PRE-/l/ VOWEL CHANGE IN AUSTRALIAN ENGLISH *POOL-PULL*

Tünde Szalay[1], Titia Benders[2], Felicity Cox[3], Michael Proctor[3]

[1]University of Sydney, [2] University of Amsterdam, [3]Macquarie University
tuende.szalay@sydney.edu.au

ABSTRACT

Contrast reduction between Australian English pre-lateral /ʉː-ʊ/ (*pool-pull*) compared to other environments might indicate an ongoing sound change resulting in merger. Our apparent-time study explores this merger by comparing the $F2$ characteristics of /ʉː-ʊ/ produced in pre-lateral and pre-obstruent contexts by 9 older and 8 younger female speakers. Differences between older and younger speakers and potential merger within younger speakers were tested using Bayesian methods. Consistent with a pre-lateral vowel change, younger speakers showed reduced pre-lateral vowel contrast compared to older speakers. There was no conclusive evidence for vowel merger, as $F2$ targets of /ʉː-ʊ/ remain contrastive in the pre-lateral environment for some young speakers.

Keywords: sound change, pre-lateral vowels, change by coarticulation

1. INTRODUCTION

Systematic and directional coarticulation leading to contrast reduction is often implicated in the initiation of sound change and merger [1, 2]. Yet, not all coarticulation leads to sound change, and not all contrast reduction leads to merger [1, 2, 3]. In the Interactive Phonetic (IP) model, a sound change may be initiated when highly coarticulated realisations of one phoneme become acoustically similar to another phoneme, and a merger is complete when contrast between two categories is lost [2].

Pre-lateral vowels show contrast reduction and/or merger due to the coarticulatory influence of /l/ in several varieties of English. For example, vowel contrast reduction between pre-/l/ /iːl-ɪl/ has been observed in some dialects of American English as /iːl/ shifts toward /ɪl/ [4]. Southern British English /uːl-ʊl/ show an acoustic merger, as pre-/l/ /uːl/ is produced with an $F2$ similar to pre-/l/ /ʊl/ [5]. In Melbourne Australian English, /el-æl/ show merger through the lowering of /el/ [6]. Members of the Australian English (AusE) vowel pairs /iː-ɪ, ʉː-ʊ, æɔ-æ,/ and /əʉ-ɔ/ (e.g., *feel-fill, pool-pull, howl-Hal, dole-doll*) show acoustic contrast reduction in the pre-lateral position, while their pre-obstruent allophones remain distinct [7, 8, 9, 10, 11]. The AusE contrast reduction is consistent with an ongoing sound change and merger brought about through coarticulation. However, only the /æɔ-æ/ (*howl-Hal, owl-Al*) contrast has been studied using apparent-time methods: sound change is shown in younger speakers producing smaller /æɔl-æl/ contrast compared to older speakers, and merger is shown in younger male speakers' increased similarity between pre-/l/ /æɔl-æl/ and pre-/d/ /æɔ/ [12].

	1967 [13]	1999 [7]	2004 [8]	2021 [10]
Gender	M	M	F	F
ʊd	950	1000	1100	1132
ʉːd	1550	1750	2200	2197
ʊl	NA	NA	1000	937
ʉːl	NA	NA	1300	983

Table 1: Changes in AusE /ʉː-ʊ/ over time: mean $F2$ for speakers of different ages. Columns: publication date and speaker gender [13, 7, 8, 10].

Sound change and merger has not been thoroughly explored for pre-lateral /ʉː-ʊ/ (*pool-pull*), although the reduced contrast may be consistent with a contextual vowel merger [9, 10]. Due to the $F2$ drop in /ʉːl/, /ʉːl/ becomes spectrally similar to /ʊl/, which may contribute to loss of spectral differentiation (Table 1). Therefore, we examine whether the reduced contrast between pre-/l/ /ʉː-ʊ/ [10] is indicative of a vowel change and merger in AusE. We hypothesised that younger speakers would (1) preserve spectral /ʉː-ʊ/ contrast in the pre-obstruent environment; (2) show smaller spectral contrast between pre-lateral allophones of the members of the vowel pair /ʉː-ʊ/ than older speakers; and (3) not spectrally differentiate between /ʉːl/ and /ʊl/.

