Rhoticity and hiatus breaking in Australian English: Associations with community diversity

Andy Gibson, Joshua Penney, Felicity Cox

Centre for Language Sciences, Department of Linguistics, Macquarie University

andy.gibson@mq.edu.au, joshua.penney@mq.edu.au, felicity.cox@mq.edu.au

Abstract

Australian English is considered non-rhotic, however increasing ethnolinguistic diversity may lead some speakers to exhibit partial rhoticity, particularly in areas of high diversity. r-sandhi is a phonological phenomenon that differentiates rhotic and non-rhotic varieties allowing realisations of /ɹ/ to arise in V_1#V_2 hiatus in non-rhotic speech whilst being absent in non-prevocalic contexts. V_1#V_2 hiatus can also be resolved by glottalisation/glottal stop insertion. We present an auditory analysis showing that children and teenagers from diverse areas are more likely to exhibit instances of non-prevocalic-r and are also more likely to use glottalisation to mark word boundaries in potential r-sandhi contexts.

Index Terms: rhoticity, vowel hiatus, linking-r, glottalisation, diversity, multicultural Australian English

1. Introduction

1.1. Multicultural Australian English

Australia is one of the most ethnically diverse countries in the world with Sydney its most multicultural city [1]. The 2016 census found that 42.9% of people in Greater Sydney were born overseas with 66.9% of people having at least one parent born overseas and 38.2% of households using a non-English language in the home [2]. These statistics do not reveal the true scope of diversity across the city which varies greatly from extremely diverse areas like Auburn where a language other than English is spoken in 83.5% of households compared with only 12.2% in Pittwater.

Across two ongoing projects we are exploring the role that ethnolinguistic diversity plays in the speech of Australian English (AusE) speaking young people. Ethnolinguistic diversity has been underexamined in AusE with the vast majority of studies focusing on the mainstream variety. These projects are designed to help fill this gap in our knowledge of variation in present-day AusE and provide a more representative picture of the variety. The Child Speech, Community Diversity (CSCD) project is longitudinally tracking the speech of children from both Sydney and New South Wales more broadly, from preschool through to their second year of school. The Multicultural Australian English (MAE) project studies the speech of teenagers from a range of areas across Sydney.

Here we focus on a feature of AusE that is typically described as one of its defining characteristics: non-rhoticity [3–5]. We examine the incidence of rhoticity according to community diversity in two controlled datasets (children and teenagers) and then explore how community diversity factors into the management of word-boundary phonological strategies.

1.2. Rhoticity and linking-r

Rhotic and non-rhotic varieties of English are differentiated by the presence of non-prevocalic-r in the former and its absence in the latter. Southern Standard British English (SSBE) and Southern Hemisphere varieties of English (such as New Zealand English (NZE) and AusE) are typically described as non-rhotic (but see [6] for NZE), although anecdotal reports suggest rhoticity may be present in the speech of some AusE speakers from diverse communities. An associated phenomenon that differentiates rhotic and non-rhotic dialects is r-sandhi, the insertion of a realisation of /ɹ/ at sites of potential vowel hiatus in non-rhotic varieties. A phonologically conditioned liaison strategy is used whereby /ɹ/ may arise in V_1#V_2 hiatus contexts (e.g., *four eggs, four o'clock*) when V_1 is non high, though /ɹ/ is absent in non-prevocalic environments (e.g., *four dogs*).

In a study of New Zealand English, [7] found that the use of linking-r at word-internal morpheme boundaries is near-categorical. At word boundaries, however, the potential V_1#V_2 hiatus can be resolved with either linking-r or glottalisation/glottal stop. Several studies have shown that the choice between linking-r and glottalisation is prosodically conditioned, primarily by the prominence of V_2 [3–5]. The choice of hiatus breaking strategy often depends on the strength of the prosodic boundary [3–5, 8–10] with glottalisation and linking-r appearing in complementary distribution. When V_2 is stressed, glottalisation is more likely and when V_2 is unstressed linking–r is more likely. [11] proposed the glottal stop to be the optimal boundary marking segment in V_1#V_2 contexts as it is maximally consonantal compared to an approximant which minimises contrast between surrounding vowels. This echoes findings from other studies (e.g., [12–14]) that show increased use of glottalisation when a stressed vowel appears to the right of a boundary.

