INTERSPEECH 2023
20-24 August 2023, Dublin, Ireland

Creak prevalence and prosodic context in Australian English

Hannah White[1], *Joshua Penney*[1], *Andy Gibson*[1], *Anita Szakay*[1], *Felicity Cox*[1]

[1]Centre for Language Sciences, Department of Linguistics, Macquarie University, Australia

hannah.white2@hdr.mq.edu.au, joshua.penney; andy.gibson; anita.szakay; felicity.cox@mq.edu.au

Abstract

Creaky voice has been found to mark phrase-finality in many varieties of English, as well as in other languages. The present study aims to investigate whether this is also true for Australian English (AusE), a variety that is understudied in creaky voice research. Using automatic creak detection methods, the need for manual annotation of creak is reduced, and we are able to analyse a large dataset of Australian teenagers' speech. As in other varieties, creak is found to be a marker of finality in AusE. Additionally, we find that males use higher rates of creaky voice than females, challenging the widely held assumption that creak is a feature of female speech.

Index Terms: creaky voice, automatic methods, prosodic context, Australian English

1. Introduction

Creaky voice is often used as an "umbrella" term to encompass a number of voice qualities, each with their own set of distinct acoustic characteristics [1, 2]. Keating et al. [2] provide a thorough overview of the various phonetic realisations of creaky voice and their acoustic properties including prototypical creaky voice (characterised by low fundamental frequency (f0), irregular f0 and high levels of glottal constriction), vocal fry (produced with low but regular f0 and high damping of glottal pulses) and multiply pulsed creak (characterised by two sets of harmonics which creates an irregular and rough pitch). Despite these various realisations, studies have shown that low f0 is a consistent and salient cue to creak and, when asked to identify creaky voice, naive listeners do not differentiate depending on acoustic properties (i.e., a range of different realisations are perceived as creaky voice) [1, 3, 4].

The present study will investigate how creak patterns according to prosodic context in Australian English (AusE), a relatively understudied variety of English. Several studies have found that creak is more likely to occur at the ends of sentences/utterances/phrases across a number of English varieties (RP and Modified Northern English: [5]; American English: [6, 7, 8]) and other languages (Spanish: [9]; Finnish: [10]). As a result, creak has been labelled a pre-pausal or finality marker [5]. Podesva [6] offers a potential physiological explanation for this finding, relating it to declination (i.e., the gradual lowering of f0 across an utterance) [11]. Research has shown that prosodic domain also impacts creak rate. Creak has been found to be more common at the ends of phrases when they are utterance-final rather than utterance-medial in both males and females [7, 8]. Garellek and Keating [8] found that over half of utterance-final vowels occurred with phrase-final creak. González et al. [9] investigated creak rate in word-medial and word-final vowels at the ends of intermediate intonational phrases and full intonational phrases in Spanish. They found that creak was significantly more common word-finally compared to word-medially and that, in both word-positions, creak was significantly more frequent at the end of full intonational phrases than at the end of intermediate intonational phrases. Perception studies have also shown that listeners find creak easier to identify in some prosodic positions compared to others: as creaky voice is commonly used as a marker of phrase-finality, listeners may be habituated to creak in sentence-final position making it less salient in this position compared to creak in non-final positions which may carry social meaning [3, 5, 6, 7, 12].

Creak prevalence has been explored according to binary categorisations of speaker sex in many studies, and the results have been mixed [5, 6, 13, 14, 15, 16]. Early studies of UK English found creak more in male speech than in female speech [5, 15]. More recently creak has been associated with the speech of women in the US [6, 16]; however, Dallaston and Docherty [17] note there is a bias in that creak prevalence studies tend to focus on young American women, indicating that more evidence is required to support this association. Abdelli-Beruh et al. [13] found no difference in creak rates between American males and females. Work on creak prevalence in AusE is limited, however, recent research has suggested that it may occur more often in the speech of males than females [14]. Research in AusE suggests that creak is more prevalent in younger speakers compared to older speakers [14, 18].

As well as marking phrase-finality, creaky voice has been associated with other linguistic and paralinguistic functions such as hiatus resolution between adjacent vowels at word boundaries [19, 20, 21, 22] and a cue to coda stop voicelessness [7, 18, 23], as well as several social functions such as indexing high levels of education and genuineness [16] or, conversely, as signalling boredom [24] and low levels of competence and trustworthiness [25]. Henton and Bladon [5] link creak to marking hyper-masculinity and Yuasa [16] has suggested that women may be using creak to lower the pitch of their voices to take advantage of the positive associations of male voices (e.g., high levels of education and authority).