2. METHODS

2.1. Speakers

Data were extracted from AusTalk, an AusE speech corpus recorded between 2011 to 2015 [14]. Recordings of 8 younger (age = 20 – 29, mean = 24.4) and 9 older (age = 54 – 80, mean = 66.6) native female speakers of AusE were selected. Speakers were born and educated in the Greater Sydney Metro Region with at least one Australian-born parent. The speakers did not report any reading, speaking, or hearing difficulties.

2.2. Material and procedure

The stressed vowels /ʉː-ʊ/ were produced in two monosyllabic paradigms, /hVd/ and /pVl/ (*who'd-hood, pool-pull*), in a single-word production task. Speakers read 322 isolated words, including the four target words, as they were presented orthographically on a computer monitor in random order. The task was recorded during three separate sessions, each using a different order of words. Each speaker produced up to three repetitions of each lexical item; the number of repetitions differs between participants, as not all participants attended all sessions.

2.3. Phonetic analysis

200 tokens were analysed (4 target words × 17 speakers × 3 maximal repetitions - 4 missing repetitions). Segment boundaries were located automatically using the MAUS forced aligner with the AusE grapheme-to-phoneme converter [15, 16, 17], and corrected manually in a Praat interface when a boundary was misplaced by more than 20 ms [18, 19]. Vowel onset was determined based on voicing onset and sudden increase in amplitude. Vowel offset in the /d/ context was determined based on amplitude drop. Rime offset in the /l/ context was determined based on voicing offset. As there is no discernible boundary between the vowel and the following /l/ in /pVl/ words, the entire /pVl/ rime was analysed (Fig. 1). Vowel duration was not measured due to the lack of discernible boundary.

Formant trajectories in pre-/d/ vowels and lateral-final rimes were extracted automatically and corrected manually in Praat [18]. Formant frequencies were estimated at every 5 ms throughout a 25 ms formant analysis window using a 50 ms Gaussian window with 75% overlap and with 50 dB dynamic range and a pre-emphasis filter increasing spectral slope above 100 Hz by 6 dB/octave. To optimise formant settings, four to five formants were tracked up to 4500 Hz ceiling in tokens with comparatively lower $F2$ and $F3$, or up to a maximum frequency of 6000 Hz in tokens with a comparatively higher $F2$ or $F3$ trajectory. Formant trajectories were manually corrected using a Praat-based interface that superimposed formant estimates over a broadband spectrogram calculated over 5 ms windows with 40% overlap, allowing for corrections of estimates that did not align with the visible formants. After hand-correction, $F1$–$F3$ trajectories for every word were visually inspected; values 1.5 times above or below the interquartile range for each formant in each vowel × coda × age group were rechecked.

Acoustic targets were located automatically in the corrected $F2$ trajectories using the $F2$ maxima for central /ʉː/ and the $F2$ minima for back /ʊ/. Pre-/d/ targets were located in the vowel trajectory and pre-/l/ targets in the first half of the rime (Fig. 2). Phonetic analysis was conducted by the first author.

2.4. Statistical analysis

The $F2$ difference between coda contexts and age groups was tested using a Bayesian Multilevel Model (BMLM). In Bayesian modelling, our certainty regarding the estimate of each factor prior to modelling the data is quantified by a prior probability distribution of the estimate [20, 21]. The prior distributions are combined with the data, generating a posterior distribution of estimates for each factor [20]. The interval containing 95% of all posterior estimates is referred to as the Credible Interval (CI). The CI is compared to the Region Of Practical Equivalence (ROPE), defined by the researcher as the "effect smaller in magnitude than the minimally interesting effect" [21]. When the CI falls outside the ROPE, the two groups are considered to be different; when it falls within the ROPE, the two groups are considered to be equivalent; and when the 95% CI spans values both outside and within the ROPE, the results are considered to be inconclusive [21].

