitive to glottalisation in perception if this is a production-led change in its early stages; older speakers may also be sensitive to glottalisation in perception if this is a perception-led change.

3.2 Method

3.2.1 Participants

We initially recruited 80 listeners to take part in this task, who received either course credit or payment for their participation. Data for three participants were discarded due to these participants failing to engage in the task (one participant) or failing to identify any phonological distinction during the task (two participants), leaving data for 77 listeners remaining. The participants were recruited from two age groups: the younger group were aged between 18 and 36 years of age (n = 46 (mean: 21); female: 38; male: 8), and the older group were aged 50 years and above (n = 31(mean: 61); female: 22; male: 9). All participants were native speakers of AusE. With the exception of one participant who was born overseas and who migrated to Australia at 12 months, all participants were born in Australia, and all completed their school education (primary and secondary) exclusively in Australia. All participants reported normal hearing for their age.

3.2.2 Task

The participants took part in a two-alternative forced choice word identification task in which they were presented with audio stimuli through Sennheiser HD 380 pro headphones, while orthographic representations of minimal pairs differing only in coda voicing were displayed on an Apple MacBook notebook monitor. The minimal pairs included in this task were CVC *bead/beat*, *bard/bart,*[3] *bid/bit*, and *bud/but*. All of the words are real English words, though *bard* and *bart* are less frequent than the other words. In addition, *bart* is a proper noun (as is *bard*, when used to refer to William Shakespeare). All of the words were presented orthographically in lower case. In order to familiarise participants with the words and to mitigate against possible frequency effects, all of the target words were included in a production exercise that took place immediately prior to the perception task, in which participants were recorded reading 396 randomly presented sentences from the notebook screen while wearing a headset condenser microphone. In this exercise the target words were presented in a variety of phonetic environments and phrase positions. Full details and results of the production task are reported in Penney, Cox, & Szakay (in prep [Chapter 5]). Headphone volume was set at a standard level for all participants, measured at 85dBA using a BK4153 artificial ear headphone coupler and a BK2231 Sound Level Metre. For each audio stimulus, a fixation cross was first presented for 600 ms in the centre of the screen. An orthographic representation of one of the minimal pairs (i.e.

[3] Note that Australian English is a non-rhotic dialect so these words did not contain [ɹ] or a rhoticised vowel.

two words) was then displayed, and after 500 ms a single word audio stimulus was presented. Participants identified which word they heard by responding with a key press on the notebook keyboard. Following the participant's response, the sequence began anew for the next stimulus item. Prior to beginning the test phase of the task, participants took part in a familiarisation phase in which they were presented with non-manipulated auditory stimuli of each test word (produced by a different speaker than the test stimuli) paired with an orthographic representation of the word. They also took part in a training phase, in which they were presented with (non-manipulated) auditory stimuli which they were required to match with one item from an orthographically presented minimal pair, after which the correct word was displayed on the screen.

All of the auditory stimuli presented were of the form /bVC/, where /V/ represents one of the four vowels /iː, ɪ, ɐː, ɐ/ and /C/ represents an alveolar stop with ambiguous voicing. Due to the length of the task, listeners were randomly assigned to either an inherently long vowel condition, in which the vowels /iː, ɐː/ were presented, or an inherently short vowel condition, in which the vowels /ɪ, ɐ/ were presented. This was necessary as the task would have been too long and taxing for participants if they were to have engaged with the full set of stimuli. Each participant was presented with a total of 648 stimulus tokens, organised by vowel into two blocks. Listeners were presented with randomised stimuli in which vowel duration, coda closure duration, and C/V ratio had been manipulated to produce 9-step continua of equally spaced steps (see 3.2.3 below for details). For each step of each continuum, listeners were randomly presented with both a non-glottalised token and a glottalised token. Three repetitions of each item were presented. Listeners were therefore presented with: 3 repetitions X 9 steps X 2 conditions (glottalised/non-glottalised) X 3 manipulations (vowel duration, coda closure duration, and C/V ratio) X 2 vowels (high/low) X 2 voicing categories (source voiced/source voiceless) = 648 tokens. Although we collected responses to both source voiced and source voiceless stimuli, in the preliminary stages of our data analysis we identified a discrepancy, such that listeners exhibited a strong bias to voiceless responses for the stimuli created from the source token *but*. Investigation of H1*-H2* values revealed a slight difference in voice quality in the final portion of the vowel in these items. In order to avoid compromising our analysis, we therefore separated the responses to the source voiced and source voiceless stimuli. The results for both were very similar, apart from the effects of the short low vowel (i.e. the *but* items) in responses to the source voiceless stimuli. Here we report only on the analysis of the stimuli created from the voiced source tokens (i.e. 324 items per participant).

3.2.3 Stimuli

Stimuli were created from natural speech tokens, recorded in a sound treated recording studio in the Department of Linguistics at Macquarie University, using an AKG C535 EB microphone, through a Presonus StudioLive 16.2.4 AT mixer to an Apple iMac computer at a 48kHz sampling rate. All of the tokens were produced phrase medially in carrier sentences by a phonetically trained, female AusE speaker, aged 25 years, who was born in Sydney and received all of her schooling within the greater Sydney region. The words recorded for the task were: *bard, bart, bead, beat, bud, but, bid, bit*. These words were selected to enable a comparison of high and low vowels (/iː, ɪ/ vs /ɐː, ɐ/) as well as inherently long and inherently short vowels (/iː, ɐː/ vs /ɪ, ɐ/). Each word was produced multiple times with a modally voiced falling intonation pattern. The example of each word with the longest steady state vowel portion (i.e. the least amount of onglide or offglide) was chosen as the stimuli source. The speaker also produced sustained instances of each vowel with creaky voice, which were used for the purposes of creating the glottalised stimuli (see below for details).

Using Praat (Boersma & Weenik, 2018), we manipulated the source tokens to remove cues to the voicing of the coda stop. First, coda stop bursts (including any following frication and/or aspiration) were replaced with a release burst from a low intensity voiced stop taken from the unstressed syllable in the word *stated*, which was shown in our piloting to produce an ambiguous percept with regard to coda voicing. All coda closure periods were replaced with silence to ensure no voice bar was present. Additionally, all of the vowels were truncated immediately prior to the onset of F1 formant transitions into the following consonant, to remove the F1 voicing cue.

Using the intensity tier function in Praat (Boersma & Weenik, 2018), we then manipulated the intensity contours to ensure consistency between tokens. Each point in the intensity tier was removed and six points were added at various acoustic landmarks throughout the word. These points and their respective intensity levels are shown in Table A1 in Appendix A.

Following this, three separate manipulations were carried out on each token to create three separate continua for each source token: vowel duration manipulation, coda closure duration manipulation, and C/V ratio manipulation.

3.2.3.1 Vowel duration manipulation

For each source token we created a vowel duration continuum using the PSOLA function in Praat (Boersma & Weenik, 2018). Nine equally spaced vowel duration steps were created, with minimum and maximum durational values based on the mean values for each of the vowels, as reported in Penney et al. (2018) for young (< 36 years) female speakers' production data in the Austalk corpus (Burnham et al., 2011). The shortest vowel duration (i.e. step 1) was two standard deviations shorter than the reported mean duration of the relevant vowel as produced in a voiceless coda context, and the longest vowel duration (i.e. step 9) was two standard deviations greater than the mean duration of the vowel in a voiced coda context. Therefore, the first step was markedly shorter than the mean duration of that particular vowel in a voiceless context, and the ninth step was markedly longer than the mean duration in a voiced context. The duration of the coda closure at each step in the continuum was kept stable and was derived from the mean coda closure duration for the relevant vowel across voiceless and voiced coda contexts combined, as reported in Penney et al. (2018). Table 1 lists the durations of the first and ninth step stimulus for each source vowel, as well step size, and the coda closure duration.

Table 1. Duration of vowel in shortest and longest steps, step size, and coda closure duration for each vowel continuum.

Source vowel	First step duration (ms)	Ninth step duration (ms)	Step size (ms)	Coda closure duration (ms)
iː	104	340	29.5	82
ɪ	65	166	12.625	94
ɐː	165	370	25.625	75
ɐ	64	201	17.125	102

3.2.3.2 Coda closure duration manipulation

We also created a closure duration continuum for each source token using the same methods as described above. Nine equally spaced coda closure duration steps were created for each source token, with minimum and maximum durational values based on mean values reported in Penney et al. (2018). The shortest closure duration (i.e. step 1) was two standard deviations shorter than the mean coda closure duration for voiced stops preceded by the particular vowel, and the longest closure duration (i.e. step 9) was two standard deviations greater than the mean coda closure duration for voiceless stops preceded by the particular vowel. The duration of the vowel at each step in the continuum was kept stable and was derived from the mean vowel duration of the relevant vowel in voiceless and voiced coda contexts combined as reported in Penney et al. (2018). Table 2 lists the durations of the first and ninth step stimulus for each source vowel, as well step size, and the vowel duration.

Table 2. Duration of coda closure in shortest and longest steps, step size, and vowel duration for each vowel continuum.

Source vowel	First step duration (ms)	Ninth step duration (ms)	Step size (ms)	Vowel duration (ms)
iː	15	164	18.625	204
ɪ	38	160	15.25	116
ɐː	16	144	16	263
ɐ	38	170	16.5	129

3.2.3.3 C/V ratio manipulation

We then created a nine-step continuum for each of the source tokens in which the overall duration of the rhyme remained consistent but the proportions of the rhyme belonging to the vowel and the coda closure were manipulated. It was necessary to maintain a consistent overall rhyme duration to ensure that rhyme duration itself would be ambiguous with regard to coda voicing. The mean values of vowel duration and coda closure duration across both voiced and voiceless coda contexts (as reported in Penney et al., 2018) were taken as the default rhyme duration, that is, the ambiguous rhyme duration (e.g. for /iː/ the mean vowel duration was 204 milliseconds and the mean coda closure duration was 82 milliseconds therefore the rhyme for the /iː/ C/V continuum was 286 milliseconds). For the first step in each continuum the vowel and coda closure both occupied an equal share of the rhyme. In each of the subsequent steps the vowel proportion increased by 5% while the closure proportion of the rhyme decreased by 5%. In the ninth step the vowel occupied 90% of the rhyme and the closure occupied 10%. Table 3 lists the overall rhyme duration for each source vowel, as well as the duration of vowel and coda closure at the first and ninth steps.

Table 3. Overall duration of rhyme, and vowel and closure durations at shortest and longest steps each vowel continuum.

Source vowel	Rhyme duration (ms)	First step vowel/closure duration (ms)	Ninth step vowel/closure duration (ms)
iː	286	143/143	257/29
ɪ	210	105/105	189/21
ɐː	338	169/169	304/34
ɐ	231	115.5/115.5	208/23

3.2.3.4 F0 manipulation and intensity scaling

In each of the three manipulations, the F0 at the onset of the vowel was set at 265hz, and the F0 at the offset of the vowel was set at 203hz. This strategy maintained consistency across tokens and across continua, and ensured a standardised F0. The F0 values were calculated according to the means for the speaker over all of her tokens in both voiced and voiceless coda contexts. Finally, the intensity of all tokens was scaled to 70dB.

3.2.3.5 Glottalised stimuli

In order to test the effects of glottalisation we then created a second ‘glottalised’ set of each of the continua described above. These glottalised stimuli were identical to the original stimuli but we manually spliced glottalisation into the final portion of each vowel. The glottalised portions were taken from the sustained vowels produced by our speaker with creaky voice,[4] and were matched according to vowel (e.g. the creaky production of sustained /iː/ was used as the source for the glottalisation in the word *bead*). Following proportions reported in Penney et al. (2018), the final 25% of each inherently long vowel (at each step) and the final 35% of each inherently short vowel (at each step) was replaced with glottalisation. Figure 1 displays the waveforms and spectrograms in both non-glottalised and glottalised conditions for the ninth step (i.e. longest vowel duration) of the source *bard* vowel duration continuum. Examples of the auditory stimuli in both non-glottalised and glottalised conditions can be found in the supplementary materials.

3.2.4 Data analysis

We fitted mixed effects logistic regression models (GLMER) to the results of each of the manipulations using the *lme4* package (Bates, Maechler, Bolker, & Walker, 2015)

[4] We trialled a number of methods to create the glottalised stimuli, including manipulation of open quotient, spectral tilt, flutter and double pulsing as in Chong & Garellek (2018), and manipulation of F0 as in Crowhurst (2018). These trials produced stimuli in which the ‘glottalisation’ was either not sufficiently salient or sounded too artificial. We also tried manually splicing in examples of coda glottalisation, rather than creaky voice (as differences may exist between the two, see for example Garellek & Seyfarth, 2016); however, such examples were necessarily taken from words ending with voiceless codas and as such provided too strong a bias towards coda voicelessness. In addition, it was difficult to find natural examples of coda glottalisation that were long enough to take up the necessary 25 or 35% of the vowels with the longest durations. Thus we chose this method as it resulted in stimuli with the most salient and natural sounding glottalisation, it produced glottalised portions long enough for the required portion of the longest vowel duration steps, and it did not contain cues to coda voicing.

in R (R Core Team, 2016). That is, we a fitted separate models for the vowel duration manipulation, the coda closure duration, and the C/V ratio manipulation. The dependent variable in each of the models was listener response: selecting either a word with a voiced or a voiceless coda stop (i.e. /d/ or /t/). The fixed factors included in each of the models were: *continuum step* (coded as continuous and centred on 0), *condition* (whether a stimulus was non-glottalised or glottalised with non-glottalised as reference level), *inherent vowel length* (*IVL*; short or long with long as reference level), *vowel height* (high or low with high as reference level), and *age group* (older or younger with older as reference level). In order to test our specific hypotheses (outlined above) we initially included the following three-way interaction terms (which also include all lower order two-way interactions) in each model: *step*condition*age*, *step*condition*IVL*, *step*condition*vowel height*, *condition*age* IVL*, *condition*age*vowel height*, *step*age* IVL*. Random intercepts were included in each model for *participant*. Random slopes were included for all within-subjects factors by participant. In all cases this maximal random effects structure produced a better fit than a reduced random effects structure. We then identified the most parsimonious model in each case. First, using the *drop1* function, we identified the terms in each model that would improve model fit if removed. We proceeded to remove these terms (one at a time and beginning with the term that would improve model fit the most) and compared the resulting model with the previous model using the *anova* function with reference to AIC values until we identified the model with the best fit. For the vowel duration manipulation model the best fitting model was one in which the three-way interaction terms for *step*age* IVL* and *condition*age* IVL* were removed (AIC: 6923.2 vs 6928.2 for the full model).[5] For the closure duration manipulation model the best fitting model was one in which the interaction term for *condition*age* IVL* was removed (AIC: 7510.4 vs 7512.3 for the full model).[6] For the C/V ratio manipulation model the best fitting model was one in which the interaction term for *step*age* IVL* was removed (AIC: 7160.9 vs 7162.9 for the full model).[7]

We conducted power analyses using the simR package (Green & MacLeod, 2016) in R (R Core Team, 2016). The results indicate that our study is sufficiently powered, based on 1000 Monte Carlo simulations with a power threshold of 80%, with effect sizes set at moderate values substantially below those found in our preliminary results. In the results section below we focus only on the highest order significant effects that involve each particular factor. Model summaries for each of the mixed effect regression models are provided in Tables A2-A4 in Appendix B. In addition, the data sheet, code, and model summaries are available in the supplementary materials. All post-hoc analyses were conducted using the *lsmeans* package (Lenth, 2016), using Tukey HSD corrections to *p*-values for multiple comparisons.

---

[5] The syntax for the final vowel duration manipulation model was: glmer(response ~ step*condition*age + step*condition*vowel length + step*condition*vowel height + age*condition*vowel height + (1 + step + condition + vowel height | participant)).

[6] The syntax for the final coda closure duration manipulation model was: glmer(response ~ step*condition*age + step*condition*vowel length + step*condition*vowel height + age*condition*vowel height + step*vowel length*age + (1 + step + condition + vowel height | participant).

[7] The syntax for the final C/V ratio manipulation model was: glmer(response ~ step*condition*age + step* condition*vowel length + step*condition*vowel height + age*condition*vowel height + age*condition*vowel length + (1 + step + condition + vowel height | participant)).

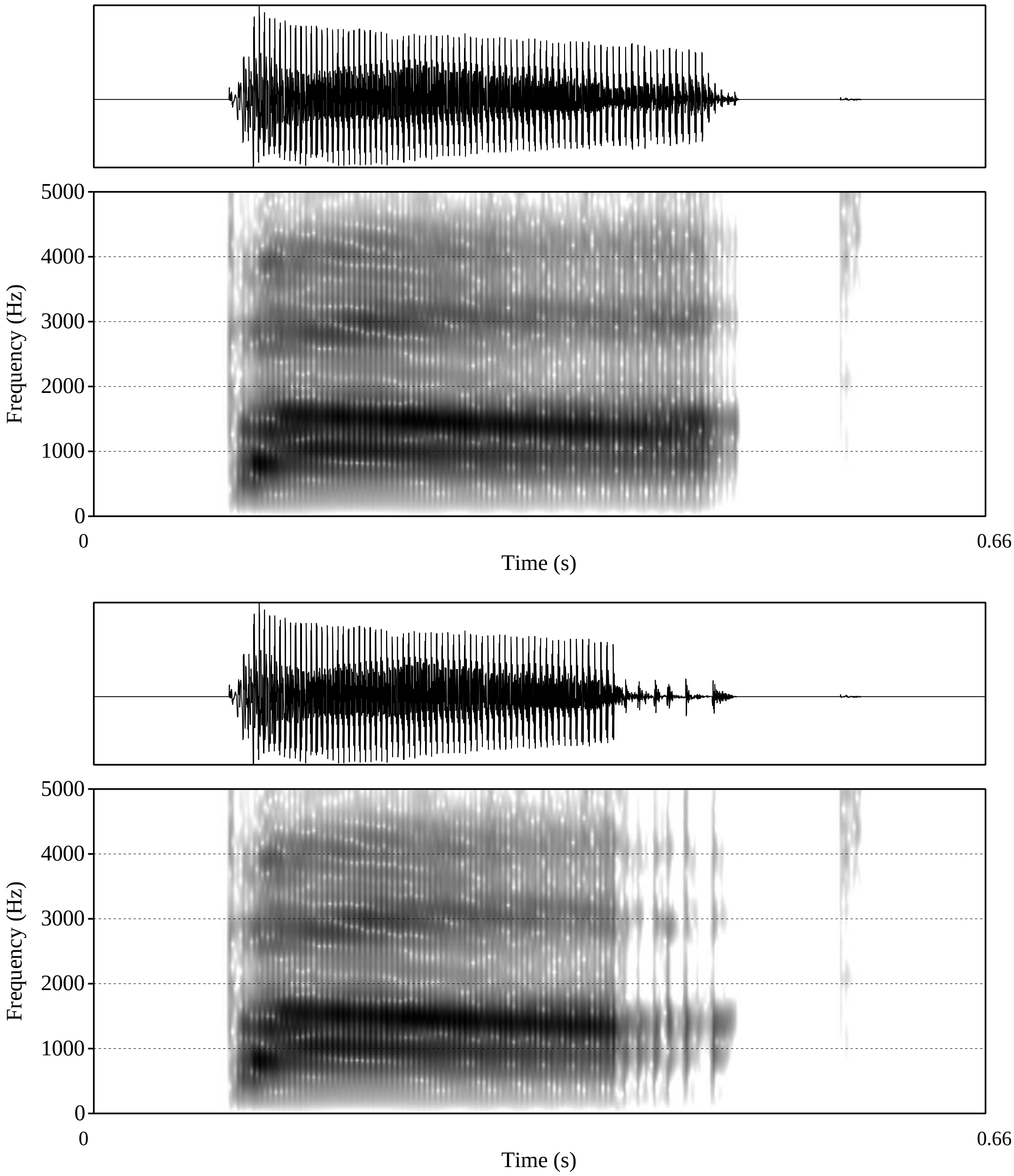

Figure 1. Waveform and spectrogram of ninth step stimuli for source 'bard' continuum in non-glottalised condition (upper panel) and glottalised condition (lower panel) in vowel manipulation.

3.3 Results

3.3.1 Manipulation of vowel duration

Figure 2 shows the proportion of voiced responses to the vowel duration manipulation stimuli in the non-glottalised and glottalised conditions at each step according to age group (averaged across inherent vowel length and height tokens). Overall, in the non-glottalised condition, for both age groups the proportion of voiced responses is low at the lower steps and this increases as vowel duration increases, with a high proportion of voiced responses at the higher steps. This shows that, consistent with our first hypothesis, listeners in both age groups utilise preceding vowel duration as a cue to coda stop voicing, with shorter vowel duration cueing perception of voiceless coda stops, and longer vowel duration cueing perception of voiced coda stops, albeit with some differences according to age. In the glottalised condition, however, it can be seen that the proportion of voiced responses is lower, even at the higher steps. This suggests that listeners are sensitive to the presence of glottalisation and utilise it as a cue to coda voicelessness, providing strong support for our fourth hypothesis that glottalisation facilitates the perception of coda voicelessness.

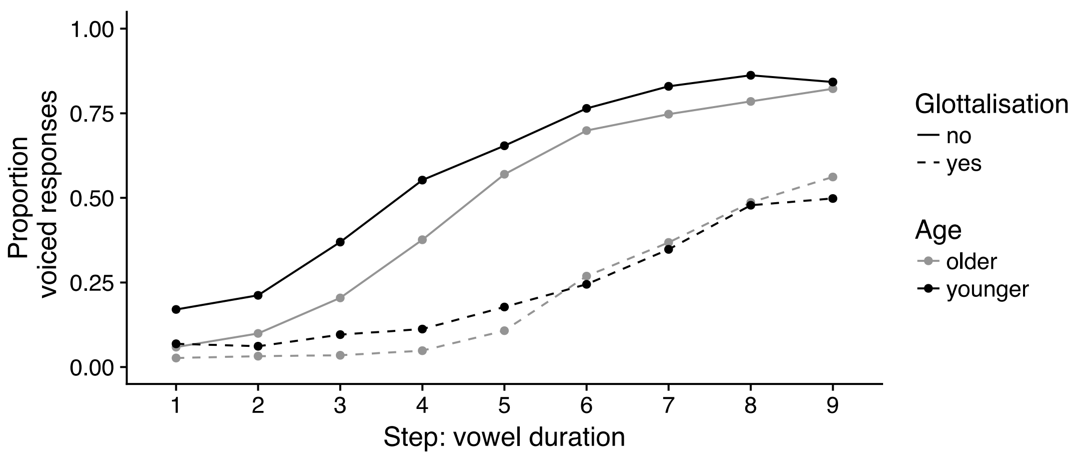

Figure 2. Proportion of voiced responses to vowel duration manipulation for older (grey lines) and younger listeners (black lines) in non-glottalised (solid lines) and glottalised (dashed lines) conditions.

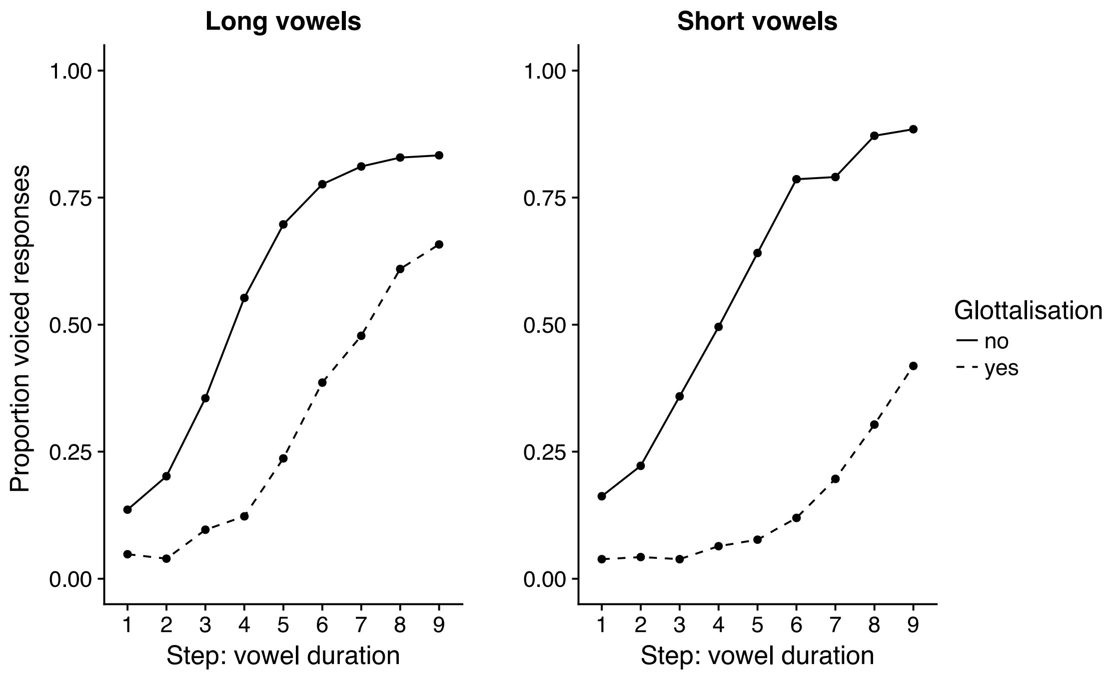

Figure 3. Proportion of voiced responses to vowel duration manipulation for inherently long vowels (left panel) and inherently short vowels (right panel) in non-glottalised (solid lines) and glottalised (dashed lines) conditions.

Figure 3 shows the proportion of voiced responses for inherently long and inherently short vowels in the glottalised and non-glottalised conditions. As can be seen, the effect of glottalisation in facilitating a greater proportion of voiceless responses appears to be stronger for the inherently short vowels than for the inherently long vowels; the proportion of voiced responses remains low throughout the short vowel continua, and even at the highest step (i.e. the longest vowel duration) approximately half of the responses were for voiceless codas. The vowel duration manipulation model returned a significant two-way interaction between *condition* and *IVL* (β = 1.205; OR = 3.34; SE = 0.336; z = 3.581; p = 0.0003). Post-hoc comparisons did not find evidence that the inherently short and inherently long vowels differed in the non-glottalised condition, but they did differ significantly in the glottalised condition (p < 0.0001).

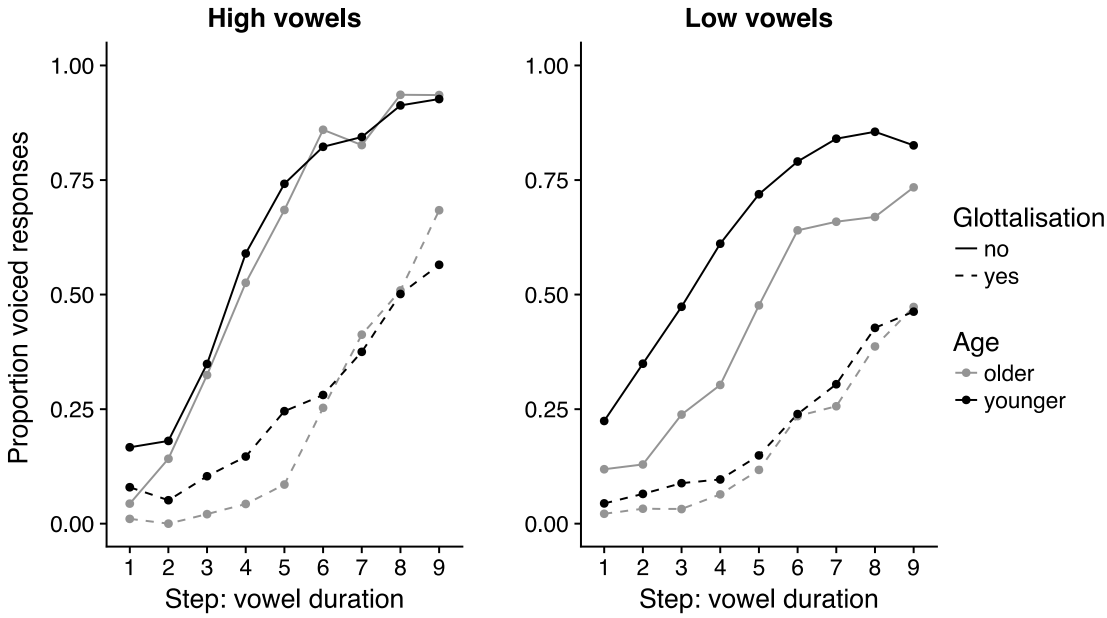

Figure 4. Proportion of voiced responses to vowel duration manipulation for high vowels (left panel) and low vowels (right panel) for older (grey lines) and younger listeners (black lines) in non-glottalised (solid lines) and glottalised (dashed lines) conditions.

The model also showed a significant three-way interaction between *condition*, *age*, and *vowel height* (β = 1.150; OR = 3.16; SE = 0.273; z = 4.220; p < 0.0001). Figure 4 shows that both age groups produce fewer voiced responses in the glottalised condition compared to the non-glottalised condition, and post-hoc comparisons confirm that both high and low vowels differ significantly between conditions within each age group (all p < 0.0001). However, it can also be seen from Figure 4 that the older speakers produce fewer voiced responses for the low vowels (source *bard* and *bud*) in the non-glottalised condition (proportion averaged over all steps: older: 0.44; younger: 0.63). Post-hoc tests revealed no significant differences between the older and younger listeners for either high or low vowels in the glottalised condition, nor for high vowels in the non-glottalised condition; on the other hand, the older and younger listeners differ significantly from each other for low vowels in the non-glottalised condition (p = 0.0017).

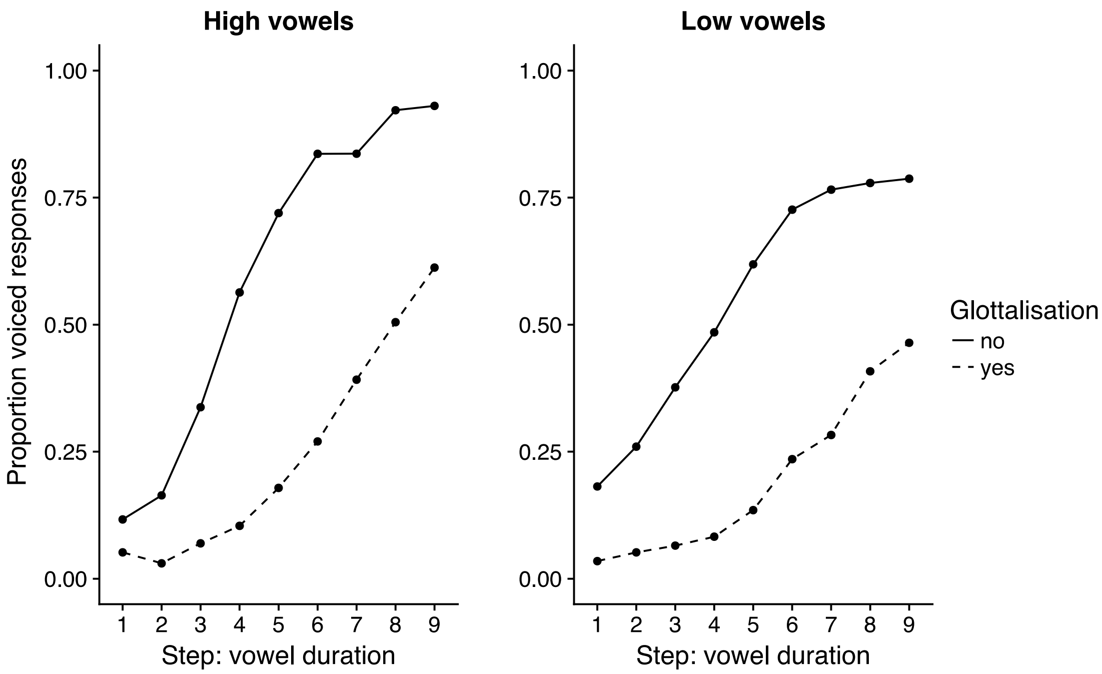

Figure 5. Proportion of voiced responses to vowel duration manipulation for high vowels (left panel) and low vowels (right panel) in non-glottalised (solid lines) and glottalised (dashed lines) conditions.

We also found a significant three-way interaction between *step*, *condition*, and *vowel height* (β = -0.155; OR = 0.86; SE = 0.060; z = -2.589; p = 0.010), suggesting differences between the high and low vowels at particular steps in the two conditions. Figure 5 illustrates the proportion of voiced responses in the non-glottalised and glottalised conditions for high and low vowels. Differences can be seen between high and low vowels in both conditions. Post-hoc tests reveal that in the non-glottalised condition high and low vowels differ significantly at all steps except steps 1-3 (4 (p = 0.019), 5-9 (p < 0.0001)). In the glottalised condition high and low vowels differ significantly at steps 5 to 9 (5 (p = 0.033), 6 (p = 0.002), 7 (p = 0.0002), 8-9 (p = 0.0001). The differences in the non-glottalised condition are presumably due to the older listeners showing a stronger preference for voiceless responses (as discussed above). In the glottalised condition the differences are primarily in the higher continuum steps, where the low vowels show lower proportions of voiced responses compared to the high vowels.

3.3.2 Manipulation of coda closure duration

Figure 6 shows the proportion of voiced responses to the coda closure duration manipulation stimuli in the non-glottalised and glottalised conditions at each step according to age group (averaged across inherent vowel length and height tokens). As can be seen, in both conditions there is a slight decrease in the proportion of voiced responses as closure duration increases. The coda closure duration manipulation model showed a significant effect of *continuum step* (β = 0.115; OR = 1.12; SE = 0.034; z = 3.440; p = 0.001), confirming that increasing the duration of the coda closure results in a lower proportion of voiced responses. Nevertheless, the proportion of voiced responses never falls below 50 per cent in either of the age groups in the non-glottalised condition, which may suggest that coda closure duration alone is not a salient cue to voicing for listeners. In the glottalised condition, the proportion of voiced responses is substantially lower at every step compared to the non-glottalised condition, providing further support for our prediction that the presence of glottalisation cues coda voicelessness.

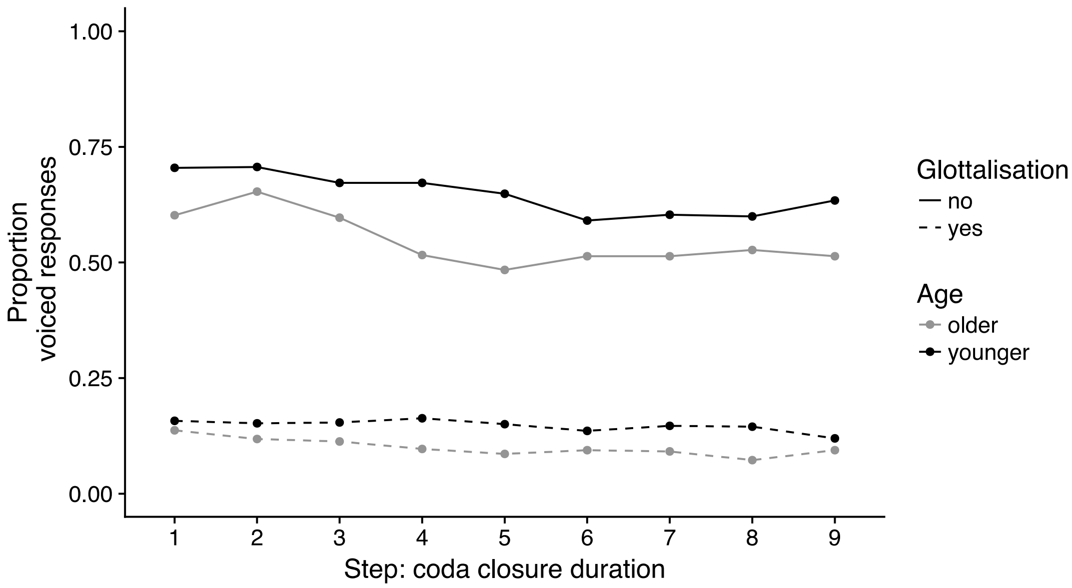

Figure 6. Proportion of voiced responses to coda closure duration manipulation for older (grey lines) and younger listeners (black lines) in non-glottalised (solid lines) and glottalised (dashed lines) conditions.