The use of glottalisation rather than linking-r may also be sociolinguistically conditioned. Studies have found that speakers from some diverse communities in Britain are more likely to use glottalisation than linking-r in hiatus contexts. [15] and [16] showed that amongst speakers from Tower Hamlets in London, boys with a Bangladeshi background used a greater percentage of glottalisation in linking-r contexts than Anglo-background boys and girls from the same area, who had high rates of linking-r. Other studies of hiatus resolution (not specific to the linking-r context, e.g. [17–19]) have also found that speakers from high diversity areas make greater use of glottalisation to resolve hiatus. In a corpus of pop and hip hop songs, [20] found that linking-r occurred at lower rates amongst African American performers, who tended to use glottalisation, while American performers of European descent used high

rates of linking-r. As in other studies, a stressed V_2 favoured the use of glottalisation while an unstressed V_2 favoured linking-r across all social groups.

1.3. Research questions

If some AusE speakers in diverse communities are partially rhotic, as anecdotal evidence suggests, the question arises as to whether and how word boundary marking strategies might also vary across communities. We explore this question through an auditory analysis of speech data collected from children and teenagers, from areas that vary widely in their levels of ethnolinguistic diversity, elicited through a picture naming task designed to sample various phonological contexts including potential environments for both non-prevocalic-r and linking-r.

Our research aims to systematically assess whether rhoticity is present in the speech of the Australia-born children and teenagers in our two studies, and if so, whether it occurs more in areas with greater diversity, and in particular vocalic environments.

We then look at the way potential vowel hiatus at word boundaries is managed by the speakers in this dataset. We expect to replicate the prosodic effect described above: greater use of glottalisation/glottal stops at strong boundaries (i.e., when the vowel on the right edge of the potential hiatus is stressed) and greater use of linking-r at weaker boundaries (i.e., when the vowel on the right of the potential hiatus is unstressed). We will also explore whether the management of potential vowel hiatus at word boundaries varies with respect to community diversity as has been suggested in [15–19].

2. Analysis of rhoticity

2.1. Rhoticity Methods

For both the CSCD and MAE datasets, an impressionistic analysis of all potential non-prevocalic-r tokens was conducted for the presence/absence of rhoticity (e.g., in words like *car, shirt, water)*. All non-final schwa environments and function words were removed. The analysis was conducted in a quiet environment utilising headphones and Praat software [21], with reference made to spectrograms and waveforms where necessary.

2.1.1. CSCD methods

Here we include only data from the first timepoint of our longitudinal project, when most children were in their last year of preschool. The median age was 5;0 (ranging between 4;9 and 6;3, with a total of 57 females and 78 males). All children included in this analysis were born in Australia and live in New South Wales, with a large proportion from Sydney, and a cluster from the Mid North Coast. For this analysis a total of 135 children from a range of areas varying according to their level of ethnolinguistic diversity were included.

Children engaged in a self-recorded picture naming task delivered via the Gorilla online platform [22], framed as a game where the child helped a cartoon alien find its friends and its spaceship. 150 single words and short phrases were elicited but incidental items were often recorded as the children engaged with the task. Self-recorded in their own homes, the recordings come from a wide range of devices of varying quality, with a sample rate of either 44.1kHz or 48kHz.

The 1st author and a research assistant coded the majority of the tokens, with the remainder coded by the 3rd author. In order to check for reliability, 7% of tokens were coded by more than one analyst. While the results from later timepoints are not presented in this paper, our inter-coder reliability rate from the full dataset (three timepoints) was 96%.

2.1.2. MAE methods

The data analysed here were collected via a picture naming task in which 225 single words and short phrases were elicited through presentation of images on a computer monitor. Speakers were recruited from schools in disparate areas of Sydney that differ both in terms of the level of diversity as well as the dominant non-English languages spoken. Data from 117 adolescent speakers (female: n = 65; male: n = 52) aged between 15–17 years are included here. 69 of the speakers were recorded face-to-face in a quiet room in their school using a Zoom H6 recorder with a sample rate of 44.1kHz; 48 speakers were recorded remotely via a supervised video call while at school using a browser-based recorder [23] with a sample rate of 48kHz. These files were subsequently resampled to 44.1kHz.

The third author coded the majority of tokens, with the remainder coded by the second author. 10% of the items were coded by both analysts, with an inter-coder agreement rate of 96%.

2.2. Rhoticity Results

8643 tokens (children: n = 4548; teenagers: n = 4095) of potential non-prevocalic-r were analysed. A total of 694 tokens (8%) of potential occurrences were deemed to be rhotic (7.2% rhoticity in children's data and 8.9% for teenagers). 39% (53/135) of children and 33% (39/117) of teenagers produced at least one token of non-prevocalic-r. 19% of the children (26/135) and 19% of the teenagers (22/117) had a rate of 10% rhoticity or more. In each of the two groups three speakers had a rate of rhoticity of over 80%.

