Interspeech 2025
17-21 August 2025, Rotterdam, The Netherlands

Variability in intervocalic /t/ and community diversity in Australian English

Hannah White[1], Joshua Penney[1], Felicity Cox[1]

[1] Department of Linguistics, Macquarie University, Australia
hannah.white@mq.edu.au, joshua.penney@mq.edu.au, felicity.cox@mq.edu.au

Abstract

The voiceless alveolar stop /t/ exhibits considerable variation in English. Realisations of /t/ vary depending phonetic context and social factors such as gender, age and socioeconomic status. Generally, studies on Australian English have focused on the "mainstream" variety, without acknowledging the wide range of linguistic diversity speakers are exposed to in contemporary multicultural Australian society. In the present paper, we explore intervocalic /t/ variation in data collected from 183 speakers as part of the Multicultural Australian English – Voices of Sydney corpus. Results show that, in certain phonetic contexts, exposure to community linguistic diversity can affect intervocalic /t/ realisation, with speakers from more diverse areas showing a preference for a single variant (the tap) compared to those from less diverse areas. We interpret this as an example of simplification that can occur in diverse communities where there is extreme variability in ambient language exposure.

Index Terms: intervocalic /t/, Australian English, community diversity, sound change

1. Introduction

In Australian English (AusE), as in other varieties of English, the voiceless alveolar stop /t/ exhibits considerable phonetic variation [1, 2, 3, 4, 5]. Aspirated /t/ occurs in onset position preceding a stressed vowel, whereas an unaspirated variant occurs in onset clusters and preceding unstressed vowels [6]. Additionally, a wide range of coda /t/ realisations can be identified, many of which are more likely to occur in particular phonetic contexts. For example, coda /t/ can be aspirated (i.e. produced with closure and audible release – often referred to as 'canonical /t/'), preaspirated (produced like canonical /t/ but with high frequency noise at the end of the preceding vowel), fricated/spirantised (produced with an incomplete closure and presenting with frication), affricated (produced with closure followed by an intense and extended period of frication during the closure interval), and unreleased (produced with no audible release) [7]. Coda /t/ may also be glottalised (i.e. replaced or 'reinforced' with a glottal stop or a period of laryngealisation). Glottalised variants often occur before sonorants across a word boundary and occur word internally preceding syllabic /n/ (but not syllabic /l/) [5, 8], whereas unreleased variants are more likely to occur before obstruents [6, 9]. Ingram [9] reports that glottal variants appear more constrained in their environments in AusE compared to some varieties of British English, where they also commonly occur word internally before sonorants other than syllabic /n/ [10].

AusE is also a 'tapping' variety of English. Taps are short ballistic stop-like gestures created as the tongue tip moves rapidly upward to strike (or approximate) the alveolar ridge [11]. A distinction is sometimes made between taps and flaps, particularly in American English, e.g., [12]; however, for the purposes of the present study, we do not make a distinction between these two realisations. Tapping of /t/ often occurs in intervocalic environments both within words and across word boundaries in many varieties of English when the following syllable is unstressed [5, 8, 10, 12, 13, 14]. In addition, /t/ may be tapped when it precedes a syllabic consonant (e.g. *battle* [bæɾl̩], *bottom* [bɔɾm̩]), although as mentioned above glottal stops are thought to surface more commonly preceding syllabic /n/ (e.g. *button* /bɐʔn̩/ or /bɐtˀn̩/) [5, 6].

The various realisations of coda /t/ are also known to be associated with socioindexical variation; in particular, region, gender, age, socioeconomic status (SES), and speaking style have all been reported to be important factors in coda /t/ realisation in mainstream AusE [3, 5, 9, 13, 15, 16]. Tapped and glottal variants have been associated with less formal speech styles and with speakers with lower SES [5, 13], while aspirated and affricated variants have been associated with high SES female speakers [13], and fricated variants have been associated with higher SES [2] and older speakers [5]. Penney et al. [16] found that younger speakers (aged 18–26 years) produced more glottalised variants than older speakers (aged over 50 years). In one of the few studies of /t/ to examine speakers other than the mainstream, Loakes et al. [3] compared speakers of Aboriginal AusE and Mainstream AusE from two regions of Victoria, Australia, showing variation in the frequency of canonical variants according to region. Differences were also found by dialect with speakers of mainstream AusE producing more canonical, fricated and affricated variants compared to speakers of Aboriginal AusE who produced more ejectives, glottal stops and taps [3].