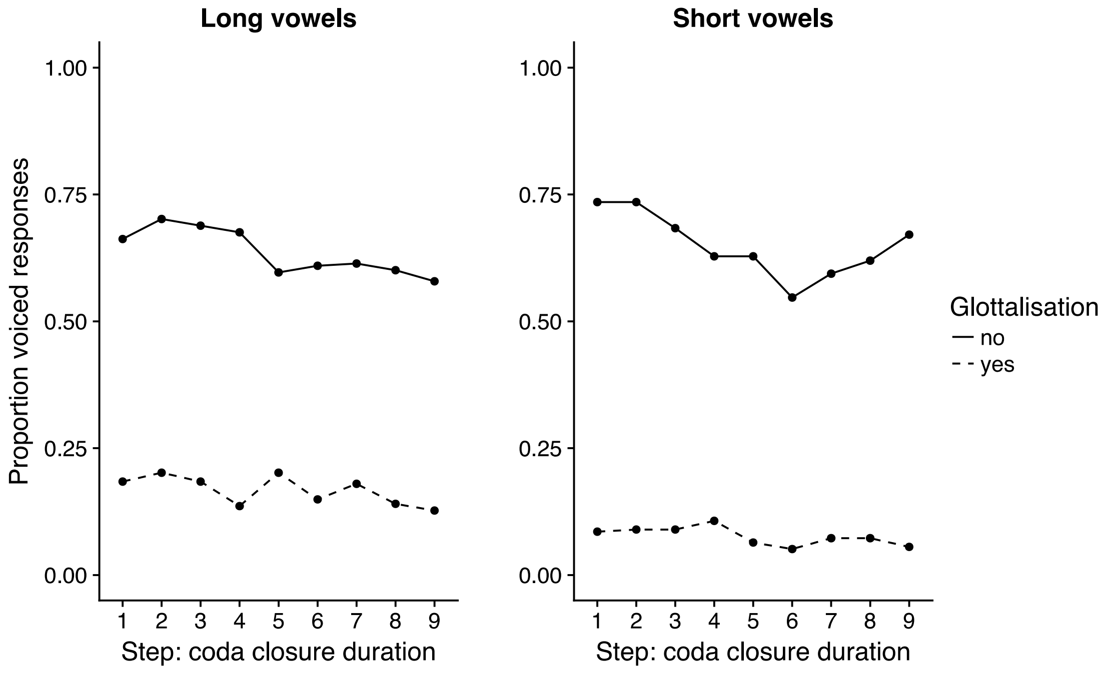

Figure 7. Proportion of voiced responses to coda closure duration manipulation for inherently long vowels (left panel) and inherently short vowels (right panel) in non-glottalised (solid lines) and glottalised (dashed lines) conditions.

As in the vowel duration manipulation model, for the coda closure manipulation we found a significant two-way interaction between condition and IVL (β = 0.943; OR = 2.57; SE = 0.330; z = 2.854; p = 0.004). Figure 7 displays the proportion of voiced responses for inherently long and inherently short vowels in both of the conditions, and shows that the inherently short vowels were less likely to be classified as voiced in the glottalised condition than the inherently long vowels. Post-hoc tests show that the inherently long and inherently short vowels do not differ from each other in the non-glottalised condition, but do differ from each other in the glottalised condition (p = 0.019).

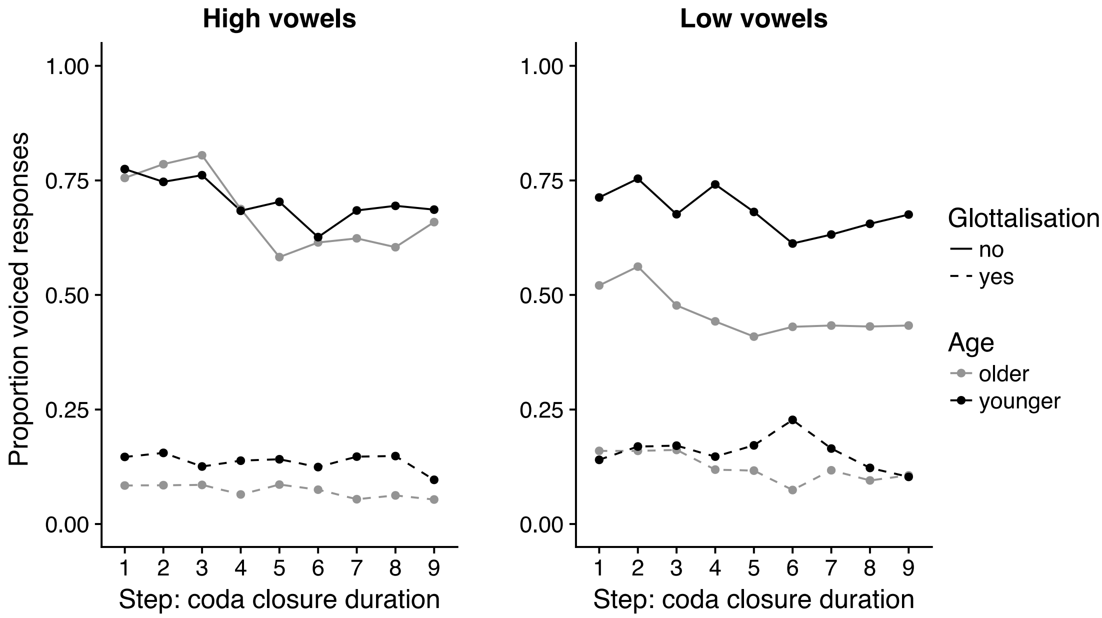

Figure 8. Proportion of voiced responses to coda closure duration manipulation for high vowels (left panel) and low vowels (right panel) for older (grey lines) and younger listeners (black lines) in non-glottalised (solid lines) and glottalised (dashed lines) conditions.

We also found a significant three-way interaction between *condition*, *age group*, and *vowel height* (β = 1.634; OR = 5.12; SE = 0.284; z = 5.758; p < 0.0001). Figure 8 shows the proportion of voiced responses in each condition for older and younger listeners for high vowels and low vowels. The older and younger listeners differ markedly in their responses to the low vowels in the non-glottalised condition (proportion averaged over all steps: older: 0.46; younger: 0.68). Post-hoc tests verify that in the non-glottalised condition younger and older listeners differ for the low vowels ($p = 0.001$) but not for the high vowels, and in the glottalised condition for neither the high nor the low vowels.

3.3.3 Manipulation of C/V ratio

Figure 9 shows the proportion of voiced responses to the C/V ratio manipulation stimuli in the non-glottalised and glottalised conditions at each step according to age group (averaged across inherent vowel length and height tokens). The figure shows a similar overall pattern to the vowel duration manipulations shown in Figure 2 above: in the non-glottalised condition, both age groups show low proportions of voiced responses at the lower steps, but increasing voiced responses as the proportion of the rhyme occupied by the vowel increases relative to the coda closure); the proportion of voiced responses is lower in the glottalised condition compared to the non-glottalised at each step, providing further support that glottalisation facilitates the perception of voicelessness.

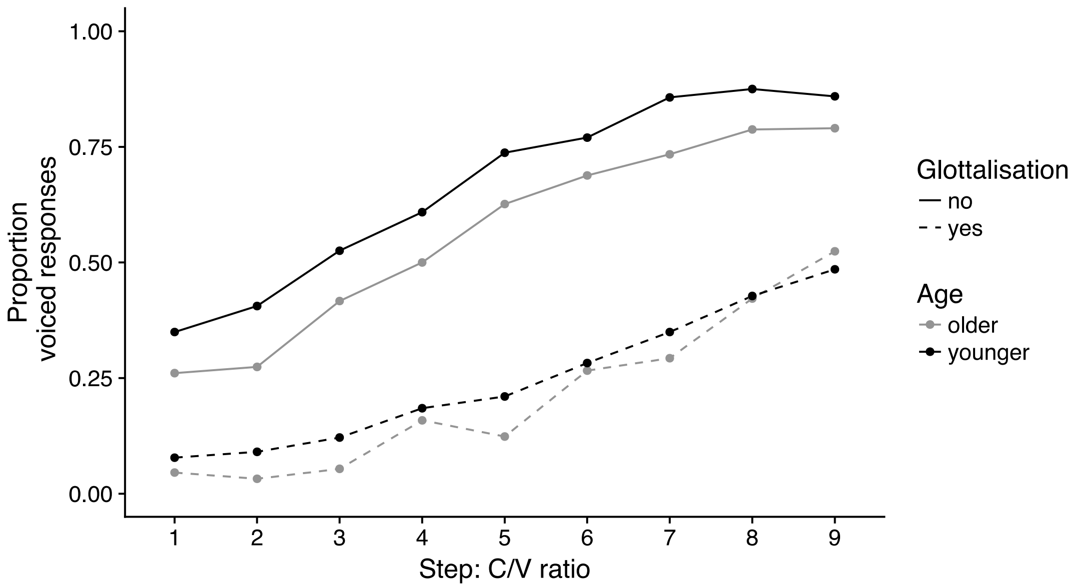

Figure 9. Proportion of voiced responses to C/V ratio manipulation for older (grey lines) and younger listeners (black lines) in non-glottalised (solid lines) and glottalised (dashed lines) conditions.

In contrast to the vowel duration manipulation (see Figure 2), Figure 9 shows that at the lowest step (i.e. where vowel and coda closure occupy equivalent portions of the rhyme) in the non-glottalised condition, over a quarter of responses from listeners in both age groups were nevertheless for a voiced coda, whereas in the vowel duration manipulation the proportion of voiced responses was lower. This is likely due to the fact that the absolute duration of the rhyme did not change in the C/V ratio manipulations and was based on mean vowel and closure duration across both voiced and voiceless coda contexts, whereas the vowel duration manipulations had more extreme durations at the first and ninth steps. The result of this was vowel durations for the inherently short vowels in the C/V ratio manipulation that were substantially longer than in the vowel duration manipulation at the lowest step. For example, vowel duration for /ɪ/ in the vowel duration manipulation extended from 65 milliseconds at the lowest step to 166 milliseconds at the highest step; in the C/V ratio, on the other hand, vowel duration at the lowest step was 105 milliseconds and 189 milliseconds at the highest step. Thus, unsurprisingly, the results of the C/V ratio manipulation model showed a significant interaction between *condition* and *IVL* (β = 1.015; OR = 2.76; SE = 0.456; z = 2.227; p = 0.026), demonstrating that the proportion of voiced responses was higher for the inherently short vowels in the non-glottalised condition.

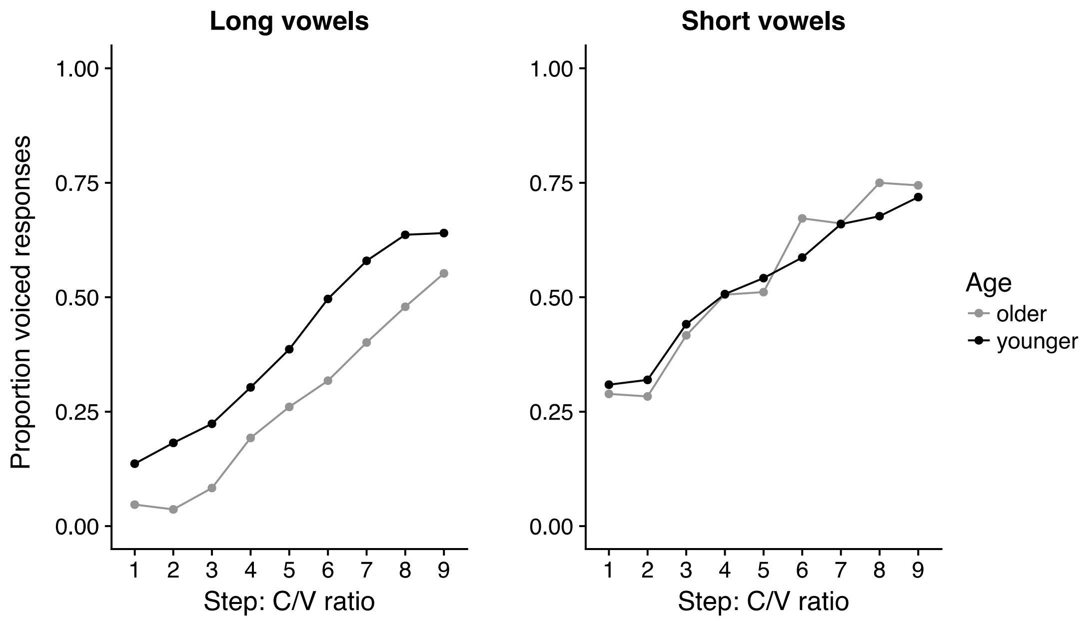

Figure 10. Proportion of voiced responses C/V ratio manipulation for inherently long vowels (left panel) and inherently short vowels (right panel) for older (grey lines) and younger (black lines) age group.

We also found a significant two-way interaction between *age group* and *IVL* (β = 1.100; OR = 3.0; SE = 0.420; z = 2.621; p = 0.009). Figure 10 illustrates the proportion of voiced responses at each step according to inherent vowel length. While both older and younger listeners' responses differed between inherently long and inherently short vowels (with higher proportions of voiced responses to the short vowels), the older listeners had a lower proportion of voiced responses to the inherently long vowels than the younger listeners (proportion averaged over all steps: older: 0.26; younger: 0.40). This was due primarily to the older listeners showing a preference for voiceless responses for the long low vowel (source *bard*). Post-hoc analysis confirmed that the old and young listeners did not differ from each other for the inherently short vowels, but differed significantly for the inherently long vowels (p = 0.0003).

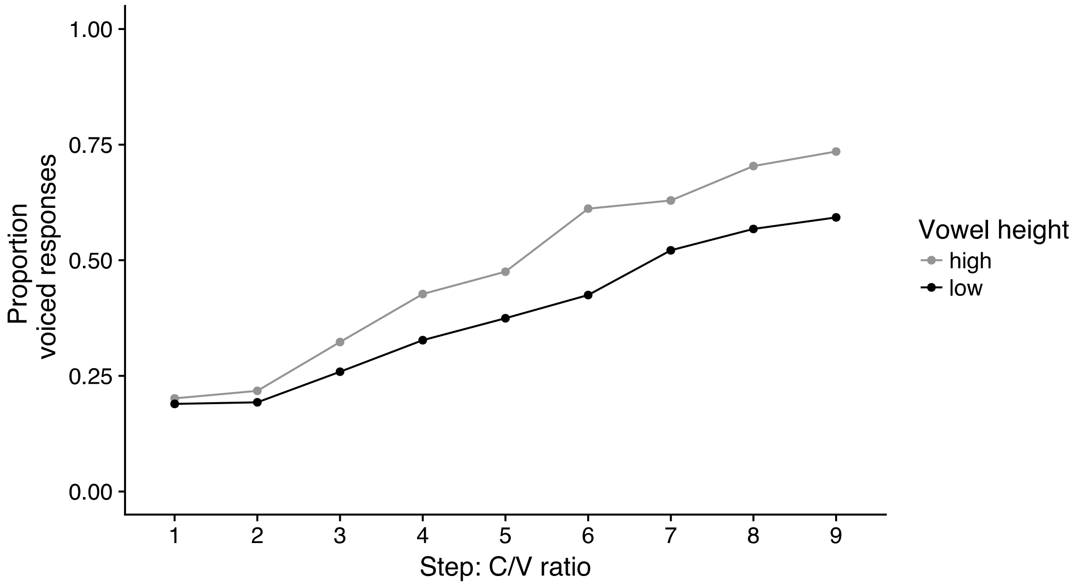

Figure 11. Proportion of voiced responses to C/V ratio manipulation for high vowels (grey lines) and low vowels (black lines) at each continuum step.

The C/V ratio manipulation model also showed a significant two-way interaction between *continuum step* and *vowel height* (β = 0.124; OR = 1.13; SE = 0.038; z = 3.300; p = 0.001). Figure 11 shows the proportion of voiced responses at each step according to vowel height. As can be seen, low vowels were more likely to be perceived as having voiceless codas than were high vowels, especially at the higher steps. We also found a significant three-way interaction between *condition*, *age*, and *vowel height* (β = 0.766; OR = 2.15; SE = 0.270; z = 2.837; p = 0.005). Figure 12 displays the proportion of voiced responses in each condition for older and younger listeners for high vowels and low vowels. For both older and younger listeners the effect of glottalisation is a decrease in voiced responses for both high and low vowels. As in the vowel duration and coda closure duration manipulations, older listeners produced more voiceless responses than younger listeners in the non-glottalised condition for low vowels. Post-hoc tests confirm that the younger and older listeners differ significantly from each other for low vowels in the non-glottalised condition ($p = 0.0001$), but are not significantly different for high vowels. Younger and older listeners did not differ from each other for either low or high vowels in the glottalised condition. In addition, for the older listeners high vowels differed significantly from low vowels in both conditions (non-glottalised: $p < 0.0001$; glottalised: $p = 0.012$), whereas for the younger listeners there was no significant difference in either condition.

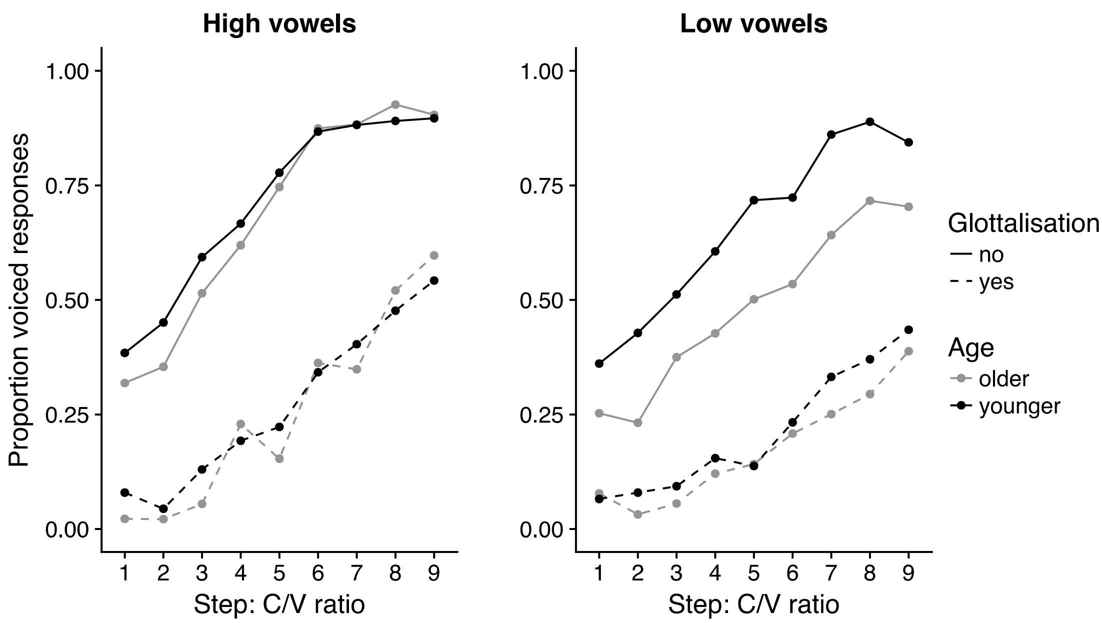

Figure 12. Proportion of voiced responses to C/V ratio manipulation for high vowels (left panel) and low vowels (right panel) for older (grey lines) and younger listeners (black lines) in non-glottalised (solid lines) and glottalised (dashed lines) conditions.

3.3.4 Comparison of modally voiced portions of vowels in non-glottalised and glottalised conditions

As glottalisation necessarily results in non-modal phonation, it may be possible that, rather than glottalisation facilitating listeners' perception of voicelessness, listeners are simply tuning in to the shorter modally voiced portion of the vowels in the glottalised stimuli rather than being sensitive to the glottalised portion of the vowel.[8] If this were the case, each stimulus in the glottalised condition could be considered to have a shorter vowel duration compared to the corresponding stimulus in the non-glottalised condition. In order to test the possibility that listeners are responding to the shorter modally voiced part of the vowel, we calculated the duration of the modally voiced portions of the vowels in the glottalised stimuli and compared these to the stimuli with the closest corresponding vowel durations in the non-glottalised condition.[9] If the listeners' responses were based on the modally voiced portions of the glottalised vowels without regard for glottalisation, the proportion of voiced responses should be similar to the stimuli in the non-glottalised condition with corresponding vowel durations.

We recoded the data to compare only the tokens with corresponding modally voiced vowel durations and fitted a mixed effects logistic regression model to the newly coded data. For /iː/, six steps approximately corresponded to modally voiced vowel durations; for /ɪ/ there were four approximately corresponding steps; for /ɐː/ and /ɐ/ there were five corresponding steps. Table A5 in Appendix C lists the durations of the modally voiced portions of the glottalised stimuli for each vowel and the corresponding continuum step with similar vowel durations. As above, the dependent variable was the listener's response (voiced or voiceless coda). Fixed factors were *recoded continuum step*, *age group*, and *condition* (non-glottalised vs modally voiced portion of glottalised stimuli). We also included a two-way interaction between *age group* and *condition*. We included random intercepts for *participant* and

[8] We thank Jonathan Harrington for bringing this to our attention.

[9] We are grateful to Eva Reinisch, who suggested this approach.

random slopes for (recoded) *continuum step* and *condition*. Table A6 in Appendix C contains the summary of this model.

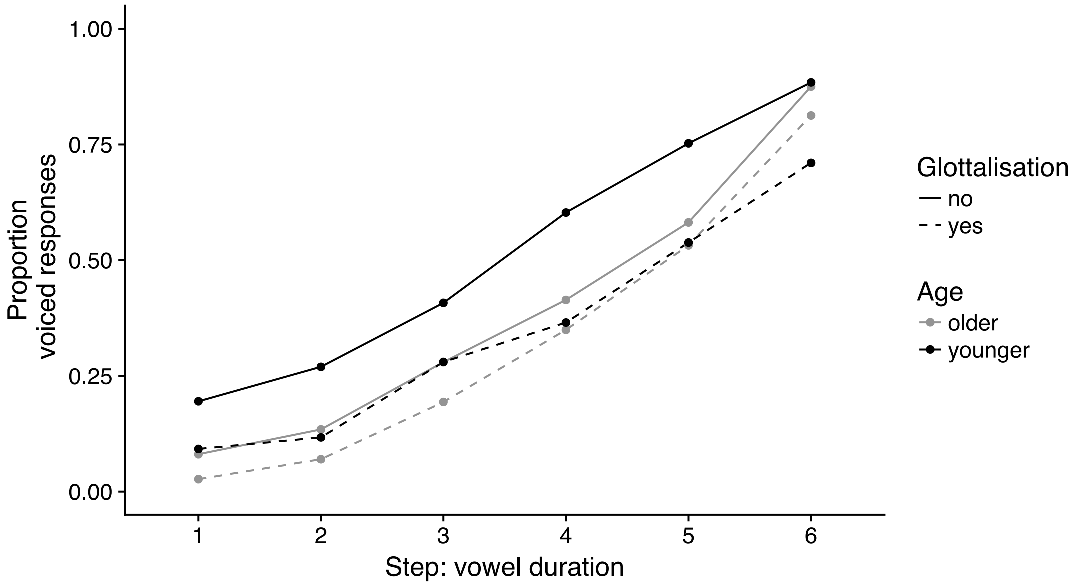

Figure 13. Proportion of voiced response to non-glottalised stimuli (solid lines) and modally voiced portions of glottalised stimuli (dashed lines) by older (grey lines) and younger listeners (black lines).

Figure 13 illustrates the differences in proportion of voiced responses between the two conditions for older and younger listeners (averaged across inherent vowel length and height tokens). As can be seen, there is a difference between the responses to the non-glottalised stimuli and the modally voiced portions of the glottalised stimuli for both age groups, though the difference is smaller for the older group, suggesting that younger listeners utilise glottalisation to a greater degree. The results of the model showed significant effects for recoded *continuum step* (β = -0.820; OR = 0.44; SE = 0.051; z = -16.073; p < 0.0001)*, age group* (β = -0.833; OR = 0.44; SE = 0.207; z = -4.023; p < 0.0001), and *condition* (β = 0.580; OR = 1.79; SE = 0.217; z = 2.671; p = 0.008). The results show that that, as expected, the proportion of voiced responses increased as the steps increased (i.e. as vowel duration increased), that there were differences between the age groups in their responses, and that there were differences between the conditions (non-glottalised vs the modally voiced portion of the glottalised stimuli). Importantly, we found no significant interaction between *condition* and *age group*, demonstrating that although the older listeners differed from the younger listeners, both age groups nevertheless differed between the two conditions. These results show that the responses of listeners in both age groups differed between the modally voiced portions of the glottalised stimuli and the non-glottalised tokens of corresponding duration and thereby demonstrate that listeners did not simply perceive shorter vowels in the glottalised stimuli. This supports our interpretation of the results above that listeners utilise glottalisation to cue coda voicelessness.

3.4. Discussion

The main aim of this paper was to investigate how AusE listeners utilise preceding vowel duration, coda closure duration, and glottalisation as cues to coda voicing. The results of both the vowel duration and C/V ratio manipulations support our first and third hypotheses that listeners use preceding vowel duration and changes in the C/V ratio of the rhyme as cues to coda stop voicing (though see below regarding the role of vowel duration in C/V ratio). The proportion of voiced responses was low when listeners were presented with shorter vowels and increased as the vowels became longer, as can be seen in the non-glottalised condition in Figures 2 and 9. The results also suggest that the duration of the coda closure is a much weaker voicing cue, supporting our second hypothesis that listeners would not be as sensitive to this cue as they are to the vowel duration cue. While we did find a significant effect of longer coda closures leading to more voiceless coda stop percepts, the differences in listener responses between the lowest and the highest steps were small. This finding suggests that coda closure duration on its own may not be a sufficient to cue coda voicing, but rather its duration relative to the duration of the preceding vowel is important (Port, 1981; Port & Dalby, 1982; Luce & Charles-Luce, 1985). Indeed, our manipulation of C/V ratio showed greater differences over the continua than manipulating only the coda closure. However, the effects that we found in the C/V ratio manipulation largely mirrored those found in the vowel duration manipulation, suggesting that vowel duration may be the more important element of the rhyme for the voicing contrast, in line with previous findings by Luce and Charles-Luce (1985). That is, the increase in vowel duration may drive the effects in the C/V ratio manipulation, rather than changes in the relative proportions of the rhyme per se.

Our fourth hypothesis that the presence of glottalisation would facilitate the perception of coda voicelessness was also strongly supported by the results. In each of the three manipulations the proportion of voiceless responses in the glottalised condition was significantly higher than in the non-glottalised condition. Note that this effect held even in the presence of extended vowel duration (i.e. at the higher steps in the vowel duration and C/V ratio continua), which is a strong cue to a voiced coda stop. That is, glottalisation resulted in an increase in voiceless percepts despite the presence of a competing cue. These results indicate that AusE listeners associate glottalisation with voiceless codas, in line with findings for AmE (Chong & Garellek, 2018). However, Chong and Garellek (2018) did not find that glottalisation improved AmE listeners' perception of voicelessness, but rather that it only inhibited their perception of a coda being voiced. In contrast, our results suggest that glottalisation promotes voicelessness even in the face of a competing cue to voicing (i.e. extended preceding vowel duration). We interpret this as support for the hypothesis that glottalisation facilitates perception of coda voicelessness, in addition to hindering perception of coda stop voicing as found by Chong and Garellek (2018). Why did Chong and Garellek (2018) not find evidence of facilitation of voicelessness though we have? This may be due to the availability of other cues to voicing being present in their stimuli (e.g. differences in F0 or F1), whereas in this study we controlled for additional voicing cues. Another possibility is that AmE and AusE listeners differ in their use of glottalisation as a cue to coda stop voicing. It would be interesting to examine speaker/listeners of different English varieties to explore possible dialectal differences in the perception of voicing contrasts.

Coda voicing-related vowel duration differences in AusE are greater for inherently long vowels than for inherently short vowels (Cox & Palethorpe, 2011; Penney et al., 2018); thus our fifth hypothesis was that the listener response to preceding vowel duration would be weaker for inherently short vowels. The results do not support this hypothesis; listeners made use of vowel duration to cue coda voicing in both inherently short and inherently long vowel contexts, and we found no significant difference between vowel length categories in the non-glottalised condition. This suggests that although vowel duration is a weaker cue for inherently short vowels in production, listeners can nonetheless use it reliably in perception. We did find, however, that the effect of glottalisation was stronger for inherently short vowels compared to inherently long vowels. In the glottalised condition of the vowel duration manipulation the proportion of voiced responses was lower for the inherently short vowels than for the inherently long vowels at the higher steps (i.e. when vowel duration was at its greatest) and in the coda closure duration manipulation the inherently short vowels also showed significantly lower proportions of voiced responses than the inherently long vowels, thus supporting our sixth hypothesis that the effect of glottalisation in facilitating the perception of coda voicelessness is stronger for inherently short vowels than for inherently long vowels.

The results also provided some support for our seventh hypothesis that the effect of glottalisation would be stronger for low vowels than for high vowels. In the glottalised condition of both the vowel duration manipulation and the C/V ratio manipulation, we found a greater proportion of voiceless responses for the low vowels than for the high vowels at the higher steps, suggesting that the effect of glottalisation in cueing coda voiceless is greater for low vowels. In our previous research we found that the most robust use of vowel duration differences to cue coda voicing occurred in the production of high vowels, and that complementarily, speakers used glottalisation less in high vowel contexts than non-high vowel contexts (Penney et al., 2018). In addition, other studies have also found that glottalisation tends to favour low vowels (Brunner & Żygis, 2011; Hejná & Scanlon, 2015; Malisz, Żygis, & Pompino-Marschall, 2013; Pompino-Marschall & Żygis, 2010). Therefore, it is not surprising that listeners may use glottalisation more in low vowel contexts, and this result may provide evidence for a trading relationship between vowel duration and glottalisation as cues to coda voicing. This result, however, should be taken with caution, given that we found older listeners had a strong bias towards voiceless codas for the stimuli with low vowels in the non-glottalised condition (especially for *bard/bart*, but also to a lesser extent for *bud/but*) in each of the manipulations (in the glottalised condition listeners in both age groups responded similarly). At first glance this may appear to be a simple frequency effect, as *bart* and *but* are both more frequent than *bard* and *bud* respectively. Nevertheless, if this were the case we would expect younger listeners to also show a similar bias, which is not evident in the results. In addition, as noted above, listeners took part in a number of tasks prior to the experiment including a production task in which all of the words were included, as well as familiarisation and training phases. These preliminary tasks would have raised the activation of each word; therefore, it is not clear why the older listeners in particular display this bias to voiceless codas for low vowels.

Our final hypothesis was that the effect of glottalisation would not be as prominent in the older listeners, based on previous findings that older speakers glottalise less than younger speakers in production. This hypothesis was not supported by the results. We found no age effects for glottalisation (though we did find age effects relating to low vowels in the non-glottalised condition; see discussion above); both younger and older listeners showed the same pattern of increased perception of coda stop voicelessness when glottalisation was present. That is, older listeners appear to utilise glottalisation in perception in much the same way as do younger listeners, despite utilising this feature less in production. As already discussed, glottalisation appears to be a recent change to AusE, and consistent with leading the change, younger speakers employ glottalisation at higher rates than older speakers (Penney et al., 2018; Penney et al., 2019 [Chapter 2]; Tollfree, 2001). This result may therefore be interpreted in line with studies of sound change that posit a shift in listeners' perceptions prior to a change in speakers' productions (Harrington et al., 2008; Kleber et al., 2012; Kuang, 2018; Ohala, 1981). That is, it may be the case that glottalisation is a change led by perception, which would explain why both older and younger listeners show uniformity in utilising this cue. In production, however, which in such a perception-led change would lag behind a shift in perception, the younger speakers demonstrate a more advanced stage in the change process — consistent with leading the change —whereas the older speakers display a less innovative utilisation of this feature.

Alternatively, it may be possible that the older listeners are utilising cues in perception that they do not make use of themselves in production, or at least make less use of, but that they are aware of as being present in the community. That is, older listeners may be aware that younger speakers utilise glottalisation more frequently in production than older speakers do, and therefore associate glottalisation with younger speakers based on their experience as listeners. Many studies have shown that listeners are sensitive to social information, and may vary their perception based on who they are listening to (or who they believe they are listening to) (e.g. Drager, 2010; Hay, Warren, & Drager, 2006; Jannedy & Weirich, 2014). The stimuli in this experiment were produced by a young speaker, so perhaps the older listeners identified the speaker as sounding young, and therefore as more likely to produce glottalisation. When interacting with younger speakers, older listeners would then associate glottalisation with voicelessness, though they may not use this cue in their own production or when listening to an older speaker's voice, at least not to the same extent. Based on these speculations, a natural extension of this study would be to run a similar task using an older speaker's voice for the stimuli. A further possibility is that listeners, both older and younger, are aware that there are multiple cues to coda voicing, and make use of whichever of these cues is available in the signal, despite their own production preferences.

We found that listeners reacted differently to a non-glottalised vowel than they did to a glottalised vowel with a modally voiced portion of similar duration. We interpret this as evidence that listeners do not merely perceive a shorter modally voiced vowel when presented with glottalised stimuli but appear to be sensitive to the glottalisation itself. Nevertheless, it may be possible that one of the reasons why glottalisation is employed in voiceless coda contexts is precisely because it ensures that the modally voiced portion of the vowel is kept short. Listeners may then experience it occurring in voiceless contexts and thus associate it with this context. Alternatively, perhaps listeners are sensitive to the sudden change in F0 that commonly occurs with glottalisation, which may signal that the vowel will end presently, and thereby increase the saliency of the shorter vowel duration. It would be interesting to examine whether L2 English listeners whose L1 does not employ glottalisation to cue the voicing contrast, or whose L1 does not contrast coda stop voicing, also interpret glottalisation as a cue to coda voicelessness. While these questions are beyond the scope of this study, they do offer interesting directions for possible future research.

3.5. Conclusion

In this study we examined how AusE listeners utilise acoustic cues in the perception of coda stop voicing. The results show that listeners use preceding vowel duration as a cue to the voicing of coda stops, but that the presence of glottalisation facilitates increased perception of coda voicelessness. This effect was stronger for inherently short vowels than for inherently long vowels, and for low vowels than high vowels. The vowel duration cue to coda voicing is weaker for inherently short vowels and for low vowels in production; thus, this finding may provide support for a trading relationship between glottalisation and vowel duration. Despite previous findings showing differences between younger and older speakers in the use of glottalisation in production, we found that older and younger listeners behave similarly in their perception of glottalisation. These results may provide support for theories of sound change that suggest a shift in perception occurs prior to a shift in production. Future work will explore links between individual speaker/listeners with regard to this change.

3.6 Acknowledgements

This work was supported by an Australian Government Research Training Program scholarship and an Australian Government Endeavour Research Fellowship to the first author. Preliminary analyses of this work were presented at the 16th Conference of the Association for Laboratory Phonology (LabPhon16) in Lisbon, Portugal, the 2nd Workshop on Sociophonetic Variability in the English Varieties of Australia (SocioPhonAus 2018) in Brisbane, Australia, and the 19th Annual Conference of the International Speech Communication Association (Interspeech 2018) in Hyderabad, India; we thank participants at these events for their helpful comments and suggestions. We would also like to thank Louise Ratko and Rebecca Holt for supplying their voices for the stimuli, Shuting Liu for assistance with Psyscope, Ulrich Reubold, Marc Garellek, Maria Paola Bissiri and members of the MQ phonetics lab for their suggestions and/or feedback regarding stimuli design, and lab members of IPS Munich, particularly Jonathan Harrington and Eva Reinisch, for their thoughts, feedback and comments on the interpretation of the results. We are also grateful to Taehong Cho and three anonymous reviewers for their insightful suggestions that have helped to improve this manuscript, and to Peter Humburg for statistical advice.

3.7 Appendix A

Table A1. Manipulated intensity contour values according to acoustic landmarks.

Landmark	Intensity (dB)
Onset of word	15
Onset of vowel	65
Offset of vowel	60
Coda closure midpoint	15
Coda burst	25
Offset of coda burst	15

3.8 Appendix B

Table A2. Summary of mixed effects logistic regression model for vowel duration manipulation.

	β	OR	SE	*z*	*p*
(Intercept)	-0.965	0.38	0.213	-4.541	<0.0001
Step	-0.895	0.41	0.072	-12.371	<0.0001
Condition	2.639	14.00	0.326	8.095	<0.0001
Age	-0.035	0.97	0.232	-0.152	0.880
Vowel length	0.235	1.27	0.225	1.048	0.295
Vowel height	1.300	3.67	0.228	5.708	<0.0001
Step: Condition	0.010	1.01	0.070	0.141	0.888
Step: Age	0.133	1.14	0.074	1.798	0.072
Condition: Age	-0.617	0.54	0.359	-1.721	0.085
Step: Vowel length	0.052	1.05	0.073	0.710	0.478
Condition: Vowel length	1.205	3.34	0.336	3.581	0.0003
Step: Vowel height	0.250	1.29	0.040	6.298	<0.0001
Condition: Vowel height	-0.790	0.45	0.236	-3.349	0.001
Age: Vowel height	-1.257	0.29	0.286	-4.395	<0.0001
Step: Condition: Age	0.103	1.11	0.067	1.537	0.124
Step: Condition: Vowel length	0.084	1.09	0.061	1.383	0.167
Step: Condition: Vowel height	-0.155	0.86	0.060	-2.589	0.010
Condition: Age: Vowel height	1.150	3.16	0.273	4.220	<0.0001

Table A3. Summary of mixed effects logistic regression model for closure duration manipulation.

	β	OR	SE	*z*	*p*
(Intercept)	-0.669	0.51	0.229	-2.925	0.003
Step	0.115	1.12	0.034	3.440	0.001
Condition	3.570	35.51	0.338	10.572	<0.0001
Age	-0.418	0.66	0.291	-1.435	0.151
Vowel length	-0.448	0.64	0.306	-1.464	0.143
Vowel height	1.247	3.48	0.204	6.124	<0.0001
Step: Condition	-0.030	0.97	0.056	-0.547	0.585
Step: Age	-0.040	0.96	0.039	-1.025	0.305
Condition: Age	-0.776	0.46	0.373	-2.083	0.037
Step: Vowel length	-0.003	1.00	0.041	-0.070	0.944
Condition: Vowel length	0.943	2.57	0.330	2.854	0.004
Step: Vowel height	-0.023	0.98	0.028	-0.816	0.414
Condition: Vowel height	-1.713	0.18	0.238	-7.199	<0.0001
Age: Vowel height	-1.133	0.32	0.261	-4.341	<0.0001
Age: Vowel length	0.626	1.87	0.367	1.709	0.087
Step: Condition: Age	-0.011	0.99	0.052	-0.213	0.832
Step: Condition: Vowel length	0.009	1.01	0.050	0.174	0.862
Step: Condition: Vowel height	0.038	1.04	0.049	0.763	0.446
Condition: Age: Vowel height	1.634	5.12	0.284	5.758	<0.0001
Step: Age: Vowel length	0.001	1.00	0.050	0.013	0.989

Table A4. Summary of mixed effects logistic regression model for C/V ratio manipulation.