Factor	Level	Estimate	SD
Age	Younger	100	100
Vowel	/ʉː/	750	100
Coda	/l/	-200	100
Age:Vowel	Younger:/ʉː/	250	100
Age:Coda	Younger:/l/	-100	100
Age:Vowel:Coda	Younger:/ʉː/:/l/	-300	300

Table 2: Informed priors in the Bayesian models

We constructed a BMLM with the dependent variable $F2$ Target, and the independent variables Vowel, Coda, and Age (interacting, dummy coded, comparing /ʉː/ to the baseline /ʊ/, /l/ to /d/, younger to older) using the brms library [22, 23, 24]. The model included a random intercept for Speaker with slope for Coda. Informed priors are reported in Table 2 [7, 8, 10]. The ROPE was defined as -76 to +76 Hz: a factor of two larger than the perceptual threshold of noticeable vowel difference, comparable to the smallest distance between two adjacent vowels in American English [25]. Pillai-scores were calculated for each participant using MANOVA [26, 27]. All data analysis was conducted in R [28].

3. RESULTS

Older speakers produced /ʉː/ with a higher $F2$ than /ʊ/ (β = 1115, est. error = 33.34, CI = 1051 – 1182), with the CI exceeding the ROPE, providing strong evidence for distinct pre-/d/ vowels. Older speakers produced /ʊ/ with a lower $F2$ in the /l/ than in the /d/-context (β = -136, est. error = 47.16, CI = -232 – -45). The evidence for older speakers backing

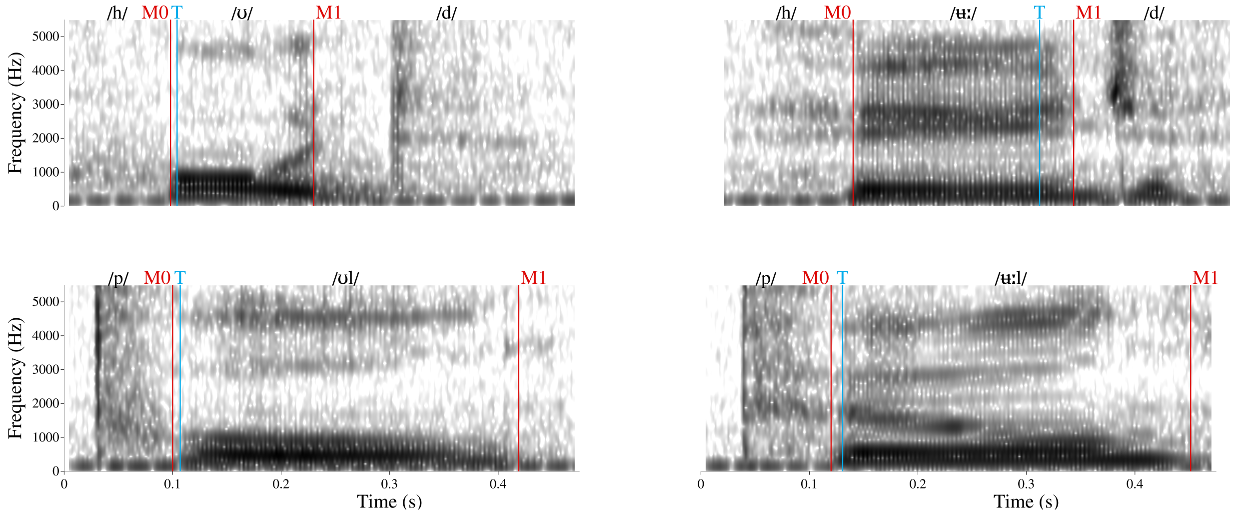

Figure 1: *Vowel (top) and rime (bottom) boundaries (start:M0, end: M1) with vowel target (T) by an older speaker.*

pre-/l/ /ʊ/ is inconclusive, as the CI spans interesting values and the ROPE. There is strong evidence for older speakers backing /ʉː/ more than /ʊ/ (β = -670, est. error = 44.4, CI = -760 – -585), as the CI exceeds the ROPE.