To examine the role of community diversity, we use a postcode-based measure, obtained from census data using Table Builder [2]. Modeling the variation in our results according to the percentage of households in a given postcode area that speak a language other than English, we treat community diversity as a continuous variable. Note that gender differences will not be examined in this paper.

We fit a binomial generalised linear mixed effects regression model predicting the log-odds of non-prevocalic-r, with speaker and word as random intercepts, and three fixed effects:

- diversity (the proportion of households in the speaker's postcode area that speak a language other than English, centred and scaled);
- cohort (children vs. teenagers);
- vocalic context (a six-level factor distinguishing the lexical sets NEAR, SQUARE, NURSE, START, NORTH, and LETTER, i.e., word-final instances of schwa with orthographic <r>, e.g., *water*).

An interaction between diversity and cohort was tested and was not significant. Results show that speakers from more diverse areas were more likely to produce non-prevocalic-r than speakers from more homogeneous areas (Estimate=2.067; $p<0.001$). The vowel context was highly significant ($p<0.001$), with three r-favouring contexts (NEAR, SQUARE and NURSE) all being significantly more likely to exhibit rhoticity than three r-disfavouring contexts (START, NORTH and LETTER). Finally, there was a significant effect of cohort, such that the children were more likely to produce non-prevocalic-r than the teenagers (Estimate=- 1.510; $p=0.004$). The effect of community diversity is plotted in Figure 1, which shows the log-odds of rhoticity,

back-transformed to percentages according to diversity (also back-transformed) along with points for each speaker's rate of rhoticity, coloured according to cohort.

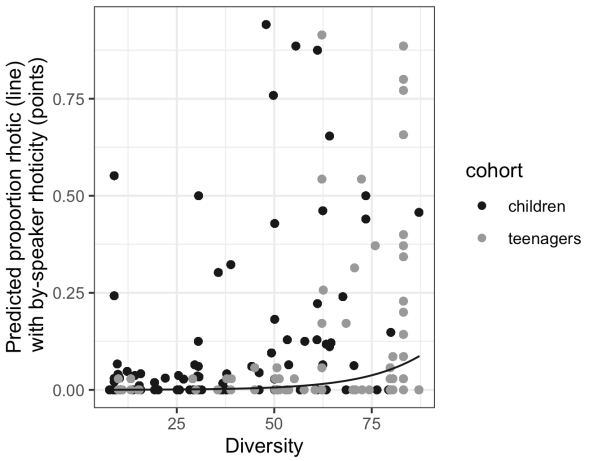

Figure 1: *Predicted rhoticity proportions according to community diversity (line), with raw by-speaker rhoticity proportions for children (black points) and teenagers (grey points).*

2.3. Rhoticity Discussion

The results provide evidence of variable rhoticity in Australia-born speakers. There is a clear effect of community diversity with higher rates of rhoticity in more diverse areas, defined as those areas where there are fewer households where only English is spoken, with children being more likely to produce rhoticity than teenagers, once the community diversity effect is held constant.

The likelihood of non-prevocalic-r also varies according to the vocalic context, with greater rhoticity in NEAR, SQUARE and NURSE than other environments.

It is crucial to note here that these data come from a picture naming task. Items are therefore in citation form in a performative/game-like context which may promote the use of clear speech or perhaps exaggerated productions in some speakers. Future work will assess rhoticity rates in the spontaneous speech data collected for the teenagers.

We have defined community diversity according to the proportion of households in a given area that speak languages other than English. While all participants were born in Australia, English is not the dominant language used in the home for many of these speakers. We acknowledge the effect of individuals' language backgrounds in addition to these area-based measures. However, since language background and our diversity measure are highly correlated, it is not statistically feasible to include both measures within the same analysis. Ultimately, it is important to disentangle the effects of an individual's language background and the ethnolinguistic diversity of the area in which they live. Future analyses will consider the role of language background and whether rhoticity can be attributed to transfer from a heritage language or dialect. The finding that children are more likely than teenagers to produce rhoticity may be due to greater influence of language used in the home on children's speech prior to starting school, after which peer influence gains momentum.