	β	OR	SE	*z*	*p*
(Intercept)	-0.130	0.88	0.259	-0.501	0.617
Step	-0.679	0.51	0.056	-12.115	<0.0001
Condition	2.747	15.60	0.354	7.754	<0.0001
Age	-0.557	0.57	0.330	-1.686	0.0919
Vowel length	-2.670	0.07	0.352	-7.594	<0.0001
Vowel height	1.616	5.03	0.224	7.210	<0.0001
Step: Condition	0.101	1.11	0.066	1.524	0.127
Step: Age	0.106	1.11	0.057	1.865	0.062
Condition: Age	-0.420	0.66	0.437	-0.961	0.336
Step: Vowel length	0.068	1.07	0.055	1.238	0.216
Condition: Vowel length	1.015	2.76	0.456	2.227	0.026
Step: Vowel height	0.124	1.13	0.038	3.300	0.001
Condition: Vowel height	-0.751	0.47	0.231	-3.250	0.001
Age: Vowel height	-1.123	0.33	0.281	-3.993	<0.0001
Age: Vowel length	1.100	3.00	0.420	2.621	0.009
Step: Condition: Age	-0.007	0.99	0.061	-0.114	0.909
Step: Condition: Vowel length	-0.058	0.94	0.056	-1.027	0.304
Step: Condition: Vowel height	-0.052	0.95	0.055	-0.946	0.344
Condition: Age: Vowel height	0.766	2.15	0.270	2.837	0.005
Condition: Age: Vowel length	0.180	1.20	0.547	0.329	0.742

3.9 Appendix C

Table A5. Modally voiced vowel duration of glottalised steps and approximately corresponding non-glottalised steps (rounded to nearest millisecond).

Vowel	Glottalised step	Duration of modally voiced portion of vowel (ms)	Approximately corresponding non-glottalised step	Duration of vowel (ms)
/iː/	2	100	1	104
/iː/	3	122	2	134
/iː/	5	166	3	163
/iː/	6	189	4	193
/iː/	8	233	5	222
/iː/	9	255	6	251
/ɪ/	4	67	1	65
/ɪ/	5	75	2	78
/ɪ/	7	92	3	90
/ɪ/	8	100	4	103
/ɐː/	3	162	1	165
/ɐː/	4	181	2	190
/ɐː/	6	220	3	216
/ɐː/	7	239	4	241
/ɐː/	8	258	5	267
/ɐ/	3	64	1	64
/ɐ/	5	86	2	81
/ɐ/	6	97	3	98
/ɐ/	8	120	4	115
/ɐ/	9	131	5	133

Table A6. Summary of mixed effects logistic regression model to compare modally voiced portions of vowels in non-glottalised and glottalised conditions.

	β	OR	SE	z	p
(Intercept)	3.656	38.70	0.235	15.569	<0.0001
Step (recoded)	-0.820	0.44	0.051	-16.073	<0.0001
Condition	0.580	1.79	0.217	2.671	0.008
Age	-0.833	0.44	0.207	-4.023	<0.0001
Condition: Age	0.505	1.66	0.284	1.780	0.075

3.10 References

Abramson, A. (1977). Laryngeal timing in consonant distinctions. *Phonetica, 34,* 295–303.

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

Beddor, P. S. (2015). The relation between language users' perception and production repertoires. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow, Scotland.*

Boersma, P., & Weenink, D. (2018). *Praat: Doing phonetics by computer* (Version 6.0.37). http://www.praat.org/.

Brunner, J., & Żygis, M. (2011). Why do glottal stops and low vowels like each other? In W. Sum Lee & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS XVII), Hong Kong* (pp. 376–379).

Burnham, D., Estival, D., Fazio, S., Viethen, J., Cox, F., Dale, R., Cassidy, S., Epps, J., Togneri, R., Wagner, M., Kinoshita, Y., Göcke, R., Arciuli, J., Onslow, M., Lewis, T., Butcher, A., & Hajek, J. (2011). Building an audio-visual corpus of Australian English: Large corpus collection with an economical portable and replicable black box. In P. Cosi, R. De Mori, G. Di Fabbrizio & R. Pieraccini (Eds.), *Proceedings of 12th Annual Conference of the International Speech Communication Association (INTERSPEECH 2011), Florence, Italy* (pp. 841–844).

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

Cho, T., Whalen, D. H., & Docherty, G. J. (2019). Voice onset time and beyond: Exploring laryngeal contrast in 19 languages. *Journal of Phonetics, 72*, 52–65.

Chong, A., & Garellek, M. (2018). Online perception of glottalized coda stops in American English. *Laboratory Phonology*, *9,* 1–24.

Coetzee, A. W., Beddor, P. S., Shedden, K., Styler, W., & Wissing, D. (2018). Plosive voicing in Afrikaans: Differential cue weighting and tonogenesis. *Journal of Phonetics, 66*, 185–216.

Cox, F. (2006). The acoustic characteristics of /hVd/ vowels in the speech of some Australian teenagers. *Australian Journal of Linguistics, 26*, 147–179.

Cox, F., & Palethorpe, S. (2007). Illustrations of the IPA: Australian English. *Journal of the International Phonetic Association, 37*, 341–350.

Cox, F., & Palethorpe, S. (2011). Timing differences in the VC rhyme of standard Australian English and Lebanese Australian English. In W. Sum Lee & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS XVII), Hong Kong* (pp. 528–531).

Crowhurst, M. J. (2018). The influence of varying vowel phonation and duration on rhythmic grouping biases among Spanish and English speakers. *Journal of Phonetics, 66*, 82–99.

Davidson, L. (2016). Variability in the implementation of voicing in American English obstruents. *Journal of Phonetics, 54,* 35–50.

Denes, P. (1955). Effect of Duration on the perception of voicing. *Journal of the Acoustical Society of America, 27*, 761–764.

Docherty, G. J. (1992). *The timing of voicing in British English obstruents*. Berlin: Foris.

Docherty, G. J., & Foulkes, P. (1999). Derby and Newcastle: Instrumental phonetics and variationist studies. In P. Foulkes & G. J. Docherty (Eds.), *Urban voices: Accent studies in the British Isles* (pp. 47–71). London: Arnold.

Drager, K. (2010). Sociophonetic Variation in speech perception. *Language and Linguistics Compass, 4*, 473–480.

Fitch, H. L. (1981). Distinguishing temporal information for speaking rate from temporal information for intervocalic stop consonant voicing. *Haskins Laboratories Status Report of Speech Research, SR-65*, 1–32.

Flege, J. E., & Brown, W. S. (1982). The voicing contrast between /p/ and /b/ as a function of stress and position in utterance. *Journal of Phonetics, 10*, 334–345.

Foulkes, P., & Docherty, G. J. (2006). The social life of phonetics and phonology. *Journal of Phonetics, 34,* 409–438.

Fowler, C. A. (1992). Vowel duration and closure duration in voiced and unvoiced stops: There are no contrast effects here. *Journal of Phonetics, 20,* 143–165.

Garellek, M. (2011). The benefits of vowel laryngealization on the perception of coda stops in English. *UCLA Working Papers in Phonetics, 109*, 31–39.

Garellek, M. (2015). Perception of glottalization and phrase-final creak. *Journal of the Acoustical Society of America, 137*, 822–831.

Gordeeva, O. B., & Scobbie, J. M. (2013). A phonetically versatile contrast: Pulmonic and glottalic voicelessness in Scottish English obstruents and voice quality. *Journal of the International Phonetic Association, 43,* 249–271.

Green, P., & MacLeod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, 7, 493–498.

Gruenenfelder, T. M., & Pisoni, D. B. (1980). Fundamental frequency as a cue to postvocalic consonantal voicing: Some data from speech perception and production. *Perception & Psychophysics, 28,* 514–520.

Hanson, H. 2009. Effects of obstruent consonants on fundamental frequency at vowel onset in English. *Journal of the Acoustical Society of America, 125*, 425–441.

Harrington, J., Cox, F., & Evans, Z. (1997). An acoustic phonetic study of broad, general, and cultivated Australian English vowels. *Australian Journal of Linguistics, 17,* 155–184.

Harrington, J., Kleber, F., & Reubold, U. (2008). Compensation for coarticulation, /u/-fronting, and sound change in standard southern British: An acoustic and perceptual study. *Journal of the Acoustical Society of America, 123,* 2825–2835.

Haslerud, V. C. D. (1995). *The variable (t) in Sydney adolescent speech: A sociolinguistic study of phonological variation* (Cand. Philol dissertation, University of Bergen).

Hay, J., Warren, P., & Drager, K. (2006). Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics, 34*, 458–484.

Hejná, M., & Scanlon, J. (2015). New laryngeal allophony in Manchester English. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow, Scotland*.

Higginbottom, E. (1964). Glottal reinforcement in English. *Transactions of the Philological Society, 63*, 129–142.

Hogan, J. T., & Rozsypal, A. J. (1980). Evaluation of vowel duration as a cue for the voicing distinction in the following word-final consonant. *Journal of the Acoustical Society of America, 67,* 1764–1771.

House, A. S., & Fairbanks, G. (1953). The influence of consonant environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America, 25,* 105.

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

Ingram, J. C. L. (1989). Connected speech processes in Australian English. *Australian Journal of Linguistics, 9*, 21–49.

Jannedy, S., & Weirich, M. (2014). Sound change in an urban setting: Category instability of the palatal fricative in Berlin. *Laboratory Phonology, 5*, 91–122.

Keyser, S. J., & Stevens, K. N. (2006). Enhancement and overlap in the speech chain. *Language, 82*, 33–62.

Kingston, J., & Diehl, R. L. (1994). Phonetic Knowledge. *Language, 70,* 419–454.

Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, *59*, 1208–1221.

Kohler, K. J. (1982). F0 in the production of lenis and fortis plosives. *Phonetica, 39*, 199–218.

Kleber, F., Harrington, J., & Reubold, U. (2012). The relationship between the perception and production of coarticulation during a sound change in progress. *Language and Speech, 55,* 383–405.

Kuang, J., & Cui, A. (2018). Relative cue weighting in production and perception of an ongoing sound change in Southern Yi. *Journal of Phonetics, 71*, 194–214.

Lenth, R. V. (2016). Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software, 69*, 1–33.

Liberman, A.M., Delattre, P.C., & Cooper, F. S. (1958). Some cues for the distinction between voiced and voiceless stops in initial position. *Language and Speech, 1*, 153–167.

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

Lisker, L. (1974) On explaining vowel duration variation. *Glossa, 8*, 233–246.

Lisker, L. (1975). Is it VOT or a first-formant transition detector? *Journal of the Acoustical Society of America, 57*, 1547–1551.

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

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word, 20*, 384–422.

Lisker, L., & Abramson, A. S. (1967). Some effects of context on voice onset time in English stops. *Language and Speech, 10*, 1–28.

Lisker, L., & Abramson, A.S. (1970). The voicing dimension: Some experiments in comparative phonetics. In B. Halá, M. Romportl, & P. Janota (Eds.), *Proceedings of the 6th International Conference of Phonetic Sciences* (pp. 563–567). Prague: Academia.

Luce, P. A., & Charles-Luce, J. (1985). Contextual effects on vowel duration, closure duration, and the consonant/vowel ratio in speech production. *Journal of the Acoustical Society of America, 78*, 1949–1957.

Malécot, A. (1970). The lenis–fortis opposition: Its physiological parameters. *Journal of the Acoustical Society of America, 47*, 1588–1592.

Malisz, Z., Żygis, M., & Pompino-Marschall, B. (2013). Rhythmic structure effects on glottalisation: A study of different speech styles in Polish and German. *Laboratory Phonology, 4*, 119–158.

Mathisen, A. G. (1999). Sandwell, West Midlands: Ambiguous perspectives on gender patterns and models of change. In P. Foulkes & G. J. Docherty (Eds.), *Urban voices: Accent studies in the British Isles* (pp.107–123). London: Arnold.

Ohala, J. J. (1981). The listener as a source of sound change. In C.S. Masek, R.A. Hendrick, & M.F. Miller (Eds.), *Papers from the parasession on language and behaviour* (pp. 178–203). Chicago: Chicago Linguistic Society.

Ohala, J. (1993). The phonetics of sound change. In C. Jones (Ed.), *Historical linguistics: Problems and perspectives* (pp. 237–278). London: Longman.

Ohde, R. N. (1984). Fundamental frequency as an acoustic correlate of stop consonant voicing. *Journal of the Acoustical Society of America, 75*, 224–230.

Penney, J., Cox, F., Miles, K., & Palethorpe, S. (2018). Glottalisation as a cue to coda consonant voicing in Australian English. *Journal of Phonetics, 66*, 161–184.

Penney, J., Cox, F., & Szakay, A. (2019). Glottalisation of word-final stops in Australian English unstressed syllables. *Journal of the International Phonetic Association*. https://doi.org/10.1017/S0025100319000045.

Penney, J., Cox, F., & Szakay, A. (in prep). *Production of coda glottalisation in Australian English*.

Peterson, G., & Lehiste, I. (1960). Duration of syllabic nuclei in English. *Journal of the Acoustical Society of America, 32*, 693–703.

Pierrehumbert, J. B. (1994). Knowledge of variation. In K. Beals, J. Denton, R. Knippen, L. Melnar, H. Suzuki, & E. Zeinfeld (Eds.), *Papers from the Parasession on Variation, 30th meeting of the Chicago Linguistic Society Chicago* (pp. 232–256). Chicago: Chicago Linguistic Society.

Pierrehumbert, J. B. (1995). Prosodic effects on glottal allophones. In O. Fujimura & M. Hirano (Eds.), *Vocal fold physiology: Voice quality control* (pp. 39–60). San Diego: Singular.

Pinget, A.-F. (2015). *The actuation of sound change* (Doctoral dissertation, Utrecht University).

Pompino-Marschall, B., & Żygis, M. (2010). Glottal marking of vowel-initial words in German. *ZAS Papers in Linguistics, 52*, 1–17.

Port, R. F. (1979). The influence of tempo on stop closure duration as a cue for voicing and place. *Journal of Phonetics, 7*, 45–56.

Port, R. F. (1981). Linguistic timing factors in combination. *Journal of the Acoustical Society of America, 69*, 262–274.

Port, R. F., & Dalby, J. (1982). Consonant/vowel ratio as a cue for voicing in English. *Perception and Psychophysics, 32*, 141–152.

R Core Team (2016). *R: A language and environment for statistical computing* (Version 3.3.1). https://www.R-project.org/.

Raphael, L. J. (1972). Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English. *Journal of the Acoustical Society of America, 51*, 1296–1303.

Raphael, L. J. (1981). Durations and contexts as cues to word-final cognate opposition in English. *Phonetica, 38*, 126–147.

Redi, L., & Shattuck-Hufnagel, S. (2001). Variation in the realization of glottalization in normal speakers. *Journal of Phonetics, 29*, 407–429.

Repp, B. H. (1982). Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception. *Psychological Bulletin, 92*, 81–110.

Roach, P. J. (1973). Glottalization of English /p/, /t/, /k/ and /ʧ/ – a reexamination. *Journal of the International Phonetic Association, 3*, 10–21.

Seyfarth, S., & Garellek, M. (2015). Coda glottalization in American English. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow, Scotland.*

Song, J. Y., Demuth, K., & Shattuck-Hufnagel, S. (2012). The development of acoustic cues to coda contrasts in young children learning American English. *Journal of the Acoustical Society of America, 131*, 3036–3050.

Tollfree, L. (1999). South East London English: Discrete versus continuous modelling of consonantal reduction. In P. Foulkes & G. J. Docherty (Eds.), *Urban voices: Accent studies in the British Isles* (pp.163–184). London: Arnold.

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

Trudgill, P. (1986). *Dialects in contact.* Oxford: Blackwell.

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

Westbury, J. R., & Keating, P.A. (1986) On the naturalness of stop consonant voicing. *Journal of Linguistics*, *22*, 145–166.

Wright, R. (2004). A review of perceptual cues and cue robustness. In B. Hayes, R. Kirchner, & D. Steriade (Eds.), *Phonetically-based phonology* (pp. 34–57). New York: Cambridge University Press.

Zlatin, M. (1974). Voicing contrast: perceptual and productive voice onset time characteristics of adults. *Journal of the Acoustical Society of America, 56,* 981–994.

Chapter 4: Perception of coda voicing: Glottalisation, vowel duration, and silence

This chapter is based on the following published proceedings paper:

Penney, J., Cox, F., & Szakay, A. (2019). Perception of coda voicing: Glottalisation, vowel duration, and silence. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS 2019), Melbourne, Australia* (pp. 1863–1867).

I certify that I was responsible for the development of the concept of this paper, in discussion with my supervisors/co-authors. I took leadership in conducting the research, and was responsible for all data collection, the construction of the stimuli, the analyses, and the writing of all parts of the paper. My co-authors provided advice to improve the research protocol, the analyses, the interpretation of the data, as well as the presentation of the written component.

4.0 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.

4.1 Introduction

Glottalisation occurs in conjunction with voiceless coda stops in many varieties of English (Docherty & Foulkes, 1999; Gordeeva & Scobbie, 2013; Huffman, 2005; Pierrehumbert, 1994; Roach, 1973; Seyfarth & Garellek, 2015), 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 (Penney, Cox, Miles, & Palethorpe, 2018). 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 (Keyser & Stevens, 2006; Pierrehumbert, 1994), as glottalisation appears to be more common when voiceless coda stops precede sonorants, an environment in which anticipatory coarticulatory voicing is likely (Huffman, 2005; Pierrehumbert, 1995). 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, Chong and Garellek (2018) 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.

Penney, Cox, and Szakay (2018a, 2018b, submitted [Chapter 3]) 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 (Cox & Palethorpe, 2011; Penney, Cox, Miles, & Palethorpe, 2018). 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 (Penney, Cox, & Szakay, 2018b, submitted [Chapter 3]). 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 (Chen, 1970, Lisker, 1986; Raphael, 1972).

As glottalisation necessarily results in non-modal phonation affecting the vowel preceding the coda stop, it may be possible that listeners in Penney et al. (2018a, 2018b, submitted [Chapter 3]) 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 (Chen, 1970; Lisker, 1986; Raphael, 1972), 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 (Hogan & Rozsypal, 1980; Lisker, 1986).

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.

4.2 Method

4.2.1 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.

4.2.2 Stimuli

The stimuli (the same as reported in Penney et al., submitted [Chapter 3]) 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 Penney, Cox, Miles, and Palethorpe (2018) 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.

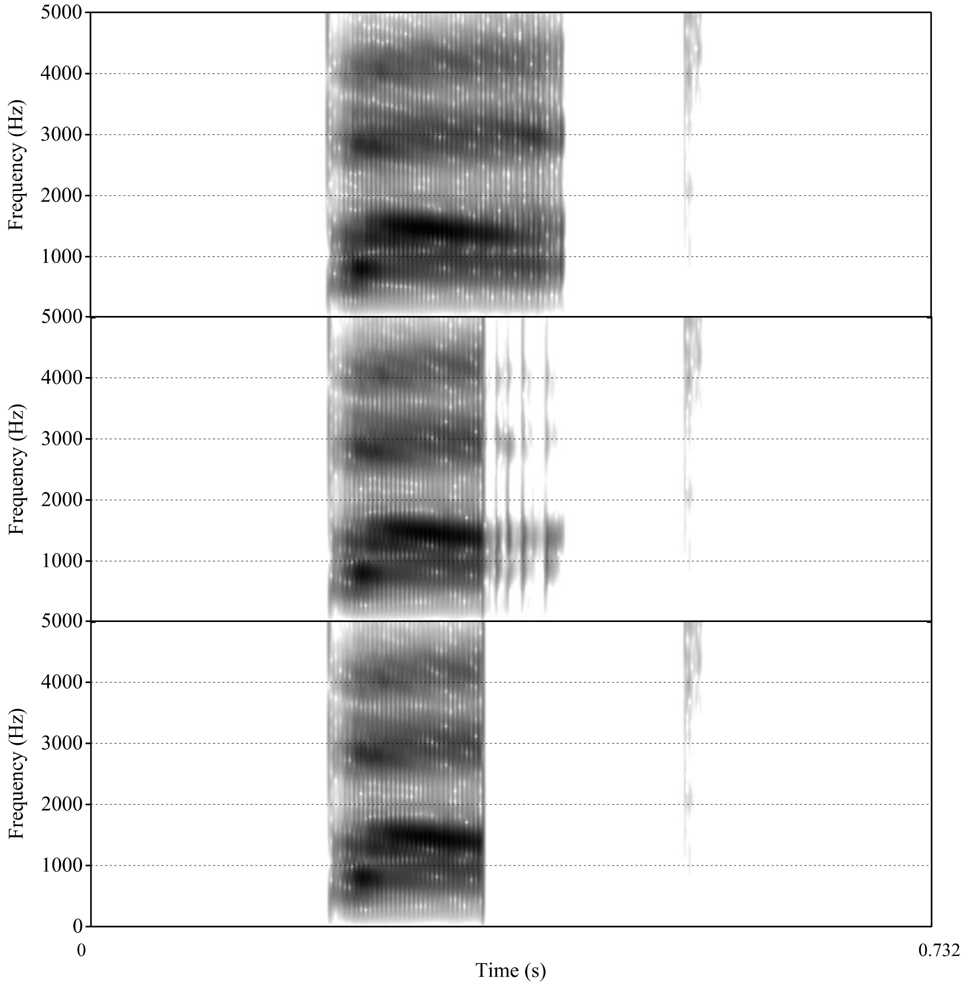

Figure 1: Spectrograms of ninth-step (i.e. longest vowel duration) 'bud' stimuli. Upper panel shows control condition; middle panel shows glottalised condition; lower panel shows silence condition.

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 Penney, Cox, Miles, & Palethorpe (2018).

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.

4.2.3 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).

4.2.4 Analysis

Using *lme4* (Bates, Maechler, Bolker, & Walker, 2015), 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.

4.3 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 (Penney et al., 2018a, 2018b, submitted [Chapter 3]). 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 (Penney et al., 2018a, 2018b, submitted [Chapter 3]). 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 ($p = 0.0113$).

4.4 Discussion

The results presented above confirm previous findings that AusE listeners use vowel duration to cue coda stop voicing (Penney et al., 2018a, 2018b, submitted [Chapter 3]). 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 (Penney et al., 2018a, 2018b, submitted [Chapter 3]), 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: Chong and Garellek (2018) 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.

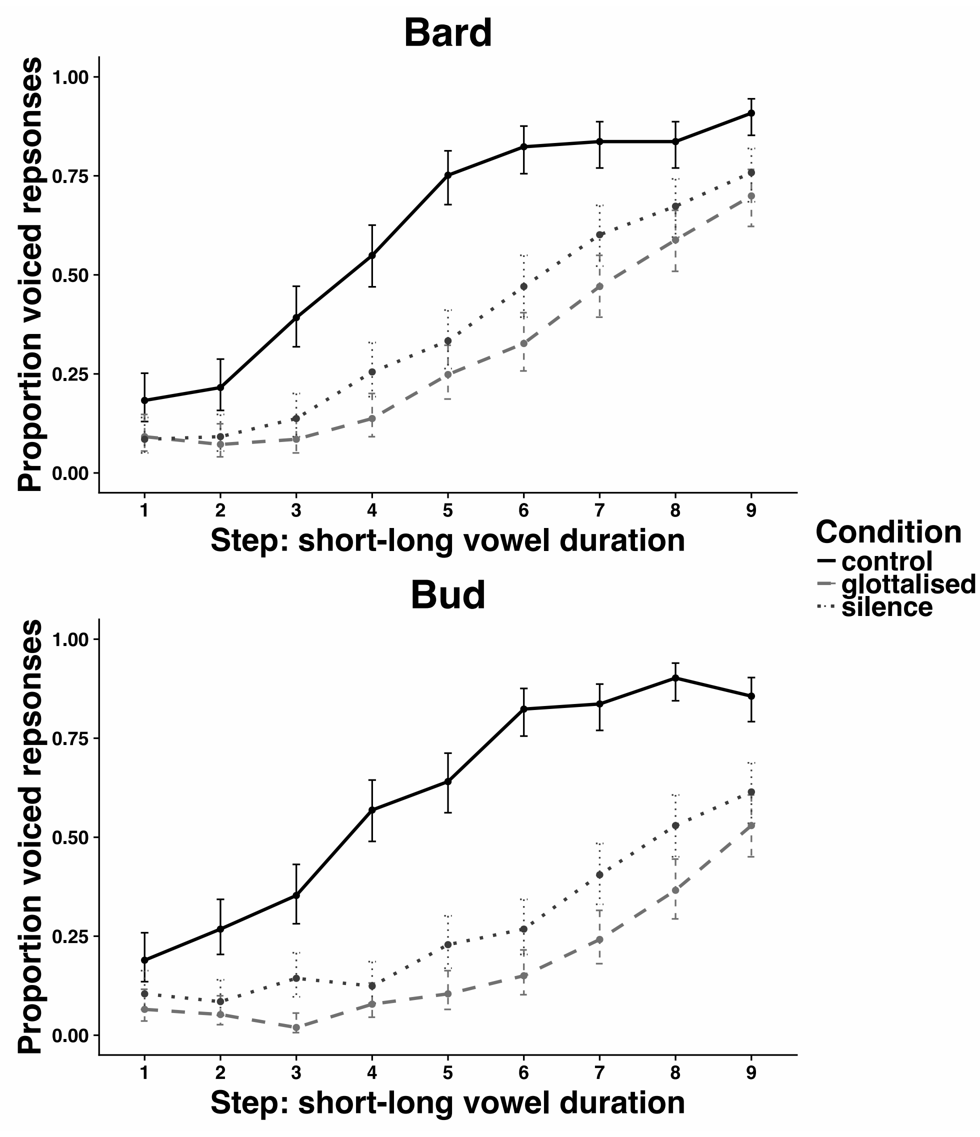

Figure 2: Proportion of voiced coda responses in control (solid lines), glottalised (dashed lines), and silence (dotted lines) conditions.

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. Penney et al. (2018b, submitted [Chapter 3]) 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 (Cox & Palethorpe, 2011;

Penney, Cox, Miles, & Palethorpe, 2018). 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.

4.5 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.

4.6 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.

4.7 References

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

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

Chong, A., & Garellek, M. (2018). Online perception of glottalized coda stops in American English. *Laboratory Phonology,* 9, 1–24.

Cox, F., & Palethorpe, S. (2011). Timing differences in the VC rhyme of standard Australian English and Lebanese Australian English. In W. Sum Lee & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS XVII), Hong Kong* (pp. 528–531).

Docherty, G. J., & Foulkes, P. (1999). Derby and Newcastle: Instrumental phonetics and variationist studies. In P. Foulkes & G. J. Docherty (Eds.), *Urban voices: Accent studies in the British Isles* (pp. 47–71). London: Arnold.

Gordeeva, O. B., & Scobbie, J. M. (2013). A phonetically versatile contrast: Pulmonic and glottalic voicelessness in Scottish English obstruents and voice quality. *Journal of the International Phonetic Association*, *43*, 249–271.

Hogan, J. T., & Rozsypal, A. J. (1980). Evaluation of vowel duration as a cue for the voicing distinction in the following word-final consonant. *Journal of the Acoustical Society of America, 67*, 1764–1771.

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

Keyser, S. J., & Stevens, K. N. (2006). Enhancement and overlap in the speech chain. *Language*, *82*, 33–62.

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

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

Penney, J., Cox, F., Miles, K., & Palethorpe, S. (2018). Glottalisation as a cue to coda consonant voicing in Australian English. *Journal of Phonetics, 66*, 161–184.

Penney, J., Cox, F., & Szakay, S. (2018a). *Weighting of vowel duration and glottalisation in the perception of coda stop voicing*. Poster presented at the 16th Conference on Laboratory Phonology (LabPhon 16), Lisbon, Portugal.

Penney, J., Cox, F., & Szakay, S. (2018b). Weighting of coda voicing cues: Glottalisation and vowel duration. In B. Yegnanarayana, C.C. Sekhar, S. Narayanan, S. Umesh, S.R.M. Prasanna, H.A. Murthy, P. Rao, P. Alku & P.K. Ghosh (Eds.), *Proceedings of the 19th Annual Conference of the International Speech Communication Association (INTERSPEECH 2018), Hyderabad, India* (pp. 1422-1426).

Penney, J., Cox, F., & Szakay, S. (submitted). *Effects of glottalisation, preceding vowel duration, and coda closure duration on the perception of coda stop voicing*.

Pierrehumbert, J. B. (1994). Knowledge of variation. In K. Beals, J. Denton, R. Knippen, L. Melnar, H. Suzuki, & E. Zeinfeld (Eds.), *Papers from the Parasession on Variation, 30th meeting of the Chicago Linguistic Society* (pp. 232–256). Chicago: Chicago Linguistic Society.

Pierrehumbert, J. B. (1995). Prosodic effects on glottal allophones. In O. Fujimura & M. Hirano (Eds.), *Vocal fold physiology: Voice quality control* (pp. 39–60). San Diego: Singular.

Raphael, L. J. (1972). Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English. *Journal of the Acoustical Society of America, 51*, 1296–1303.

Roach, P. J. (1973). Glottalization of English /p/, /t/, /k/ and /ʧ/ – a reexamination, *Journal of the International Phonetic Association, 3*, 10–21.

Seyfarth, S., & Garellek, M. (2015). Coda glottalization in American English. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow, Scotland*.

Chapter 5: Production of coda glottalisation in Australian English

This chapter is based on the following paper, which is being prepared for submission to a specialist phonetics journal (such as *Journal of Phonetics* or *Phonetica*) as:

Penney, J., Cox, F., & Szakay, A. (in preparation). *Production of coda glottalisation in Australian English*.

I certify that I was responsible for the development of the concept of this paper, in discussion with my supervisors/co-authors. I took leadership in conducting the research, and was responsible for all data collection, the preparation and annotation of the bulk of the data, the analyses, and the writing of all parts of the paper. My co-authors provided advice to improve the research protocol, the analyses, the interpretation of the data, as well as the presentation of the written component. Additional annotation of some of the data was conducted by Dr Kimiko Tsukada. Ms Linda Buckley and Mr Sakib Islam also assisted with annotation for reliability.

5.0 Abstract

Glottalisation has been demonstrated to cue coda stop voicelessness in Australian English. This is reported to be a recent change and, accordingly, younger speakers have been shown to utilise glottalisation more than older speakers. However, most previous Australian English studies have been based on single word utterances. Therefore, it is possible that the identified glottalisation may be due to the acoustically similar phrase final creaky voice instead of coda-related glottalisation. To address this issue we explore the differential effects of phrase position on the production of glottalisation. The results show that rates of glottalisation increase phrase finally, but phrase medially (where phrase final creaky voice does not occur), the results reflect previous findings that glottalisation cues coda stop voicelessness and younger speakers glottalise more frequently than older speakers. In a second analysis, we examine coda glottalisation in multiple heterosyllabic phonetic environments to determine how glottalisation is environmentally conditioned. The results show that glottalisation occurs more frequently preceding heterosyllabic consonants than vowels. Additionally, the results show that younger speakers employ glottalisation more frequently than older speakers in pre-vocalic environments, which may provide insight into the progression of change.

5.1 Introduction

This paper examines coda glottalisation, in which a glottal stop or increased glottal (and/or laryngeal) constriction is produced in conjunction with an oral coda stop (Docherty & Foulkes, 1999; Esling, Fraser, & Harris, 2005; Moisik, Esling, Crevier-Buchman, Amelot, & Halimi, 2015; Seyfarth & Garellek, 2015). Glottalisation results in irregular phonation on the vowel preceding the coda stop, producing the percept of a glottal stop and/or creakiness. In some cases, the supralaryngeal closing gesture of the oral stop may be 'lost' (or obscured); in such cases glottalisation is often referred to as glottal replacement (Docherty & Foulkes, 1999; Roach, 1973; Seyfarth & Garellek, 2015; Wells, 1982). Where the oral gesture is retained, glottalisation is sometimes referred to as glottal reinforcement or pre-glottalisation (Andrésen, 1968; Esling, Fraser, & Harris, 2005; Roach, 1973; Seyfarth & Garellek, 2015). Glottalisation occurs predominately in voiceless coda contexts in multiple varieties of English (e.g. Docherty & Foulkes, 1999; Gordeeva & Scobbie 2013; Higginbottom, 1964; Huffman, 2005; Pierrehumbert, 1994, 1995; Roach 1973, 2004) and may serve as a perceptual cue to coda voicelessness for English listeners (Chong & Garellek, 2018; Penney, Cox, & Szakay, 2018, submitted [Chapter 3]). As is the case for other varieties, glottalisation has been found to occur in Australian English (AusE) primarily in voiceless coda stop contexts, and only rarely in voiced coda stop contexts: in an analysis of 2427 items produced by an older and a younger group of speakers, Penney, Cox, Miles, and Palethorpe (2018) found glottalisation to be present in 55% of items with voiceless codas, but only in 6% of items with voiced codas. Penney, Cox, & Szakay (2019 [Chapter 2]) also found a similar pattern of glottalisation in unstressed syllables in an analysis of 754 items produced by young female AusE speakers, with glottalisation present in 83% of items with voiceless codas compared to 10% of items with voiced codas. Yet these analyses were based on single items produced in citation form. In such cases each single item represents a separate intonational phrase, and hence there is a possibility that the observed irregular phonation may be due to effects of prosody and not specifically associated with the coda stop. Creaky voice is commonly found in phrase final position and, like coda-related glottalisation, is manifested as irregular phonation (Garellek, 2015; Garellek & Keating, 2015; Henton & Bladon 1988; Kreiman, 1982; Ogden, 2001; Redi & Shattuck-Hufnagel, 2001). Phrase final creaky voice differs from coda glottalisation in that it is prosodically conditioned and that it can occur across multiple (voiced) segments within an utterance or phrase (Garellek, 2015; Garellek & Keating, 2015), whereas coda glottalisation is localised to the vowel (or in some cases sonorant consonant) preceding a coda stop. Very little work has been conducted on creaky voice in AusE, though it is reported to be present in the variety and anecdotal comments suggest that its occurrence may be increasing (Dallaston & Docherty, 2019). Although the potential for phrase final creaky voice would presumably apply equally to both voiced and voiceless coda contexts in single word productions, to disentangle coda glottalisation from phrase final creaky voice it is necessary to extend the examination of coda glottalisation to items produced in positions other than phrase final.

The presence of glottalisation is context-specific. In American English (AmE), glottalisation of voiceless stops[1] is commonly found in environments with following non-vocalic sonorants, particularly nasals and laterals (e.g. *atlas*, *gate number*,

[1] Note that although the voiceless stop is phonologically present the oral stop gesture may not be present in such cases, particularly in the case of /t/ which may undergo glottal replacement (Pierrehumbert, 1994; Seyfarth & Garellek, 2015).

lightning) (Huffman, 2005; Pierrehumbert, 1994, 1995; Seyfarth & Garellek, 2015). Glottalisation also occurs, albeit less frequently, in environments with following voiced and voiceless obstruents (Eddington & Channer, 2010; Huffman, 2005; Seyfarth & Garellek, 2015). Furthermore, glottal replacement of /t/ is reported to alternate with tapping/flapping in word final intervocalic environments in AmE (Eddington & Channer, 2010; Eddington & Taylor, 2009). The environments favouring glottalisation in British English (BrE) are complex, and can vary substantially across dialects.[2] Glottalised variants are commonly associated with London English, where reinforcement and replacement of voiceless stops frequently occurs preceding obstruents, nasals, and vowels, both word-internally (including intervocalically) and across word boundaries (Mott, 2012; Schleef, 2013; Tollfree, 1999; Wells, 1982, 1997). In Received Pronunciation (RP), glottal reinforcement (glottalisation) of voiceless stops is also common word finally in environments with following obstruent and sonorant onset consonants, though generally this does not occur before a vowel. Glottal replacement may occur both within and across word boundaries preceding a following obstruent or sonorant consonant, particularly before syllabic nasals, and to a lesser extent word finally preceding a vowel or a pause (Fabricius, 2002; Higginbottom, 1964; Roach, 1973, 2004; Trudgill, 2001; Wells, 1997).[3] In Liverpool English glottalisation is reportedly rare, but glottal replacement may occur word-internally preceding laterals or syllabic consonants (Watson, 2007). In Tyneside English glottal reinforcement can occur before following obstruents or vowels, both word-internally and across word boundaries, and glottal replacement

---

[2] Note that we do not here attempt to provide a comprehensive summary of glottalisation in all British dialects, which would require a study of its own.