Younger speakers might have produced /ʊ/ with a higher $F2$ than older speakers (β = 71, est. error = 50.94, CI = -33 – 168). The Age:Vowel interaction (β = -16, est. error = 49.27, CI = -112 – 81) indicates that younger speakers, compared to older speakers, might not have fronted /ʉː/ as much as they fronted /ʊ/ (Fig. 2). The Age:Coda interaction (β = -13, est. error = 66.61, CI = -136 – 125) indicates that the difference between pre-obstruent and pre-lateral /ʊ/ might have been similar for younger and older speakers. The evidence for age-related vowel change in pre-/d/ vowels and pre-/l/ /ʊ/ is inconclusive as the CIs span both interesting values and the ROPE (Fig. 2). The negative three-way interaction (β = -344, est. error = 66.63, CI = -475 – -212) provides strong evidence that the $F2$ decrease in /ʉː/ associated with /l/ is greater for younger than older speakers, as the CI exceeds the ROPE.

Planned comparisons show a categorical difference for older speakers with the CI exceeding the ROPE in both coda contexts (/d/: β = 1117, CI = 1050 – 1181; /l/: β = 446, CI = 382 – 513). For younger speakers, $F2$ difference exceeded the ROPE before /d/ only (β = 1099, CI = 1027 – 1171). Pre-/l/ results are inconclusive, as the CI spans interesting values and the ROPE (β = 85, CI = 11 – 158).

Pillai-scores show distinct vowel categories in both contexts for older speakers (/d/-context: mean = 0.97, sd = 0.05; /l/-context: mean = 0.90, sd = 0.14). Younger speakers maintain contrast in the /d/-context (mean = 0.99, sd = 0.004), but reduce it before /l/ (mean = 0.45, sd = 0.312) (Fig. 4).

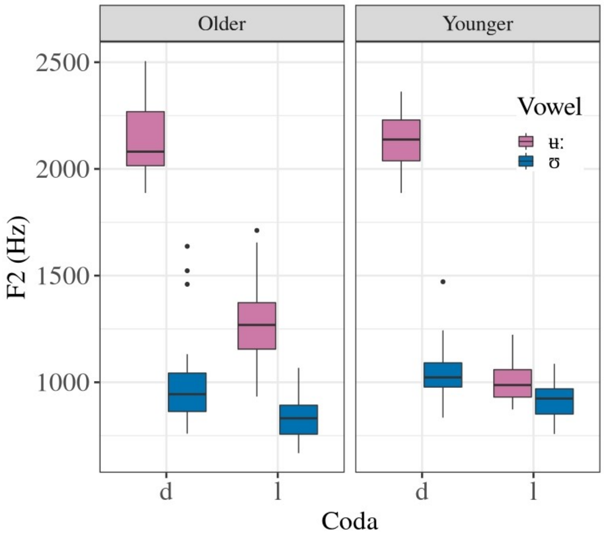

Figure 2: *F2 at vowel target*

4. DISCUSSION

H(1) predicted that vowel contrast would be preserved in the /d/-context. H(1) holds, as both younger and older speakers maintain /ʉː-ʊ/ contrast in the pre-/d/ context through higher $F2$ in /ʉː/ and distinct /ʉː-ʊ/ categories in the vowel space (Figs. 3-4). Although young speakers may produce /ʊ/ with a marginally higher $F2$ than older speakers, results on an ongoing /ʊ/-fronting are inconclusive.