Dallaston and Docherty [17] also note the variety of different methods used to conduct creak prevalence research. In the above literature, methods of identifying creaky voice included auditory analysis [5], auditory-visual analysis [6, 7, 16] and acoustic voice quality measures (e.g., H1*-H2* and f0) [14]. Creak has been measured by the proportion of creaky to non-creaky syllables [5, 6], vowels [8, 9] and lexical items [7]. Perception study results have suggested that listeners can struggle to correctly identify creak in some circumstances, for example in lower-pitched voices, where there is little differentiation between the pitches of modal and creaky voice [3, 12, 26, 27]. Due to this lack of consistency across studies and the subjective and

10.21437/Interspeech.2023-1454

tedious nature of manual annotation, Dallaston and Docherty [17] suggest the exploration of automatic creaky voice detection methods [28, 29, 30, 31].

The present study aims to examine creak prevalence in different prosodic contexts using automatic creaky voice detection methods applied to a large database of conversational speech from AusE [31]. Based on previous findings [8, 7, 9], we predict that creak prevalence should increase over the course of an utterance, and that turn-final utterances will have higher creak prevalence than turn-medial utterances. We also investigate whether speaker gender has an impact on creak prevalence in AusE, which has been largely unexplored (although see [14]). The use of automatic creak detection methods in this study enables us to efficiently process a large quantity of speech while avoiding issues that commonly occur with manual annotation such as annotator bias and fatigue.

2. Methods

2.1. Data

Data are taken from recordings of conversations between pairs of teenage speakers collected in high schools across Sydney. These data are a subset of recordings collected as part of a larger data collection, which is currently ongoing. A research assistant (RA) guided the topic matter of each conversation. The teenagers knew each other prior to recording. Each teenage speaker was recorded to a separate channel of a Zoom H6 portable recorder through a Rode HS2 headset microphone, with a sample rate of 44.1 kHz and 16 bit resolution. Only the speech of the teenage pairs is included in the present analysis.

Thirty-two conversations were analysed in total, ranging in length from 9.4 to 34.6 minutes (total recording time across all conversations: 577 minutes). All conversations apart from one were between same-gender pairs for a total of 33 female speakers and 31 male speakers. The speakers were aged between 15 and 16 years old (mean age: 15.5 years).

2.2. Automatic creak detection

To identify instances of creaky voice, we used the optimised Union method described in White et al. [31]. This method combines two separate tools which rely on different acoustic cues for identifying creak to improve creak detection compared to when the tools are used on their own [28, 29, 31, 32].

To improve creak detection accuracy, only sonorant segments (i.e., vowels, nasals, liquids and glides) were included in the analysis to ensure the exclusion of instances where automatic methods may identify creak erroneously in periods of silence due to background noise, or in speech segments where creak does not occur such as voiceless fricatives [31].

Recordings were orthographically transcribed using the Speech to Text API from IBM Watson (`https://www.ibm.com/cloud/watson-speech-to-text`). Resulting transcriptions were manually corrected by trained RAs and then checked by a trained phonetician to ensure the content was true to the recording and speech was attributed correctly to each speaker. Recordings were then processed through the MAUS automatic forced-aligner, set to an AusE model, to extract phoneme segment boundaries [33]. Due to the amount of data, phoneme boundaries were not manually corrected. Phoneme boundaries were required for extracting sonorant segments rather than detailed phonetic analysis in this study; therefore, manual boundary correction was determined to be less critical to the present study than other phonetic analyses and we focused instead on efficiently extracting as much data as possible. Before being processed through the Union method, all recordings were resampled to 16 kHz [29].