3. Analysis of glottalisation and linking-r at sites of potential vowel hiatus

Having found evidence for some rhoticity in Australia-born teenagers and children, particularly in diverse areas, we are drawn to the question of hiatus management. While both rhotic and non-rhotic speakers may produce /ɹ/ at word boundaries in phrases such as *four apples*, they may do so for different reasons, with the /ɹ/ having special status as part of a hiatus management system for non-rhotic speakers. We now turn to an analysis of the strategies employed by speakers at sites of potential vowel hiatus, exploring first the full range of realisations that occur, and then turning to a systematic analysis of the two primary variants: linking-r and glottalisation.

3.1. Hiatus methods

Both CSCD and MAE employed a counting task in the picture naming game where the speakers were encouraged to produce two-word phrases. For the children, this included four phrases that contain $V_1\#V_2$ hiatus (*four* + *apple/egg/alien/o'clock*), three with a stressed V_2 and one with an unstressed V_2 (*o'clock*). Similar two-word phrases were also elicited from the teenagers, as well as a task designed to elicit the same nouns paired with the possessive determiner *her*. Ten separate contexts were elicited: six with a stressed vowel on the right edge of the hiatus (*ear, eye, arm, eagle, oar, uber*) and 4 with a weak V_2 (*alarm, exam, award, o'clock*). Both sets of data were coded for the presence/absence of linking-r and the presence/absence of glottalisation/glottal stop. Glottalisation and full glottal stops are treated as equivalent for this analysis [24]. An important difference in methods between the cohorts is that for the children's data we included cases of labiovelar or labiodental glides within the linking-r category. In these tokens, the child appears to be enacting 'glide insertion' for the purposes of breaking the hiatus despite not being able to produce an adult-like /ɹ/.

All tokens were coded by at least two analysts, with discrepancies checked and resolved by a third analyst.

3.2. Hiatus results

For the children, a total of 359 tokens were analysed, revealing three main variants: true vowel hiatus (where there was no evidence of either linking-r or glottalisation) occurred in 29.5% of tokens; linking-r occurred in 34%, and glottalisation in 35.4% of tokens.

For the teenagers, a total of 2050 tokens were analysed (*four*: 1142; *her*: 908). As no differences were found according to the strength of the left-hand environment (strong *four* vs. weak *her*), these contexts are combined to focus on differences according to the strength of the right-hand environment. The two primary variants were linking-r and glottalisation. In addition, there were a number of other realisations, including a combination of rhoticity and glottalisation, true vowel hiatus, glide insertion ([w] or [j]), and elision of the right edge vowel (which only occurred in the tokens including *o'clock*). These realisations are excluded from the statistical analysis presented below, which focuses on the binary distinction between glottalisation and linking-r.

For teenagers from ethnolinguistically homogeneous areas (those where less than 20% of households speak a language other than English) there is noticeable systematicity to the resolution of potential vowel hiatus at word boundaries. When the right-edge vowel is stressed, glottalisation is used in 88% of

tokens. When the right-edge vowel is unstressed, linking-r is the preferred variant (72%).

This is not the case in more diverse areas (postcodes where more than 20% of households speak a language other than English). At stressed boundaries, glottalisation is near-categorical (95%). At weak boundaries, however, the use of both glottalisation (46%) and linking-r (40%) occur at similar rates, and this is the only environment where true vowel hiatus occurs in a sizeable number of tokens (13%). It was also in this environment where most of the variants excluded from the analysis (such as elision) occurred.

Figure 2 summarises these raw results, showing the patterning of the three main variants (glottalisation, linking-r and true vowel hiatus) for each cohort (children and teenagers) according to the stress of the right edge vowel and whether the speaker lives in a more homogeneous area or a more diverse area.

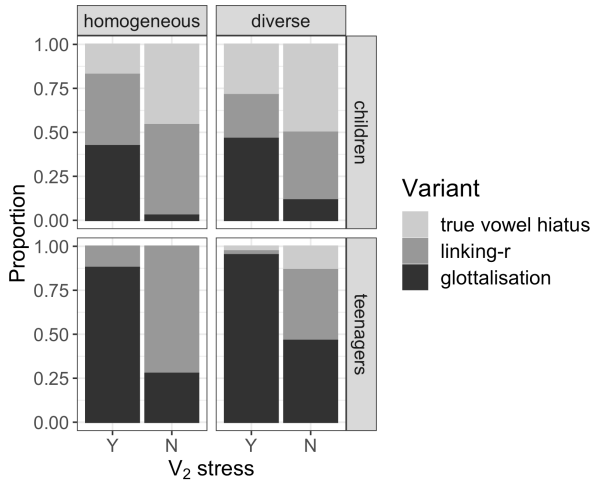

Figure 2: *Raw proportions of the three main variants at sites of potential vowel hiatus according to cohort, community diversity and the stress of the right-edge vowel.*