H(2) predicted that younger speakers would produce pre-/l/ allophones of /ʉː-ʊ/ with smaller contrast than older speakers. H(2) holds, as younger speakers produce pre-/l/ /ʉː-ʊ/ with a smaller $F2$ contrast and a larger overlap in the $F1$-$F2$ space (Figs. 2–4). Younger speakers' larger overlap is primarily attributed to their larger $F2$ drop in /ʉːl/ compared to older speakers'. Younger speakers may produce /ʊ/ with an overall higher $F2$ in both contexts, further reducing the pre-/l/ /ʉː-ʊ/ contrast.

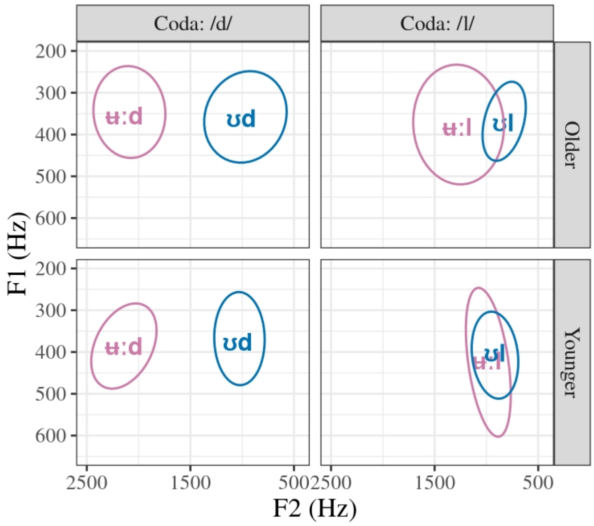

Figure 3: *F1 and F2 at vowel target*

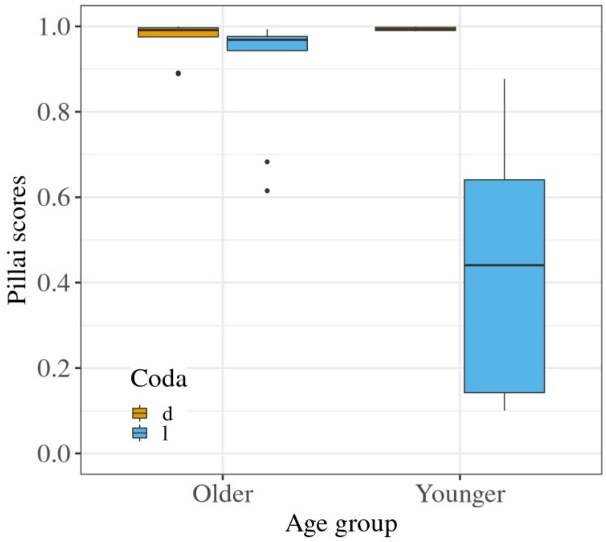

Figure 4: *Pillai-scores for the /ʊ-ʉː/ contrast*

Increased contrast reduction is consistent with an ongoing pre-lateral sound change in AusE during which the pre-/l/ allophone of /ʉː/ shifts towards pre-/l/ /ʊ/ in the acoustic space. That is, contrast reduction caused by coarticulatory variation in the /l/-context has become a sound change in AusE. This sound change can be represented in the IP model [2]: as listeners and speakers interact over time, coarticulated /ʉːl/ realisations are incorporated into listeners' representation of /ʊl/, shifting /ʉːl/ closer to /ʊl/ [2].

Pre-/l/ vowel contrast reduction due to $F2$ drop and coarticulatory vowel backing is observed in many varieties of English [29, 6, 30]. In Melbourne Australian English, $F2$ drop in /e/ before /l/, together with an $F1$ increase, contribute to an /el-æl/ merger [6]. *Pool* shows an $F2$ lowering in Ohio-, British-, and Yorkshire English, reducing *pool-pull* contrast [5, 31, 30, 32]. However, tongue body contrast is maintained for British English /uː-ʊ/, showing that an $F2$ drop does not always indicate tongue backing [5]. Thus, an articulatory study on AusE *pool-pull* contrast is required.