The Union method combines the AntiMode (AM) method, a tool which identifies creak based on f0 distributions [28, 32] and the Creak Detector (CD) algorithm, which uses several different acoustic measures such as H2-H1 and residual peak prominence to detect creaky voice [29]. The AM method uses Robust Epoch and Pitch EstimatoR (REAPER) [34], a pitch tracker that is able to reliably estimate f0 through creaky portions of speech, which are generally low in f0 [2]. An antimode is then generated for each speaker based on the f0 distributions of their sonorant segments, with any segments in the distribution below the antimode labelled as creak (see [31] for more details). We ran the data through the CD algorithm, optimised for speaker gender [29, 30, 31]. As the CD algorithm assigns a binary creak decision in 10 ms intervals, the Union method operates at this level: if a 10 ms sonorant interval was labelled as creak by either the AM method or the CD algorithm, it was coded as creak by the Union method [31].

Creak prevalence was calculated for each speaker by dividing the number of 10 ms creaky sonorant intervals used by that speaker by the total number of 10 ms sonorant intervals. One speaker (S056, a male speaker) creaked in 80% of 10 ms sonorant intervals, substantially more than any other speaker (the next highest prevalence was 52.5%). The Union method output of S056 was examined manually and compared to the output of other speakers. It was determined that the Union method was over-identifying creaky voice in the speech of S056. As a result, S056 was excluded from the analysis, resulting in a total of 63 speakers (33 female; 30 male). For eight of the speakers, an issue with the microphone meant that background noise interfered with creak detection periodically throughout the conversations. Portions of speech where this occurred were removed from analysis (total of 0.9% data removed).

2.3. Data preparation

For each speaker, all speech was segmented into utterances. The utterance-level was chosen for the analysis due to the use of automatic methods making phrase-level labelling impossible. Utterances were defined as stretches of speech between pauses, with pauses having been automatically identified by MAUS. According to Fletcher [35], pauses that are less than 200 ms are generally considered to be inaudible; therefore, when MAUS identified a pause of less than 200 ms, it was removed and speech on either side of the pause was counted as being part of the same utterance. This allowed us to avoid misidentifying articulatory pauses as between-utterance pauses [35, p. 573]. Once utterances were identified (n = 14002), they were labelled as either long utterances or minor utterances. Minor utterances were defined as utterances containing less than three words (n = 5461). All other utterances were long utterances (henceforth we will refer to long utterances as *utterances* and minor utterances as *minor utterances*). Minor utterances were removed as over half consisted of filler words and discourse markers (56%); research has suggested that creaky voice tends to co-occur with filler words [36] and with discourse markers, particularly “yeah” and “like” which make up 37% of our minor utterances [37, 38]. One female speaker was exluded at this point as they had many more utterances than the rest of the speakers (484 utterances compared to the next largest at 284 utterances). There were no outliers with small numbers of utterances. This resulted in a total of 62 speakers and 8057 utterances with a mean of 130 per speaker (range per speaker: 24 – 284).

To investigate whether creak prevalence changed over the course of an utterance, each utterance was divided equally into thirds (each third was labelled as first third, second third and third third). In 88 utterances, 10 ms sonorant intervals were identified in only two thirds of the utterances (i.e., data to calculate creak prevalence was absent from one third). These utterances were removed from analysis so that all three thirds in each included utterance provided an opportunity for creak to occur, leaving 8453 utterances. To ensure that calculations of creak prevalence were sensible, only thirds with at least 100 ms of sonorant segments were included in the analysis. Within each third, creak prevalence was calculated by dividing the number of 10 ms creaky sonorant intervals by the total number of 10 ms sonorant intervals within that third. A total of 23088 observations were included in the final analysis.

Mean utterance length was 2.05 seconds (length range: 0.29 – 16.05 seconds). The position of each utterance within the number of utterances produced by the speaker (henceforth called "position-in-conv") was calculated by dividing the order of occurrence of the utterance in the speaker's speech by the total number of utterances of that speaker.

Utterances were labelled as either turn-medial or turn-final (henceforth referred to as "prosodic-position"). Utterances were turn-medial if the same speaker maintained the turn with another utterance (including minor utterances) after the completion of an utterance. Utterances were labelled turn-final if the other speaker took the floor or if followed by RA speech. RA speech was identified by the first author in Praat [39].

2.4. Analysis

We used linear mixed effects regression (LMER) modelling using the lme4 package [40] and lmerTest package [41] in R [42] for our analysis. First, a maximal model was built with creak prevalence as the dependent variable and three three-way interactions as independent variables. Interactions were between:

- Utterance third (three-level ordered factor: first [reference], second or third), prosodic-position (two-level factor: turn-medial [reference] or turn-final) and speaker gender (two-level factor: female [reference] or male);
- Utterance third, prosodic-position and utterance length (scaled continuous variable);
- Utterance third, prosodic-position and position-in-conv (scaled continuous variable indicating the position of an utterance within the total number of utterances of the speaker).