Most studies of AusE coda /t/ variation have only considered the mainstream variety. Research has established that exposure to and experience with different languages/accents can impact speech processing and perception, e.g., [17, 18, 19], and studies have shown that variation across a range of phonetic and phonological features may vary according to the level of linguistic diversity in a speaker's community [20, 21, 22, 23, 24]. For example, in a study on the realisation of the definite article '*the*', Cox et al. [21] found differences between non-mainstream AusE speakers of Lebanese heritage and mainstream AusE speakers in terms of the use of allomorphic variation: while both groups produced the definite article as [ðə] before consonants, the non-mainstream group also produced this allomorph in prevocalic environments rather than [ðiː] to a greater extent than mainstream AusE speakers. Cox et al. [21] suggest this provides some evidence for a process of simplification described by [25] whereby varieties spoken in communities with great linguistic diversity may tend towards grammatical simplification because

10.21437/Interspeech.2025-664

complex patterns are less robustly realised in such highly diverse environments due to extensive linguistic variability. In non-diverse communities, more robust exposure to complex but consistent patterns occurs through greater linguistic homogeneity. This process of apparent simplification was also observed by Clements [20] who found that speakers from more culturally and linguistically diverse areas of Sydney showed a reduction in TRAP vowel durational complexity in certain lexical and phonological conditions compared to speakers from more homogeneous areas. Similar scenarios have been described in other multicultural English varieties (e.g. in London [26]).

In the present paper, we focus on variation of intervocalic /t/ in words produced by speakers who vary according to the levels of the community linguistic diversity they are exposed to. Understanding how exposure to diversity impacts speech production helps us understand the forces that shape sound systems and provides more nuance to our understanding of variation, both in AusE and in language more generally.

2. Methods

2.1. Data

The data included in this study have been drawn from 183 (94 female; 89 male) adolescent speakers collected as part of the Multicultural Australian English – Voices of Sydney (MAE-VoiS) project who had attended primary and secondary schooling exclusively in Australia [27]. Data were collected in schools from areas of Sydney that vary according to their levels of ethnic/linguistic diversity. Although Sydney is very multicultural, this is not reflected equally across all areas of the city. For example, in some of the areas selected for the MAE-VoiS corpus more than 80% of the population speak only English at home, whereas at the other end of the spectrum there are areas where only 21.9% of residents speak only English at home [28]. In the MAE-VoiS corpus 68.3% of the participants spoke a language other than English (LOTE) in addition to English or lived in a household where a LOTE was spoken. Thirty-seven different LOTEs are represented in the corpus.

Speakers participated in a picture naming task and a conversation in pairs, guided by a research assistant. One hundred and thirty-five participants were recorded in person using a RODE HS2 headset microphone and a Zoom H6 portable audio recorder with a 44.1kHz sample rate and 16-bit resolution. In addition, 48 participants were recorded remotely during COVID-19 lockdowns through an online digital recorder (https://mmig.github.io/speech-to-flac/) with a 44.1kHz or 48kHz sample rate using a local laptop while engaging in a video conference call with a research assistant. For the present study we examine only data elicited in the picture-naming task where participants produced single word/short phrase items. Ten items within this task contained environments where phonemically intervocalic /t/ between a preceding stressed vowel and a following unstressed syllable was expected (i.e., intervocalic /t/ in within-foot trochees). Eight items (all of which are relatively high in frequency) are analysed here and were categorized according to three different contexts: items in which /t/ is produced before an unstressed vowel (*water*, *city*, *tomato*, *potato*), items in which a syllabic /l/ is likely to be produced following /t/ (*bottle*, *turtle*), and items in which a syllabic /n/ is likely to be produced following /t/ (*button*, *kitten*). The two additional items not included in the analysis were the polysyllabic words *caterpillar* and *alligator*; these were excluded since the /t/ in each of these occurs in feet that differ in their strength (strong vs weak) and therefore cannot be considered structurally equivalent. Although most speakers produced all eight items, four speakers produced only seven, resulting in a total number of 1460 tokens for analysis.