[3] Fabricius (2002) suggests that glottal replacement in pre-vocalic and pre-pausal environments in RP may be due to the influence of regional variants from London and the Home Counties.

occurs frequently word-internally before laterals (Docherty & Foulkes, 1999; Docherty, Foulkes, Milroy, Milroy, & Walshaw, 2007; Watt & Allen, 2003). In Derby glottal replacement is very common (almost categorical) word-finally before consonants. Pre-vocalic glottal replacement, both word-internally and finally, is also common for younger speakers, but occurs less frequently for older speakers (Docherty & Foulkes, 1999).

Previous descriptions of glottalisation in AusE suggest some commonalities with those mentioned above for other English varieties, though it should be noted that these descriptions were largely based on impressionistic analyses. Cox and Palethorpe (2007) note that glottal reinforcement of voiceless coda stops occurs commonly, in particular in non-pre-vocalic syllable-final position. Tollfree (2001) found glottalisation to be common preceding obstruents and consonantal sonorants, both within and across word boundaries, and near-categorical before syllabic /n/, though glottalisation did not occur preceding other syllabic consonants. Glottal stops are reported to be present as allophones of /t/ preceding alveolar nasals, whether syllabic or non-syllabic (e.g. *kitten, catnip*), as well as before other (non-syllabic) consonantal sonorants in some lexical items (e.g. *battler*) and across word boundaries (e.g. *might rain*) (Cox & Fletcher, 2017; Cox & Palethorpe, 2007; Haslerud, 1995; Ingram, 1989). It is not clear whether an oral gesture is retained in these items; that is, it is not clear whether these are examples of glottal reinforcement or replacement.[4] Tollfree (2001) found that none of the auditorily-identified glottal stops in her data (i.e. stops that produced the percept of a glottal stop) appeared to be a canonical glottal stop when analysed acoustically, but rather involved "a glottal gesture in addition to an

[4] Ingram (1989) and Haslerud (1995) both consider glottalisation in their data to be examples of glottal replacement, though this cannot be confirmed based on impressionistic analysis.

alveolar closure gesture for plosive [t]" (p. 51). On the other hand, Tollfree (2001) reports that the most common realisation of glottalised /t/ in her data exhibited a lack of formant transitions into the stop, which would suggest that no oral gesture was present. Unlike in some varieties of BrE and AmE, intervocalic glottalisation (either reinforcement or replacement) is considered uncommon or even unusual in AusE, particularly word-medially (e.g. *butter*) (Cox & Palethorpe, 2007; Tollfree, 2001), although Haslerud (1995), Ingram (1989), and Tollfree (2001) all report occasional glottalisation in pre-vocalic environments, particularly across word boundaries (e.g. *get out*).[5] Pre-pausal glottalisation is also reported: Ingram (1989) found occasional glottalisation in this environment, whereas Haslerud (1995) found glottalisation to occur less frequently pre-pausally than before obstruents or nasals, yet Tollfree (2001) found glottalisation to occur frequently pre-pausally, especially in informal speech. Note that pre-pausal occurrences may represent examples of phrase final creaky voice.

A further point of interest is that glottalisation appears to be a relatively recent change to AusE. Descriptions from the 1980s remarked on the lack of glottalisation in the variety (Trudgill, 1986; Wells, 1982); however, shortly thereafter its presence has been increasingly noted (Cox & Palethorpe, 2007; Haslerud, 1995; Ingram, 1989; Tollfree, 2001). In addition, recent studies have shown that both older and younger speakers produce glottalisation primarily in voiceless contexts but younger speakers use it significantly more often (Penney Cox, Miles, & Palethorpe, 2018; Penney et al., 2019 [Chapter 2]), supporting the suggestion of a change.

[5] Note though that glottal stop insertion is employed in vowel hiatus environments in AusE (see Yuen, Cox, & Demuth, 2018).

Penney, Cox, Miles, and Palethorpe (2018) suggest that younger speakers may utilise glottalisation to signal coda stop voicelessness more than older speakers as they make less use of other cues to coda stop voicing. Generally, vowels preceding voiced coda stops (e.g. *bud*) are longer than vowels preceding voiceless coda stops (e.g. *but*) and this is a strong voicing cue, particularly as stops in English may be variably devoiced (Davidson, 2016; Docherty, 1992; Westbury & Keating, 1986). Vowel duration is a heavily weighted cue in voicing perception (Fowler, 1992; Klatt, 1976; Luce & Charles-Luce, 1985; Raphael, 1972). However, Penney, Cox, Miles, and Palethorpe (2018) found that younger speakers exhibited less difference in vowel duration between coda voicing contexts than older speakers while at the same time they used glottalisation more than older speakers. Similarly, both older and younger speakers in their study used vowel duration most robustly with high vowels and exhibited less glottalisation in the high vowel context. These findings may point to a trading relationship between decreased use of vowel duration and increased use of glottalisation, where glottalisation may signal the contrast between voiced and voiceless coda stops in the absence of strong vowel duration cues. There is also evidence that AusE listeners' perception of coda voicelessness is strengthened when glottalisation is present, even when this occurs in conjunction with a vowel duration that would otherwise cue a voiced coda (Penney, Cox, & Szakay, 2018, submitted [Chapter 3]). Therefore, listeners may weight glottalisation more heavily than vowel duration in some cases. However, these results were based on studies of isolated words, raising questions about the impact of prosodic context.

In this paper we will therefore examine preceding vowel duration in voiced and voiceless coda contexts in both phrase final and phrase medial positions to explore these two prosodic contexts in comparison to our previous findings for single words (Penney, Cox, Miles, & Palethorpe, 2018). As F0 differences at vowel offset may also signal coda voicing, with relatively higher F0 before voiceless stops and relatively lower F0 before voiced stops, we will also include an analysis of F0.[6] Penney et al. (2019 [Chapter 2]) found that F0 differences were reduced in their analysis of unstressed syllables, which may support the suggestion that the increasing use of glottalisation as a cue to coda voicing may reduce the utility of other ancillary voicing cues.

In this paper we present two separate analyses. In section 5.2 below we present an examination of glottalisation in voiced and voiceless coda stops contexts in both phrase medial and phrase final position for speakers in two age groups: older and younger. In section 5.3 we present an analysis of glottalisation in voiceless coda stop contexts with multiple following heterosyllabic phonetic environments for both age groups.

The aims of this paper are:

- to examine whether reported patterns of glottalisation (i.e. voicing context and age-based differences) are found in both phrase medial and phrase final position, to tease apart the possible effects of phrase final creaky voice;
- to explore whether there are differences between older and younger speakers in their relative use of glottalisation, preceding vowel duration, and F0 as cues

[6] There are of course a number of other cues to coda stop voicing in addition to vowel duration, glottalisation, and F0, such as the presence of voicing during the closure, the amplitude, intensity and frequency of the release burst, the duration of the stop closure, F1 transitions, etc. (e.g. Lisker, 1986). We focus here on vowel duration, which has been suggested to be a primary cue to voicing in coda position (Klatt, 1976), glottalisation, which may operate in a trading relationship with vowel duration (Penney, Cox, Miles, & Palethorpe, 2018), and F0, as F0 differences across voicing contexts may be reduced in AusE (Penney et al., 2019 [Chapter 2]).

to coda stop voicing, amid claims that glottalisation may be increasingly employed to signal the voicing contrast;

- to assess the extent to which the occurrence of glottalisation is conditioned by the following phonetic environment. This may provide insights into the progression of the change in this variety.

In line with these aims, we predict that:

- glottalisation will occur primarily in voiceless contexts (as has been shown previously in single items utterances) in both phrase medial and phrase final position. Note however, that it is also possible that glottalisation found in voiceless contexts in previous work was due to phrase final creaky voice, thus we may expect to find a lesser degree of glottalisation as a voicing cue in medial position;
- older speakers will utilise glottalisation less than younger speakers, consistent with previous findings, and we expect this to be the case in all environments;
- voicing-related vowel duration and F0 differences may be reduced in younger speakers, following previous findings of the increasing importance of glottalisation as a voicing cue;
- we may expect higher levels of glottalisation preceding nasals than other following environments, given previous descriptions of increased coda stop glottalisation in pre-nasal contexts.

There are multiple possible non-canonical realisations of AusE coda stops, which may alternate or occur in conjunction with glottalisation (Penney et al., 2019 [Chapter 2]). For example, unreleased stops, in which the occlusion phase is produced but no release burst is present, often occur in final position or preceding another stop and are commonly associated with glottalisation (Blevins, 2006; Ingram, 1989; Kahn 1976; Selkirk, 1982); spirantised (fricated) stops are produced with a weak or incomplete occlusion causing frication (Jones & McDougall, 2009; Tollfree, 2001); stops may exhibit preaspiration, whereby the offset of the vowel contains aspiration prior to the stop closure (Hejná, 2015; Helgason, 2002, Nance & Stuart-Smith, 2013); stops may be realised as voiced taps in intervocalic environments preceding an unstressed vowel (Cox & Palethorpe, 2007; Derrick & Gick, 2011; Herd, Jongman, & Sereno, 2010; Tollfree, 2001). In addition, glottal squeaks, in which low amplitude (relatively) high frequency periodicity is visible in the stop closure may occur in conjunction with glottalisation for some speakers, possibly as a "mechanical consequence of glottalisation" (Hejná, Palo & Moisik, 2016: 1139; see also Penney et al., 2019 [Chapter 2]; Redi & Shattuck-Hufnagel, 2001). This analysis provides an opportunity to further illuminate the non-canonical stop realisations in the data in order to advance our understanding of variation in stop consonants in general.

5.2 Glottalisation in phrase final and phrase medial positions

5.2.1 Methods

5.2.1.1 Participants

We collected production data from 77 speakers allocated to two age groups: an older group aged 65 and above (n = 33 speakers; 23 female; 10 male), and a younger group aged between 18 and 36 years (n = 44; 37 female; 7 male). We originally enlisted 83 participants; three of these did not fall into the relevant age group categories; data for a further three (all young) were discarded as they were variably rhotic (AusE is a non-rhotic dialect). All participants were native speakers of AusE, were born in Australia (apart from one participant who migrated to Australia as an infant), and completed all of their primary and secondary schooling exclusively in Australia. All received either course credit or were paid for their participation.

5.2.1.2 Data collection

The majority of the data were recorded in a sound attenuated room in the Department of Linguistics at Macquarie University. For a subset of the participants (*n* = 7), the recordings took place in a quiet office at a local church, and a further 4 participants were recorded in a quiet room of their home. Regardless of recording venue, all participants were recorded to a Marantz PMD661 MK II solid-state recorder with an AKG C520 headset condenser microphone with a sampling rate of 44.1 kHz and 16 bit quantisation.

Participants were seated comfortably in front of a notebook computer from which they read 396 randomly presented sentences. The sentences were presented in two blocks and participants' progression through the task was self paced. The sentences contained a set of 12 target words all of the forms /hVt/ and /bVC/, in which V represents a vowel and C represents either a /t/ or /d/. The target words were embedded in carriers of the form: *SAY* <*TARGET*> <*FOLLOWING*> *ONE MORE TIME*, where the word directly following the target was one of the words: *again*, *only*, *town*, *down*, *now*, *sack*, *zack*, *long*, *once*, *right*. The <*FOLLOWING*> words were chosen to enable an analysis of following phonetic environments, including stressed and unstressed vowels, voiced and voiceless obstruents, and approximants. For each target word an additional carrier was also included to elicit the target words in phrase final position: *NOW ONE MORE TIME SAY* <*TARGET*>. To control for prosody, participants were instructed to accent the word directly following the target in the phrase medial context, or the initial *now* in the phrase final carrier, which was indicated by upper case letters in the orthography (e.g. *say bead NOW one more time*). The following two carrier environments will be examined in this analysis: *SAY <TARGET> ONLY ONE MORE TIME* AND *NOW ONE MORE TIME SAY <TARGET>*. The first carrier was selected to produce the target words in phrase medial position preceding an accented vowel initial syllable, which we anticipated would be the most likely environment to elicit released stops. The second carrier was selected to ensure the target word was produced in phrase final position.

Eight target words with the form /bVC/ were selected for examination here: *bead*, *beat*, *bard*, *bart*, *bid*, *bit*, *bud, but*. The selected target words were chosen to enable an examination of high/low, short/long vowel pairs (/ɐː, ɐ, iː, ɪ/), and the voiced/voiceless stop pair to enable an examination of coda stop voicing on the presence of glottalisation. Participants produced 3 repetitions of each sentence. This provided 3696 items for analysis (3 repetitions X 8 target words X 2 phrase contexts X 77 participants = 3696). Speakers were given instructions on how to produce the phrases but were not corrected if they made mistakes during the task. 124 items were discarded due to mispronunciations or other anomalies (e.g. yawning, laughter). A further 283 items in the phrase medial context were discarded due to speakers inserting a phrase break between the target word and the following vowel initial word or for accenting the target word. Items in which phrase breaks were inserted or the incorrect word was accented were identified by the first author. A second trained annotator then examined 10 per cent of the phrase medial position data and labelled these for the presence of phrase breaks/incorrect accents. We calculated a Cohen’s Kappa score for inter-rater reliability with the *irr* package (Wolak, Fairbairn, & Paulsen, 2012), which showed high agreement between the annotators (Kappa = 0.81;

$z = 11.6$; $p < 0.0001$). Finally, 38 items (all in the words *bud* and *but)* were excluded as they were produced with a schwa, rather than the short, low vowel [ɐ]. These segments were labelled as schwa by the automatic aligner (described below) and were subsequently checked auditorily. After exclusions there were 3251 items remaining for analysis (older: 1453; younger: 1798).

5.2.1.3 Acoustic analysis

The data were first labelled and segmented into phonemes by WebMAUS (Kisler, Reichel, & Schiel, 2017) utilising an AusE model which returned Praat textgrids (Boersma & Weenik, 2018). The segmentations were hand checked and corrected by a trained annotator with reference to wideband spectrograms and aligned waveforms. In addition, the target words were hand labelled for the offset of the vowel/onset of coda stop closure, the coda stop release burst, non-canonical stop variants (where present), the presence of glottalisation, and the presence of creaky voice according to the criteria specified below.

The presence of glottalisation was visually determined by reference to the waveform and wideband spectrogram. Glottalisation is characterised by irregularity in F0 and amplitude (Batliner et al., 1994; Dilley, Shattuck-Hufnagel, & Ostendorf, 1996; Garellek, 2015; Huffman, 2005), thus items were labelled as glottalised when irregularity in the peaks of the waveform (representing irregular amplitude) and/or irregularity in the spacing of striations in the spectrogram and waveform (representing irregularity of F0) were visible in the second half of the vowel. As in our previous studies (Penney, Cox, Miles, & Palethorpe, 2018; Penney et al., 2019 [Chapter 2]), items in which evidence of glottalisation extended more than halfway through the preceding vowel segment were labelled as creaky rather than glottalised. Creaky voice can extend across multiple voiced segments within a word or an utterance, whereas glottalisation associated with coda stops is localised to a single segment (Garellek, 2015). It may be possible that some longer examples of glottalisation associated with a coda stop were categorised as creaky under this approach. This issue only potentially applies to a small number of items, as most of the items for which there was evidence of glottalisation/creaky voice had irregular F0 and amplitude that were either clearly limited to the second half of the vowel (and would therefore be categorised as coda glottalisation) or extended throughout the entire vowel (and would therefore be categorised as creaky). Of course, this method cannot ensure that items labelled as glottalised are not simply produced with very short portions of creaky voice, though we note that, if there are indeed items with short portions of creaky voice that have been labelled as glottalised, this should apply equally to both voiced and voiceless coda contexts.

Although it was not our intention in this study to examine in detail the duration of glottalisation or the proportion of the vowel that was glottalised, we did observe substantial variation in this regard, both within and across speakers. Figures 1 and 2 provide examples of glottalised items that exemplify this variation. In Figure 1 glottalisation appears to be brief; there is only a single irregular pulse present at the offset of the vowel. In Figure 2, on the other hand, glottalisation appears to make up a larger proportion of the vowel, with several irregular pulses visible. This demonstrates the fact that individual speakers' production of glottalisation may be variable, and that a range of realisations may be present in the community. This point will be explored in section 5.3.4 below where we explore whether the glottalisation in our data constitutes glottal reinforcement or rather glottal replacement.

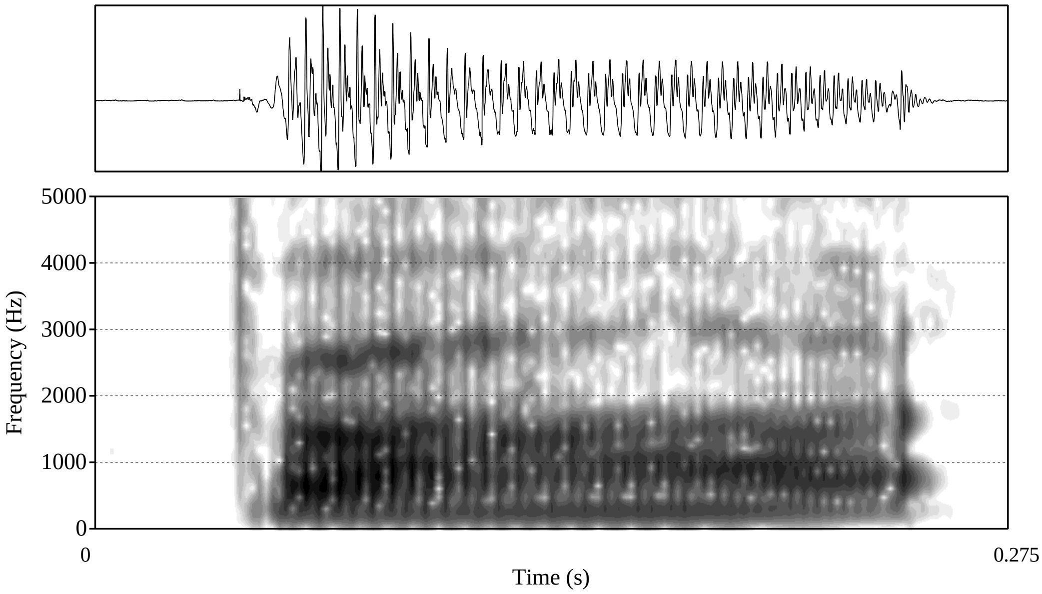

Figure 1. Spectrogram and waveform of the word *bart* with a single irregular pulse.

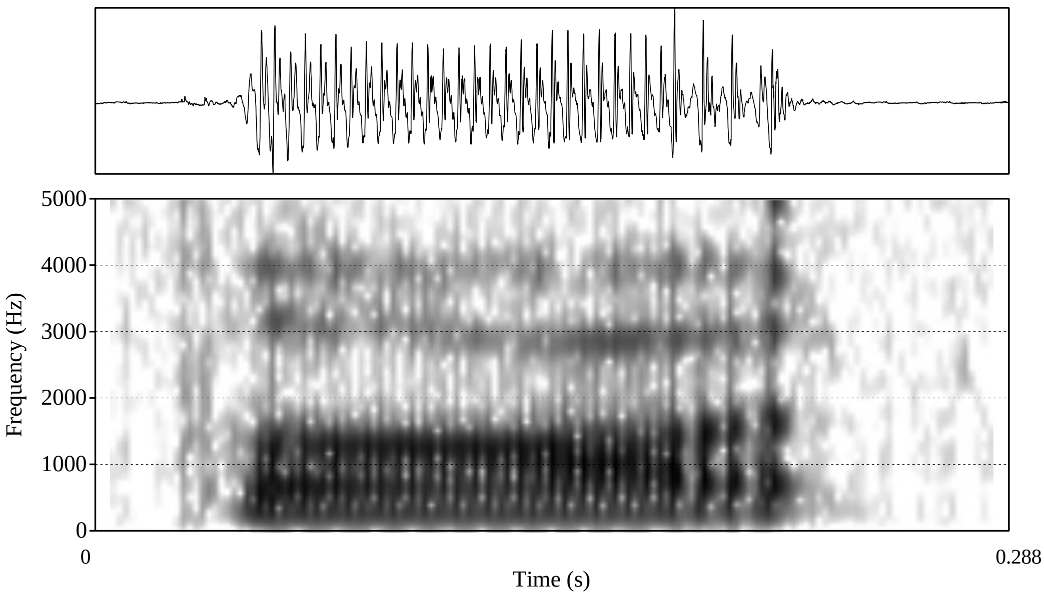

Figure 2. Spectrogram and waveform of the word *bart* with multiple irregular pulses.

In addition to visual classification of glottalisation, we also extracted H1*-H2* measures across the vowel for each item. H1*-H2* is a measure of difference between the amplitude of the first (H1) and second (H2) harmonics. The asterisks signal that the amplitudes of the harmonics have been corrected for effects of formant frequencies. H1*-H2* is correlated with glottal constriction with lower H1*-H2* values indicative of increased glottal constriction (Garellek & Seyfarth, 2016; Holmberg, Hillman, Perkell, Guiod, Goldman, 1995; Keating, Garellek, & Kreiman, 2015; Seyfarth & Garellek, 2015). Therefore, items labelled as glottalised should exhibit lower H1*-H2* values compared to modally voiced items. Glottalisation and phrase final creaky voice both manifest as 'prototypical creaky voice' (Garellek & Seyfarth, 2016), a type of creaky voice that is characterised by glottal constriction as well as low and irregular F0 (Keating, Garellek, & Kreiman, 2015). Consequently, items labelled as creaky should also exhibit reduced H1*-H2* values compared to modally voiced items. However, differences in the timing of the two phenomena can be observed, with glottalisation demonstrating a more sudden change towards the end of the vowel (Garellek & Seyfarth, 2016), consistent with a constriction gesture that is localised to the coda stop.

H1*-H2* measurements were extracted from each of the vowels using VoiceSauce (Shue, Keating, Vicenik, & Yu, 2011). Note that reliable F0 estimation is required for the accurate estimation of harmonics and measures derived from harmonic estimation such as H1*-H2* (Seyfarth & Garellek, 2018); see below for a description of steps taken to improve F0 estimation. Harmonic amplitudes corrected for the effect of formant frequencies (Iseli, Shue, & Alwan, 2007) were automatically calculated to enable a comparison of different vowels, using formant frequencies estimated with the Snack Sound Toolkit (Sjölander, 2004). H1*-H2* measurements were smoothed with the default moving average filter of 20 milliseconds. To ensure accurate estimation of formant frequencies, we plotted and visually inspected for each speaker all formant measurements for each vowel (Seyfarth & Garellek, 2018). H1*-

H2* measurements were excluded for 38 items which showed F1 or F2 values that were obvious outliers for that speaker and that vowel. We then calculated average H1*-H2* values over each of three equal subsections of each vowel, enabling an examination of change in H1*-H2* throughout the vowel. Figure 3 illustrates the mean H1*-H2* values for items classified by visual inspection as glottalised and non-glottalised in phrase medial and phrase final positions. Note that the non-glottalised set contains items produced with phrase final creaky voice. It can be seen that in phrase medial position glottalised tokens show a drop in H1*-H2*, particularly from the second to the third subsection of the vowel, as expected. By contrast, the non-glottalised tokens show very little change in the second half of the vowel. In phrase final position the drop in H1*-H2* in the final subsection can also be seen in the glottalised items; in the non-glottalised items there is again not a great change; however, in the non-glottalised items in this position H1*-H2* is low throughout the entire vowel. This is likely due to the increased presence of creaky voice in phrase final position (see 5.2.2.2 below).

In order to support our visual determination of coda glottalisation versus creaky voice we fitted a linear mixed effects model with the dependent variable the difference in H1*-H2* between the second and final subsections of the vowel. Fixed effects included were *Glottalisation, Phrase position* and *Age group*. We also looked for all two- and three-way interactions between these factors. Random intercepts were included for *Participant* and *Word* and random slopes for all within-participant factors.[7] The results of the model are shown in Table 1. We found a significant effect

[7] The syntax for this model was:

lmer (H1*-H2* difference ~ (Glottalisation + Phrase position+ Age group)^3 + (1+ Glottalisation + Phrase position|Participant) + (1|Word) for *Glottalisation*, confirming that the lowering of H1*-H2* between the second and final subsections of the vowel was significantly greater for the glottalised items, consistent with a glottal constriction gesture timed such that it occurs just prior to the coda stop. No other significant effects were found.

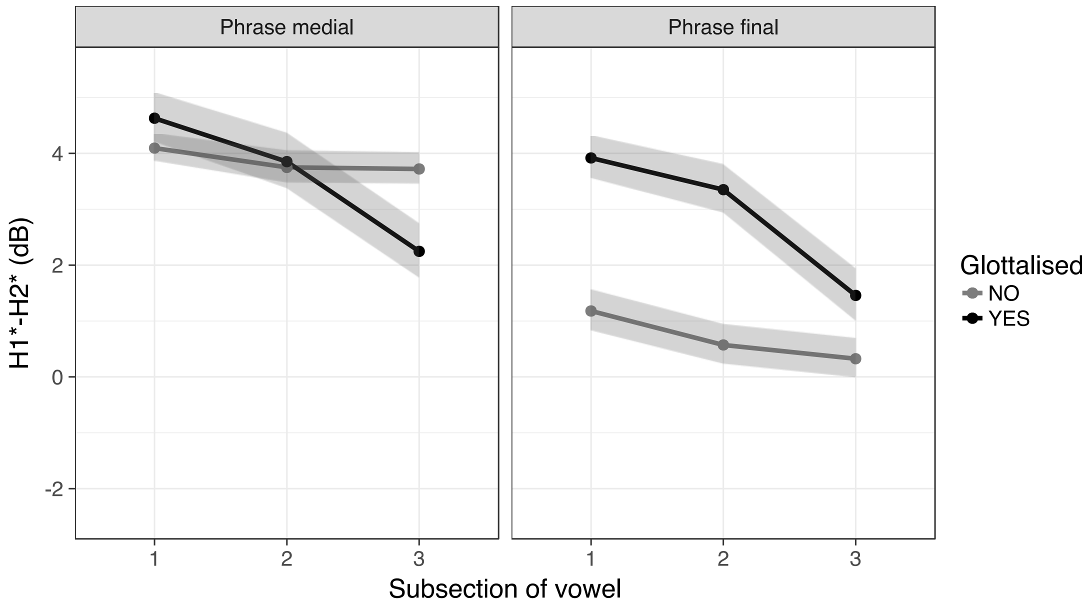

Figure 3. Mean H1*-H2* values from three equal subsections of the vowel in glottalised and non-glottalised items according to phrase position. Error ribbons represent 95% confidence intervals.

Table 1. Summary of linear mixed effects model to investigate effects of glottalisation, phrase position, and age group on changes in H1*-H2* values in the second and third subsections of the vowel. Significant effects (at α = 0.05) are marked with asterisks.

	β	SE	*t*	*p*
(Intercept)	0.277	0.317	0.872	0.392
Glottalisation	2.059	0.375	5.492	<0.0001*
Phrase position	-0.402	0.274	-1.467	0.145
Age group	0.008	0.317	0.024	0.981
Glottalisation: Phrase position	- 0.476	0.550	-0.865	0.387
Glottalisation: Age group	-0.929	0.487	-1.91	0.058
Phrase position: Age group	0.148	0.376	0.393	0.695
Glottalisation: Phrase position: Age group	1.27	0.689	1.839	0.066

F0 measurements were extracted from the vowels in each item using the STRAIGHT pitch tracker (Kawahara, Masuda-Katsuse, & de Cheveigné, 1999) in VoiceSauce (Shue et al., 2011) at 1 millisecond intervals. In order to reduce the likelihood of pitch tracking errors we used a two-pass method (Al-Tamini & Khattab, 2015, 2018; Hirst, 2011; Seyfarth & Garellek, 2018). In the first pass, F0 was estimated for each speaker using the program's default settings (Min: 40 Hz; Max: 500 Hz). Following Seyfarth and Garellek (2018), we then determined the apparent normal F0 range for each speaker by visually inspecting a histogram of their F0 estimates taken from each 1ms window in each vowel, and excluded any obvious outliers. In the second pass, we used each speaker's observed range (excluding outliers) to limit the range within which F0 was estimated (Al-Tamini & Khattab, 2015, 2018; Seyfarth & Garellek, 2018). Any item in which an increase or decrease in the F0 estimate from one window to the next was greater than half an octave was then inspected so as to eliminate items which contained pitch doubling or halving. Seven items containing octave jumps were excluded in this manner. Once we were confident in the accuracy of the F0 measurements, we then averaged F0 over three equal subsections of each vowel. That is, we calculated the mean F0 per token for the first, second, and final thirds of each vowel. In the analysis of F0 below we examine F0 in the final subsection, as this represents F0 at the offset of the vowel.

The following non-canonical stop realisations were labelled: items with no visible coda release burst were classified as unreleased; items in which frication was visible through the stop closure suggesting an incomplete or weak closure were labelled as spirantised; items with a portion of voiceless aspiration visible prior to the stop closure were classified as preaspirated (Hejná, 2015; Helgason, 2002; Nance & Stuart-Smith, 2013); items that did not provide evidence of a stop closure/burst, exhibited a short drop in amplitude, and were voiced throughout were labelled as taps (as is to be expected, taps were only present in intervocalic environments and were hence not found in final position); items which exhibited a low amplitude periodic waveform in the stop closure in conjunction with a sudden increase in fundamental frequency were classified as glottal squeaks (Hejná, Palo, & Moisik, 2016; Redi & Shattuck-Hufnagel, 2001). Figures containing example spectrograms of non-canonical stop realisations can be found in Penney, Cox, Miles, and Palethorpe (2018) and Penney et al. (2019 [Chapter 2]).

5.2.1.4 Annotation reliability

Ten per cent of the data were relabelled by the annotator, to ensure intra-annotator reliability. An additional 10 per cent of the data were labelled by the first author to assess inter-annotator reliability. Using the *irr* package (Wolak, Fairbairn, & Paulsen, 2012), we calculated intra-class correlations and 95% confidence intervals for the measure of vowel duration, and a Cohen's Kappa score for the binary labelling of glottalisation. The results can be seen in Table 2. In both cases reliability was high.

Table 2. Results inter- and intra-annotator reliability tests. Intra-class correlations and 95% confidence intervals are provided for vowel duration. Cohen's Kappa scores are provided for presence of glottalisation.

	Inter-annotator	Intra-annotator		Inter-annotator	Intra-annotator
ICC	0.996	0. 997	Kappa	0.877	0.862
95% CI	0.995, 0.997	0.996, 0.997	*z*	17.5	18.5
p	<0.0001	<0.0001	*p*	<0.0001	<0.0001

5.2.1.5 Statistical analyses

We conducted statistical analyses using the *lme4* package (Bates, Maechler, Bolker, & Walker, 2015) in R (R Core Team, 2017). For the analyses of the presence of glottalisation and of creaky voice we fitted separate generalised linear mixed effects models, utilising the BOBYQA optimizer (Powell, 2009). For the analyses of F0, H1*-H2*, and coda voicing-related vowel duration differences we fitted separate linear mixed effects models using the *lmerTest* (Kuznetsova, Brockhoff, & Christensen, 2017) package to estimate *p*-values. Post-hoc analyses using Tukey HSD corrections to *p*-values for multiple comparisons were conducted using the *lsmeans* package (Lenth, 2016). Specific details regarding dependent variables, fixed factors and random effects structures will be given separately below for each model.

5.2.2 Results

5.2.2.1 Presence of glottalisation

In phrase medial position, speakers produced more glottalised items with voiceless codas compared to voiced codas (older: 24% voiceless, 2% voiced; younger: 56% voiceless, 2% voiced). Similarly, in phrase final position there were more items glottalised with voiceless codas than with voiced (older: 47% voiceless, 10% voiced; younger: 48% voiceless, 14% voiced). Figure 4 below illustrates the proportion of items that were glottalised within each coda voicing context and phrase position according to age group. An effect of voicing context is clearly visible in both phrasal positions, with glottalisation primarily occurring in the voiceless coda context in each position. Glottalisation in the voiced coda context is essentially only present in the phrase final position; in phrase medial position there are very few items in the voiced context with glottalisation.

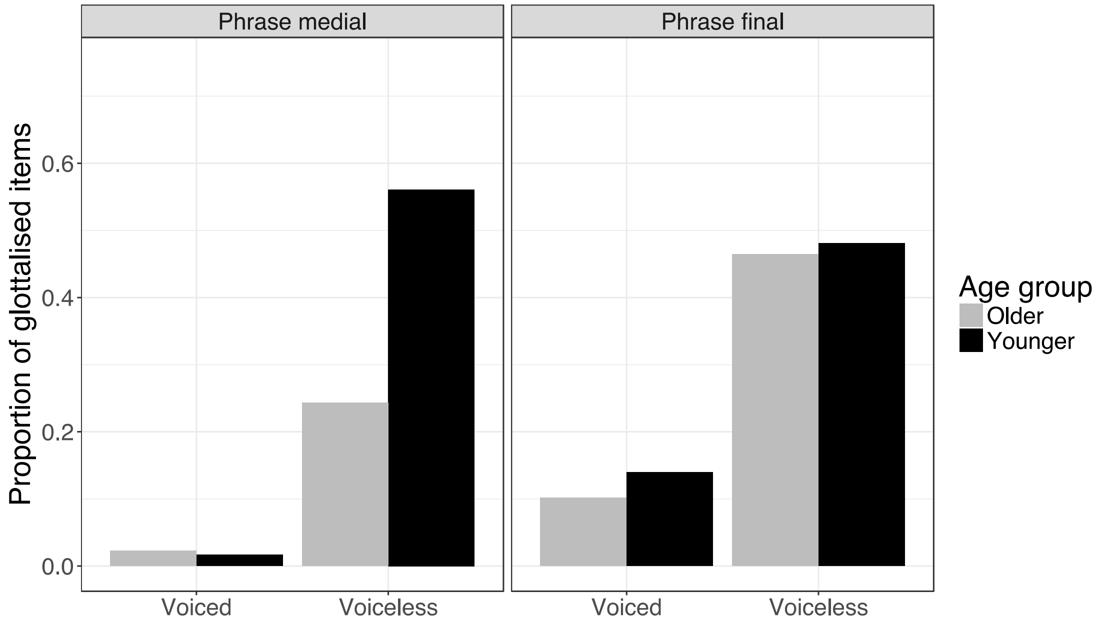

Figure 4. Proportion of glottalised items according to coda voicing context, phrase position, and age group.

We fitted a generalised linear mixed effect model to analyse the effects of coda voicing, phrase position, and age group on the presence of glottalisation. The dependent variable was the presence or absence of glottalisation. Fixed factors were *Voicing context* (voiced/voiceless), *Phrase position* (final/medial), and *Age group* (older/younger). We also included the factors *Vowel length* (short/long) and *Vowel height* (low/high) to investigate the effects these features may have on glottalisation. We included all two- and three-way interactions between the fixed factors. We included random intercepts for *Participant*, and random slopes for all within-participant factors.[8]

[8] The syntax for this model was:

glmer (Glotalisation ~ (Phrase position + Coda voicing + Age group + Vowel height + Vowel length)^3 + (1+ Phrase position + Coda voicing + Vowel height + Vowel length | Participant))

Table 3. Summary of generalised linear mixed effects model to investigate effects of phrase position, voicing context, age group, vowel height, and vowel length on the presence of glottalisation. Significant effects (at α = 0.05) are marked with asterisks.

	β	SE	z	p
(Intercept)	-2.062	0.399	-5.171	< 0.0001*
Phrase position	-1.678	0.645	-2.602	0.009*
Voicing context	1.838	0.415	4.429	< 0.0001*
Age group	-0.033	0.502	-0.065	0.948
Vowel height	-0.894	0.471	-1.898	0.058
Vowel length	-0.843	0.447	-1.887	0.059
Phrase position: Voicing context	0.340	0.619	0.55	0.583
Phrase position: Age group	-0.486	0.784	-0.62	0.535
Phrase position: Vowel height	-0.282	0.788	-0.358	0.721
Phrase position: Vowel length	-0.422	0.742	-0.569	0.569
Voicing context: Age group	0.037	0.513	0.072	0.943
Voicing context: Vowel height	0.838	0.517	1.622	0.105
Voicing context: Vowel length	1.017	0.500	2.035	0.042*
Age group: Vowel height	0.421	0.528	0.797	0.425
Age group: Vowel length	0.516	0.511	1.01	0.313
Vowel height: Vowel length	-0.013	0.552	-0.023	0.981
Phrase position: Voicing context: Age group	2.692	0.727	3.705	0.0002*
Phrase position: Voicing context: Vowel height	0.153	0.79157	0.194	0.847
Phrase position: Voicing context: Vowel length	0.419	0.736	0.569	0.569
Phrase position: Age group: Vowel height	-1.022	0.494	-2.07	0.038*
Phrase position: Age group: Vowel length	0.503	0.486	1.036	0.300
Phrase position: Vowel height: Vowel length	-0.316	0.481	-0.656	0.512
Voicing context: Age group: Vowel height	0.123	0.559	0.219	0.826
Voicing context: Age group: Vowel length	-0.913	0.544	-1.678	0.093
Voicing context: Vowel height: Vowel length	-0.094	0.550	-0.172	0.864
Age group: Vowel height: Vowel length	0.105	0.442	0.237	0.813

The summary of the generalised linear mixed effects model can be seen in Table 3. The model showed significant effects of *Phrase position*, indicating that glottalisation was more likely to be present in phrase final position, and *Voicing context*, demonstrating that glottalisation occurred primarily in the context of voiceless codas, as expected. In addition, we found a significant interaction between *Voicing context* and *Vowel length*. In the voiced context slightly more glottalisation was present for the long vowels compared to the short vowels, though this does not appear to be a strong effect. Post-hoc tests confirm that this is a weak effect, with no significant differences between short and long vowels in either the voiced or voiceless contexts.