A random intercept was included for speaker with random slopes for utterance length, position-in-conv and the interaction between utterance third and prosodic-position. Continuous variables (position-in-conv and utterance length) were transformed to *z*-scores.

The maximal model did not converge. We took a stepwise elimination approach, reducing non-significant interactions by order of least significance, then removing non-significant fixed effects one-by-one, each time running an ANOVA to compare models, ensuring the removal of terms did not significantly worsen the model. Finally we reduced the random effects structure until the model converged without warnings.

The final model had a two-way interaction between utterance third and prosodic-position, simple effects of speaker gender and position-in-conv and a random intercept for speaker. The model produced singular fit and convergence warnings until all random slopes were removed.

3. Results

The output of the final model is shown in Table 1. A type 3 mixed model ANOVA was run on the final model with the LRT-method using the afex package [43] in R. The ANOVA showed a significant simple effect of speaker gender (df = 1, chisq = 5.24, *p*-value = 0.022), presented in Figure 1. The significant simple effect of position-in-conv is shown in Figure 2 (df = 1, chisq = 40.05, *p*-value <0.001). The interaction between utterance third and prosodic-position was also significant and is shown in Figure 3 (df = 2, chisq = 16.49, *p*-value <0.001).

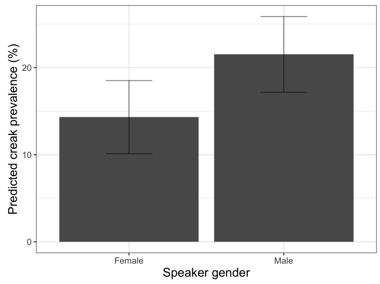

Figure 1: *LMER predicted difference in creak prevalence by speaker gender.*

Figure 1 shows predicted creak prevalence by speaker gender. We see that, overall, males in our data have significantly greater prevalence of creaky voice than females.

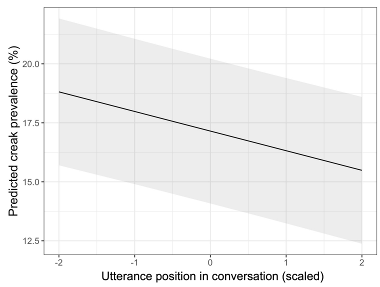

Figure 2: *LMER predicted creak prevalence across the course of a conversation.*

The significant effect of position-in-conv on creak prevalence is shown in Figure 2. It shows that speakers became less creaky over the course of their total utterances in the conversation (i.e., creak decreased over the course of the conversation).

Figure 3 shows the predicted effect of utterance third and prosodic-position on creak prevalence. Although speakers became less creaky over the course of a conversation, they became more creaky over the course of an utterance. *Post hoc* pairwise comparisons were carried out using the emmeans package [40] in R, with Tukey HSD corrections to *p*-values for multiple comparisons. These showed that, in both turn-medial and turn-final position, creak prevalence differed significantly between each third (all $p<0.001$). Additionally, within each third of an utterance, creak prevalence was significantly higher in turn-final position compared to turn-medial (all $p<0.001$).

4. Discussion

The results of our analysis of prosodic position on creaky voice prevalence are consistent with previous literature, showing that

Table 1: *LMER model for effect of position-in-conv, prosodic-position and speaker gender on creaky voice prevalence. Speaker is included as a random intercept. Reference levels: utterance third = first; prosodic-position = medial; speaker gender = female.*

	Estimate	Std. error	df	*t* value	*p*-value
(Intercept)	9.52	2.16	63.77	4.40	<0.001 ***
uttThird (second)	2.37	0.46	23025.54	5.17	<0.001 ***
uttThird (third)	7.29	0.46	23025.53	15.91	<0.001 ***
prosodicPos (final)	2.00	0.46	23040.01	4.32	<0.001 ***
gender (male)	7.21	3.08	61.52	2.34	0.023 *
posConv	-0.83	0.13	23026.00	-6.33	<0.001 ***
uttThird (second): prosodicPos (final)	0.89	0.64	23025.52	1.39	0.164
uttThird (third): prosodicPos (final)	2.57	0.64	23025.52	4.00	<0.001 ***