H(3) predicted that younger speakers would not contrast pre-lateral /ʉː-ʊ/. The evidence for H(3) is inconclusive: younger speakers may or may not produce pre-/l/ /ʉː-ʊ/ with a contrastive $F2$. Our results regarding a pre-/l/ /ʉː-ʊ/ merger may be inconclusive due to the large inter-speaker variation in younger speakers. Pillai scores, calculated on a speaker-by-speaker basis, reveal that younger speakers exhibit a range from near-maximal to near-minimal contrast (Fig. 4). Variation among younger speakers is consistent with the spread of a sound change during which some speakers are innovators merging pre-lateral /ʉːl-ʊl/, while others lag behind and maintain a contrast similar to older speakers [33]. However, younger speakers producing small contrast at the target may preserve duration contrast or contrast at some other point in the formant trajectories.

The inconclusive results on the /ʉːl-ʊl/ merger are consistent with the variation in young listeners' ability to distinguish /ʉːl/ and /ʊl/ in perception [11]. As spectral difference may fall above or below the just-noticeable difference, listeners are still able to discriminate between the two vowels, albeit less accurately [11]. Speakers producing /ʉːl-ʊl/ with a smaller spectral contrast may be less accurate at perceiving the same contrast, as speakers who produce a contrast less robustly are less likely to perceive it accurately [30]. More research is required to explore the relationship between listeners' perceptual accuracy with respect to the size of the spectral difference and its relation to just noticeable difference. Perceptual contrast reduction, a key marker of vowel merger [1, 2], may also be caused by contrast reduction in $F1$, in dynamic formant properties, and/or in durational differences [12]. Future research may explore the link between acoustic and perceptual contrast reduction using a combination $F1$, $F2$, durational values, and formant trajectories.

5. CONCLUSION

Pre-lateral /ʉːl-ʊl/ show an ongoing sound change in AusE, as younger speakers produce the pre-/l/ allophones with smaller $F2$ contrast than older speakers. Contrast reduction is caused by /ʉːl/ shifting towards /ʊl/. Yet, it is not clear if the $F2$ contrast reduction reaches the threshold of a merger or if the vowels are still contrastive.

6. ACKNOWLEDGEMENTS

This research was supported in part by the Australian Research Council DE150100318 and FT180100462 grants.

7. REFERENCES

[1] Ohala, J. J. 1989. Sound change is drawn from a pool of synchronic variation. In: Breivik, L. E., Jahr, E. H. (eds), *Language change: Contributions to the study of its causes*. Mouton de Gruyter, 173–198.

[2] Harrington, J., Kleber, F., Reubold, U., Schiel, F., Stevens, M. 2018. Linking cognitive and social aspects of sound change using agent-based modeling. *Topics in cognitive science* 10(4), 707–728.

[3] Garrett, A., Johnson, K. 2013. Phonetic bias in sound change. In: Yu, A. C. L. (ed), *Origins of Sound Change*. Oxford, UK: Oxford University Press, 51–97.

[4] Labov, W., Ash, S., Boberg, C. 2008. *The Atlas of North American English, Phonetics, Phonology and Sound Change*. Berlin, Boston: Gruyter Mouton.

[5] Strycharczuk, P., Scobbie, J. M. 2017. Fronting of Southern British English high-back vowels in articulation and acoustics. *JASA* 142(1), 322–331.

[6] Loakes, D., Hajek, J., Fletcher, J. 2017. Can you t[æ]ll I'm from M[æ]lbourne? *English World-Wide* 38(1), 29–49.

[7] Cox, F. 1999. Vowel change in Australian English. *Phonetica* 56, 1–27.

[8] Cox, F., Palethorpe, S. 2004. The border effect: Vowel differences across the NSW-Victorian border. Moskovsky, C. (ed), *Proc. Conference of Australian Linguistics Society* Newcastle, Australia.

[9] Palethorpe, S., Cox, F. 2003. Vowel modification in pre-lateral environments. *Int. Seminar on Speech Production* Macquarie University, Sydney.