We also found a significant three-way interaction between *Phrase position*, *Voicing context* and *Age group*, showing that although glottalisation occurred more frequently in the voiceless context compared to the voiced context for both age groups, there were nevertheless differences between the age groups across phrase positions. Post-hoc analyses confirm that for both age groups there was a significant difference between voiced and voiceless contexts in both phrase final and phrase medial position (all $p < 0.0001$). They also show that the younger and older groups differed from each other only in phrase medial position in the voiceless context ($p = 0.0001$). As Figure 4 above illustrates, glottalisation in the voiced context occurred primarily in phrase final position for both age groups, and occurred infrequently in phrase medial position. Correspondingly, for both age groups, in the voiced context, phrase position showed a significant effect (older: $p = 0.\ 0197$; younger: $p = 0.0001$). Phrase position in the voiceless context also affected the degree of glottalisation for the older group ($p = 0.0017$), as they produced less glottalisation phrase medially than phrase finally. The younger group did not show a significant difference between phrase medial and phrase final position in the voiceless context.

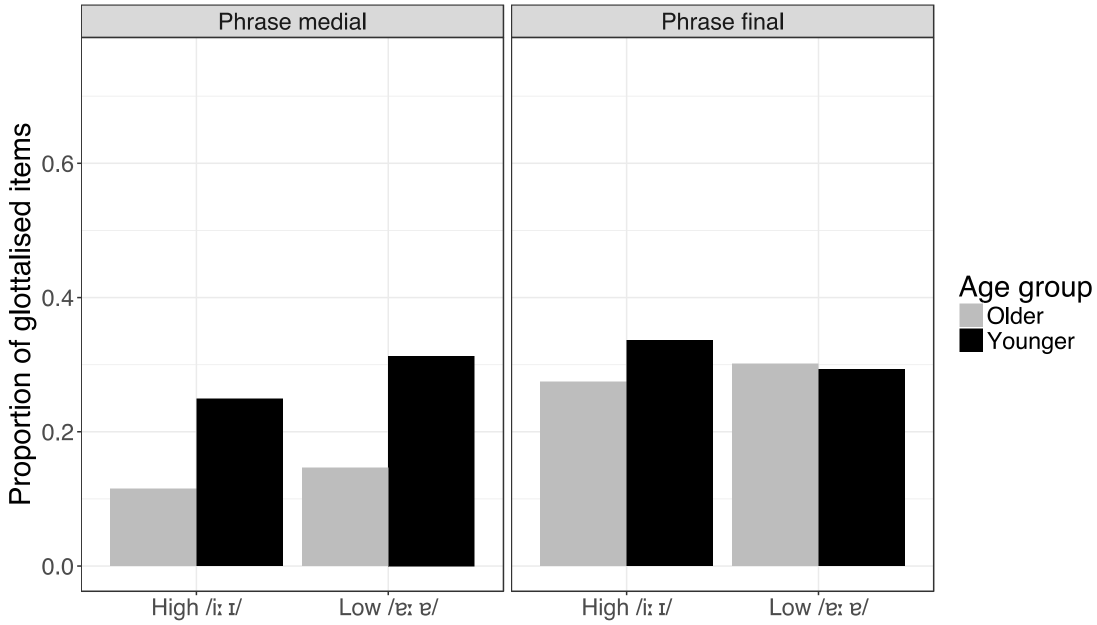

Figure 5. Proportion of glottalised items according to vowel height, phrase position, and age group. Note that this figure includes glottalisation occurring in both voiceless and voiced contexts.

Finally, the model also showed a significant interaction between *Phrase position*, *Age group*, and *Vowel height*, indicating that vowel height has an effect on glottalisation in phrase medial position, particularly for younger speakers. Figure 5 shows the proportion of items that were glottalised within each phrase position according to age group and vowel height. It should be borne in mind that glottalisation occurring in both voiceless and voiced contexts is included in this figure. As can be seen, in phrase final position glottalisation occurred at similar rates for both the older and younger groups. In phrase medial position the younger group glottalised more than the older group, and more so on low vowels compared to high vowels. Post-hoc analyses show that within each phrase position neither the older nor the younger group showed a significant difference between high and low vowels, though there was a strong trend for the younger speakers to have increased glottalisation on low vowels in phrase medial position (p = 0.0519). In addition, final and medial positions differed significantly for the older group for both low (p = 0.0073) and high vowels (p = 0.0051), with older speakers producing more glottalisation in the phrase final position in both cases. Similarly the younger group showed less glottalisation of high vowels in medial position compared to final position, and this difference was significant (p = 0.0064). However, they glottalised a similar proportion of low vowels in both phrase positions with no significant difference found. To summarise, the older speakers produced more glottalisation in final position and there was no difference between high and low vowels in either position. The younger speakers produced similar rates of glottalisation on low vowels in both positions, but less glottalisation on high vowels in medial position than final position. They also showed a tendency to glottalise low vowels more than high vowels in medial position.

This analysis demonstrates that glottalisation occurs primarily in the voiceless context, that younger speakers show more use of glottalisation than older speakers in phrase medial position but not phrase final position, and that for younger speakers low vowels are more affected by glottalisation.

5.2.2.2 Presence of creaky voice

As discussed above, we used a binary classification system whereby items were classified as creaky rather than glottalised when irregularity in F0 and amplitude extended leftward beyond the second half of the vowel but as glottalised if irregularity was localised to the last half of the vowel. In the majority of creaky voice cases the entire vowel segment was affected. Creaky voice was rarely present in phrase medial position (older: 3% voiceless, 2% voiced; younger: 5% voiceless, 2% voiced). As expected, substantially more creaky voice was observed in phrase final position (older: 24% voiceless, 16% voiced; younger: 41% voiceless, 45% voiced). Figure 6 illustrates the proportion of items that were produced with creaky voice within each coda voicing context and phrase position according to age group.

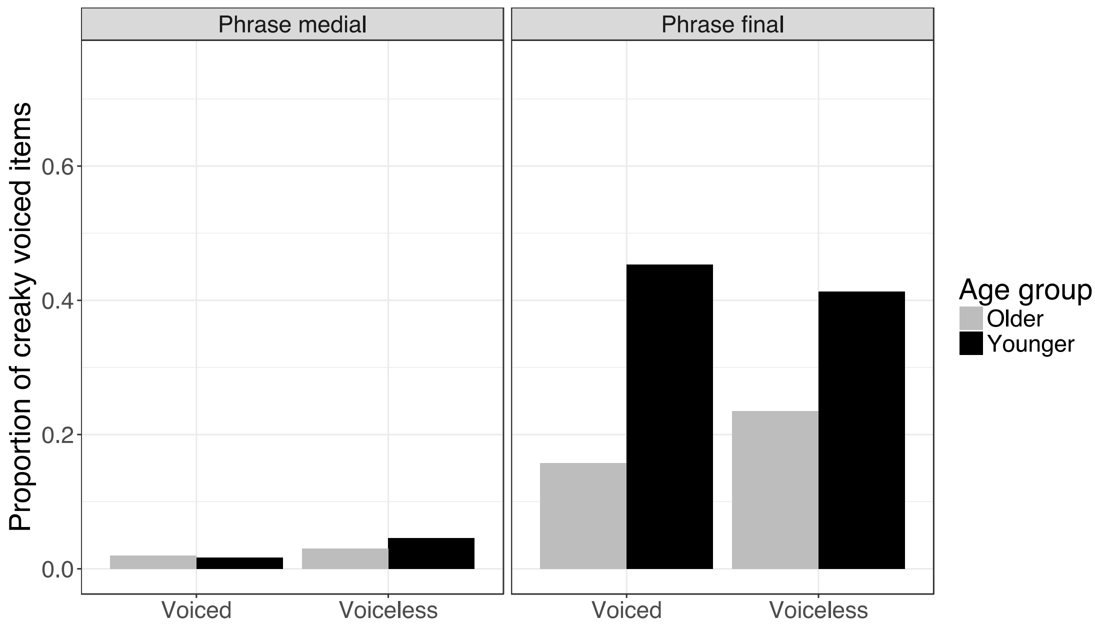

Figure 6. Proportion of items produced with creaky voice according to coda voicing context, phrase position, and age group.

In order to explore possible effects of coda voicing context and age group on the presence of creaky voice, we fitted a generalised linear mixed effects model to the data. As there were so few creaky voiced items in medial position, we included only items produced in final position in the model (1758 items; older: 770; younger: 988). The dependent variable was the presence of creaky voice. Fixed factors were *Voicing context* and *Age group*. We also included a two-way interaction between these factors. Random intercepts were included for *Participant* and *Word*, and random slopes for *Voicing context* by participant.[9]

[9] The syntax for this model was:

glmer(Creaky voice ~ Voicing context * Age group + (1+ Voicing context|Participant) + (1 | Word))

The model summary is shown in Table 4. We found a significant effect of *Voicing context*, with creaky voice found more in the voiceless than in the voiced coda context overall. The model also returned a significant effect of *Age group*, showing that the younger speakers produced more creaky voice than the older speakers. In addition, there was a significant interaction between *Voicing context* and *Age group*. Post-hoc tests confirmed that the older and younger speakers differed from each other in both voicing contexts, with the younger producing more creaky voice, though the difference was greater in the voiced ($p < 0.0001$) than in the voiceless ($p = 0.0304$) context. There was no significant difference between voicing contexts within either age group.

Table 4. Summary of generalised linear mixed effects model to investigate effects of voicing context and age group on the presence of creaky voice. Significant effects (at α = 0.05) are marked with asterisks.

	β	SE	*z*	*p*
(Intercept)	-2.747	0.441	-6.232	< 0.0001*
Voicing context	0.864	0.364	2.375	0.018*
Age group	2.337	0.518	4.512	< 0.0001*
Voicing context: Age group	-1.095	0.328	-3.34	0.001*

5.2.2.3 F0

Figure 7 shows the mean values of F0 at the offset of the vowel (i.e. averaged over the final third of the vowel) for all items (i.e. glottalised and non-glottalised) according to coda voicing context, phrase position, and age group. As both males and female speakers were present in both age groups, and differences in F0 are to be expected due to gender based physiological differences, we provide the means for males and females within these age groups separately. Mean F0 values along with standard deviations and standard errors are provided in Table A1 in Appendix A. There was little difference in F0 across coda voicing contexts within each age and gender group (< 5Hz). The exception to this is the younger female group, who appear to maintain a small voicing-related difference in F0 in both phrase final (~15 Hz) and phrase medial (~9 Hz) position. However, for the younger females in phrase medial position F0 is *lower* in the voiceless context than in the voiced context; that is, the difference is in the opposite direction to that which is expected. The young males also show a lower F0 in the voiceless compared to the voiced context in medial position, though the difference is small (~4.5 Hz). It is possible that that the lower F0 values in the voiceless context may be due to the increased likelihood of glottalisation, which may result in a drop in F0 (Dilley, Shattuck-Hufnagel, & Ostendorf, 1996; Garellek & Seyfarth, 2016).

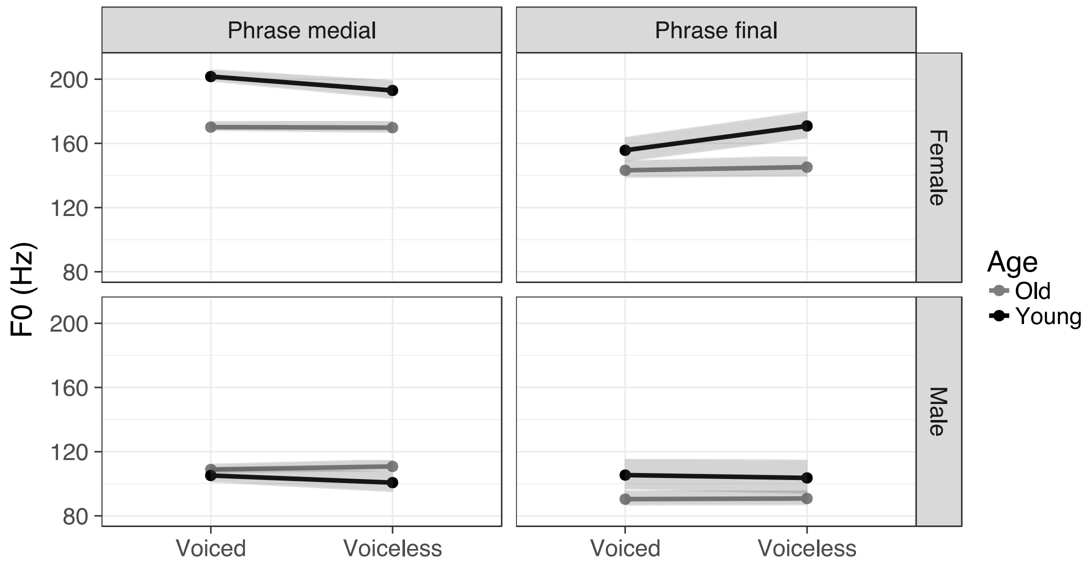

Figure 7. Mean F0 (Hz) in final third of the vowel according to voicing context, phrase position, age group and gender. Error ribbons represent 95% confidence intervals.

We fitted linear mixed effects models for the females and males separately, both with the dependent variable F0, and fixed effects *Voicing context*, *Phrase position* and *Age group*. We included all two- and three-way interactions between these factors. Random intercepts were included for *Participant* and *Word*, and random slopes for all within-participant factors.[10]

Table 5. Summary of linear mixed effects model to investigate effects of glottalisation, phrase position, and age group on F0 for female speakers. Significant effects (at α = 0.05) are marked with asterisks.

	β	SE	*t*	*p*
(Intercept)	142.368	10.783	13.202	<0.0001*
Voicing context	2.543	7.251	0.351	0.732
Phrase position	28.014	9.929	2.821	0.006*
Age group	15.900	12.801	1.242	0.219
Voicing context: Phrase position	-3.011	5.583	-0.539	0.590
Voicing context: Age group	13.883	5.726	2.425	0.017*
Phrase position: Age group	15.718	12.740	1.234	0.222
Voicing context: Phrase position: Age group	-23.237	7.227	-3.215	0.001*

The model summary for the female speakers is given in Table 5 and shows a significant effect of *Phrase position*, a significant two-way interaction between *Voicing context* and *Phrase position*, and a three-way interaction between *Voicing context*, *Phrase position* and *Age group*. Post-hoc analyses showed no significant difference between voiced and voiceless contexts in either phrase position for both the younger and older speakers. The older speakers also showed no difference across phrase positions for either voiced or voiceless contexts; the younger speakers showed no differences between phrase positions for the voiceless context, but did show a significant difference between phrase positions in the voiced context ($p < 0.0001$), with a higher mean F0 in medial (201 Hz) compared to final (155 Hz) position.

[10] The syntax for both of these models was:

lmer (F0 ~ (Voicing context + Phrase position + Age group)^3 + (1 + Voicing context + Phrase position | Participant) + (1 | Word)

Table 6. Summary of linear mixed effects model to investigate effects of glottalisation, phrase position, and age group on F0 for male speakers. Significant effects (at α = 0.05) are marked with asterisks.

	β	SE	*t*	*p*
(Intercept)	90.040	8.183	11.004	<0.0001*
Voicing context	0.755	3.990	0.189	0.853
Phrase position	19.034	7.539	2.525	0.022*
Age group	15.127	12.339	1.226	0.239
Voicing context: Phrase position	1.437	3.618	0.396	0.692
Voicing context: Age group	-3.579	4.251	-0.842	0.402
Phrase position: Age group	-16.866	11.851	-1.423	0.172
Voicing context: Phrase position: Age group	-4.165	5.903	-0.706	0.481

The model summary for the male speakers is shown in Table 6. For the male speakers the model showed a significant effect of *Phrase position*, demonstrating that F0 was higher in phrase medial position than in phrase final position.

5.2.2.4 Non-canonical stop realisations

As was the case in our previous studies (Penney, Cox, Miles, & Palethorpe, 2018; Penney et al., 2019 [Chapter 2]), we observed a number of non-canonical stop realisations apart from glottalisation. Figure 8 illustrates the various stop realisations according to age group and voicing context. Note that the proportions shown in this figure represent the proportion of all items (i.e. including glottalised and creaky tokens) that were produced with a non-canonical stop realisation, averaged across phrase position. Unreleased stops were the most frequent non-canonical stop realisation, particularly in the voiceless context. This was especially the case for the younger speakers, who produced unreleased stops more frequently than the older speakers. Consistent with previous reports, many of the unreleased stops co-occurred with glottalisation (Blevins, 2006; Ingram, 1989; Kahn, 1976; Penney, Cox, Miles, &

Palethorpe, 2018; Penney et al., 2019 [Chapter 2]; Selkirk, 1982): glottalisation occurred in conjunction with approximately half of all cases of unreleased stops. In addition, we found a substantial proportion of stops that were spirantised. The older speakers showed a tendency to produce spirantised stops, whereas the younger group produced relatively little spirantisation. Glottalisation co-occurred with spirantisation in some cases, though not as frequently as it did with unreleased stops. Pre-aspirated realisations, taps and glottal squeaks were rare. Specific details regarding each of the non-canonical realisations can be found in Appendix B. As some of the non-canonical stop realisations have in the past been shown to vary according to gender, Table A2 in Appendix B provides details regarding gender, age group, phrase position and voicing context.

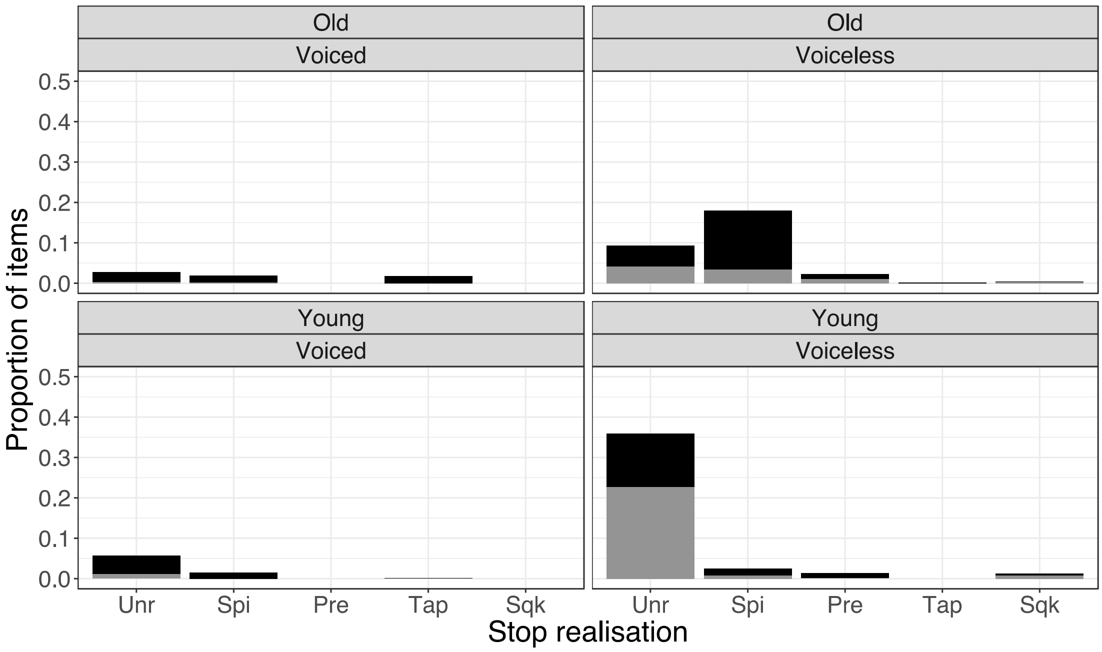

Figure 8. Proportions of items produced with non-canonical stop realisations according to coda voicing context and age group. Grey portions represent proportion of items co-occurring with glottalisation. Unr = unreleased; Spi = spirantised; Pre = preaspirated; Tap = tapped; Sqk = glottal squeaks.

5.2.2.5 Coda voicing-related duration difference

In this section we analyse potential differences between the age groups in their use of vowel duration to cue coda voicing. We here focus only on the data produced in phrase medial position. This comprises 1493 items (older: 683; younger: 810). The rationale for limiting the scope of this analysis in this way is that in doing so we can be reasonably certain that glottalisation observed in this position is less likely to be due to creaky voice.

In order to account for possible vowel duration differences between older and younger speakers being related to differences in speaking rate (e.g. Sole, 2007), we calculated a normalised vowel duration measure. For each item, we first calculated the ratio of the vowel duration of the target item relative to the duration of the portion of the phrase from the onset of the initial fricative /s/ in ‘say’ to the offset of the vowel in the target item, which totalled four segments (i.e. ***say bi****t* /sæɪ bV/). We used the target vowel offset as many items did not include a coda stop release making it difficult to establish the final consonant duration.

We then calculated the ratio of mean normalised vowel duration in the voiceless coda context relative to mean normalised vowel duration in the voiced coda context per vowel for each speaker (e.g. the ratio of normalised vowel duration in *beat* relative to *bead*). A durational difference ratio close to 1 would show that normalised vowel duration in the voiceless context was of similar duration to normalised vowel duration in the voiced context, whereas a durational difference ratio of 0.5 would indicate that a speaker produced a vowel in the voiced context that was double the length of the vowel in the voiceless context. Figure 9 shows the mean durational difference ratio according to age group, vowel height, and vowel length. The figure illustrates that the younger speakers have higher ratios than the older speakers for all of the vowels (i.e. a reduced difference in vowel duration in voiced compared to voiceless coda contexts), though the difference between age groups appears to be reduced for the long, low vowel /ɐː/.

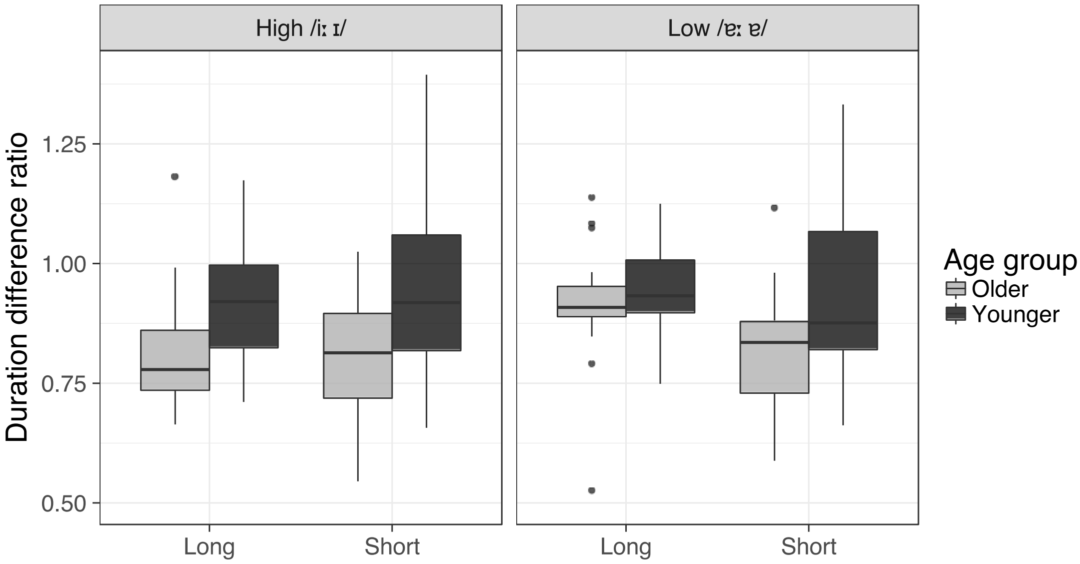

Figure 9. Mean duration difference ratio according to age group, vowel height, and vowel length.

We fitted a linear mixed model with the dependent variable of the duration difference ratio. We included *Age group* as a fixed factor, and in order to further explore whether coda voicing-related durational differences may be vowel-specific, we also included *Vowel height* and *Vowel length* as fixed factors. We included interaction terms for all two- and three-way interactions and random intercepts for *Participant*. This was the maximal random effects structure to converge.[11]

The model summary is shown in Table 7. The results of the model included significant effects for *Vowel height* and *Vowel length*, and a significant interaction

[11] The syntax for this model was:

lmer(Duration difference ratio ~ (Age group + Vowel height +Vowel length)^3 + (1|Participant) between *Vowel height* and *Vowel length.* Post-hoc tests show that for the short vowels there is no difference between the low and high vowels. For the long vowels, however, there was a significant difference (p = 0.010), with a higher ratio (i.e. less difference in duration across voicing contexts) found for the long low vowel compared to the long high vowel, as can be seen in Figure 9 above. In addition, there was a significant interaction between *Age group* and *Vowel length*. Post-hoc tests show that, although the older group showed lower ratios (i.e. more difference in duration across voicing contexts) than the younger group for both short and long vowels, the difference was greater for the short vowels (p < 0.0001) than it was for the long vowels (p = 0.019).

Table 7. Summary of linear mixed effects model to investigate effects of age group, vowel height, and vowel length on coda voicing-related vowel duration differences. Significant effects (at α = 0.05) are marked with asterisks.

	β	SE	t	p
(Intercept)	0.917	0.024	38.359	<0.0001*
Age group	0.033	0.032	1.032	0.303
Vowel height	-0.105	0.031	-3.363	0.0009*
Vowel length	-0.112	0.033	-3.459	0.0006*
Age group: Vowel height	0.078	0.042	1.857	0.065
Age group: Vowel length	0.092	0.043	2.131	0.034*
Vowel height: Vowel length	0.105	0.045	2.336	0.021*
Age group: Vowel height: Vowel length	-0.056	0.060	-0.934	0.352

5.2.3 Discussion

The results presented above confirm the prediction that glottalisation would occur primarily in the voiceless coda context. In both phrase medial and phrase final position glottalisation occurred significantly more frequently in words with voiceless codas than in words with voiced codas for both age groups. These results are consistent with previous findings (Penney, Cox, Miles, & Palethorpe, 2018; Penney et al., 2019 [Chapter 2]) and confirm that glottalisation can be considered an acoustic correlate of coda voicelessness in AusE, and this is not limited to single item utterances. In addition, although the overall rates of glottalisation were slightly lower than previous reports (/t/ in medial position – older: 24%, younger: 56%; /t/ in stressed syllables in single items (Penney, Cox, Miles, & Palethorpe, 2018) – older: 36%, younger: 71%; /t/ in unstressed syllables in single items (Penney et al., 2019 [Chapter 2]) – older: 43%, younger: 74%), the patterns of glottalisation observed in phrase medial position here (where creaky voice is neither generally expected nor was it observed) closely reflect results previously reported for single item utterances; namely, that glottalisation primarily occurs in voiceless coda contexts and that younger speakers glottalise more than older speakers. This supports the interpretation of those previous data and the conclusions drawn, in particular that the observed glottalisation was due to coda glottalisation and not to phrase final creaky voice.

We also predicted that older speakers would utilise glottalisation less than younger speakers, in line with previous findings. Indeed, in phrase medial position the results showed that more glottalisation was present for the young speakers, similar to previous reports (Penney, Cox, Miles, & Palethorpe, 2018). However, in phrase final position we found comparable rates of glottalisation in the two age groups: in the voiceless context the older speakers glottalised 47% of items and the younger speakers 48%. We suggest that there are two possible explanations for this finding. First, it may be that glottalisation is employed to a similar degree by older and younger speakers to cue coda voicelessness in phrase final position. This is a plausible explanation, given that rates of glottalisation may increase in final position, possibly as a mechanism to support the voicing contrast, which may be reduced by weak voicing in final position (Huffman, 2005, Keyser & Stevens, 2006; Seyfarth & Garellek, 2015). This would suggest that glottalisation is more likely to occur in final position, but that the younger speakers employ glottalisation in a more generalised fashion such that they glottalise at high rates in both final and medial position. However, the fact that both age groups showed a significant increase in glottalisation in the voiced context in final position may suggest that an alternative explanation is more reasonable: namely, that some of the examples of glottalisation in final position are actually short examples of phrase final creaky voice that were labelled as glottalised according to our annotation criteria. The target items in phrase final position came at the end of a relatively long phrase; therefore, speakers were likely to be near the end of their breath cycle, an environment that may favour creaky voice (Aare, Lippus, Włodarczak, & Heldner, 2018; Slifka, 2006). On the other hand, if this were the case, we would expect to observe an increase in glottalisation in both voiced and voiceless contexts and for both younger and older speakers. While we do see an increase for both age groups in the voiced context phrase finally and for older speakers in the voiceless context phrase finally, for the older speakers the increase in the voiceless context is much greater than in the voiced context, and for the younger speakers we observe a slight decrease in the proportion of glottalised items in the voiceless context phrase finally compared to medially. Further examination is necessary to explore this issue in more detail.

We found some evidence to support glottalisation being utilised more on low vowels than high vowels; in medial position the younger speakers produced more glottalisation on low vowels but there was no significant effect for the older speakers. In a previous study glottalisation was found to be more prevalent on non-high vowels than on high vowels (Penney, Cox, Miles, & Palethorpe, 2018), and glottalisation tends to favour low vowels cross linguistically (Brunner & Żygis, 2011; Malisz, Żygis, & Pompino-Marschall, 2013). In final position, rates of glottalisation were similar across vowel height for both age groups.

In the analysis of creaky voice, we found very few creaky items in phrase medial position, but considerable creaky voice in phrase final position, as is to be expected. In final position younger speakers produced more items with creaky voice than the older speakers in line with findings that creaky voice occurs frequently in the speech of young people, particularly AmE women (Abdelli-Beruh, Wolk, & Slavin, 2014; Wolk, Abdelli-Beruh, & Slavin, 2012; Yuasa, 2010). Anecdotal reports about increasing prevalence of creaky voice in AusE are common (Dallaston & Docherty, 2019). This result might suggest that phrase final creaky voice is increasingly being used by young AusE speakers and may also point to a more general difference in laryngeal strategies being employed by older and younger speakers.

We found no F0 contrast between voiced and voiceless contexts for either females or males in either phrase position for either age group, which supports the suggestion that glottalisation is increasingly a more reliable cue to coda voicing in the face of reduction in the use of other voicing cues. For both males and females we found a strong effect of phrase position, reflecting lower F0 values phrase finally, as would be expected (although the younger males had similar values in both phrase positions). The younger females showed a large drop in F0 in final position, which is consistent with the use of phrase final creaky voice.

Our analysis of vowel duration differences showed that this cue to coda voicing may be reduced in younger speakers, further supporting the idea that glottalisation is becoming increasingly important to maintain the voicing contrast. Overall, vowel duration differences across voicing contexts were reduced in younger compared to older speakers. Penney, Cox, Miles, and Palethorpe (2018) found that younger speakers made less use of vowel duration in conjunction with increased use of glottalisation, and the results here correspond with this finding: younger speakers appear to utilise glottalisation more and vowel duration less than older speakers. Penney, Cox, Miles, and Palethorpe (2018) also found that for both older and younger speakers, vowel duration differences were greater for high vowels, and correspondingly glottalisation was used less on high vowels. Here we also found that voicing-related vowel duration differences were greater on high vowels (in particular for the long high vowel /iː/) and, as discussed above, there was a strong trend for the younger speakers to produce more glottalisation on low vowels, though we did not find an effect of vowel height on rates of glottalisation for the older speakers. This may reflect differences in the methodology employed in the various studies.

5.3 Analysis of following heterosyllabic environments

5.3.1 Methods

In this section we analyse whether the following heterosyllabic environment has an effect on the presence of glottalisation. The participants and details of the data collection are as described in 5.2.1 above. We examined target words produced in the *SAY <TARGET> <FOLLOWING> ONE MORE TIME* carrier where the word directly following the target was one of the words: *again*, *only*, *town*, *down*, *now*. This enabled an examination of items in phrase medial position with the following environment of a stressed vowel (*only*; STRV), an unstressed vowel (*again*; UNSTRV), a voiceless stop (*town*; VLST), a voiced stop (*down*; VOIST), or a nasal (*now*; NAS). For this analysis we examined only target words with voiceless codas, as this is the voicing context in which glottalisation is generally present (see 5.2.2.1 above). The target words included in this analysis were: *bart, beat, but, bit.* Note that the data in the STRV environment were included in the analyses in section 5.2 above, though in this section we only examine items with voiceless codas, whereas items with both voiceless and voiced codas were examined in 5.2.

As above, each of the 77 participants produced 3 repetitions of each of the 4 target words in each of the 5 environments, providing 4620 items. 222 items were discarded due to mispronunciations or other issues. A further 631 items were excluded due to participants inserting a phrase break after the target word or accenting the target word. In addition, 55 items (all *but*) were excluded for being produced with a schwa instead of /ɐ/ (labelled as schwa by WebMAUS and confirmed auditorily). This left 3712 items (older: 1713; younger: 1999) remaining for this analysis (*beat*: 912; *bart*: 943; *bit*: 975; *but*: 882).

5.3.1.2 Annotation reliability

To ensure intra-annotator reliability, 10 per cent of the data were relabelled. To ensure inter-annotator reliability, an additional 10 per cent of the data were labelled by another trained annotator. We calculated a Cohen's Kappa score for the labelling of glottalisation as present or absent, using the *irr* package (Wolak et al., 2012). Table 8 shows that reliability was high for both inter- and intra-annotator tests.

Table 8. Results inter- and intra-annotator reliability tests. Cohen's Kappa scores are provided for presence of glottalisation.

	Inter-annotator	Intra-annotator
Kappa	0.856	0.916
z	17.1	18.3
p	< 0.0001	< 0.0001

5.3.1.3 Statistical analysis

We fitted a generalised linear mixed effects model to the data to investigate whether the following environment has an effect on the presence of glottalisation. The dependent variable was the presence of glottalisation. Fixed factors included were *Age group*, *Environment, Vowel height* and *Vowel length*. We also included terms for all two- and three-way interactions. We included random intercepts for *Participant* and random slopes for all within-participant factors.[12] As a multilevel factor (*Environment*) was included in this model, we then conducted Type III Wald tests in order to better interpret the results.

5.3.2 Results

Table 9 shows the summary of the generalised linear mixed effects model. We found significant effects for *Age group*, illustrating that the younger speakers use glottalisation more than the older speakers in each environment, and *Environment*, confirming that the following environment has an effect on whether glottalisation is employed. We also found a significant two-way interaction between *Environment* and *Vowel height*, and a significant three-way interaction between *Age group*, *Environment*, and *Vowel height*. Post-hoc analyses indicate that for the older group there is a difference between the pre-vocalic environments (STRV and UNSTRV) and the other three pre-consonantal environments (VLST, VOIST, NAS) for both high and low vowels (p = 0.0062 and below), though not between the two pre-vocalic

[12] The syntax for this model was:

glmer(Glottalisation ~ (Age group + Environment +Vowel height + Vowel length)^3 + (1+ Environment + Vowel height + Vowel length|Participant) environments or between the three pre-consonantal environments. That is, for the older speakers there is a difference in the use of glottalisation between pre-vocalic and pre-consonantal environments, with more glottalisation present in the pre-consonantal environments. This can be seen below in Figure 10, which displays the proportion of glottalised items in each of the following environments according to vowel height and age group.

Table 9. Summary of Type III (Wald) test of generalised linear mixed effects model to investigate effects of age group, following environment, vowel height and vowel length the presence of glottalisation. Significant effects (at α = 0.05) are marked with asterisks.

	χ^2	df	*p*
(Intercept)	23.065	1	< 0.0001*
Age group	18.798	1	< 0.0001*
Environment	35.495	4	< 0.0001*
Vowel height	1.810	1	0.179
Vowel length	1.916	1	0.166
Age group: Environment	5.183	4	0.269
Age group: Vowel height	1.128	1	0.288
Age group: Vowel length	0.016	1	0.901
Environment: Vowel height	11.054	4	0.026*
Environment: Vowel length	2.318	4	0.678
Vowel height: Vowel length	0.100	1	0.753
Age group: Environment: Vowel height	14.577	4	0.006*
Age group: Environment: Vowel length	2.352	4	0.671
Age group: Vowel height: Vowel length	0.130	1	0.719
Environment: Vowel height: Vowel length	2.868	4	0.580

For the younger group there is also a difference between the STRV environment and each of the pre-consonantal environments for both high and low vowels ($p = 0.0158$ and below) and, as with the older group, no differences between the pre-consonantal environments. The UNSTRV vowel environment did not differ significantly from the STRV environment for either high or low vowels, nor did it differ from any of the pre-consonantal environments for low vowels; however, the UNSTRV environment did differ from the NAS environment for the high vowels (p = 0.0015). Figure 10 shows that there is generally less glottalisation for the high vowels, but in the NAS environment glottalisation rates remain high (particularly for the younger speakers). To summarise, for the younger speakers there is a difference in the use of glottalisation between stressed pre-vocalic and pre-consonantal environments, with glottalisation more frequent in pre-consonantal environments, but in unstressed pre-vocalic environments glottalisation occurs at rates between those in the stressed pre-vocalic and pre-consonantal environments. This may suggest that glottalisation is spreading from pre-consonantal environments to pre-vocalic environments, with the UNSTRV environment initially affected.