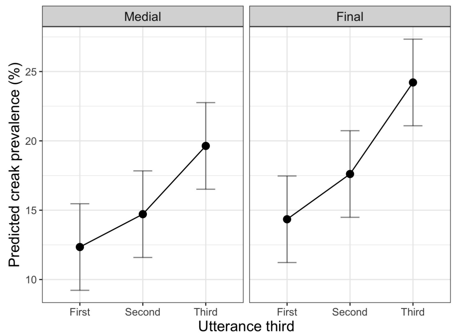

Figure 3: *LMER predicted change in creak prevalence across utterance in turn-medial and turn-final utterances.*

creaky voice increases throughout the course of an utterance [5, 6, 7, 8, 9] and creaky voice prevalence is greatest when an utterance is turn-final compared to turn-medial, suggesting a turn-yielding function in AusE.

Our results show a greater prevalence of creaky voice in male speech compared to female speech. This is in line with Loakes and Gregory's [14] study of AusE; however, it contradicts many studies of American English and stereotypical gendered associations [6, 16, 44]. This finding and the fact that most research on creak prevalence has focused on American speakers, particularly women [17], emphasises further the need for continuing investigation of creak prevalence in different communities of speakers. It would be interesting to investigate whether Australian listeners hold the same stereotypical gendered associations with creak as those shown to be held by American listeners in a perception study [25]. The speakers in our dataset are all of a homogeneous age group so we do not know whether creak has always been more prevalent in male speech or whether this is a recent change. Our future work will aim to investigate historical use of creak in AusE.

Results also show that creak prevalence significantly decreases over the course of a speaker's total utterances (i.e., over the course of the conversation) in the present task. Although we did not have any specific prior predictions as to how creak prevalence would pattern over the course of a conversation, creak has previously been associated with boredom and low energy levels, which could be two side-effects of the conversation task [24, 45]. It is possible that speakers did not experience boredom throughout the conversation task and perhaps became more engaged and enthusiastic as the conversation progressed, particularly as they were conversing with a peer rather than the interviewer alone. Higher creak rates at the beginning of the conversations could suggest hesitancy as the speakers settled in to the task [16]. In a different analysis of this data we found evidence of convergence in creaky voice levels between the pairs, which is an additional possible explanation for the decrease in creak over time [46]. In order to better understand the source of this effect of utterance-position, a qualitative approach would be beneficial, which is beyond the scope of the present study but would be an interesting area for future work.

This study has provided evidence that in AusE, as in many other varieties of English and languages, creaky voice is a marker of finality when it occurs at the ends of utterances, particularly when those utterances are turn-final. We have also presented evidence that, for these speakers in this task, creak prevalence decreases over the course of a conversation, a finding that would greatly benefit from further qualitative exploration. This study has provided support for a higher prevalence of creak among males compared to females in AusE. This finding is particularly relevant given recent discussion around how accurately empirical data actually reflects the stereotypical association of creak with female speech [17, 47].

5. Acknowledgements

This research was supported by a Macquarie University Research Excellence Scholarship to the first author and by Australian Research Council Grant DP190102164 and Australian Research Council Future Fellowship Grant FT180100462 to the fifth author.

6. References

[1] L. Davidson, "Perceptual coherence of creaky voice qualities," in *Proc. 19th ICPhS*, S. Calhoun, P. Escudero, M. Tabain, and P. Warren, Eds. ASSTA Inc, 2019, pp. 196.1–5.

[2] P. Keating, M. Garellek, and J. Kreiman, "Acoustic properties of different kinds of creaky voice," in *Proc. 18th ICPhS*, The Scottish Consortium for ICPhS 2015, Ed. University of Glasgow, 2015, Conference Proceedings, pp. 821.1–5.

[3] L. Davidson, "The effects of pitch, gender, and prosodic context on the identification of creaky voice," *Phonetica*, vol. 76, no. 4, pp. 235–262, 2019.

[4] Y. Huang, "Different attributes of creaky voice distinctly affect mandarin tonal perception," *J. Acoust. Soc. Am.*, vol. 147, no. 3, pp. 1441–1458, 2020.