2.2. Analysis

All items were labelled for /t/ realisation using a combination of auditory and visual spectrographic methods [7, 29, 30]. The following realisations were identified:

- tap – identified by short closure duration and burst release or alternatively by intensity change [29] – by far the most frequent at 70% of items,
- canonical – identified by the presence of a long closure phase followed by a release and aspiration – 17% of items,
- fricated – identified by the presence of high frequency energy throughout the closure phase – 6% of items,
- preaspirated – identified by the presence of high frequency noise at the end of the preceding vowel, followed by a stop closure and release burst – 4% of items,
- glottalised – identified by the presence of irregular or low frequency glottal pulses and/or full glottal closure [16] –3% of items. Here we do not distinguish between glottal stop (glottal replacement of the /t/) and glottal reinforcement (laryngealisation along with oral stop).

Census data provided the proportion of households where a LOTE was spoken in each postcode [28]. From this data, and following [39], a measure of community diversity was created, with speakers who lived in areas where LOTEs are spoken in less than 25% of households categorised as "non-diverse", those where LOTEs are spoken in 25-65% of households categorised as "diverse", and those where LOTEs are spoken in more than 65% of households categorised as "super-diverse". It should be noted that the diverse and super-diverse areas are not ethnic enclaves but heterogeneously diverse communities [27]. Linguistic diversity is highly correlated with SES in the present dataset. For this reason, it is not statistically viable to examine both SES and linguistic diversity in a single model, Previous work (using these data) has suggested that linguistic diversity is a better predictor of linguistic variation [40]; therefore, we do not consider SES in the present analysis.

Due to the high proportion of taps and small numbers of other variants in this data, we ran a generalised linear mixed effects regression (GLMER) model in R to analyse the probability of the occurrence of a tap versus any other /t/ variant [31, 32]. A three-way interaction between speaker gender (female, male), diversity group (non-diverse, diverse, super-diverse) and context (unstressed vowel, syllabic /l/, syllabic /n/) was included in the initial model. This was not significant, and a stepwise reduction approach was taken, removing non-significant interactions, followed by non-significant fixed effects in order of highest *p*-values [33]. Each subsequent model was compared to the previous using ANOVAs to ensure that the reduction process did not significantly affect the model.

The final, best-fitting model included an interaction between diversity group and context, with random intercepts for

speaker and word. Gender was not a significant predictor of the presence of taps.

The following sections present the results of this model, as well as a descriptive discussion of the distribution of all variants.

3. Results

3.1 Taps vs other variants

The output of the model is shown in Table 1. It suggests a significant decrease in the probability of tapping for the non-diverse group in the pre-syllabic /n/ context when compared to the reference levels of super-diverse and following unstressed vowel. This relationship between diversity group and context is shown in Figure 1 [34].

Table 1: *Output of GLMER model showing effect of diversity group and context on presence of tap. Reference levels: diversity group (dgrp) = super-diverse; context (ctx) = unstressed vowel.*

	Est.	Std. Error	*z*-value	*p*-value
(Int.)	2.90	0.50	5.78	<0.001
dgrpDiv	-0.14	0.51	-0.28	0.782
dgrpNonDiv	-0.91	0.59	-1.54	0.123
ctxL	-0.30	0.68	-0.44	0.663
ctxN	-2.71	0.68	-3.98	<0.001
dgrpDiv:ctxL	-0.47	0.47	-1.01	0.315
dgrpNonDiv:ctxL	0.16	0.56	0.29	0.774
dgrpDiv:ctxN	-0.85	0.47	-1.79	0.074
dgrpNonDiv:ctxN	-1.83	0.63	-2.89	0.004

Post hoc pairwise comparisons were run using the emmeans package [35] with *p*-values adjusted using the Tukey method for multiple comparisons. For /t/ occurring in the unstressed vowel or syllabic /l/ context, the probability of tapping was high and did not differ significantly between any of the diversity groups (all $p>0.2$). However, in contexts where /t/ occurred before a possible syllabic /n/, speakers from non-diverse areas were significantly less likely to produce taps compared to those from diverse ($p=0.026$) and super-diverse ($p<0.001$) areas. There was no significant difference in tap likelihood between diverse and super-diverse speakers for this context ($p=0.133$).