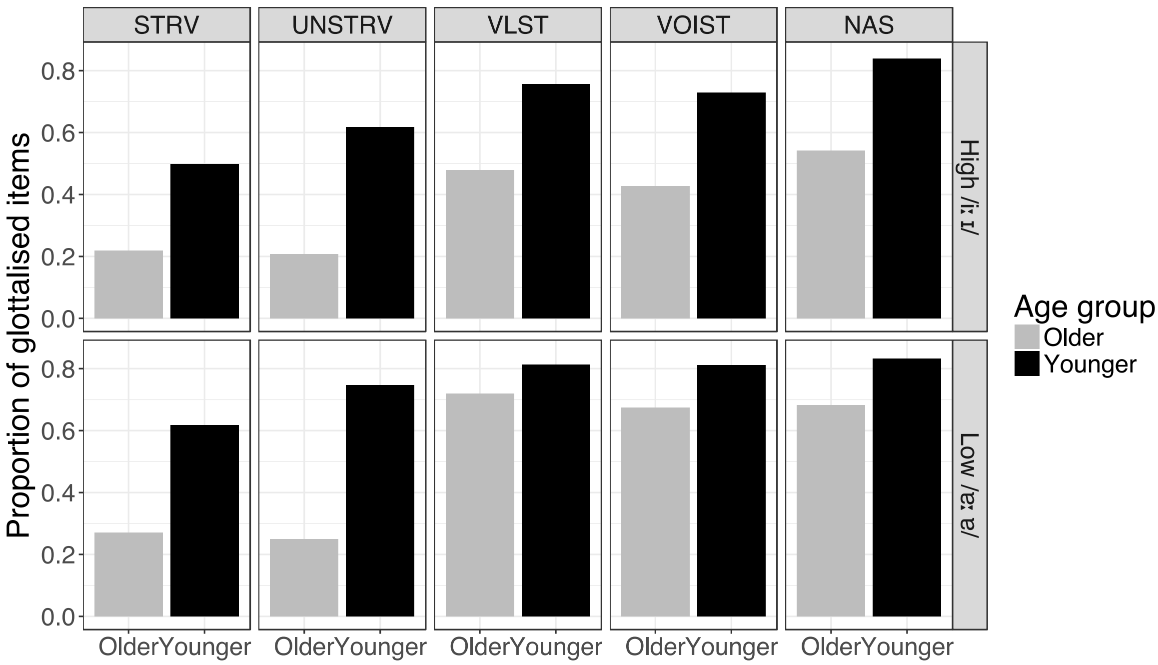

Figure 10. Proportion of glottalised items according to following environment, vowel height, and age group. STRV = stressed vowel; UNSTRV = unstressed vowel; VLST = voiceless stop; VOIST = voiced stop; NAS = nasal.

Post-hoc tests also showed that the differences between the younger and older groups were significant for high vowels in the UNSTRV (p = 0.0006), VOIST (p =

0.0353), and NAS (p = 0.0019) environments, and for low vowels in the STRV (p = 0.0008) and UNSTRV (p < 0.0001) environments. That is, for low vowels there is no significant difference between the age groups in the pre-consonantal environments, possibly suggesting that glottalisation initially affected low vowels pre-consonantally, but is now spreading to high vowels as well. Within the age groups, the younger group did not show a difference between high and low vowels in any of the environments. The older group, on the other hand, showed a significant difference between high and low vowels in the VLST (p = 0.0009) and VOIST (p = 0.0010) environments, but not in the STRV, UNSTRV, or NAS environments.

To summarise, glottalisation was generally used more in pre-consonantal than pre-vocalic environments, although the younger speakers used similarly high rates of glottalisation in the unstressed pre-vocalic environment; the younger speakers glottalised more than the older group, but in low vowel pre-consonantal contexts glottalisation occurred at similar rates for both age groups; older speakers show increased glottalisation on low vowels in pre-obstruent compared to pre-vocalic environments.

5.3.3 Non-canonical stop realisations

As was the case in section 5.2 above, there were a number of other non-canonical stop realisations present in the data. Figure 11 illustrates the various stop realisations according to age group and following environment. Note that as glottal squeaks were rare in this analysis they are not displayed. As in section 5.2, unreleased stops were the most frequent non-canonical stop realisation. In both age groups unreleased stops occurred frequently in pre-consonantal environments, at much higher rates than those observed in section 5.2; however, it should be noted that the unreleased category includes items in which there was only evidence of a single (geminate) stop release associated with a following onset stop in the VOIST and VLST environments, and items that were nasally released in the NAS environment. Unreleased stops occurred less frequently in pre-vocalic environments for both age groups, though the younger group produced substantially more unreleased stops pre-vocalically than the older group. As in 5.2 above and in previous studies, (Blevins, 2006; Ingram, 1989; Kahn, 1976; Penney, Cox, Miles, & Palethorpe, 2018; Penney et al., 2019 [Chapter 2]; Selkirk, 1982), unreleased stops occurred frequently in conjunction with glottalisation. The older speakers showed increased use of spirantised realisations in pre-vocalic environments, particularly preceding an unstressed vowel. Consistent with section 5.2 above, some spirantised items were also produced with glottalisation, though this occurred much less frequently than with unreleased stops. Preaspirated realisations and taps occurred infrequently (taps occurred only in intervocalic environments), and only three glottal squeaks were observed. Specific details regarding each of the non-canonical realisations can be found in Appendix C. Table A3 in Appendix C lists the various stop realisations according to age group, gender and environment.

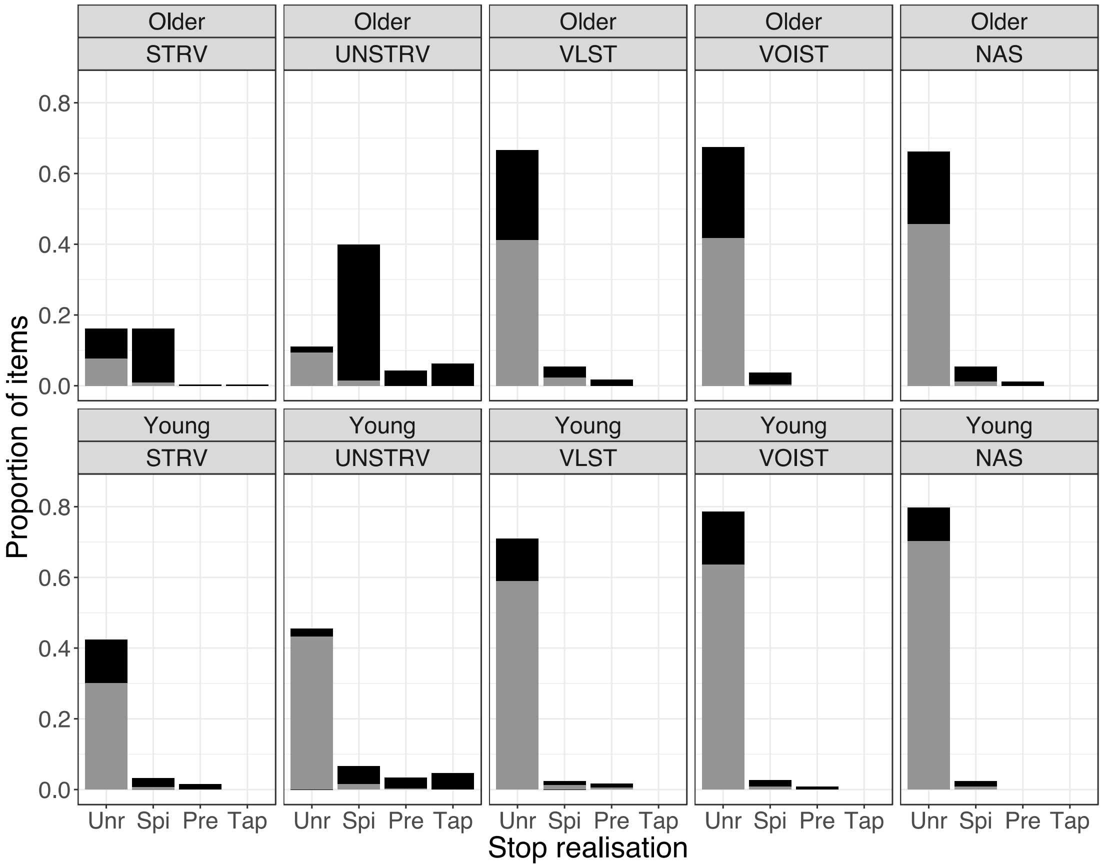

Figure 11. Proportion of non-canonical stop realisations according to following environment and age group. Grey portions represent proportion of items co-occurring with glottalisation. Unr = unreleased; Spi = spirantised; Pre = preaspirated; Tap = tapped. STRV = stressed vowel; UNSTRV = unstressed vowel; VLST = voiceless stop; VOIST = voiced stop; NAS = nasal.

5.3.4 Glottal reinforcement vs glottal replacement

In order to analyse whether the glottalisation identified here represents glottal reinforcement or rather glottal replacement, we carried out an additional examination of F2 transitions. As discussed above, glottalisation can take the form of glottal reinforcement, in which a glottal gesture occurs in conjunction with a supralaryngeal oral stop gesture, or of glottal replacement, where the oral stop gesture is 'lost' and replaced with a glottal gesture in a process of debuccalisation (O’Brien, 2012). Glottalised items that have an obvious stop release necessarily have an oral gesture; however for glottalised items that are unreleased it can be difficult to ascertain whether an oral gesture is present or not. If an oral gesture is retained, that is, if glottalisation reinforces the oral stop, we should expect to see an F2 transition at the offset of the vowel indicative of an alveolar closure (i.e. towards the locus for alveolar place of articulation (Delattre, Liberman, & Cooper, 1955; Modarresi, Sussman, Lindblom, & Burlingame, 2005); recall that in this data all of the coda stops are alveolar). For the high (front) vowels (/iː, ɪ/) we should therefore expect evidence of a drop in F2 from the vowel target to the vowel offset, and for the low (central) vowels (/ɐː, ɐ/) we would expect to see a rise in F2. On the other hand, if glottal replacement is present and hence the oral gesture has been lost, such transitions would not be expected and the F2 should remain relatively flat from the vowel target to the offset of the vowel. The pre-consonantal environments all contain onset consonants in the following word with alveolar place of articulation (VLST: *town*; VOIST: *down*; NAS: *now*). Therefore, we would expect transitions indicative of an alveolar stop to be present in these environments even in the case of glottal replacement. We thus restrict the analysis of F2 transitions to the pre-vocalic environments (STRV and UNSTRV), where we can be reasonably certain that no alveolar gesture associated with the following environment will be present.

For all items in the pre-vocalic environments that were identified as glottalised we extracted and checked formant values using VoiceSauce (Shue et al., 2011) as described in 5.2.1.3. F2 measurements were calculated at the target of each item (identified as the peak of F2 for high (front) vowels and the trough of F2 for low (central) vowels) and at the offset of the vowel, as identified by the cessation of F2. We examined 654 glottalised items (307 released; 347 unreleased). As differences in formant measurements are expected between male and female speakers, we plot the F2 means separately below. Note that our participants were not balanced for gender, and there were more female than male speakers in each age group. Moreover, as we examine only glottalised items here, it should be borne in mind that the means for the male speakers are based on relatively small numbers of items (see Table A4 in Appendix D for details).

Figure 12 illustrates the F2 transitions in released and unreleased glottalised items for the older and younger age groups according to vowel height, separated by gender. As expected, F2 values are lower for males than for females. Additionally, transitions towards an alveolar locus are visible in all cases, suggesting that an oral gesture is present in glottalised items with no coda stop burst, indicating glottal reinforcement rather than glottal replacement. Note that more variation is present for the male speakers (particularly for the high vowels); however, as mentioned above, there were fewer male speakers and hence relatively few items included here.

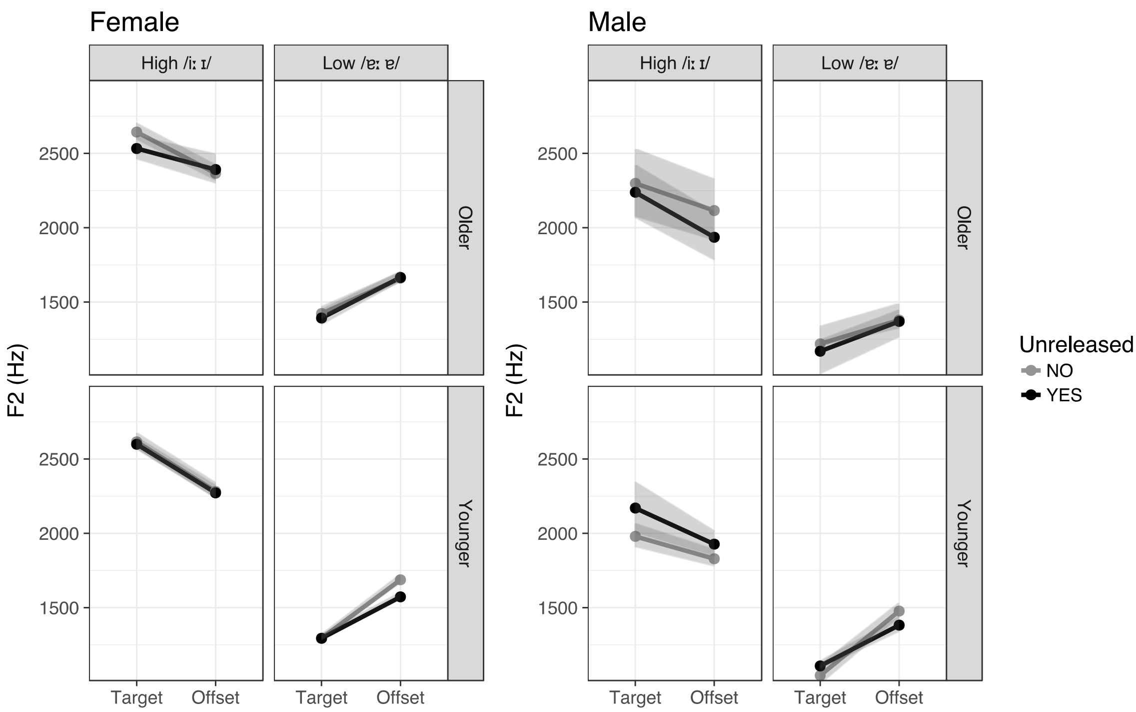

Figure 12. F2 transitions (Hz) in released and unreleased glottalised items according to vowel height and age group for female (left panel) and male (right panel) speakers. Error ribbons represent 95% confidence intervals.

Although the group means shown above indicate that an oral coda stop was generally present in unreleased glottalised items, visual inspection of the individual items revealed 33 items in which no F2 transition into a stop were visible, suggesting that these were examples of glottal replacement (all of these items produced the auditory percept of a glottal stop). Although restricted to relatively few items, this finding nevertheless suggests that some AusE speakers may employ glottal replacement in pre-vocalic environments; that is, the voiceless coda stop may be realised solely by the presence of glottalisation. Figure 13 shows an example of a glottalised item in which no F2 transition to an alveolar locus is evident. The items with glottal replacement were produced by 14 speakers; some individuals produced a single example whereas others produced multiple examples (the maximum produced by any one speaker was six), but none of these speakers produced examples of glottal replacement for all of the glottalised stops they produced. All but one of the examples of glottal replacement were produced by speakers in the younger group. Vowel height did not appear to exhibit any influence on whether glottal replacement occurred – 17 examples occurred in conjunction with low vowels compared to 16 with high vowels – though short vowels did seem to favour their occurrence (short vowels: 22; long vowels: 11).

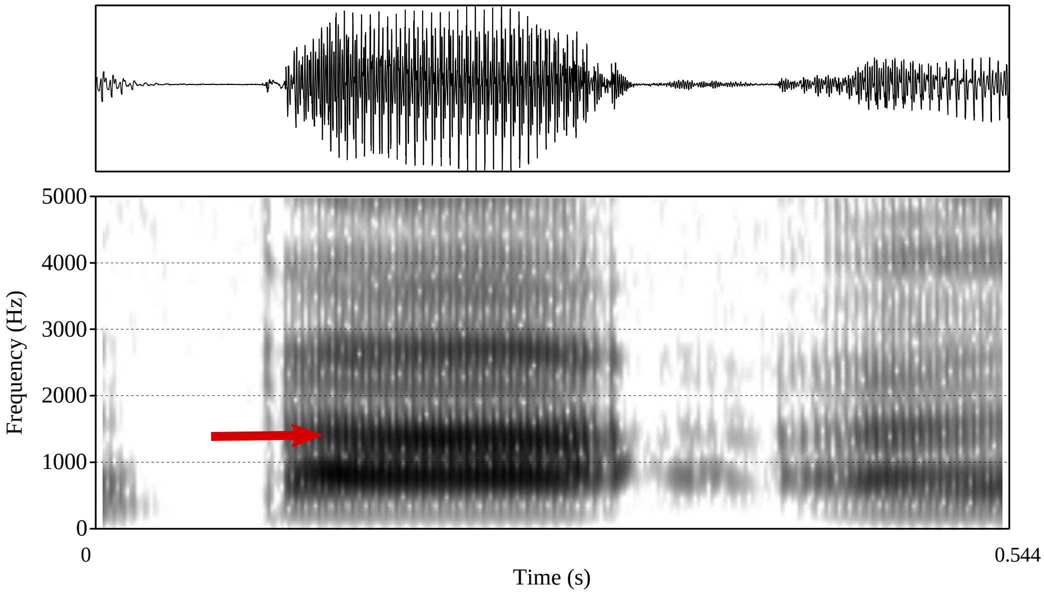

Figure 13. Spectrogram and waveform of the word *bart* containing an unreleased, glottalised coda stop with no visible (rising) F2 transition (approximate location of F2 indicated by arrow).

5.3.5 Discussion

The results presented above are consistent with previous findings of age-based differences in the use of glottalisation, namely, that younger speakers glottalise more than older speakers. This has been found both in previous studies (Penney, Cox, Miles, & Palethorpe, 2018; Penney et al., 2019 [Chapter 2]) and in section 5.2 above. In the analysis here, younger speakers produced more glottalisation than older speakers in each of the environments, though this was particularly evident in the pre-vocalic environments, where older speakers produced generally low rates of glottalisation. In the pre-consonantal environments, glottalisation was employed more frequently by both age groups, particularly on low vowels where there was no significant difference in rates of glottalisation between the age groups. This finding may suggest that glottalisation originally affected pre-consonantal environments, and is now spreading to be used in pre-vocalic environments as well, with younger speakers leading the change. This also supports previous findings that glottalisation is used less on high vowels (Penney, Cox, Miles, & Palethorpe, 2018); although we did not find a significant effect of vowel height on glottalisation in section 5.2 above, we did find a strong trend for the younger speakers to produce more glottalisation on low vowels in medial (pre-vocalic) position. Here we found that rates of glottalisation were generally lower for high vowels, in particular for the older speakers who showed a significant difference between high vowels and low vowels when followed by a stop. Perhaps the effect of vowel height has been reduced in pre-consonantal environments for the younger speakers as glottalisation becomes more generalised.

We predicted that higher levels of glottalisation might be found preceding nasals, based on previous descriptions of AusE that reported pre-nasal glottalisation (Cox & Fletcher, 2017; Cox & Palethorpe, 2007; Ingram, 1989; Tollfree, 2001) and as glottalisation commonly occurs pre-nasally in other varieties of English (Huffman, 2005; Pierrehumbert, 1994, 1995; Roach, 2004; Seyfarth & Garellek, 2015, Tollfree, 1999). We found high rates of glottalisation pre-nasally in both age groups. In addition, for high vowels both age groups produced more glottalisation in the pre-nasal environment than in the other pre-consonantal environments. Furthermore, the older speakers did not show a difference between high and low vowels in the nasal environment, whereas they did show a difference between high and low vowels preceding voiced and voiceless stops (they also showed no difference in the pre-vocalic environments but their rates of glottalisation were low pre-vocalically). This could indicate that glottalisation may have entered the variety in the pre-nasal environment and has now been generalised to other environments. Of course, this suggestion would need further examination through the analysis of historical data.

As in section 5.2, we found high rates of unreleased stops, and these often occurred in conjunction with glottalisation, particularly in the pre-consonantal environments for both age groups. In the pre-vocalic environments, the younger group produced both more glottalisation and more unreleased stops than the older group. We examined the F2 transitions in pre-vocalic released and unreleased glottalised items and found in the majority of cases evidence of an oral gesture for the coda stop in addition to glottalisation (recall that in the pre-consonantal environments an alveolar gesture would be expected anyway due to the following onset consonant). This suggests that glottalisation in AusE generally occurs as glottal reinforcement, that is, glottalisation strengthens the oral stop rather than replacing it, as occurs in some other varieties. Glottal replacement of /t/ occurs in some varieties of British English in intervocalic environments, and has also been reported to occur in AmE (Eddington & Channer, 2010; Eddington & Taylor, 2009; Tollfree, 1999; Wells, 1982). As mentioned above though, this is typically not considered to be the case in AusE (Cox & Fletcher, 2017; Cox & Palethorpe, 2007). Nevertheless, we surprisingly found a small number of items in which no oral gesture appeared to be present, which may hint at more individual variability in the implementation of glottalisation, including the possibility that various types of glottalisation may be employed as sociophonetic markers by some speakers (see e.g. Docherty et al., 2007).

5.4 General discussion

One of the main aims of this study was to determine whether the patterns we have previously found in our research on glottalisation based on single word productions would also be observed in phrase medial position of longer utterances. Glottalisation has previously been found to occur in conjunction with voiceless coda stops in single word utterances in both stressed and unstressed syllables in AusE (Penney, Cox, Miles, & Palethorpe, 2018; Penney et al., 2019 [Chapter 2]), and glottalisation of voiceless coda stops is common in other varieties of English (Docherty & Foulkes, 1999; Gordeeva & Scobbie, 2013; Higginbottom, 1964; Huffman, 2005; Pierrehumbert, 1994, 1995; Roach, 1973, 2004). By comparison, glottalisation of voiced coda stops is rarely observed in English,[13] with the exception of African American English varieties, where this is reported to occur in conjunction with devoicing (Anderson & Nguyen, 2004; Fasold, 1981; Farrington, 2018; Koops & Niedielski, 2009). The results presented in section 5.2 confirm that glottalisation is employed in the context of voiceless coda stops, and that this is the case in both phrase medial and phrase final position. Some examples of glottalisation in the voiced coda context were observed, but rarely in medial position. Although somewhat more glottalisation of voiced coda stops was found in final position compared to medial position, there was nevertheless significantly more glottalisation in the voiceless than in the voiced context. As noted above, it is possible that some of the glottalisation found in final position may be due to phrase final creaky voice. Nevertheless, even in final position there was a clear difference between voiceless and voiced coda contexts. In addition, as AusE listeners have been shown to associate glottalisation with coda voicelessness in perception (Penney, Cox, & Szakay, 2018, submitted [Chapter 3]), we suggest that glottalisation is a strong cue to coda voicelessness which appears to apply more generally, not only in final position or in single word utterances.

[13] Note that other studies have reported some glottalisation in conjunction with voiced stops, although such occurrences are limited to a small number of items which are not comparable to those which occurred in conjunction with voiceless stops, similar to what we have found here and elsewhere (e.g. Foulkes & Docherty, 2007; Gordeeva & Scobbie, 2013; Penney, Cox, Miles, & Palethorpe, 2018).

We have previously found that younger speakers utilise glottalisation more frequently than older speakers and, consistent with previous findings from single word utterances (Penney, Cox, Miles, & Palethorpe, 2018; Penney et al., 2019 [Chapter 2]), the results in sections 5.2 and 5.3 show that younger speakers glottalise voiceless coda stops more than older speakers in phrase medial position in all of the environments examined. These results provide addition evidence that glottalisation is a recent change to AusE. Interestingly, we found that in phrase final position the older speakers glottalised voiceless coda stops at approximately the same level as the younger speakers. This finding differs substantially from the previous results found for single word utterances. As discussed in 5.2.3 above, either: 1) older and younger speakers use glottalisation similarly in phrase final position though younger speakers use more glottalisation phrase medially and in single word utterances; or 2) that some of the glottalisation produced by the older speakers in final position is due to phrase final creaky voice. Previous studies show that rates of glottalisation increase phrase finally (Huffman, 2005, Keyser & Stevens, 2006; Redi & Shattuck-Hufnagel, 2001; Seyfarth & Garellek, 2015); however, creaky voice is also to be expected phrase finally. Separating the effects of creaky voice and coda glottalisation is challenging due to their similar acoustic effects (Garellek & Seyfarth, 2016; Keating, Garellek, & Kreiman, 2015). On the other hand, creaky voice generally extends across more than a single segment (Garellek 2015; Garellek & Seyfarth, 2016). If the higher rates of glottalisation by the older speakers in final position were solely due to creaky voice we should perhaps expect to see an equal increase in glottalisation in the voiced coda context for the older speakers in this position as well. We therefore tentatively suggest that older speakers utilise glottalisation more similarly to younger speakers in phrase final position, but further research is required to identify the extent to which this may be due to phrase final creaky voice.

In addition to making greater use of glottalisation than older speakers in medial position, we also found that younger speakers made less use of preceding vowel duration to cue the coda stop voicing. Penney, Cox, Miles, and Palethorpe (2018) found that younger speakers used the vowel duration cue to coda voicing less robustly than older speakers in single word utterances and suggested that a trading relationship may exist between vowel duration and glottalisation. The results found here provide support for such a trading relation. Our analysis of F0 also revealed no contrast between voiced and voiceless coda contexts in either of the age groups. In an analysis of unstressed syllables Penney et al. (2019 [Chapter 2]) also found no significant difference in F0 across coda voicing contexts alongside high rates of glottalisation. This may suggest that glottalisation is becoming increasingly important in signalling the coda stop voicing contrast alongside a concurrent weakening of other voicing cues. Penney, Cox, Miles, and Palethorpe (2018) also found a relationship between vowel height, vowel duration and glottalisation, such that the vowel duration cue to coda voicing was maximised on high vowels, where less glottalisation was present, and reduced on non-high vowels, which had higher rates of glottalisation. Although we found evidence that vowel duration differences were greater on high vowels, we did not find a clear link between glottalisation and vowel height in section 5.2: we did find a trend for younger speakers to glottalise more on low vowels but no height effect for the older speakers. However, the following environment sampled in section 5.2 was vocalic, and in section 5.3 we found that the older speakers glottalised relatively infrequently in pre-vocalic environments. In section 5.3 we found that the older speakers glottalised less on high vowels when the following environment was a stop. In addition, the older and younger groups glottalised at similar rates in the pre-consonantal environments for low vowels but showed significant differences for the high vowels. Taken together these results support previous claims that glottalisation appears to interact with vowel height, with low vowels favouring glottalisation (Brunner & Żygis, 2011; Hejná & Scanlon, 2015; Malisz et al., 2013; Penney, Cox, Miles, & Palethorpe, 2018).

Our analysis of glottalisation in multiple following environments in section 5.3 may provide an indication about the spread of glottalisation through the variety. Tollfree (2001) raised the possibility that glottalisation may have arisen in a pre-sonorant environment in AusE, before spreading to other environments. As discussed above, the high rates of glottalisation for both age groups for high and low vowels in the pre-nasal environment suggest that this environment may have been the catalyst for the appearance of glottalisation in AusE. Glottalisation preceding nasals is common in other varieties of English (Huffman, 2005; Pierrehumbert 1994, 1995; Roach, 2004; Seyfarth & Garellek, 2015; Tollfree, 1999), possibly due an enhancement strategy to avoid the perceptual confusion that may occur when a voiceless stop borders a nasal (Pierrehumbert, 1995; though see also Huffman, 2001, and Seyfarth & Garellek, 2015, who explain that this explanation cannot account for the occurrence of glottalisation in all environments). It is possible that glottalisation then spread to the other pre-consonantal environments, which for the older speakers show less frequent glottalisation for high vowels. The differences between younger and older speakers in the pre-vocalic environments may then indicate that the change is now spreading further, with younger speakers leading the spread to pre-vocalic environments. For younger speakers low vowels in the unstressed vowel environment patterns with the pre-consonantal environments, but for high vowels this environment shows levels of glottalisation between those found before stressed vowels or pre-consonantally. It should be remembered that glottalisation also occurs phrase finally in AusE, both in single word utterances and at the end of longer utterances, and it is not clear at what point in this possible progression glottalisation would have spread to this environment. Future analysis of archival data may be able to provide further insights into the progression of glottalisation through the variety.

Finally, our analysis of F2 transitions in glottalised items suggests that for the most part glottalisation in AusE can be considered a form of glottal reinforcement: evidence of F2 transitions suggest that an oral closure gesture is retained for the coda stop when glottalisation occurs. Penney et al. (2019 [Chapter 2]) report formant transitions indicative of an oral stop gesture in almost all of their glottalised items in voiceless codas of unstressed syllables. Nonetheless, we found a small number of items in which no F2 transitions were observed, suggesting that a few individuals occasionally produced glottalisation as glottal replacement. Though this only applied to relatively few items in our data, it may be possible that additional variation exists that could index social characteristics. Although glottal replacement and glottal reinforcement may be considered to be related, Docherty et al. (2007) have shown that in Tyneside English the two glottal variants are associated with different social groups in the community. Further research on a wider range of AusE speakers may therefore provide interesting insights into sociophonetic variation in the community.

5.5 Conclusion

In this paper we have shown that glottalisation occurs with voiceless coda stops in phrase medial position as well as in phrase final position, and thereby confirmed that that previous reports of coda glottalisation in AusE cannot (solely) be attributed to phrase final creaky voice. In addition, we have found evidence to support the suggestion that glottalisation is increasingly used by younger speakers to cue coda voicelessness amid a reduction of other cues to coda stop voicing. An analysis of glottalisation in a range of following environments suggests that coda glottalisation occurs most frequently preceding nasals and stops for both older and younger speakers, but that younger speakers also utilise glottalisation in pre-vocalic environments, which may offer some insight into how glottalisation has spread through this variety of English.

5.6 Acknowledgements

This research was supported by a Research Training Scheme scholarship from the Australian Government to the first author. We thank Kimiko Tsukada, Linda Buckley, and Sakib Islam for their assistance with annotation and reliability.

5.7 Appendix A

Table A1. Mean F0 (Hz) in final subsection of vowel in voiced and voiceless coda contexts according to phrase position, age group, and gender.

Gender	Voicing context	Phrase position	Age group	N	F0 (Hz)	SD	SE
Female	Voiced	Final	Older	254	143.22	42.01	2.64
Female	Voiceless	Final	Older	271	145.18	53.68	3.26
Female	Voiced	Final	Younger	386	155.68	75.67	3.85
Female	Voiceless	Final	Younger	410	170.87	85.75	4.24
Female	Voiced	Medial	Older	238	170.13	22.40	1.45
Female	Voiceless	Medial	Older	225	169.79	27.93	1.86
Female	Voiced	Medial	Younger	353	201.62	35.35	1.88
Female	Voiceless	Medial	Younger	330	192.98	52.19	2.87
Male	Voiced	Final	Older	117	90.40	24.57	2.27
Male	Voiceless	Final	Older	121	90.82	26.00	2.36
Male	Voiced	Final	Younger	80	105.42	42.45	4.75
Male	Voiceless	Final	Younger	81	103.62	47.37	5.26
Male	Voiced	Medial	Older	111	108.86	14.74	1.40
Male	Voiceless	Medial	Older	108	110.80	19.13	1.84
Male	Voiced	Medial	Younger	60	105.17	18.83	2.43
Male	Voiceless	Medial	Younger	61	100.70	24.65	3.16

5.8 Appendix B

Details of non-canonical stop realisations in section 5.2

Unreleased stops

14% of the items (n = 462) contained unreleased coda stops, produced by 59 of the 77 participants: (older group: 6%; younger group: 21%). Older males produced slightly more unreleased stops compared to older females (8% vs 5%); 33% of stops produced by the younger males were unreleased, compared to 19% by younger females. The vast majority of unreleased items occurred in the voiceless context. 53% occurred in conjunction with glottalisation, although glottalisation occurred more frequently in conjunction with unreleased stops in the voiceless compared to voiced context (60% vs 17%). Unreleased realisations were common in both phrase positions (medial: 16%; final: 13%).

Spirantisation

6% of the items (n = 182) contained spirantised coda stops. These were produced by 38 of the 77 participants: (older group: 10%; younger group: 2%). The older females spirantised the most, at 11%, with older males producing spirantisation in 8% of items. The younger females spirantised 2% of items but only one spirantised item was produced by the younger males. As was the case for unreleased stops, the majority of spirantised variants were in the voiceless context. 21% of the spirantised stops in the voiceless context occurred in conjunction with glottalisation compared to only one in the voiced context. Spirantised realisations occurred in both phrase positions (medial: 5%; final: 6%).

Preaspiration

Less than 1% of the items (n = 29) contained preaspiration. The preaspirated stops were produced by 19 of the 77 speakers (older group: 1%; younger group: less than 1%). The older females preaspirated 2% of stops, and the younger females preaspirated less than 1%. There was only one preaspirated item produced by an older male and none produced by the younger male speakers. All of the preaspirated stops occurred in the voiceless context: with the majority occurring in final position. 31% of the preaspirated stops were produced in conjunction with glottalisation.

Taps

Less than 1% of the items (n = 15) contained tapped stop realisations. These were produced by 7 of the participants. All but one of the tapped items were produced by the older speakers with the older males producing the majority of the taps. Three tapped items were produced by the older females and one was produced by a younger female. All but one of the tapped items occurred in the voiced context. None of the taps occurred in conjunction with glottalisation. Unsurprisingly, all of the taps occurred in phrase medial position.

Glottal squeaks

Less than 1% of the items (n = 14) contained glottal squeaks. These were produced by 11 of the participants (11 by younger speakers – 10 females, 1 male; 3 by older female speakers). Nine of the squeaks were produced in conjunction with glottalisation with the remainder in conjunction with creaky voice. All of the squeaks occurred in the voiceless context in phrase final position.

Table A2. Non-canonical stop realisations according to age group, gender, phrase position and voicing context. Figures in brackets show number of items produced in conjunction with glottalisation.

Age group	Gender	Phrase position	Voicing context	UNR	SPI	PRE	TAP	SQK
Older	Female	Final	Voiced	7/257 (0)	3/257 (1)	0/257 (0)	0/257 (0)	0/257 (0)
Older	Female	Final	Voiceless	3/275 (1)	60/275 (16)	15/275 (8)	0/275 (0)	3/275 (2)
Older	Male	Final	Voiced	4/117 (1)	6/117 (0)	0/117 (0)	0/117 (0)	0/117 (0)
Older	Male	Final	Voiceless	13/121 (5)	18/121 (7)	1/121 (0)	0/121 (0)	0/121 (0)
Younger	Female	Final	Voiced	12/402 (3)	11/402 (0)	0/402 (0)	0/402 (0)	0/402 (0)
Younger	Female	Final	Voiceless	115/422 (66)	10/422 (4)	6/422 (1)	0/422 (0)	10/422 (6)
Younger	Male	Final	Voiced	22/80 (2)	1/80 (0)	0/80 (0)	0/80 (0)	0/80 (0)
Younger	Male	Final	Voiceless	44/84 (22)	0/84 (0)	0/84 (0)	0/84 (0)	1/84 (1)
Total		**Final**	**Voiced**	**45/856 (6)**	**21/856 (1)**	**0/856 (0)**	**0/856 (0)**	**0/856 (0)**
Total		**Final**	**Voiceless**	**175/902 (94)**	**88/902 (27)**	**22/880 (9)**	**0/880 (0)**	**14/880**
Older	Female	Medial	Voiced	7/238 (1)	4/238 (0)	0/238 (0)	3/238 (0)	0/238 (0)
Older	Female	Medial	Voiceless	36/223 (21)	42/223 (1)	1/223 (0)	0/223 (0)	0/223 (0)
Older	Male	Medial	Voiced	2/112 (0)	1/112 (0)	0/112 (0)	10/112 (0)	0/112 (0)
Older	Male	Medial	Voiceless	16/110 (3)	11/110 (1)	0/110 (0)	1/110 (0)	0/110 (0)
Younger	Female	Medial	Voiced	11/356 (3)	2/356 (0)	0/356 (0)	1/356 (0)	0/356 (0)
Younger	Female	Medial	Voiceless	143/333 (108)	13/333 (3)	6/333 (0)	0/333 (0)	0/333 (0)
Younger	Male	Medial	Voiced	6/60 (2)	0/60 (0)	0/60 (0)	0/60 (0)	0/60 (0)
Younger	Male	Medial	Voiceless	21/61 (8)	0/61 (0)	0/61 (0)	0/61 (0)	0/61 (0)
Total		**Medial**	**Voiced**	**26/766 (6)**	**7/766 (0)**	**0/766 (0)**	**14/766 (0)**	**0/766 (0)**
Total		**Medial**	**Voiceless**	**216/727 (140)**	**66/727 (5)**	**7/727 (0)**	**1/727 (0)**	**0/727 (0)**

5.9 Appendix C

Details of non-canonical stop realisations in section 5.3

Unreleased stops

Overall, 55% of the items (n = 2035) were produced with unreleased coda stops. That is, no release burst was visible for the coda /t/ of the target word. This figure is high due to the inclusion of pre-consonantal environments: In the case of items with a following stop (VLST, VOIST), there was often only evidence of a single (geminate) stop release associated with the following onset stop, and in the case of following nasals (NAS) there was generally no evidence of an oral release and these were likely nasally released. Unreleased stops were produced by 71 of the participants (older group: 45%; younger group: 63%; older females: 43%; older males: 50%; younger females: 64%; younger males: 60%). Unreleased stops occurred frequently in all of the environments for the younger group, though they were less frequent in the pre-vocalic compared to pre-consonantal environments. Unreleased stops were also frequent for the older group in the pre-consonantal environments, but occurred less frequently in the pre-vocalic environments. Glottalisation frequently co-occurred with unreleased stops: older group: 64%; younger group: 84%.