[5] C. Henton and A. Bladon, "Creak as a sociophonetic marker," in *Language, speech, and mind: Studies in honour of Victoria A. Fromkin*, L. M. Hyman and C. N. Li, Eds. Routledge, 1988, pp. 3–29.

[6] R. J. Podesva, "Gender and the social meaning of non-modal phonation types," *Ann. Meet. Berkeley Linguistics Soc.*, vol. 37, no. 1, pp. 427–448, 2013.

[7] L. Redi and S. Shattuck-Hufnagel, "Variation in the realization of glottalization in normal speakers," *J. Phonetics*, vol. 29, no. 4, pp. 407–429, 2001.

[8] M. Garellek and P. Keating, "Phrase-final creak: Articulation, acoustics, and distribution," in *Ann. Meet. Linguistic Soc. Am.*, 2015.

[9] C. González, C. Weissglass, and D. Bates, "Creaky voice and prosodic boundaries in Spanish: An acoustic study," *Studies in Hispanic and Lusophone Linguistics*, vol. 15, no. 1, pp. 33–65, 2022.

[10] R. Ogden, "Turn transition, creak and glottal stop in finnish talk-in-interaction," *J. Int. Phonetic Assoc.*, vol. 31, no. 1, pp. 139–152, 2001.

[11] D. R. Ladd, "Declination: A review and some hypotheses," *Phonology Yearbook*, vol. 1, pp. 53–74, 1984.

[12] A. Li, W. Lai, and J. Kuang, "How do listeners identify creak? The effects of pitch range, prosodic position and creak locality in Mandarin," in *Proc. Speech Prosody 2022*, 2022, pp. 480–484.

[13] N. Abdelli-Beruh, T. Drugman, and R. H. Red Owl, "Occurence frequencies of acoustic patterns of vocal fry in American English speakers," *J. Voice*, vol. 30, no. 6, pp. 759.e711–759.e720, 2016.

[14] D. Loakes and A. Gregory, "Voice quality in Australian English," *JASA Express Letters*, vol. 2, no. 8, p. 085201, 2022.

[15] J. Stuart-Smith, "Glasgow: Accent and voice quality," in *Urban voices: Accent studies in the British Isles*, P. Foulkes and G. Docherty, Eds. Arnold, 1999, pp. 203–222.

[16] I. P. Yuasa, "Creaky Voice: A New Feminine Voice Quality for Young Urban-Oriented Upwardly Mobile American Women?" *American Speech*, vol. 85, no. 3, pp. 315–337, 2010.

[17] K. Dallaston and G. Docherty, "The quantitative prevalence of creaky voice (vocal fry) in varieties of English: A systematic review of the literature," *PloS ONE*, vol. 15, no. 3, p. e0229960, 2020.

[18] J. Penney, F. Cox, and A. Szakay, "Glottalisation, coda voicing, and phrase position in Australian English," *J. Acoust. Soc. Am.*, vol. 148, no. 5, pp. 3232–3245, 2020.

[19] F. Cox, S. Palethorpe, L. Buckley, and S. Bentink, "Hiatus resolution and linking /r/ in Australian English," *J. Int. Phonetic Assoc.*, vol. 44, no. 2, pp. 155–178, 2014.

[20] F. Cox, J. Penney, and S. Palethorpe, "Fifty years of change to definite article allomorphy in Australian English," *J. Int. Phonetic Assoc.*, in press, 2022.

[21] L. Davidson and D. Erker, "Hiatus resolution in American English: The case against glide insertion," *Language*, pp. 482–514, 2014.

[22] J. Peña, L. Davidson, and S. Orosco, "The independence of phrasal creak and segmental glottalization in American English," *JASA Express Letters*, vol. 1, p. 075205, 2021.

[23] S. Seyfarth and M. Garellek, "Physical and phonological causes of coda /t/ glottalization in the mainstream American English of central Ohio," *Laboratory Phonology*, vol. 11, no. 1, p. 24, 2020.

[24] C. Gobl and A. Ní Chasaide, "The role of voice quality in communicating emotion, mood and attitude," *Speech Communication*, vol. 40, pp. 189–212, 2003.

[25] R. C. Anderson, C. A. Klofstad, W. J. Mayew, and M. Venkatachalam, "Vocal fry may undermine the success of young women in the labor market," *PLoS ONE*, vol. 9, no. 5, p. e97506, 2014.