[10] Szalay, T., Benders, T., Cox, F., Palethorpe, S., Proctor, M. 2021. Spectral contrast reduction in Australian English /l/-final rimes. *JASA* 149(2), 1183–1197.

[11] Szalay, T., Benders, T., Cox, F., Proctor, M. 2021. Perceptual vowel contrast reduction in Australian English /l/-final rimes. *LabPhon* 12(1).

[12] Szalay, T., Benders, T., Cox, F., Proctor, M. 2022. Vowel merger in Australian English lateral-final rimes: /æɔ-æ/. Billington, R. (ed), *Proc. 18th Australasian Int. Conf on Speech Science and Technology* Canberra, Australia, 106–110.

[13] Bernard, J. R. 1967. *Some measurements of some sounds of Australian English*. PhD thesis.

[14] Burnham, D., Estival, D., Fazio, S., Viethen, J., Cox, F., Dale, R., Cassidy, S., Epps, J., Togneri, R., Wagner, M., others, 2011. Building an audio-visual corpus of Australian English: large corpus collection with an economical portable and replicable black box. ISCA.

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

[16] Schiel, F. August 1999. Automatic Phonetic Transcription of Non-Prompted Speech. Ohala, J. J., Hasegawa, Y., Ohala, M., Granville, D., Bailey, A. C. (eds), *Proceedings 14th ICPhS* San Francisco, CA, USA, 607–610.

[17] Schiel, F. 2015. A statistical model for predicting pronunciation. *Proc. ICPhS*.

[18] Boersma, P., Weenink, D. 2021. Praat 6.1.39.

[19] Cosi, P., Falavigna, D., Omologo, M. 1991. A preliminary statistical evaluation of manual and automatic segmentation discrepancies. *EuroSpeech*.

[20] Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., Selker, R., Gronau, Q. F., Šmíra, M., Epskamp, S., others, 2018. Bayesian inference for psychology. part I: Theoretical advantages and practical ramifications. *Psychonomic bulletin & review* 25(1), 35–57.

[21] Dienes, Z. 2021. How to use and report Bayesian hypothesis tests. *Psychology of Consciousness: Theory, Research, and Practice* 8(1), 9.

[22] Bürkner, P.-C. 2017. brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software* 80(1), 1–28.

[23] Bürkner, P.-C. 2018. Advanced Bayesian multilevel modeling with the R package brms. *The R Journal* 10(1), 395–411.

[24] Bürkner, P.-C. 2021. Bayesian item response modeling in R with brms and Stan. *Journal of Statistical Software* 100(5), 1–54.

[25] Kewley-Port, D., Zheng, Y. 1999. Vowel formant discrimination: Towards more ordinary listening conditions. *JASA* 106(5), 2945–2958.

[26] Pillai, K. C. S. 1955. Some new test criteria in multivariate analysis. *The Annals of Mathematical Statistics* 26(1), 117 – 121.

[27] Hay, J., Warren, P., Drager, K. 2006. Factors influencing speech perception in the context of a merger-in-progress. *Journal of phonetics* 34(4), 458–484.

[28] R Core Team, 2021. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing Vienna, Austria.

[29] Giles, S., Moll, K. 1975. Cinefluorographic study of selected allophones of English /l/. *Phonetica* 31(3-4), 206–227.

[30] Harrington, J., Kleber, F., Reubold, U. 2008. Compensation for coarticulation, /u/-fronting, and sound change in standard southern British: An acoustic and perceptual study. *JASA* 123(5), 2825–2835.

[31] Wade, L. 2017. The role of duration in the perception of vowel merger. *LabPhon* 8(1).

[32] Gorman, E., Kirkham, S. 2020. Dynamic acoustic-articulatory relations in back vowel fronting: Examining the effects of coda consonants in two dialects of British English. *JASA* 148(2), 724–733.

[33] Rogers, E. M., Singhal, A., Quinlan, M. M. 2014. Diffusion of innovations. In: *An integrated approach to communication theory and research*. Routledge, 432–448.