Spirantisation

Spirantisation was present in 9% of the items (n = 314) overall. Spirantised stops were produced by 42 of the participants. The older group produced spirantisation more frequently than the younger group (14% vs 3%). The older females spirantised 16% of items; the older males spirantised 11%, and the younger females spirantised 4%. The younger males did not produce any spirantised realisations. Although spirantised stops were present in all of the environments, these were most frequent in the pre-vocalic environments. Only small numbers of spirantised stops were present pre-consonantally. 29% of the spirantised items produced by the younger group co-occurred with glottalisation, whereas 9% of the spirantised stops produced by the older group co-occurred with glottalisation.

Preaspiration

2% of the items (n = 55) contained preaspiration. Preaspirated stops were produced by 23 participants. Both the younger and older speakers produced preaspiration 2% of items. Both the older and younger females preaspirated in 2% of items, and the older males preaspirated in 1% of items. Younger males did not produce preaspiration. Preaspiration occurred most frequently in the UNSTRV environment, though even in this environment it did not occur frequently. The younger group also produced some preaspirated items in the STRV, VOIST, and VLST environments. The older group produced items in the STRV, VLST, and NAS environments. Three of the preaspirated items produced by the younger speaker were also glottalised. None of the preaspirated items produced by the older group occurred in conjunction with glottalisation.

Taps

Taps were present in 1% of the data (n = 41) and were produced by 16 participants. 1% of items produced by both the younger and older group were tapped. Both the older and younger females produced taps in less than 1% of items, whereas taps were present in 3% of items produced by both older and younger males. Unsurprisingly, all of the examples of taps occurred in the pre-vocalic environments, and all but one item (by an older speaker) were in the UnstrV environment. None of the taps co-occurred with glottalisation.

Squeaks

There were 3 examples of glottal squeaks, each produced by a separate participant, all in younger group: 2 by females and 1 by a male. Two of these were produced in the VlSt environment, and the other was produced in the Nas environment. All 3 squeaks were produced in conjunction with glottalisation.

Table A3. Non-canonical stop realisations according to age group, gender, and environment. Figures in brackets show number of items produced in conjunction with glottalisation.

Age group	Gender	Environment	UNR	SPI	PRE	TAP	SQK
Older	Female	STRV	40/235 (23)	46/235 (2)	1/235 (0)	0/235 (0)	0/235 (0)
Older	Female	UNSTRV	28/243 (23)	111/243 (2)	13/243 (0)	7/243 (0)	0/243 (0)
Older	Female	VLST	148/241 (109)	14/241 (5)	6/241 (0)	0/241 (0)	0/241 (0)
Older	Female	VOIST	139/218 (97)	4/218 (1)	0/218 (0)	0/218 (0)	0/218 (0)
Older	Female	NAS	146/230 (119)	11/230 (4)	3/230 (0)	0/230 (0)	0/230 (0)
Older	Male	STRV	17/117 (4)	11/117 (1)	0/117 (0)	1/117 (0)	0/117 (0)
Older	Male	UNSTRV	11/107 (10)	29/107 (3)	2/107 (0)	15/107 (0)	0/107 (0)
Older	Male	VLST	85/109 (35)	5/109 (3)	0/109 (0)	0/109 (0)	0/109 (0)
Older	Male	VOIST	81/108 (39)	8/108 (0)	0/108 (0)	0/108 (0)	0/108 (0)
Older	Male	NAS	76/105 (34)	7/105 (0)	1/105 (0)	0/105 (0)	0/105 (0)
Younger	Female	STRV	154/350 (116)	13/350 (3)	6/350 (0)	0/350 (0)	0/350 (0)
Younger	Female	UNSTRV	156/330 (148)	26/330 (6)	13/330 (1)	9/330 (0)	0/330 (0)
Younger	Female	VLST	246/355 (217)	10/355 (5)	7/355 (2)	0/355 (0)	1/355 (1)
Younger	Female	VOIST	259/332 (219)	10/332 (3)	3/332 (0)	0/332 (0)	0/332 (0)
Younger	Female	NAS	267/330 (239)	9/330 (3)	0/330 (0)	0/330 (0)	1/330 (1)
Younger	Male	STRV	21/63 (8)	0/63 (0)	0/63 (0)	0/63 (0)	0/63 (0)
Younger	Male	UNSTRV	23/63 (22)	0/63 (0)	0/63 (0)	9/63 (0)	0/63 (0)
Younger	Male	VLST	49/61 (28)	0/61 (0)	0/61 (0)	0/61 (0)	1/61 (1)
Younger	Male	VOIST	45/55 (27)	0/55 (0)	0/55 (0)	0/55 (0)	0/55 (0)
Younger	Male	NAS	44/60 (35)	0/60 (0)	0/60 (0)	0/60 (0)	0/60 (0)
Total			**2035/3712 (1552)**	**314/3712 (41)**	**55/3712 (3)**	**41/3712 (0)**	**3/3712 (3)**

5.10 Appendix D

Table A4. Number of released and unreleased glottalised items in pre-vocalic environments according to gender, age group, and vowel height.

Gender	Age group	Unreleased	Vowel height	N
Female	Older	NO	High	47
Female	Older	YES	High	15
Female	Younger	NO	High	86
Female	Younger	YES	High	112
Female	Older	NO	Low	34
Female	Older	YES	Low	29
Female	Younger	NO	Low	91
Female	Younger	YES	Low	148
Total female				**562**
Male	Older	NO	High	5
Male	Older	YES	High	7
Male	Younger	NO	High	8
Male	Younger	YES	High	11
Male	Older	NO	Low	17
Male	Older	YES	Low	6
Male	Younger	NO	Low	19
Male	Younger	YES	Low	19
Total male				**92**

5.11 References

Aare, K., Lippus, P., Włodarczak, M., & Heldner, M. (2018). Creak in the respiratory cycle. In B. Yegnanarayana, C.C. Sekhar, S. Narayanan, S. Umesh, S.R.M. Prasanna, H.A. Murthy, P. Rao, P. Alku & P.K. Ghosh (Eds.), *Proceedings of the 19th Annual Conference of the International Speech Communication Association (INTERSPECCH 2018*), *Hyderabad, India* (pp. 1408–1412).

Abdelli-Beruh, N.B., Wolk, L., & Slavin, D. (2014). Prevalence of vocal fry in young adult male American English speakers. *Journal of Voice, 28,* 185–190.

Al-Tamimi, J., & Khattab, G. (2015). Acoustic cue weighting in the singleton vs geminate contrast in Lebanese Arabic: The case of fricative consonants. *Journal of the Acoustical Society of America, 138,* 344–360.

Al-Tamimi, J., & Khattab, G. (2018). Acoustic correlates of the voicing contrast in Lebanese Arabic singleton and geminate stops. *Journal of Phonetics, 71,* 306–325.

Anderson, B., & Nguyen, J. (2004). *The social stratification of glottalized variants of /d/ among Detroit African American speakers*. Paper presented at New Ways of Analyzing Variation (NWAV) 33, Ann Arbor.

Andrésen, B. S. (1968). *Pre-glottalization in English Standard Pronunciation*. Oslo: Norwegian Universities Press.

Batliner, A., Burger, S., Johne, B., & Kießling, A. (1993). MÜSLI: A classification scheme for laryngealizations. In D. House & P. Touati (Eds.), *Proceedings of European Speech Communication Association Workshop on Prosody*, *Lund, Sweden* (pp. 176–179).

Blevins, J. (2006). A theoretical synopsis of evolutionary phonology. *Theoretical Linguistics*, *32*, 117–165.

Boersma, P., & Weenink, D. (2018). *Praat: Doing phonetics by computer* (Version 6.0.37). http://www.praat.org/.

Brunner, J., & Żygis. M. (2011). Why do glottal stops and low vowels like each other? In W. S. Lee & E. Zee (eds.), *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS XVII), Hong Kong* (pp. 376–379).

Chong, A. J., & Garellek, M. (2018). Online perception of glottalized coda stops in American English. *Laboratory Phonology, 9*, 1–24.

Cox, F., & Fletcher, J. (2017). *Australian English: Pronunciation and transcription* (2nd ed.). Melbourne: Cambridge University Press.

Cox, F., & Palethorpe, S. (2007). Illustrations of the IPA: Australian English. *Journal of the International Phonetic Association, 37*, 341–350.

Dallaston, K., & Docherty, G. J., (2019). Estimating the prevalence of creaky voice: A fundamental frequency-based approach. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS 2019), Melbourne, Australia* (pp. 532–536).

Davidson, L. (2016). Variability in the implementation of voicing in American English obstruents. *Journal of Phonetics, 54*, 35–50.

Delattre, P. C., Liberman, A. M., & Cooper, F. S. (1955). Acoustic loci and transitional cues for consonants. *Journal of the Acoustical Society of America, 27*, 769–773.

Derrick, D., & Gick, B. (2011). Individual variation in English flaps and taps: A case of categorical phonetics. *Canadian Journal of Linguistics, 56*, 307–319.

Dilley, L., Shattuck-Hufnagel, S., & Ostendorf, M. (1996). Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics, 24*, 423–444.

Docherty, G. J. (1992). *The timing of voicing in British English obstruents*. Berlin: Foris.

Docherty, G. J., & Foulkes, P. (1999). Derby and Newcastle: Instrumental phonetics and variationist studies. In P. Foulkes & G. J. Docherty (Eds.), *Urban voices: Accent studies in the British Isles* (pp. 47–71). London: Arnold.

Docherty, G. J., Foulkes, P., Milroy, J., Milroy, L., & Walshaw, D. (1997). Descriptive adequacy in phonology: A variationist perspective. *Journal of Linguistics, 33*, 275–310.

Eddington, D., & Channer, C. (2010). American English has goʔ a loʔ of glottal stops: Social diffusion and linguistic motivation. *American Speech, 85*, 338–351.

Eddington, D., & Taylor, M. (2009). T-glottalization in American English. *American Speech, 84*, 298–314.

Esling, J. H., Fraser, K. E., & Harris, J. G. (2005). Glottal stop, glottalized resonants, and pharyngeals: A reinterpretation with evidence from a laryngoscopic study of Nuuchahnulth (Nootka). *Journal of Phonetics, 33*, 383–410.

Fabricius, A. (2002). Ongoing change in modern RP: Evidence for the disappearing stigma of t-glottalling. *English World-Wide, 23*, 15–136.

Farrington, C. (2018). Incomplete neutralization in African American English: The case of final consonant voicing. *Language Variation and Change, 30*, 361–383.

Fasold, R. W. (1981). The relation between black and white speech in the South. *American Speech, 56*, 163–189.

Foulkes, P., & Docherty, G. J. (2007). Phonological variation in England. In D. Britain (Ed.), *Language in the British Isles* (pp. 52–74). Cambridge: Cambridge University Press.

Fowler, C. (1992). Vowel duration and closure duration in voiced and unvoiced stops: There are no contrast effects here. *Journal of Phonetics, 20*, 143–165.

Garellek, M. (2015). Perception of glottalization and phrase-final creak. *Journal of the Acoustical Society of America, 137*, 822–831.

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

Garellek, M., & Seyfarth, S. (2016). Acoustic differences between English /t/ glottalization and phrasal creak. In N. Morgan (Ed.), *Proceedings of the 17th Annual Conference of the International Speech Communication Association (INTERSPEECH 2016), San Francisco* (pp. 1136–1140).

Gordeeva, O. B., & Scobbie, J. M. (2013). A phonetically versatile contrast: Pulmonic and glottalic voicelessness in Scottish English obstruents and voice quality. *Journal of the International Phonetic Association, 43*, 249–271.

Haslerud, V. C. D. (1995). *The variable (t) in Sydney adolescent speech: A sociolinguistic study of phonological variation* (Cand. Philol. dissertation, University of Bergen).

Hejná, M. (2015). *Pre-aspiration in Welsh English: A case study of Aberystwyth* (Doctoral dissertation, University of Manchester).

Hejná, M., Palo, P., & Moisik, S. R. (2016). Glottal squeaks in VC sequences. In N. Morgan (Ed.), *Proceedings of 17th Annual Conference of the International Speech Communication Association (INTERSPEECH 2016), San Francisco* (pp. 1136–1140).

Hejná, M., & Scanlon, J. (2015). New laryngeal allophony in Manchester English. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow, Scotland*.

Helgason, P. (2002). *Preaspiration in the Nordic languages: Synchronic and diachronic aspects* (Doctoral dissertation, Stockholm University).

Henton, C., & Bladon, A. (1988). Creak as a sociophonetic marker. In L. M. Hyman & C. N. Li (Eds.), *Language, speech and mind: Studies in honour of Victoria A. Fromkin* (pp. 3–29). London: Routledge.

Herd, W., Jongman, A., & Sereno, J. (2010). An acoustic and perceptual analysis of /t/ and /d/ flaps in American English. *Journal of Phonetics, 38*, 504-516.

Higginbottom, E. (1964). Glottal reinforcement in English. *Transactions of the Philological Society*, *63*, 129–142.

Hirst, D. (2011). The analysis by synthesis of speech melody: From data to models. *Journal of Speech Sciences, 1*, 55–83.

Holmberg, E. B., Hillman, R. E., Perkell, J. S., Guiod, P., & Goldman, S. L. (1995). Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice. *Journal of Speech and Hearing Research, 38,* 1212–1223.

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

Ingram, J. C. L. (1989). Connected speech processes in Australian English. *Australian Journal of Linguistics, 9*, 21–49.

Iseli, M., Shue, Y.-L., & Alwan, A. (2007). Age, sex, and vowel dependencies of acoustic measures related to the voice source. *Journal of the Acoustical Society of America, 121*, 2283–2295.

Jones, M. J., & McDougall, K. (2009). The acoustic character of fricated /t/ in Australian English: A comparison with /s/ and /ʃ/. *Journal of the International Phonetic Association*, 39, 265–289.

Kahn, D. (1976). *Syllable based generalizations in English phonology* (Doctoral dissertation, Massachusetts Institute of Technology).

Kawahara, H., Masuda-Katsuse, I., & de Cheveigné, A. (1999). Restructuring speech representations using a pitch-adaptive time–frequency smoothing and an instantaneous-frequency-based F0 extraction: Possible role of a repetitive structure in sounds. *Speech Communication, 27*, 187–207.

Keating, P., Garellek, M., & Kreiman, J. (2015) Acoustic properties of different kinds of creaky voice. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow, Scotland.*

Keyser, S. J., & Stevens, K. N. (2006). Enhancement and overlap in the speech chain. *Language, 82*, 33–62.

Kisler, T., Reichel, U. D., & Schiel, F. (2017). Multilingual processing of speech via web services. *Computer Speech & Language, 45*, 326–347.

Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America, 59*, 1208–1221.

Koops, C., & Niedielski, N. (2009). */t, d/-glottalization and vowel variation in Houston AAE*. Paper presented at New Ways of Analyzing Variation (NWAV) 38, University of Ottawa, Canada.

Kreiman, J. (1982). Perception of sentence and paragraph boundaries in natural conversation. *Journal of Phonetics, 10*, 163–175.

Kuznetsova, A., Brockhoff, P.B., & Christensen, R.H.B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*, 1–26 (version 3.0.1).

Lenth, R. V. (2016). Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software*, *69*, 1–33 (version 2.25).

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

Luce, P. A., & Charles-Luce, J. (1985). Contextual effects on vowel duration, closure duration, and the consonant/vowel ratio in speech production. *Journal of the Acoustical Society of America, 78*, 1949–1957.

Malisz, Z., Żygis, M., & Pompino-Marschall, B. (2013). Rhythmic structure effects on glottalisation: A study of different speech styles in Polish and German. *Laboratory Phonology*, *4*, 119–158.

Modarresi, G., Sussman, H. M., Lindblom, B., & Burlingame, E. (2005). Locus equation encoding stop place: Revisiting the voicing/VOT issue. *Journal of Phonetics*, *33*, 101–113.

Moisik, S. R., Esling, J. H., Crevier-Buchman, L., Amelot, A., & Halimi, P. (2015). Multimodal imaging of glottal stop and creaky voice: Evaluating the role of epilaryngeal constriction. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow, Scotland*.

Mott, B. (2012). Traditional Cockney and popular London speech. *Dialectologia, 9*, 69–94.

Nance, C., & Stuart-Smith, J. (2013). Pre-aspiration and post-aspiration in Scottish Gaelic stop consonants. *Journal of the International Phonetic Association, 43*, 129–152.

O'Brien, J. (2012). An experimental approach to debuccalization and supplementary gestures (Doctoral dissertation, University of California, Santa Cruz).

Ogden, R. (2001). Turn transition, creak and glottal stop in Finnish talk-in-interaction. *Journal of the International Phonetic Association*, *31*, 139–152.

Penney, J., Cox, F., Miles, K., & Palethorpe, S. (2018). Glottalisation as a cue to coda consonant voicing in Australian English. *Journal of Phonetics, 66*, 161–184.

Penney, J., Cox, F., & Szakay, A. (2018). Weighting of coda voicing cues: Glottalisation and vowel duration. In B. Yegnanarayana, C. C. Sekhar, S. Narayanan, S. Umesh, S. R. M. Prasanna, H. A. Murthy, P. Rao, P. Alku & P. Kumar Ghosh (Eds.), *Proceedings of the 19th Annual Conference of the International Speech Communication Association (INTERSPEECH 2018), Hyderabad, India* (pp. 1422–1426).

Penney, J., Cox, F., & Szakay, A. (2019). Glottalisation of word-final stops in Australian English unstressed syllables. *Journal of the International Phonetic Association.* https://doi.org/10.1017/S0025100319000045.

Penney, J., Cox, F., & Szakay, A. (submitted). *Effects of glottalisation, preceding vowel duration, and coda closure duration on the perception of coda stop voicing*.

Pierrehumbert, J. B. (1994). Knowledge of variation. In K. Beals, J. Denton, R. Knippen, L. Melnar, H. Suzuki, & E. Zeinfeld (Eds.), *Papers from the Parasession on Variation, 30th meeting of the Chicago Linguistic Society Chicago* (pp. 232–256). Chicago: Chicago Linguistic Society.

Pierrehumbert, J. B. (1995). Prosodic effects on glottal allophones. In O. Fujimura & M. Hirano (Eds.), *Vocal fold physiology: Voice quality control* (pp. 39–60). San Diego: Singular.

Powell, M. J. D. (2009). The BOBYQA algorithm for bound constrained optimization without derivatives, *Department of Applied Mathematics and Theoretical Physics, Cambridge England, technical report NA2009/06*.

R Core Team (2017). *R: A language and environment for statistical computing* (Version 3.3.3). https://www.R-project.org/.

Raphael, L. J. (1972). Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English. *Journal of the Acoustical Society of America, 51*, 1296–1303.

Redi, L., & Shattuck-Hufnagel, S. (2001). Variation in the realization of glottalization in normal speakers. *Journal of Phonetics, 29*, 407–429.

Roach, P. J. (1973). Glottalization of English /p/, /t/, /k/ and /ʧ/ – a reexamination. *Journal of the International Phonetic Association*, *3*, 10–21.

Roach, P. (2004). Illustrations of the IPA: British English: Received Pronunciation. *Journal of the International Phonetic Association, 34*, 239–245.

Schleef, E. (2013). Glottal replacement of /t/ in two British capitals: Effects of word frequency and morphological compositionality. *Language Variation and Change, 25*, 201–223.

Selkirk, E. (1982). The syllable. In H. van der Hulst & N. Smith (Eds.), *The structure of phonological representations*, (Part II, pp. 337-383). Dordrecht: Foris.

Seyfarth, S., & Garellek, M. (2015). Coda glottalization in American English. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow, Scotland.*

Seyfarth, S., & Garellek, M. (2018). Plosive voicing acoustics and voice quality in Yerevan Armenian. *Journal of Phonetics, 71*, 425–450.

Shue, Y.-L., Keating, P., Vicenik, C., & Yu, K. (2011). VoiceSauce: A program for voice analysis. In W. S. Lee & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS XVII), Hong Kong* (p. 1846–1849) (version 1.36).

Sjölander, K. (2004). *The snack sound toolkit*. KTH Stockholm, Sweden. http://www.speech.kth.se/snack.

Slifka, J. (2006). Some physiological correlates to regular and irregular phonation at the end of an utterance. *Journal of Voice, 20*, 171–186.

Tollfree, L. (1999). South East London English: Discrete versus continuous modelling of consonantal reduction. In P. Foulkes & G. J. Docherty (Eds.), *Urban voices: Accent studies in the British Isles* (pp. 163–184). London: Arnold.

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

Trudgill, P. (1986). *Dialects in contact*. Oxford: Blackwell.

Trudgill, P. (2001). Received pronunciation: Sociolinguistic aspects. *Studia Anglica Posnaniensia: International Review of English Studies*.

Watson, K. (2007). Illustrations of the IPA: Liverpool English. *Journal of the International Phonetic Association, 37*, 351–360.

Watt, D., & Allen, W. (2003). Illustrations of the IPA: Tyneside English. *Journal of the International Phonetic Association, 33*, 267–271.

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

Wells, J. C. (1997). Whatever happened to Received Pronunciation? In C. Medina Casado, & C. Soto Palomo (Eds.), *II Jornadas de Estudios Ingleses* (pp. 19–28). Jaén: Universidad de Jaén.

Westbury, J. R., & Keating, P.A. (1986) On the naturalness of stop consonant voicing. *Journal of Linguistics, 22*, 145–166.

Wolak, M. E., Fairbairn, D. J., &. Paulsen, Y. R. (2012). Guidelines for estimating repeatability. *Methods in Ecology and Evolution*, *3*, 129–137.

Wolk, L., Abdelli-Beruh, N.-B., & Slavin, D. (2012). Habitual Use of Vocal Fry in Young Adult Female Speakers. *Journal of Voice, 26*, e111–e116.

Yuen, I., Cox, F., & Demuth, K. (2018). Prosodic effects on the planning of inserted /ɹ/ in Australian English. *Journal of Phonetics, 69*, 29–42.

Yuasa, I. P. (2010). Creaky voice: A new feminine voice quality for young urban-oriented upwardly mobile American women. *American Speech*, *85*, 315–337.

Chapter 6: Links between production and perception of glottalisation in individual Australian English speaker/listeners

This chapter is based on the following paper, which is being prepared for submission to Interspeech 2020:

Penney, J., Cox, F., & Szakay, A. (in preparation). *Links between production and perception of glottalisation in individual Australian English speaker/listeners*.

I certify that I was responsible for the development of the concept of this paper, in discussion with my supervisors/co-authors. I took leadership in conducting the research, and was responsible for the construction of the stimuli, all data collection, the analyses, and the writing of all parts of the paper. My co-authors provided advice to improve the research protocol, the analyses, the interpretation of the data, as well as the presentation of the written component.

6.0 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 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 weight glottalisation in perception and how frequently they use glottalisation in production in for either 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.

6.1 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 (Trudgill, 1986; Wells, 1982), later studies have shown that contemporary AusE speakers utilise glottalisation frequently (Penney, Cox, Miles, & Palethorpe, 2018; Penney, Cox, & Szakay, 2019 [Chapter 2]). Glottalisation as an acoustic correlate to coda voicelessness (Penney, Cox, Miles, & Palethorpe, 2018; Penney et al., 2019 [Chapter 2]) is common in many varieties of English (Gordeeva & Scobbie, 2013; Huffman, 2005; Pierrehumbert, 1995; Roach, 1973; Seyfarth & Garellek, 2015). As is expected in the case of a recent change (Eckert, 1988; Labov, 2001), 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 (Penney, Cox, Miles, & Palethorpe, 2018; Penney et al., 2019 [Chapter 2]; Penney, Cox, & Szakay, in prep [Chapter 5]). In an analysis of monosyllabic words, Penney, Cox, Miles, and Palethorpe (2018) found that younger speakers glottalised 71% of items in voiceless coda contexts, whereas older speakers showed lower rates of glottalisation at 36%. Similarly, Penney et al. (2019 [Chapter 2]) 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 (Penney et al., 2019 [Chapter 2]). Penney et al. (in prep [Chapter 5]) 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 (Penney et al., in prep [Chapter 5]).

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. Penney, Cox, and Szakay (2018, submitted [Chapter 3]) 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 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. Ohala, 1993; Pierrehumbert, 2001). 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 (Pierrehumbert, 1995) or British English (Roach, 1973), 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 enter 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 (Ohala, 1993; Pierrehumbert, 2001; Beddor, 2015; Yu, 2013), 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 (Beddor, 2015). 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. Beddor, 2015; Beddor, Coetzee, Styler, McGowan, & Boland, 2018; Coetzee, Beddor, Shedden, Styler, & Wissing, 2018; Yu, 2019). However, other studies have failed to find support for a close alignment of individuals' production and perception repertoires (e.g. Grosvald, 2009; Grosvald & Corina, 2012; Kataoka, 2011; Schertz, Cho, Lotto, & Warner, 2015; Shultz, Francis, & Llanos, 2012; Stevens & Reubold, 2014).

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 (Penney et al., submitted [Chapter3], in prep [Chapter 5]), 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 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.

6.2 Methods

6.2.1 Participants

80 participants originally 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 Penney et al., submitted [Chapter 3], in prep [Chapter 5] 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; female: 22; male: 9) or a younger group (aged 18–36, *n* = 43; female: 36; male: 7). All participants were L1 speakers of AusE, were born in Australia (with the exception of one participant who migrated to Australia as an infant), and were schooled exclusively in Australia.

6.2.2 Production task

Participants were fitted with an AKG C520 headset condenser microphone recorded to a Marantz PMD661 MK II solid-state recorder with 44.1 kHz sampling rate and 16-bit quantisation. They read aloud 396 sentences that were 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 (e.g. *bit*). 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. Additional details and results of the production task are provided in Penney et al. (in prep [Chapter 5]). 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 (Garellek, 2015; Keating & Garellek, 2015) making it difficult to disentangle creak from coda-glottalisation. This left 3542 items with voiceless codas remaining. There was a clear effect of younger speakers producing glottalisation more frequently than the older speakers. The older group produced glottalisation in 44% of the items with voiceless codas, whereas the younger group produced glottalisation in 72% of the items.

6.2.3 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 Penney, Cox, & Szakay, 2018, submitted [Chapter 3]) 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. For each step of each continuum, participants were randomly presented with a non-glottalised and a glottalised item. We here focus 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 Penney, Cox, and Szakay (2018, submitted [Chapter 3]).

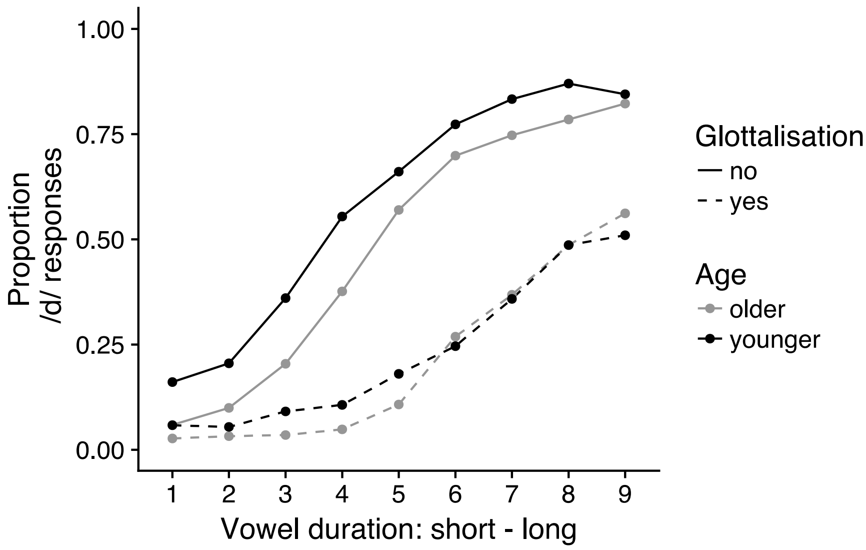

Figure 1. Proportion of items perceived as voiced in glottalised and non-glottalised conditions by older and younger listeners.

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 Penney et al., submitted [Chapter 3] for details).

6.2.4 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 then 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 duration) 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 taken into account. 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 Beddor (2015) and Coetzee et al. (2018), we then plotted each individual participant's production scores (on the x-axis) against his/her perception scores (on the y-axis), to enable us 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.

6.3 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 (Penney, Cox, Miles, & Palethorpe, 2018; Penney et al., 2019 [Chapter 2], in prep [Chapter 5]), 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.[1] Figure 2 shows, consistent with our previous findings, that the younger speakers produce glottalisation more frequently than the older speakers (Penney, Cox, Miles, & Palethorpe, 2018; Penney et al., 2019 [Chapter 2], in prep [Chapter 5]). 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 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 use 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.

[1] Note that the removal of these outliers did not change the overall results.

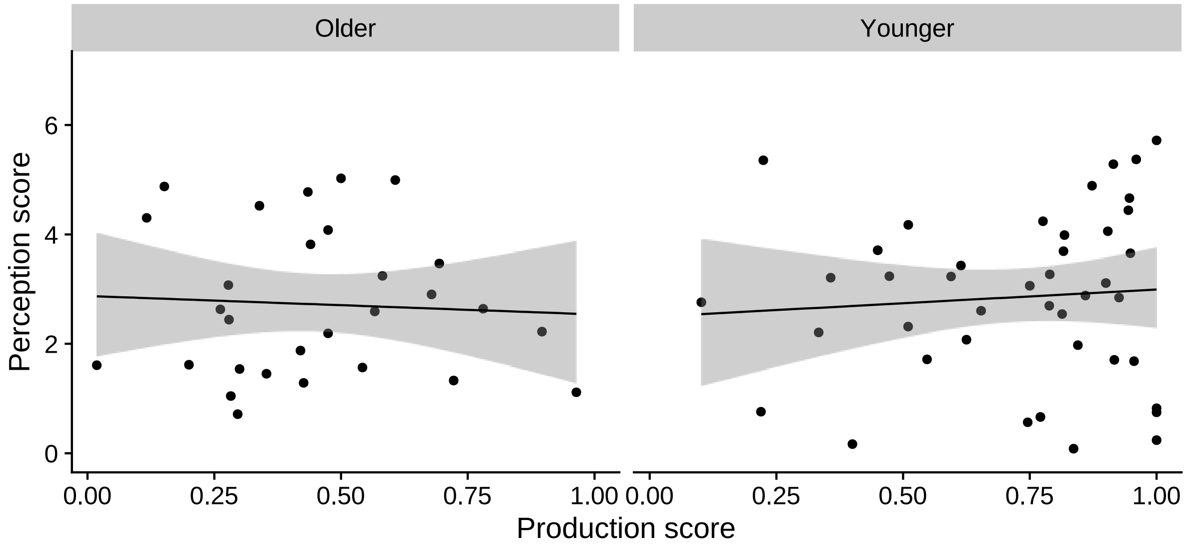

Figure 3: Scatterplot illustrating relationship between production and perception of glottalisation for each individual speaker/listener. Left panel shows older age group; right panel shows younger age group. Solid lines represent linear regression between production and perception scores within each age group. Shaded areas represent 95% confidence intervals.

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 (below 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: $R^2 = 0.006$; $p = 0.62$).[2]

6.4 Discussion

We predicted that if individuals' production and perception repertoires were linked we would see a correlation between their perception scores 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 (Beddor, 2015; Beddor et al., 2018; Coetzee et al., 2018; Yu, 2019), others have failed to find evidence of a correspondence between the two modalities in accordance with this study (Grosvald, 2009; Grosvald & Corina, 2012; Kataoka, 2011; Schertz, Cho, Lotto, & Warner, 2015; Shultz, Francis, & Llanos, 2012; Stevens & Reubold, 2014). We note that the method employed here analysed participants' (final) lexical decisions. Beddor et al.

[2] In order to be confident that the lack of a strong correlation was not due to our method of calculating the perception score we additionally tested a number of alternative methods to determine the extent to which participants weighted glottalisation in perception. For example, we calculated for each participant the increase in the proportion of voiceless responses in the glottalised condition compared to the non-glottalised condition; we identified for each participant the 50% crossover points from a voiceless to a voiced response in each condition using sigmoid curves fitted to the perception data, and 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 generalised linear mixed model on the perception data with participant included as a random factor, which we then 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.

(2018) 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 Beddor (2015), 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 (Beddor, 2015), regardless of whether they produce these themselves, particularly if these are present in the speech of some members of their community (Hay, Warren, & Drager, 2006; Jannedy & Weirich, 2014).

On the other hand, we also found a number of individuals whose behavior appeared to be what Beddor (2015) 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 (Grosvald & Corina, 2012; Stevens & Harrington, 2014). 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 (Kleber, Harrington, & Reubold, 2012).

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 of the age groups. It may be possible that an examination of participants' real time perception of glottalisation may 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.

6.5 Acknowledgements

This work was supported by an Australian Government Research Training program scholarship to the first author. We are grateful to Titia Benders and Jonathan Harrington for their comments, particularly with regard to methodology.

6.6 References

Beddor, P. S. (2015). The relation between language users' perception and production repertoires. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow, Scotland.*

Beddor, P. S., Coetzee, A. W., Styler, W., McGowan K. B., & Boland, J. B. (2018). The timecourse of individuals' perception of coarticulatory information is linked to their production: Implications for sound change. *Language*, *94*, 931–968.

Coetzee, A. W., Beddor, P. S., Shedden, K., Styler, W., & Wissing, D. (2018). Plosive voicing in Afrikaans: Differential cue weighting and tonogenesis. *Journal of Phonetics, 66*, 185–216.

Eckert, P. (1988). Adolescent social structure and the spread of linguistic change. *Language in Society, 17*, 183–207.

Garellek, M. (2015). Perception of glottalization and phrase-final creak. *Journal of the Acoustical Society of America, 137*, 822–831.

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

Grosvald, M. (2009). Interspeaker variation in the extent and perception of long-distance vowel-to-vowel coarticulation, *Journal of Phonetics, 37*, 173–88.

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

Gordeeva, O. B., & Scobbie, J. M. (2013). A phonetically versatile contrast: Pulmonic and glottalic voicelessness in Scottish English obstruents and voice quality. *Journal of the International Phonetic Association, 43*, 249–271.

Hay, J., Warren, P., & Drager, K. (2006). Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics, 34*, 458–484.

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

Jannedy, S., & Weirich, M. (2014). Sound change in an urban setting: Category instability of the palatal fricative in Berlin. *Laboratory Phonology, 5,* 91–122.

Kataoka, R. (2011). *Phonetic and cognitive bases of sound change* (Doctoral dissertation, University of California, Berkeley).

Kleber, F., Harrington, J., & Reubold, U. (2012). The relationship between the perception and production of coarticulation during a sound change in progress. *Language and Speech, 55,* 383–405.

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

Ohala, J. J. (1993). The phonetics of sound change. In C. Jones (Ed.), *Historical linguistics: Problems and perspectives* (pp. 237–278). London: Longman.

Penney, J., Cox, F., Miles, K., & Palethorpe, S. (2018). Glottalisation as a cue to coda consonant voicing in Australian English. *Journal of Phonetics*, *66,* 161–184.

Penney, J., Cox, F., & Szakay, A. (2018). Weighting of coda voicing cues: Glottalisation and vowel duration. In B. Yegnanarayana, C. C. Sekhar, S. Narayanan, S. Umesh, S. R. M. Prasanna, H. A. Murthy, P. Rao, P. Alku & P. Kumar Ghosh (Eds.), *Proceedings of the 19th Annual Conference of the International Speech Communication Association, INTERSPEECH 2018, Hyderabad, India* (pp. 1422–1426).

Penney, J., Cox, F., & Szakay, A. (2019). Glottalisation of word-final stops in Australian English unstressed syllables. *Journal of the International Phonetic Association*. https://doi.org/10.1017/S0025100319000045.

Penney, J., Cox, F., & Szakay, A. (submitted). *Effects of glottalisation, preceding vowel duration, and coda closure duration on the perception of coda stop voicing.*

Penney, J., Cox, F., & Szakay, A. (in prep). *Production of coda glottalisation in Australian English*.

Pierrehumbert, J. B. (1995). Prosodic effects on glottal allophones. In O. Fujimura and M. Hirano (Eds.), *Vocal fold physiology: Voice quality control* (pp. 39–60). San Diego: Singular.

Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In J. Bybee & P. Hopper (Eds.), *Frequency effects and the emergence of linguistic structure* (pp. 137–57). Amsterdam: Benjamins

Roach, P. J. (1973). Glottalization of English /p/, /t/, /k/ and /ʧ/ – a reexamination. *Journal of the International Phonetic Association, 3*, 10–21.

Schertz, J. Cho, T., Lotto, A., & Warner, N. (2015). Individual differences in phonetic cue use in production and perception of a non-native sound contrast. *Journal of Phonetics, 52*,183–204.