[26] J. Villegas, S. J. Lee, J. Perkins, and K. Markov, "Psychoacoustic features explain creakiness classifications made by naive and non-naive listeners," *Speech Communication*, 2023.

[27] H. White, A. Gibson, J. Penney, A. Szakay, and F. Cox, "The influence of pitch and speaker sex on the identification of creaky voice by female listeners," in *Proc. 18th Australasian Int. Conf. Speech Science and Technol.*, 2022, pp. 81–85.

[28] K. Dallaston and G. Docherty, "Estimating the prevalence of creaky voice: A fundamental frequency-based approach," in *Proc. 19th ICPhS*, S. Calhoun, P. Escudero, M. Tabain, and P. Warren, Eds. ASSTA Inc, 2019, pp. 581.1–5.

[29] T. Drugman, J. Kane, and C. Gobl, "Data-driven detection and analysis of the patterns of creaky voice," *Comput. Speech and Language*, vol. 28, no. 5, pp. 1233–1253, 2014.

[30] O. Murton, S. Shattuck-Hufnagel, J.-Y. Choi, and D. D. Mehta, "Identifying a creak probability threshold for an irregular pitch period detection algorithm," *JASA Express Letters*, vol. 145, no. 5, pp. EL379–EL385, 2019.

[31] H. White, J. Penney, A. Gibson, A. Szakay, and F. Cox, "Evaluating automatic creaky voice detection methods," *J. Acoust. Soc. Am.*, vol. 152, no. 3, pp. 1476–1486, 2022.

[32] K. Dorreen, "Fundamental frequency distributions of bilingual speakers in forensic speaker comparison," [Master's thesis, University of Canterbury], 2017.

[33] F. Schiel, C. Draxler, and J. Harrington, "Phonemic segmentation and labelling using MAUS technique," in *New tools and methods for very-large-scale phonetics research workshop*, Philadelphia, PA, USA, 2011.

[34] D. Talkin, "REAPER: Robust epoch and pitch EstimatoR," 2015. [Online]. Available: https://github.com/google/REAPER

[35] J. Fletcher, *The Prosody of Speech: Timing and Rhythm*, 2nd ed. Blackwell Publishing Ltd, 2010, pp. 521–602.

[36] M. Candea, I. Vasilescu, and M. Adda-Decker, "Inter- and intra-language acoustic analysis of autonomous fillers," in *Proceedings of DiSS'05*, 2005, pp. 47–52.

[37] T. Grivičić and C. Nilep, "When phonation matters: The use and function of yeah and creaky voice," *Colorado Research in Linguistics*, vol. 17, pp. 1–11, 2004.

[38] N. Mendoza-Denton, "The semiotic hitchhiker's guide to creaky voice: Circulation and gendered hardcore in a chicana/o gang persona," *J. Linguistic Anthropology*, vol. 21, pp. 261–280, 2011.

[39] P. Boersma and D. Weenink, "Praat: Doing phonetics by computer," 2018, [Computer program]. [Online]. Available: http://www.praat.org/

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

[41] A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen, "lmerTest package: Tests in linear mixed effects models," *J. Statistical Softw.*, vol. 82, no. 13, pp. 1–26, 2017.

[42] R Core Team, "R: A language and environment for statistical computing," 2021. [Online]. Available: https://www.R-project.org/

[43] H. Singmann, B. Bolker, J. Westfall, F. Aust, and M. S. Ben-Shachar, "afex: Analysis of Factorial Experiments," 2022, r package version 1.1-1. [Online]. Available: https://CRAN.R-project.org/package=afex

[44] N. Wolf, "Young women, give up the vocal fry and reclaim your strong female voice," *The Guardian*, 2015. [Online]. Available: https://www.theguardian.com/commentisfree/2015/jul/24/vocal-fry-strong-female-voice

[45] T. Pratt, "Affect in sociolinguistic style," *Language in Soc.*, vol. 52, no. 1, pp. 1–26, 2023.

[46] H. White, J. Penney, A. Gibson, A. Szakay, and F. Cox, "Convergence of creaky voice use in Australian English," in *Proc. 20th ICPhS*, 2023.

[47] K. Becker, S.D. Khan, and L. Zimman, "Beyond binary gender: creaky voice, gender, and the variationist enterprise," *Language Variation and Change*, vol. 34, pp. 215–238, 2022.