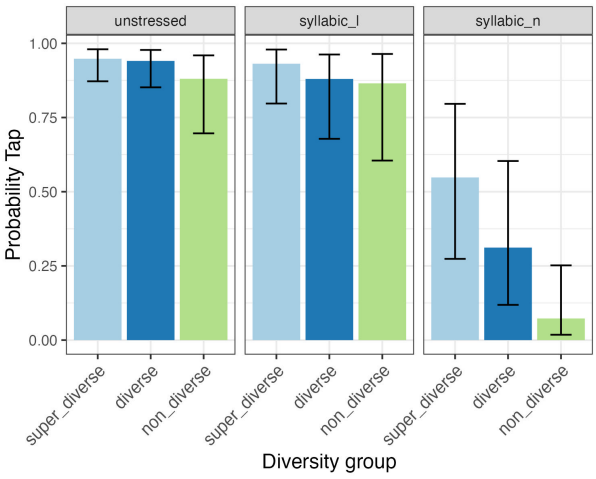


Figure 1: *Model estimates of tap probability by diversity group and context.*

Taps were used significantly less in pre-syllabic /n/ contexts compared to pre-syllabic /l/ or unstressed vowel contexts for all diversity groups (all $p<0.006$).

3.2. All variants

Turning now to a discussion of variants in addition to taps, Figure 2 shows the proportion of each of the five variants (canonical, fricated, glottal, preaspirated and tapped) for each of the eight words. It shows that, apart from taps, canonical variants are the most frequent realisation produced in all items. In those items where significantly fewer taps were produced (*button* and *kitten*), we do see a range of alternative variants being used. Notably, it is here that we observe the greatest incidence of glottalisation, as is typically described for AusE in the pre-syllablic /n/ context [5, 6, 8, 9], yet this remains a less frequently employed strategy than might have been suggested in previous literature [6, 8].

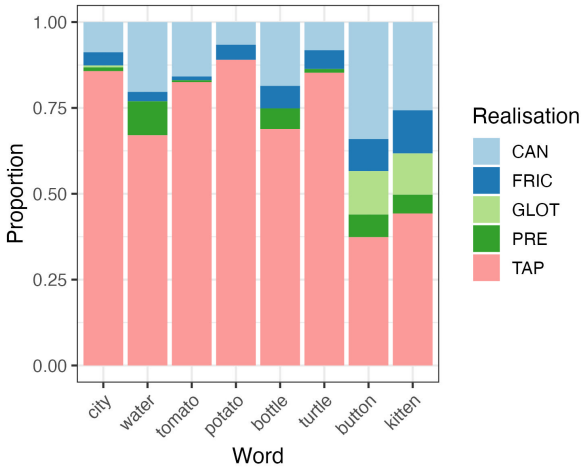


Figure 2: *Raw proportions of each realisation by word.*

Aside from in the pre-syllabic /n/ context, taps were used less frequently in the item *water* (followed closely by *bottle*). Figure 3 illustrates the breakdown of realisations by context and diversity group. Across all contexts, it can be seen that speakers in the non-diverse group produce variants other than taps and canonical /t/ to a greater extent than speakers in the super-diverse and diverse groups. This appears to be the case particularly for the fricated variant in all three contexts, and the glottalised variant in the pre-syllabic /n/ context.

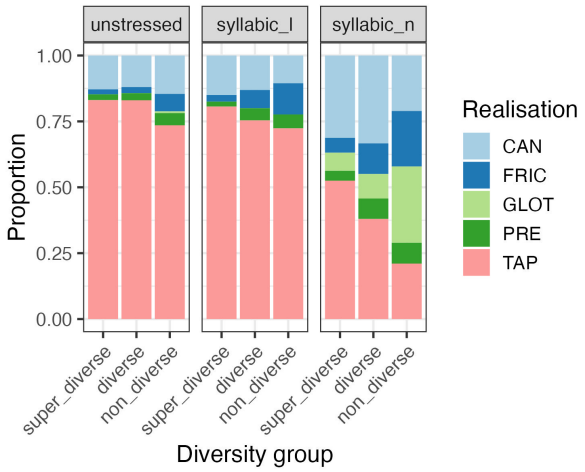


Figure 3: *Raw proportions of each realisation by context and diversity group.*

4. Discussion

In this study we investigated the relationship between variation in intervocalic /t/ realisation and a speaker's exposure to community diversity. Initial findings revealed that, overall, tapping was the most frequently produced /t/ variant. This led to a statistical analysis predicting the probability of tapping versus any other variant. Results showed that while gender did not significantly predict the presence of tapping, exposure to community diversity did have an effect when following phonetic context was taken into account.