Seyfarth, S., & Garellek, M. (2015). Coda glottalization in American English. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow, Scotland.*

Shultz, A. A., Francis, A. L., & Llanos, F. (2012). Differential cue weighting in perception and production of consonant voicing. *Journal of the Acoustical Society of America, 132*, EL95–EL101.

Stevens, M., & Reubold, U. (2014). Pre-aspiration, quantity, and sound change. *Laboratory Phonology, 5*, 455–488.

Stevens, M., & Harrington, J. (2014). The individual and the actuation of sound change. *Loquens, 1*. http://dx.doi.org/10.3989/loquens.2014.003.

Trudgill, P. (1986). *Dialects in contact*. Oxford: Blackwell.

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

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

Yu, A. C. L. (2019). On the nature of the perception-production link: Individual variability in English sibilant-vowel coarticulation. *Laboratory Phonology, 10*, 1–29.

Chapter 7: General discussion and directions for future research

7.1 Summary of results in relation to the research questions

The previous chapters of this thesis have examined the production and perception of coda glottalisation in AusE: Chapter 2 explored glottalisation in production in unstressed syllables; Chapters 3 and 4 examined listener sensitivity to glottalisation, and how glottalisation is perceived; Chapter 5 analysed the production of glottalisation in different phrase positions and phonetic environments; and Chapter 6 considered whether there are links between the production and perception of glottalisation within individuals. We now return to the general research questions of this thesis, outlined in 1.9, to ascertain how these have been addressed through the studies presented in the preceding chapters.

1. Is there evidence to support previous findings that glottalisation occurs in conjunction with voiceless rather than voiced coda stops?

The results presented in this thesis provide strong support for previous findings that glottalisation does occur in conjunction with voiceless coda stops in AusE. The first analysis of Chapter 2, in which glottalisation was examined in unstressed syllables with voiced and voiceless coda stops, demonstrated a clear effect of coda voicing context. Further support is found in the first analysis of Chapter 5, where glottalisation was shown to occur in voiceless coda contexts in both phrase medial and phrase final positions. These results reinforce and extend the findings of Penney, Cox, Miles, and Palethorpe (2018), who showed that glottalisation occurs significantly more in items with voiceless coda stops compared to voiced coda stops. The findings are also consistent with reports of other varieties of English (Docherty & Foulkes, 1999; Gordeeva & Scobbie, 2013; Huffman, 2005; Pierrehumbert, 1994, 1995; Roach, 1973, 2004; Seyfarth & Garellek, 2015; Tollfree, 1999), where glottalisation is also reported to occur with voiceless codas stops, but only rarely with voiced coda stops.

It should be noted that glottalisation was not categorically absent in voiced coda contexts: we found some examples of glottalisation occurring in conjunction with voiced coda stops in both Chapters 2 and 5. Previous studies have also reported examples of glottalisation occurring with voiced stops, though, as with the results found here, the number of glottalised items in voiceless coda contexts far exceed those in the voiced coda context (Foulkes & Docherty, 2007; Gordeeva & Scobbie, 2013; Penney et al., 2018). This suggests that glottalisation is not a categorical acoustic correlate that is physiologically impossible to produce in conjunction with voiced stops, but rather that there is a very strong tendency, within English varieties at least, for this to be associated with voiceless stops.[1]

Previous studies on glottalisation in AusE have concentrated on alveolar coda stops (Haslerud, 1995; Ingram, 1989; Penney et al., 2018; Tollfree, 2001). In Chapter 2 we found evidence for the voicing effect of glottalisation for stops with both bilabial and alveolar place of articulation. Although we could not directly test differences between voiced and voiceless coda stops with velar place of articulation (due to the lack of items with trochaic structure ending with voiced velar stops in English), in the voiceless coda context the results for velar stops reflected those for the stops at the other places of articulation. We therefore suggest that, in final position at least, glottalisation is a correlate not only of /t/, but of coda stop voicelessness more generally (Keyser & Stevens, 2006).

[1] We are grateful to Jim Scobbie and Paul Foulkes for sharing their thoughts on this point.

2. Are there differences between older and younger speakers in their production of glottalisation, and are these indicative of a recent change?

Several analyses showed that younger speakers use more glottalisation than older speakers. In the second analysis of Chapter 2 it was demonstrated that younger speakers produced glottalisation in items with voiceless codas in unstressed syllables of trochaic words more frequently than older speakers. Similarly, the first analysis in Chapter 5 showed that younger speakers used more glottalisation than older speakers in phrase medial position. The second analysis of Chapter 5 showed that younger speakers also glottalised more than older speakers in the environment of a following consonant (although both older and younger both glottalised at high rates when the vowel was low: see below). This analysis also demonstrated that younger speakers produced glottalisation in a wider range of environments than older speakers; while both younger and older speakers produced relatively high rates of glottalisation in pre-consonantal environments (particularly on low vowels), younger speakers also glottalised in pre-vocalic environments, whereas older speakers produced low rates of glottalisation pre-vocalically. Thus, our results demonstrate that younger speakers produce glottalisation at higher rates and in a wider range of environments than older speakers. These results are in accord with those described in Penney et al. (2018), where it was suggested that glottalisation is a recent change to AusE and that the change was being driven by younger speakers.

However, we did find some exceptions to this pattern of age-based difference: younger and older speakers produced similar rates of voiceless coda glottalisation in low vowel rhymes in heterosyllabic pre-consonantal environments, and in phrase final position. The pre-consonantal environment, particularly when the coda stop is in a low vowel rhyme, may provide some clues as to how coda glottalisation has progressed in AusE (see also discussion of RQ 6 below). Glottalisation is known to occur more frequently on low vowels in multiple languages (Brunner & Żygis, 2011; Hejná & Scanlon, 2015; Malisz, Żygis, & Pompino-Marschall, 2013), and Penney et al. (2018) have shown increased glottalisation on non-high vowels in AusE. Thus, it may be that environments which were initially affected by glottalisation would be most observable for low as opposed to high vowels. In Chapter 5 we found that glottalisation was most frequently observed when voiceless coda stops precede a consonant, particularly a nasal consonant, for both age groups. Thus, the lack of an age difference here may be due glottalisation being favoured in this environment.

As stated in Chapter 5, the increase in rates of glottalisation for older speakers in phrase final position may possibly be (at least partially) attributed to phrase final creaky voice, although this would suggest that the period of creaky voice was rather short, affecting only the final portion of the vowel. Perhaps more likely, older speakers favour glottalisation in final position. Indeed, previous reports have shown increases in glottalisation in final position (Huffman, 2005, Seyfarth & Garellek, 2015), possibly in order to support the coda stop voicing contrast. As devoicing (or weak voicing) of voiced stops may occur in final position, additional cues to support the voicing contrast may be necessary (Keyser & Stevens, 2006). We note though that younger speakers did not show an increase in glottalisation in phrase final position.

3. Is there evidence of a trading relationship between preceding vowel duration and glottalisation? Do speakers exhibit a reduction in the production of other cues to coda stop voicing alongside an increase in glottalisation?

Penney et al. (2018) found that in addition to producing more glottalisation, younger speakers also appeared to use voicing-related vowel duration less robustly than older speakers. Furthermore, they found that both age groups showed the most robust use of preceding vowel duration in high vowel contexts, along with the least amount of glottalisation in these contexts. They posited a trading relationship between glottalisation and preceding vowel duration. In Chapter 2 we analysed glottalisation in unstressed syllable environments, as in this environment coda voicing-related durational differences are reported to be reduced (Crystal & House, 1988; Davis & van Summers, 1989; Klatt, 1975). We predicted further reduced utility of the vowel duration cue to coda voicing and increased use of glottalisation in accordance with the hypothesised trading relationship. Although we found that the vowel duration difference across coda voicing contexts was marginally reduced compared to in Penney at al. (2018), nevertheless a significant difference was maintained between schwa preceding voiced versus voiceless codas. In addition, we did not find a significant increase in the frequency of glottalisation (when compared with analogous speakers in the study of Penney et al. (2018). However, the results were nonetheless consistent with a trading relationship: high rates of glottalisation paired with a reduction in the utility of the vowel duration cue.

In Chapter 5 we also analysed the extent to which speakers utilise preceding vowel duration as a cue to coda voicing. In accord with Penney et al. (2018), we found that younger speakers contrasted preceding vowel duration less robustly than older speakers. We calculated a normalised measure of preceding vowel duration to offset other potential effects such as speech rate differences. As discussed above, we also found that younger speakers produced more glottalisation than older speakers, providing further support that increased use of glottalisation is concomitant with a reduction in coda voicing-related vowel duration differences. Consistent with Penney et al. (2018), preceding vowel duration was differentiated most robustly for high vowels whereas both age groups showed more glottalisation on low vowels. This is in accord with previous findings that glottalisation favours low vowels (Brunner & Żygis, 2011; Hejná & Scanlon, 2015; Malisz et al., 2013; Penney et al., 2018). The results provide support for a trading relationship between glottalisation and preceding vowel duration.

Moreover, in both Chapters 2 and 5 we found no significant difference in F0 at the offset of the vowel between voiced and voiceless coda contexts, providing support for a reduction in the utility of other ancillary cues to coda voicing in conjunction with an increase in the use of glottalisation.

4. Are AusE listeners sensitive to glottalisation and do they interpret it as a cue to coda voicelessness? How do listeners weight glottalisation and preceding vowel duration as cues to coda voicing?

The results of Chapter 3 suggest that AusE listeners are indeed sensitive to glottalisation and that they associate it with coda voicelessness. For each of the manipulations tested (vowel duration, coda closure duration, C/V ratio), listeners were more likely to perceive a word with a voiceless coda when the presented item contained glottalisation. Chapter 4 replicated this finding, and also confirmed that listeners are sensitive to the presence of glottalisation in the signal, rather than responding to a shorter (modally-voiced) vowel with a longer coda closure: listeners weight a glottalised vowel differently to how they weight a shorter vowel with a closure period of the same duration as the glottalisation. Though both glottalisation and a shorter vowel/longer closure resulted in more voiceless coda percepts, glottalisation resulted in significantly more voiceless percepts than the shorter vowel/longer closure. These results provide strong evidence that AusE listeners interpret glottalisation as a cue to coda stop voicelessness. This is in line with previous research on AmE showing that glottalisation improves listener identification of /t/ (Garellek, 2011), and that the presence of glottalisation hinders the perception of voiced coda stops (Chong & Garellek, 2018).

Chapters 3 and 4 confirmed that listeners perceive shorter preceding vowels as indicative of voiceless codas and longer preceding vowels as indicative of voiced codas. However, even when vowel duration was at its longest, and thus preceding vowel duration provided a strong cue to a voiced coda, listeners nevertheless showed a reduction of voiced coda percepts when glottalisation was present. This suggests that, while listeners utilise vowel duration as a cue to coda stop voicing, when glottalisation is present in the signal the utility of vowel duration cue is diminished. In Chapters 3 and 4 we found that this reduction in the efficacy of the preceding vowel duration cue in the presence of glottalisation was most pronounced for inherently short vowels, which make less use of coda voicing-related duration differences than inherently long vowels (Cox & Palethorpe, 2011; Penney et al., 2018). In Chapter 3 we also found that glottalisation is weighted more heavily when it occurs on low vowels rather than high vowels. These results provide additional support for a trading relationship between preceding vowel duration and glottalisation as cues to the coda voicing contrast, in perception as well as in production.

5. Are the patterns of glottalisation identified in the literature found independently of phrase final creaky voice?

The data analysed in Penney et al. (2018) (as well as those in Chapter 2 of this thesis) were produced as single words in isolation. Therefore, each single word represents its own intonational phrase, raising the possibility that the glottalisation observed was not due to coda glottalisation but rather to the presence of phrase final creaky voice, which produces acoustic effects similar to those produced by coda glottalisation (see discussion of this point in 1.3 and 5.2.1.3). In order to account for this possibility, in the first analysis of Chapter 5 we examined glottalisation in phrase medial and phrase final positions. As expected, we found evidence of phrase final creaky voice in phrase final position, but not in phrase medial position. The patterns of glottalisation identified in phrase medial position in Chapter 5 can thus be considered independent of possible effects of creaky voice. The phrase medial position results largely replicate the patterns of previous findings (Penney et al., 2018), although the overall rates of glottalisation found in phrase medial position were slightly lower than those found in previous analyses (possibly as these were produced in a pre-vocalic environment: see RQ 6 below): glottalisation was found primarily in voiceless coda contexts rather than voiced coda contexts, and younger speakers produced more glottalisation than older speakers. Thus, this thesis suggests that the reported patterns of glottalisation in AusE are found independently of phrase final creaky voice.

6. Is the occurrence of glottalisation conditioned by phonetic environment?

The frequency with which coda glottalisation occurs is reported to vary according to the following phonetic environment in AmE, BrE, and in AusE (Huffman, 2005; Pierrehumbert, 1994, 1995; Roach, 2004; Seyfarth & Garellek, 2015; Tollfree, 2001). In the second analysis of Chapter 5 we examined glottalisation of voiceless coda stops produced phrase medially in a number of following heterosyllabic environments across a word boundary. Overall, we found that glottalisation was most commonly present when the voiceless coda stop was followed by a heterosyllabic consonant (i.e. voiced stops, voiceless stops, and nasals) as opposed to a following vowel for both older and younger speakers. The results also demonstrated age-based differences related to the height of the preceding vowel (i.e. the vowel preceding the affected coda stop): younger and older speakers produced glottalisation at comparable rates in VC rhymes containing low vowels followed by heterosyllabic consonants, consistent with the cross-linguistic tendency for increased glottalisation on low vowels (Brunner & Żygis, 2011; Hejná & Scanlon, 2015; Malisz et al., 2013). However, older speakers produced significantly lower rates of glottalisation on coda stops in high vowel rhymes followed by heterosyllabic consonants, apart from in the pre-nasal environment, where they produced similar levels of glottalisation across high and low vowel rhymes. Glottalisation is also reported to occur frequently in pre-nasal environments (and pre-consonantal sonorant environments more generally) in many varieties of English (Fabricius, 2002; Huffman, 2005; Pierrehumbert, 1994, 1995; Roach, 2004; Seyfarth & Garellek, 2015; Tollfree, 1999, 2001; Wells, 1982). These results may thus suggest that glottalisation entered AusE in the pre-nasal environment from which it spread to other environments (Tollfree, 2001).

In Chapter 5 we also found that younger speakers showed significantly more coda glottalisation in heterosyllabic pre-vocalic environments. This may indicate that glottalisation in the pre-vocalic environments is a more recent development than the pre-consonantal glottalisation, and that younger speakers, who appear to be more progressive in their use of glottalisation, are leading the change.

7. Are there links between an individual's use of glottalisation in production and in perception? Are speakers who use more glottalisation in production more sensitive to glottalisation in perception?

In Chapter 6 we examined whether individuals' production and perception of glottalisation were aligned. Such an alignment would mean that speakers who glottalised frequently in production would be more sensitive to glottalisation in perception, and speakers who seldom produced glottalisation would show little sensitivity to glottalisation in production (Beddor, 2015). We did not find support for a general alignment of an individual's use of glottalisation in production and sensitivity to glottalisation in perception. On the contrary, we found evidence of a 'mismatch' for many individuals, such that some speakers who glottalised frequently showed little perceptual sensitivity to glottalisation, and others who rarely produced glottalisation nevertheless appeared to weight glottalisation heavily in perception. Thus we found no correlation between how much an individual utilises glottalisation in production and how much that same individual is attuned to glottalisation in perception.

Nonetheless, a small number of individuals did appear to exhibit an alignment of production and perception: these individuals employ glottalisation frequently in production and weight glottalisation heavily in perception. Based on suggestions that only a subset of members of a speech community need to be innovative in both production and perception in order for a change to occur (Grosvald & Corina, 2012; Stevens & Harrington, 2014), we speculated that such individuals may be those who were, and perhaps still are, responsible for driving the spread of glottalisation through AusE.

7.2 Variation in Australian English stop realisation

In addition to addressing the specific research questions on glottalisation, this thesis has also provided a descriptive analysis of the range of variation in stop realisations other than canonical, released stop variants. Apart from glottalised variants, we found evidence of unreleased, preaspirated, and spirantised stops, as well as glottal squeaks. Unreleased stops were the most common non-canonical realisation in both production studies: in Chapter 2 it was shown that in final position younger speakers in particular produced unreleased stops; in Chapter 5 we also found that younger speakers produced more unreleased stops than older speakers in pre-vocalic environments, but in pre-consonantal environments both age groups produced high rates of unreleased stops (note though that this category included geminate realisations and nasally released stops). In both of these chapters we also found that glottalisation commonly occurred in conjunction with unreleased stops, consistent with previous suggestions (Blevins, 2006; Penney et al., 2018; Selkirk, 1982).

We also found a number of examples of spirantisation: in the second analysis in Chapter 2 and in Chapter 5 we found that the older speakers were primarily responsible for the spirantised stops, particularly in environments preceding unstressed vowels. Younger speakers generally produced fewer spirantised items, although we did find a slightly higher number of examples in the young speakers in the first analysis of Chapter 2. Spirantisation has been linked to female speakers and has been suggested to be a possible prestige marker (Jones & McDougall, 2009, Loakes & McDougall, 2010; Tollfree, 2001); our results do not support the gender pattern and suggest age may have an effect on the occurrence of spirantisation.

Glottal squeaks were identified in all of the production analyses, although they were most frequent in the first analysis of Chapter 2. In the second analysis of Chapter 2 and in both analyses of Chapter 5 squeaks were relatively rare. These were almost exclusively produced by younger speakers, primarily by female speakers, and occurred in all cases with either glottalisation or creaky voice (Hejná, Palo, & Moisik, 2016).

We found relatively few examples of preaspiration, though in the second analysis of Chapter 2 we found that older speakers produced more preaspirated variants than younger speakers. We also found very few examples of tapped realisations. Curiously, this was even the case in Chapter 5 in intervocalic environments where the following vowel was unstressed. This may be due to the relatively formal setting of the elicitation and may suggest that different patterns would be visible in spontaneous speech.

This work therefore contributes to our understanding of the variability that exists in the realisations within AusE stops, which has historically been understudied (with some notable exceptions, such as Haslerud, 1995; Horvath, 1985; Ingram, 1989; Tollfree, 2001), along with other recent studies on consonantal variation such as Docherty, Gonzalez, Mitchell, & Foulkes (2016), Loakes, McDougall, Clothier, Hajek, and Fletcher (2018), Penney et al. (2018), and Tait and Tabain (2016).

7.3 Limitations and directions for future research

There are some limitations to this research that should be borne in mind when interpreting the findings. Some of these limitations may provide interesting opportunities for future research on glottalisation and its contribution to the coda stop voicing contrast in AusE.

In Chapter 2 we found evidence that females produced more glottalisation than males. This finding contrasts with the results of Penney et al. (2018), who did not find evidence of a significant gender effect. We speculated that this might be due to differences in rates of glottalisation between males and females in the non-alveolar places of articulation, and may indicate that females are more progressive than males in the spread of glottalisation. For the analyses in Chapter 5, there were insufficient male participants to enable a comparison of production relating to gender. Gender may play a role in the spread of a change, with females often in the lead (Cameron & Coates, 1989; Eckert, 1989), particularly if the change is associated with prestige (or supra-local norms) (Mees, 1987; Mees & Collins, 1999; Milroy, Milroy, Hartley, & Walshaw, 1994), but also when it is below the level of conscious awareness (Labov, 1990). Gender effects have been found in relation to glottalisation in a number of varieties of English (though not always with females producing more glottalisation) (Docherty, Hay, & Walker, 2006; Holmes, 1995; Mees, 1987; Mees & Collins, 1999; Milroy, Milroy, Hartley, & Walshaw, 1994; Redi & Shattuck-Hufnagel, 2001). Therefore, future work could explore variation in the use of glottalisation according to gender.

As discussed above, as we analysed an unstressed trochaic environment in Chapter 2 we were not able to directly test whether the coda voicing effect of glottalisation holds for stops with velar place of articulation, although we showed that this does hold for stops at both alveolar and bilabial places of articulation. Furthermore, the perception experiments in Chapters 3 and 4 and the production tasks in Chapter 5 were all restricted to stops with alveolar place of articulation. Therefore, it remains possible that the results found in those studies may hold only for the alveolar stops and not for the entire English stop series. A simple extension to this work would then be to test the production and perception of glottalisation in stops at the other places of articulation, to complement the work presented here.

Although we have found consistent evidence (both in this thesis and in previous work: Penney et al., 2018) that younger speakers produce more glottalisation than older speakers, a recent change led by younger speakers may not be the only plausible explanation for these data. For example, glottalisation may be an age-graded process, whereby speakers produce glottalisation when they are young but glottalise increasingly less often as they age (e.g. Bailey, 2002). The examination of archival recordings may therefore be useful to provide illumination on this issue: if evidence of glottalisation in archival recordings of young AusE speakers exists this would suggest that glottalisation in AusE is not (such) a recent change, and may indicate age-grading of the phenomenon. Archival data could also provide information about the environments in which glottalisation previously applied (if at all), and thereby contribute to our understanding of the progression of change. Another line of enquiry may be to perform a longitudinal analysis, for example recruiting some of the younger speakers who participated in the studies in this thesis. Rerecording these speakers in a number of years may shed light on whether glottalisation is a change led by younger speakers, or whether it is simply a trait of younger speakers' speech.

All of the speakers in Chapter 2, and the vast majority of the speakers in Chapter 5 (who were the listeners in Chapter 3) were from Sydney. Therefore, it is possible that the results found in this work may not be generalisable to AusE speakers, but may be specific to speakers from Sydney. AusE is well known for its uniformity despite the size of the land in which it is spoken (e.g. Bernard, 1981), though some regional differences have been shown to exist (Bradley, 2004; Cox & Palethorpe, 1998, 2004; Loakes, Hajek, & Fletcher, 2017). A recent study by Loakes et al. (2018) reports that glottalisation is rare in AusE speakers from Warrnambool in regional Victoria, which the authors suggest might be related to differences between urban and regional areas. This may also suggest that further variation exists in the use of glottalisation throughout Australia. As mentioned in 2.4 above, it is also be possible that other factors, such as socio-economic status or non-English heritage language affiliation, may play a role in phonetic variation in AusE (Clothier & Loakes, 2018; Cox & Palethorpe, 2011; Tollfree, 2001). Chapter 5 highlighted that variation exists in the production of glottalisation: we found that individuals produced glottalisation of varying duration, and, in contrast to the majority of speakers who produced glottalisation as glottal reinforcement (i.e. where a supralaryngeal stop gesture is retained), some speakers produced glottalisation as glottal replacement. It may be possible that such variation indexes speaker identity (Docherty, Foulkes, Milroy, Milroy, & Walshaw, 1997), or perhaps the degree to which a speaker is progressive regarding the change. On the other hand, this variation may be simply attributable to individual variation. Other patterns of variation may also exist, for example in relation to the some of the non-canonical stop realisations discussed above. Therefore, in future work it may be fruitful to examine variability in stop production with a closer sociophonetic focus, to explore whether and how such characteristics may be employed to index speaker and/or group identity.

As discussed in 3.4, Chong and Garellek (2018) found that AmE listeners associate glottalisation with coda voicelessness, but they also found that the presence of glottalisation did not lead to faster identification of items with voiceless codas. Although the methods they employed were different from those used here (online perception using eye-tracking vs forced choice decision by key press), in Chapter 3 we found that the presence of glottalisation increased the perception of coda voicelessness, even when combined with extended preceding vowel duration, suggesting that glottalisation is a strong cue to voicelessness in AusE. It would therefore be interesting to employ eye-tracking with AusE listeners to examine how they use glottalisation as it unfolds in the signal. This may also prove useful for examining links between individuals' production and perception of glottalisation (Beddor, Coetzee, Styler, McGowan, & Boland, 2018). Similarly, it may be informative to extend the perception task in Chapter 3 to participants with a variety of English other than AusE, to determine whether the perception of glottalisation is variety specific, or whether English listeners perceive glottalisation similarly.

As mentioned in 1.6, glottalisation has been suggested to be a strategy to enhance voicelessness and/or to prevent coarticulatory voicing from spreading leftward (Keyser & Stevens, 2006; Pierrehumbert, 1994, 1995). In Chapter 3 we found that the presence of glottalisation strengthens listeners' perception of coda stop voicelessness; however, Chong and Garellek (2018) found no evidence that glottalisation led to faster identification of words with voiceless codas for their listeners. Despite this finding, they note that glottalisation may nevertheless be used to enhance the articulation of voicelessness in production (Chong & Garellek, 2018). That is, regardless of whether it aids listeners in the perception of voicelessness, glottalisation may be employed by speakers to enhance voicelessness. Future studies may test this proposition by examining speech in which enhancement of the voicing contrast may be motivated, for example in noisy environments or situations when clarification is sought (e.g. *did you say bet or bed?*). This would allow us to examine whether glottalisation is a strategy speakers use to achieve voiceless stop enhancement.

Finally, the production data in this thesis were analysed with acoustic measures. Future work utilising articulatory methods, such as electroglottography, ultrasound, electromagnetic articulography, and real-time MRI, would complement these analyses and could better inform our understanding of the relationship between, and the timing of, the supralaryngeal and glottal/laryngeal gestures employed in the realisation of coda glottalisation. Additionally, the production data examined in this thesis were produced as reading tasks, recorded under controlled laboratory conditions. This was of course necessary to ensure consistency among the data to enable a thorough phonetic examination and the testing of hypotheses. Nevertheless, the effects of laboratory speech are well known (e.g. Rischel, 1992) and it is possible that alternative patterns of glottalisation would be found in more natural, spontaneous speech, for example in natural conversation or sociolinguistic interviews. Future work on speech unconstrained by experimental procedures will add to our understanding of how glottalisation is used in everyday communication.

To conclude, this thesis provides an empirical examination of glottalisation, how it is produced and perceived, and how it contributes to the phonological voicing contrast in coda stops in AusE. The results suggest that glottalisation is employed to signal coda stop voicelessness, and this may be linked to a reduction in other cues to coda voicing such as preceding vowel duration and F0 differences. The findings support claims that glottalisation is a recent change to AusE, and that accordingly it is produced more frequently by younger rather than older speakers. Furthermore, glottalisation is conditioned by phonetic environment, and occurs most frequently pre-consonantally. Its effects are found independently of phrase final creaky voice. AusE listeners are sensitive to the presence of glottalisation, and interpret it as signalling coda stop voicelessness, even when presented with competing voicing cues. Although in general listeners' sensitivity to glottalisation is not found to be aligned with their use of glottalisation in production, there appear to be some speaker/listeners for whom this is the case, and these individuals may be important for the progression of glottalisation as a change through AusE.

7.3 References

Bailey, G. (2002). Real and apparent time. In J. K. Chambers, P. Trudgill, & N. Schilling-Estes (Eds.), *The handbook of language variation and change* (pp. 312–322). Malden: Blackwell.

Beddor, P. S. (2015). The relation between language users' perception and production repertoires. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow, Scotland.*

Beddor, P. S., Coetzee, A. W., Styler, W., McGowan, K. B., & Boland, J. B. (2018). The timecourse of individuals' perception of coarticulatory information is linked to their production: Implications for sound change. *Language*, *94,* 931–968.

Bernard, J. R. (1981). Australian pronunciation. In A. Delbridge (Ed.), *The Macquarie dictionary* (pp. 18-27). North Ryde: Macquarie.

Blevins, J. (2006). A theoretical synopsis of evolutionary phonology. *Theoretical Linguistics, 32*, 117–165.

Bradley, D. (2004). Regional characteristics of Australian English: Phonology. In E. W. Schneider, K. Burridge, B. Kortmann, R. Mesthrie, & C. Upton (Eds.), *A handbook of varieties of English* (Vol. 1, pp. 645-655). Berlin: Mouton de Gruyter.

Brunner, J., & Żygis, M. (2011). Why do glottal stops and low vowels like each other? In W. Sum Lee & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS XVII), Hong Kong* (pp. 376–379).

Cameron, D., & Coates, J. (1989). Some problems in the sociolinguistic explanation of sex differences. In D. Cameron & J. Coates (Eds.), *Women in their speech communities: New perspectives on language and sex* (pp. 13–26). London: Longman.

Chong, A. J., & Garellek, M. (2018). Online perception of glottalized coda stops in American English. *Laboratory Phonology, 9*, 1–24.

Clothier, J., & Loakes, D. (2018). Coronal stop VOT in Australian English: Lebanese Australians and mainstream Australian English speakers. In J. Epps, J. Wolfe, J.

Smith, & C. Jones (Eds.), *Proceedings of the Seventeenth Australasian International Conference on Speech Science and Technology, Sydney, Australia* (pp. 13–16).

Cox, F., & Palethorpe, S. (1998). Regional variation in the vowels of female adolescents from Sydney. In R. H. Mannell & J. Robert-Ribes (Eds.), *Proceedings of the 5th International Conference on Spoken Language Processing (ICSLP'98), Sydney, Australia* (pp. 2359–2362).

Cox, F., & Palethorpe, S. (2004). The border effect: vowel differences across the NSW - Victorian border. In C. Moskovsky (Ed.), *Proceedings of the 2003 Conference of the Australian Linguistic Society*. Retrieved 8 December 2014 from https://als.asn.au/Conference/Proceedings/ALS2003

Cox, F., & Palethorpe, S. (2011). Timing differences in the VC rhyme of standard Australian English and Lebanese Australian English. In W. Sum Lee & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS XVII), Hong Kong* (pp. 528–531).

Crystal, T. H., & House, A. S. (1988). Segmental durations in connected-speech signals: Syllabic stress. *The Journal of the Acoustical Society of America, 83*, 1574–1978.

Davis, S. & van Summers, W. (1989). Vowel length and closure duration in word-medial VC sequences. *Journal of Phonetics, 17*, 339–353.

Docherty, G. J., & Foulkes, P. (1999). Derby and Newcastle: Instrumental phonetics and variationist studies. In P. Foulkes & G. J. Docherty (Eds.), *Urban voices: Accent studies in the British Isles* (pp. 47–71). London: Arnold.

Docherty, G. J., Hay, J., & Walker, A. (2006). Sociophonetic patterning of phrase-final /t/ in New Zealand English. In P. Warren & C. I. Watson (Eds.), *Proceedings of the 11th Australasian International Conference on Speech Science & Technology, Auckland, New Zealand* (pp. 378–383).

Docherty, G. J., Foulkes, P., Milroy, J., Milroy, L., & Walshaw, D. (1997). Descriptive adequacy in phonology: A variationist perspective. *Journal of Linguistics, 33*, 275–310.

Docherty, G. J., Gonzalez, S., Mitchell, N., & Foulkes, P. (2016). *A study of the realisation of word-medial and word-final pre-pausal (t) in the conversational speech of young people from Perth*. *Paper presented at the 1st Workshop on Sociophonetic Variability in the English Varieties of Australia (SociophonAus), Brisbane, Australia*.

Eckert, P. (1989). The whole woman: Sex and gender differences in variation. *Language Variation and Change, 1*, 245–267.

Fabricius, A. (2002). Ongoing change in modern RP: Evidence for the disappearing stigma of t-glottalling. *English World-Wide, 23*, 15–136.

Foulkes, P., & Docherty, G. J. (2007). Phonological variation in England. In D. Britain (Ed.), *Language in the British Isles* (pp. 52–74). Cambridge: Cambridge University Press.

Garellek, M. (2011). The benefits of vowel laryngealization on the perception of coda stops in English. *UCLA Working Papers in Phonetics, 109*, 31–39.

Gordeeva, O. B., & Scobbie, J. M. (2013). A phonetically versatile contrast: Pulmonic and glottalic voicelessness in Scottish English obstruents and voice quality. *Journal of the International Phonetic Association, 43*, 249–271.

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

Haslerud, V. C. D. (1995). *The variable (t) in Sydney adolescent speech: A sociolinguistic study of phonological variation*. (Cand. Philol. dissertation, University of Bergen).

Hejná, M., & Scanlon, J. (2015). New laryngeal allophony in Manchester English. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow, Scotland*.

Hejná, M., Palo, P., & Moisik, S. R. (2016). Glottal squeaks in VC sequences. In N. Morgan (Ed.), *Proceedings of 17th Annual Conference of the International Speech Communication Association (INTERSPEECH 2016), San Francisco* (pp. 1136–1140).

Holmes, J. (1995). Glottal stops in New Zealand English: An analysis of variants of word-final /t/. *Linguistics, 33*, 433–463.

Horvath, B. M. (1985). *Variation in Australian English: The sociolects of Sydney*. Cambridge: Cambridge University Press.

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

Ingram, J. C. L. (1989). Connected speech processes in Australian English. *Australian Journal of Linguistics, 9*, 21–49.

Jones, M. J., & McDougall, K. (2009). The acoustic character of fricated /t/ in Australian English: A comparison with /s/ and /ʃ/. *Journal of the International Phonetic Association*, 39, 265–289.

Keyser, S. J., & Stevens, K. N. (2006). Enhancement and overlap in the speech chain. *Language, 82*, 33–62.

Klatt, D. H. (1975). Vowel lengthening is syntactically determined in connected discourse. *Journal of Phonetics, 3*, 129–140.

Labov, W. (1990). The intersection of sex and social class in the course of linguistic change. *Language Variation and Change, 2*, 205–254.

Loakes, D., & McDougall, K. (2010). Individual variation in the frication of voiceless plosives in Australian English: A study of twins' speech. *Australian Journal of Linguistics*, *30*, 155–181.

Loakes, D., Hajek, J., & Fletcher, J. (2017). Can you t[æ]ll I'm from M[æ]lbourne? An overview of the DRESS and TRAP vowels before /l/ as a regional accent marker in Australian English. *English World-Wide, 38*, 29–49.

Loakes, D., McDougall, K., Clothier, J., Hajek, J, & Fletcher, J. (2018). Sociophonetic variability of post-vocalic /t/ in Aboriginal and mainstream Australian English. In J. Epps, J. Wolfe, J. Smith, & C. Jones (Eds.), *Proceedings of the Seventeenth Australasian International Conference on Speech Science and Technology, Sydney, Australia* (pp. 5–8).

Malisz, Z., Żygis, M., & Pompino-Marschall, B. (2013). Rhythmic structure effects on glottalisation: A study of different speech styles in Polish and German. *Laboratory Phonology*, *4*, 119–158.

Mees, I. M. (1987). Glottal stop as a prestigious feature in Cardiff English. *English World-Wide, 8*, 25–39.

Mees, I. M., & Collins, B. (1999). Cardiff: A real-time study of glottalisation. In P. Foulkes & G. J. Docherty (Eds.), *Urban voices: Accent studies in the British Isles* (pp. 85–202). London: Arnold.

Milroy, J., Milroy, L., Hartley, S., & Walshaw, D. (1994). Glottal stops and Tyneside glottalisation: Competing patterns of variation and change in British English. *Language Variation and Change, 6*, 327–357.

Penney, J., Cox, F., Miles, K., & Palethorpe, S. (2018). Glottalisation as a cue to coda consonant voicing in Australian English. *Journal of Phonetics, 66*, 161–184.

Pierrehumbert, J. B. (1994). Knowledge of variation. In K. Beals, J. Denton, R. Knippen, L. Melnar, H. Suzuki, & E. Zeinfeld (Eds.), *Papers from the Parasession on Variation, 30th meeting of the Chicago Linguistic Society Chicago* (pp. 232–256). Chicago: Chicago Linguistic Society.

Pierrehumbert, J. B. (1995). Prosodic effects on glottal allophones. In O. Fujimura & M. Hirano (Eds.), *Vocal fold physiology: Voice quality control* (pp. 39–60). San Diego: Singular.

Redi, L., & Shattuck-Hufnagel, S. (2001). Variation in the realization of glottalization in normal speakers. *Journal of Phonetics, 29*, 407–429.

Rischel, J. (1992). Formal linguistics and real speech. *Speech Communication, 11*, 379-392.

Roach, P. J. (1973). Glottalization of English /p/, /t/, /k/ and /ʧ/ – a reexamination. *Journal of the International Phonetic Association*, *3*, 10–21.

Roach, P. (2004). Illustrations of the IPA: British English: Received Pronunciation. *Journal of the International Phonetic Association, 34*, 239–245.

Selkirk, E. (1982). The syllable. In H. van der Hulst & N. Smith (Eds.), *The structure of phonological representations* (Part II, pp. 337–383). Dordrecht: Foris.

Seyfarth, S., & Garellek, M. (2015). Coda glottalization in American English. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow, Scotland.*

Stevens, M., & Harrington, J. (2014). The individual and the actuation of sound change. *Loquens, 1*. http://dx.doi.org/10.3989/loquens.2014.003.

Tait, C., & Tabain, M. (2016). Patterns of gender variation in the speech of primary school-aged children in Australian English: the case of /p t k/. In C. Carignan & M. D. Tyler (Eds.), *Proceedings of the 16th Australasian International Conference on Speech Science and Technology, Sydney, Australia* (pp. 65–68).

Tollfree, L. (1999). South East London English: Discrete versus continuous modelling of consonantal reduction. In P. Foulkes & G. J. Docherty (Eds.), *Urban voices: Accent studies in the British Isles* (pp. 163–184). London: Arnold.

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

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

Appendix: Ethics Approval

Appendix (Ethics approval) of this thesis has been removed as it may contain sensitive/confidential content