For the purposes of modeling the data as a binary variable, all tokens of true vowel hiatus were removed from the dataset, along with the small number of other infrequent variants. A binomial linear mixed effects model was fit to the dataset (1810 tokens across 194 speakers), predicting the log-odds of glottalisation (vs. linking-r), with speaker as a random intercept and three fixed effects:

- diversity (the proportion of households in the speaker's postcode area that speak a language other than English, centred and scaled);
- cohort (children vs. teenagers);
- stress of the vowel on the right-edge of the potential hiatus.

The three main effects were significant. Speakers from more diverse areas were more likely to produce glottalisation than speakers from more homogeneous areas (Estimate=1.135; $p<0.001$). Glottalisation was significantly less likely when V_2 was unstressed (Estimate=-4.525; $p<0.001$), and teenagers were more likely to use glottalisation than children (Estimate=3.086; $p<0.001$). The interaction between diversity and stress described above with respect to the raw data was tested but was not significant.

3.3. Hiatus discussion

An analysis of potential environments for linking-r revealed different strategies for managing hiatus between the children and the teenagers. While true vowel hiatus was rare for the teenagers, it was common for children to use neither linking-r nor glottalisation at the word boundary. Linking-r and glottalisation were the primary variants for the teenagers, and across the whole dataset the choice between these variants was strongly motivated by the prosodic context consistent with previous studies [e.g., 3, 4, 5]. When the vowel on the right edge of the potential hiatus is stressed, glottalisation is preferred. Linking-r mainly occurs when V_2 is unstressed.

Additionally, we find again a significant effect of community diversity, with those in more diverse areas more likely to use glottalisation to resolve vowel hiatus than those in more homogeneous areas. While the interaction of stress with community diversity was not significant, teenagers from homogeneous areas seemed to be particularly systematic in their use of glottalisation at strong boundaries and linking-r at weak boundaries.

Finally, teenagers used more glottalisation than children. Given the high rate of true vowel hiatus for children, this difference may due to developmental factors. The three variants observed here can be viewed as falling on a scale from true vowel hiatus ($V_1\#V_2$) – the absence of a boundary, to linking-r – a medium-strength boundary, to glottalisation – a strongly marked boundary. The children may not yet have learned the norms for breaking vowel hiatus at word boundaries. Ongoing work will assess whether these children develop in the direction of the teenagers in their first years of schooling.

4. General discussion

This analysis has assessed the role of community diversity on the realisation of two distinct but related phenomena. We found evidence that partial rhoticity is used by Australia-born children and teenagers, particularly in areas with greater ethnolinguistic diversity. At sites of potential vowel hiatus, linking-r is traditionally thought to be the primary hiatus-breaking device in non-rhotic dialects [3]. Consistent with [3–5], we find that for $V_1\#V_2$ contexts, this is much more likely when V_2 is a weak syllable, and glottalisation is strongly favoured when V_2 is stressed.

We might expect that the higher incidence of rhotic speakers in more diverse areas would lead to greater use of linking-r, however the present results suggest that speakers in more diverse areas use glottalisation to mark word boundaries to a greater extent than those in homogeneous areas, where linking-r and glottalisation appear to be in complementary distribution, governed by the strength of the boundary. It may be the case that strengthening the word boundary may offer communicative enhancement in areas where speech patterns in the ambient language of the community are extremely variable. This suggestion requires empirical support.

Future analyses will examine the acoustic characteristics of the rhotic consonants to provide a more nuanced account of variation in this highly diverse sample of speakers.

Much sociophonetic work describing AusE has been focused on monolingual speakers of Anglo-Celtic heritage from relatively homogeneous communities. By systematically studying speakers from areas with greater ethnolinguistic diversity, we are able to trace the development of variation and change in AusE that reflects the increasing diversity of its speakers.