The high probability of tapping in this data is reflective of Tollfree's 2001 auditory analysis [5] which found that approximately 90% of all intervocalic /t/ realisations were tapped. For contexts where the /t/ precedes an unstressed vowel or potential syllabic /l/ environment, tapping is near categorical for all speakers regardless of the level of linguistic diversity within their community. In the pre-syllabic /n/ context, however, while speakers in the more diverse community groups are predicted to still produce relatively high proportions of taps, this probability is significantly reduced for speakers from areas with low linguistic diversity. According to previous literature on AusE, we might expect to see an increase in the use of glottalised variants in the pre-syllabic /n/ context [5, 6, 9]. When we examine the distribution of all realisations by context and diversity group (Figure 3) we see there is indeed an increase in the presence of glottalised variants before syllabic nasals. However, the proportion of glottalised variants is greatest in the non-diverse group; with glottalisation used by these speakers more than any other variant while taps are still the predominant realisation for both diverse groups of speakers. We suggest that this finding may provide additional evidence for simplification of linguistic complexity occurring among AusE speakers in the more linguistically diverse communities [20, 21, 22, 23, 24, 25]. In communities with large proportions of households where a LOTE is spoken, a wide range of English varieties will be used by a large number of speakers from many language backgrounds who have learned English as adults and may produce forms that do not display the consistently nuanced structural variability that is typical of native speakers. New generations of L1 speakers of English in such communities are therefore not exposed to the robust structural variation that occurs in more homogeneous communities where the majority of speakers have learned their L1 as children in the community [20, 24, 36]. Research shows that such a scenario in linguistically diverse communities can lead to structural simplification [25]. In this study speakers in the non-diverse group maintain a range of variables at relatively high frequencies in the pre-nasal context, but speakers in more diverse areas are more likely to use a single variant – the tap, albeit at a reduced rate than in the other contexts, consistent with simplification.

Turning to the proportions of realisations by individual words across diversity groups shown in Figure 2, the distribution of realisations in *button* and *kitten* reflect the overall reduction in tap use and increase in canonical and glottalised variants in pre-syllabic /n/ contexts. The lower proportion of tapping in the words *water* and *bottle* could be due to a level of awareness on the part of the speakers. Tapping is a salient feature of AusE and it could be that these words are particularly well-cited as examples of this phenomenon in Australian culture and therefore speakers are more conscious of not producing taps in these specific items such as in an experimental setting such as picture naming. Another possible explanation for variability according to word relates to word frequency. Although all the words analysed are relatively high in frequency, they do vary with *water* and *city* very high in frequency (according to the Australian component of the International Corpus of English [37]). Hay and Foulkes [38] found that leniting changes such as tapping occurred more commonly in frequent words in their New Zealand English corpus. The results here do not support this finding but more detailed analysis of a greater set of words from spontaneous speech will allow us to test frequency effects in the future to examine this phenomenon in more detail.

We acknowledge that this study is based on a limited number of items restricted to read speech. In future work we aim to extend the analysis to the speakers' conversational data. Future work should also include a wider variety of phonetic environments beyond the intervocalic context. The variability associated with the realisation of /t/ before syllabic nasals in the low diversity group also requires further investigation. It is possible that this group is displaying a greater level of socially structured variation compared to the higher diversity groups but more detailed investigation is required to further examine socio-indexical associations related to the variation identified here.

5. Conclusion

The present study has suggested that there is a relationship between intervocalic /t/ realisation and the level of community diversity in a speaker's area, at least in terms of the likelihood of taps being realised. Overall taps are highly prevalent in our data; however, in contexts where /t/ occurs before a possible syllabic /n/, speakers who live in areas with low levels of community diversity are significantly less likely to tap than those who live in more heterogeneous areas. This suggests a process of simplification among speakers who have increased exposure to greater community diversity [24]. This finding supports previous recent work on AusE and provides insight into the way sound change is occurring in our linguistically diverse society [19, 20, 21, 22, 23].

6. Acknowledgements

This research was supported by Australian Research Council Future Fellowship Grant FT180100462 to the third author.

7. References

[1] B. M. Horvath, *Variation in Australian English: the sociolects of Sydney*. Cambridge University Press, 1985.

[2] M. J. Jones and K. McDougall, "The acoustic character of fricated /t/ in Australian English: A comparison with /s/ and /ʃ/," *Journal of the International Phonetic Association*, vol. 39, pp. 265– 289, 2009.

[3] D. Loakes, K. McDougall, and A. Gregory, "Variation in /t/ in Aboriginal and Mainstream Australian Englishes," in *Proceedings of the 18th Australasian International Conference on Speech Science and Technology*, R. Billington, Ed., 2022, pp. 61-65.

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

[5] L. Tollfree, "Variation and change in Australian English consonants: reduction of /t/," in *English in Australia*, D. Blair and P. Collins, Eds. Benjamin, 2001, pp. 45–67.

[6] F. Cox and J. Fletcher, *Australian English: Pronunciation and Transcription*, 2nd ed. Cambridge University Press, 2017.

[7] G. Docherty, J. Hay, and A. Walker, "Sociophonetic patterning of phrase-final /t/ in New Zealand English," in *Proceedings of the 11th Australian International Conference on Speech Science & Technology*, P. Warren & C. I. Watson, Eds., 2006, pp. 378–383.

[8] F. Cox and S. Palethorpe, "Illustrations of the IPA: Australian English," *Journal of the International Phonetic Association*, vol. 37, pp. 341–350, 2007.

[9] J. C. Ingram, "Connected speech processes in Australian English," *Australian Journal of Linguistics*, vol. 9, pp. 21–49, 1989.

[10] J, C. Wells, *Accents of English 2: The British Isles*. Cambridge, 1982.

[11] P. Ladefoged and I. Maddieson, *The Sounds of the World's Languages*. Blackwell Publishers, 1996.

[12] D. Derrick and B. Gick, "Individual variation in English flaps and taps: A case of categorical phonetics," *Canadian Journal of Linguistics*, vol. 56, pp. 307-319, 2011.

[13] G. Docherty, P. Foulkes, S. Gonzalez, and N. Mitchell, "Missed connections at the junction of sociolinguistics and speech processing," *Topics in Cognitive Science*, vol. 10, no. 4, pp. 759-774, 2018.

[14] W. Herd, A. Jongman, and J. Sereno, "An acoustic and perceptual analysis of /t/ and /d/ flaps in American English," *Journal of Phonetics*, vol. 38, pp. 504-551, 2010.

[15] V. C. Haslerud, "The variable (t) in Sydney adolescent speech: A sociolinguistic study of phonological variation," [Unpublished doctoral dissertation]. University of Bergen, 1995.

[16] J. Penney, F. Cox, and A. Szakay, "Glottalisation, coda voicing, and phrase position in Australian English," *Journal of the Acoustical Society of America*, vol. 148, no. 5, pp. 3232-3245, 2020.

[17] J. Choi, M. Broersma, and A. Cutler, "Early phonology revealed by international adoptees' birth language retention," *Psychological and Cognitive Sciences*, vol. 114, no. 28, pp. 7307-7312, 342 2017.

[18] A. Cutler, *Native listening*. MIT Press, 2012.

[19] G. Hurring, J. Hay, K. Drager, R. Podlubny, L. Manhire, and A. Ellis, "Social priming in speech perception: Revisiting kangaroo/kiwi priming in New Zealand English," *Brain Sciences*, vol. 12, no. 6, p. 684, 2022.

[20] C. Clements, J. Penney, A. Gibson, A, Szakay, and F. Cox, "Phonological and lexical conditioning of TRAP vowel duration in Australian English," *Journal of the International Phonetic Association*, to appear.

[21] F. Cox, J. Penney, and S. Palethorpe, "Fifty years of change to definite article allomorphy in Australian English," Journal of the International Phonetic Association, pp. 1-31, 2022.