5. References

[1] Australian Bureau of Statistics: 'Cultural Diversity: Census of Population and Housing: Australia Revealed, 2016.', 2017

[2] Australian Bureau of Statistics: 'Census 2016', 2016

[3] Cox, F., Palethorpe, S., Buckley, L., and Bentink, S.: 'Hiatus resolution and linking 'r' in Australian English', Journal of the International Phonetic Association, 2014, 44, (2), pp. 155–178

[4] Yuen, I., Cox, F., and Demuth, K.: 'Planning of Hiatus-Breaking Inserted /ɹ/ in the Speech of Australian English–Speaking Children', Journal of Speech, Language, and Hearing Research, 2017, 60, (4), pp. 826–835

[5] Yuen, I., Cox, F., and Demuth, K.: 'Prosodic effects on the planning of inserted /ɹ/ in Australian English', Journal of Phonetics, 2018, 69, pp. 29–42

[6] Marsden, S.: 'Are New Zealanders "rhotic"?', English World-Wide, 2017, 38, (3), pp. 275–304

[7] Hay, J., Drager, K., and Gibson, A.: 'Hearing r-sandhi: The role of past experience', Language, 2018, 94, (2), pp. 360–404

[8] Foulkes, P.: 'Rule inversion in a British English dialect - a sociolinguistic investigation of [r]-sandhi in Newcastle upon Tyne', University of Pennsylvania Working Papers in Linguistics: A Selection of Papers from NWAV 25, 1997, 4, (1), pp. 259–270

[9] Mompeán, J. A., and Gómez, F. A.: 'Hiatus Resolution Strategies in Non-rhotic English: The Case of /r/-Liaison', in Lee, W.S. and Zee, E. (Eds.), Proceedings of 17th International Congress of Phonetic Sciences, Hong Kong 2011, pp. 1414–1417

[10] Mompeán, J. A.: '/r/-sandhi in the speech of Queen Elizabeth II', Journal of the International Phonetic Association, 2021, pp. 1–32

[11] Uffmann, C.: 'Intrusive [r] and optimal epenthetic consonants', Language Sciences, 2007, 29, (2), pp. 451–476

[12] Dilley, L., Shattuck-Hufnagel, S., and Ostendorf, M.: 'Glottalization of word-initial vowels as a function of prosodic structure', Journal of Phonetics, 1996, 24, (4), pp. 423–444

[13] Garellek, M.: 'Glottal stops before word-initial vowels in American English: distribution and acoustic characteristics', UCLA Working Papers in Phonetics, 2012, 110, (1), pp. 1–23

[14] Garellek, M.: 'Voice quality strengthening and glottalization', Journal of Phonetics, 2014, 45, pp. 106–113

[15] Britain, D., and Fox, S.: 'The Regularisation of the Hiatus Resolution System in British English. A Contact-Induced 'Vernacular Universal'?', in Filppula, M., Klemola, J., and Paulasto, H. (Eds.): 'Vernacular Universals and Language Contacts: Evidence from Varieties of English and Beyond.' (Routledge, 2008), pp. 177–205.

[16] Fox, S.: 'The New Cockney: New Ethnicities and Adolescent Speech in the Traditional East End of London' (Palgrave Macmillan UK, 2015)

[17] Cheshire, J., Kerswill, P., Fox, S., and Torgersen, E.: 'Contact, the feature pool and the speech community: The emergence of Multicultural London English', Journal of Sociolinguistics, 2011, 15, (2), pp. 151–196

[18] Meyerhoff, M., Birchfield, A., Ballard, E., Charters, H., and Watson, C.: 'Definite change taking place: Determiner realization in multiethnic communities in New Zealand', University of Pennsylvania Working Papers in Linguistics, 2020, 25, (2), pp. 71–78

[19] Cox, F., Penney, J., and Palethorpe, S.: 'Fifty years of change to prevocalic definite article allomorphy in Australian English', Journal of the International Phonetic Association, 2022, pp. 1–31

[20] Gibson, A.: 'Sociophonetics of Popular Music: Insights from Corpus Analysis and Speech Perception Experiments', PhD dissertation, University of Canterbury, 2019. [Online]. Available: https://ir.canterbury.ac.nz/handle/10092/17892

[21] Boersma, P. and Weenink, D.: 'Praat: doing phonetics by computer', Version 6.2.19, retrieved 21 September 2022 from http://www.praat.org

[22] Anwyl-Irvine, A.L., Massonnié, J., Flitton, A., Kirkham, N., and Evershed, J. K.: 'Gorilla in our midst: An online behavioral experiment builder', Behavior Research Methods, 2020, 52, (1), pp. 388–407

[23] German Research Centre for Artificial Intelligence (DFKI): 'speech-to-flac', retrieved 21 September 2022 from https://github.com/mmig/speech-to-flac

[24] Penney, J., Cox, F., and Szakay, A.: 'Effects of glottalisation, preceding vowel duration, and coda closure duration on the perception of coda stop voicing', Phonetica, 2021, 78, (1), pp. 29–63