[22] A. Gibson, J. Penney, and F. Cox, "Rhoticity and hiatus breaking in Australian English: Associations with community diversity," in *Proceedings of the 18th Australasian International Conference on Speech Science and Technology*, R. Billington, Ed., 2022, pp. 46-50.

[23] A. Gibson, F. Cox, and J. Penney, "Acquiring allophony: GOOSE and SCHOOL vowels in the speech of Australian children," in *Proceedings of the 20th International Congress of Phonetic Sciences – ICPhS 2023*, R. Skarnitzl and J. Volín, Eds. International Phonetic Association, 2023, pp. 3750-3754.

[24] J. Penney, F. Cox, and A. Gibson, "Variation in FACE and FLEECE trajectories in Australian English adolescents according to community language diversity," in *Proceedings of the 20th International Congress of Phonetic Sciences - ICPhS 2023*, R. Skarnitzl and J. Volín, Eds. International Phonetic Association, 2023, pp. 3522–3526.

[25] P. Trudgill, "Sociolinguistic typology: Social structure and linguistic complexity," in *The Cambridge Handbook of Linguistic Typology*, Y. A. Alexandra and R. M. W. Dixon, Eds. Cambridge University Press, 2017, pp. 124–150.

[26] D. J. Britain and S. Fox, "The regularisation of the hiatus resolution system in British English: A contact-induced 'vernacular universal'?" in *Vernacular Universals and Language Contacts: Evidence from Varieties of English and Beyond*, M. Filppula, J. Klemola, and H. Paulasto, Eds. Routledge, 2009, pp. 177–205.

[27] F. Cox and J. Penney, "Multicultural Australian English - The new voice of Sydney," *Australian Journal of Linguistics*, vol. 44, pp. 200–219, 2024.

[28] Australian Bureau of Statistics. Greater Sydney: 2021 Census All persons QuickStats. [Online]. Available: https://abs.gov.au/census/find- censusdata/quickstats/2021/1GSYD

[29] S. J. Perry, M. C. Kelley, and B. V. Tucker, "Documenting and modeling the acoustic variability of intervocalic alveolar taps in conversational Peninsular Spanish," *Journal of the Acoustical Society of America*, vol. 155, pp. 294–305, 2024.

[30] G. Docherty and P. Foulkes, "Glottal variants of (t) in the Tyneside variety of English: An acoustic profiling study," in *A figure of speech: A festschrift for John Laver*, W. Hardcastle and J. Mackenzie Beck, Eds. Lawrence Erlbaum, 2005, pp. 173–199.

[31] R Core Team, "R: A language and environment for statistical computing," R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/, 2022. [Online]. Available: https://www.R- project.org/

[32] D. Bates, M. Maechler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software*, vol. 67, no. 1, pp. 1–48, 2015.

[33] H. Matuschek, R. Kliegl, S. Vasishth, H. Baayen, and D. Bates, "Balancing Type I error and power in linear mixed models," *Journal of Memory and Language*, vol. 94, pp. 305–315, 2017.

[34] H. Wickham, *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag, 2009.

[35] R. Lenth, "emmeans: Estimated marginal means, aka least-squares means," R package version 1.8.9. https://CRAN.R-project.org/package=emmeans, 2018. [Online]. Available: https: //CRAN.R- project.org/package=emmeans

[36] W. Labov, "Transmission and diffusion," *Language*, vol. 82, no. 2, pp. 344–387, 2007.

[37] S. Greenbaum, "ICE: The International Corpus of English," English Today, vol. 7, pp 3–7, 1991.

[38] J. Hay and P. Foulkes, "The evolution of medial /t/ over real and remembered time," *Language*, vol. 92, no. 2, pp. 298–330, 2016.

[39] J. Penney, F. Cox and A. Gibson, "Hiatus resolution and linguistic diversity in Australian English" *Phonetica*, vol. 81, no. 2, pp. 119-152, 2024.

[40] J. Penney and F. Cox, "Community linguistic diversity and socioeconomic status: Modelling phonetic variation in speakers from communities with intersecting predictors?" presented at Australian Linguistic Society Conference, Canberra, Australia, 2024.