Acquiring allophony: GOOSE and SCHOOL vowels in the speech of Australian children

Andy Gibson[a,b], Felicity Cox[a] & Joshua Penney[a]

[a]Macquarie University, Sydney; [b]Queen Mary University of London
gibsonism@gmail.com; felicity.cox@mq.edu.au; joshua.penney@mq.edu.au

ABSTRACT

The fronting of GOOSE has occurred in many varieties of English. In Australian English (AusE), prelateral tokens of GOOSE (referred to here with a lexical set label, SCHOOL) are the exception, and have become backed in recent generations. This study investigates the speech of 149 Australian children sampled through an online picture naming task over three timepoints, as they transition from preschool to school. Many children in the corpus speak a language other than English (LOTE) in the home. Higher levels of LOTE-usage are associated with less fronted (non-prelateral) GOOSE vowels. With respect to age, we find that while GOOSE does not appear to change across real-time, SCHOOL backs with increasing age. Additionally, children who have an older sibling show more retracted SCHOOL vowels. These findings suggest that SCHOOL-backing is an ongoing sound change, that may be subject to incrementation as children begin vernacular reorganisation.

Keywords: Australian English, child speech, GOOSE-fronting, sound change, diversity.

1. INTRODUCTION

In mainstream Australian English (AusE), as in many other varieties of English, the high /ʉː/ vowel of the GOOSE lexical set varies in its realisation from central–fronted [ʉː] in most phonetic contexts to retracted [uː] in prelateral contexts [1, 2]. This paper focuses on this allophonic split. We refer to the non-prelateral allophone of the /ʉː/ phoneme (in words such as *two*, *shoes*, *boot* and *juice*) as GOOSE, while we will refer to the prelateral allophone (in words such as *pool*, *stool*, *school*, and *cool*) as SCHOOL.

In the 1980s, [3] found that SCHOOL was retracted for speakers from Adelaide (and to a certain extent Melbourne) but not for speakers from Sydney or Brisbane, suggesting that the feature was strongly regional. Even as recently as 2008, [4:114] described retracted prelateral GOOSE as a feature of South Australian English, stating that 'In most regional varieties, similar vowel qualities occur for [GOOSE] vowels with or without a following lateral'. More recent research [1, 2, 5] suggests that the retraction of SCHOOL is a rapidly spreading change in progress for speakers from New South Wales (NSW). This SCHOOL-backing change, along with GOOSE-fronting, a change now nearing completion, leads to a strong split between the prelateral and non-prelateral allophones.

For preschoolers who speak languages other than English at home, the situation may be different. If a child's primary caregiver produces high back [uː] vowels, either in their L1 or in their (L2) English, this may impact upon a child's realisation of /ʉː/ in two ways: a backer GOOSE and possibly a less retracted SCHOOL. It should be noted, however, that even preschoolers do not always follow the model set forth by their caregivers: [6] found that, for 4–5 year old non-Anglo children growing up in London with parents born outside the U.K., there was no correlation between the children's and their principal caregivers' degree of GOOSE-fronting. They suggested that even at this age, children might reject variants used by their parents that are 'saliently non-local' [6:168]. Rather than focusing on ethnolinguistic background, in this study we look at the extent to which a given child uses a language other than English (LOTE) in the home, and consider whether this predicts degree of GOOSE-fronting.

1.1. Predictions for analysis

We assume based on recent analyses of AusE [1, 2, 5] that SCHOOL will be backer overall than GOOSE, but that the degree of this allophonic split will vary according to social factors. We have three specific predictions:

1. The more time a child spends speaking a language other than English in the home, the less fronted their GOOSE will be. This difference should not apply to SCHOOL which, if anything, may be more retracted for monolingual AusE speakers than for those who frequently use a LOTE at home;
2. If the backing of SCHOOL is an ongoing sound change, as shown in [5], we expect to see signs of incrementation [7], with retraction of SCHOOL over real-time;
3. If the backing of SCHOOL is an ongoing sound change, we also expect that children who have an older sibling will have a more

retracted SCHOOL vowel, having received early exposure to incremented variants.

The first prediction rests on the assumption that most of the LOTEs represented in the sample have a more retracted high-back vowel in their phoneme inventory than AusE. The children in the sample speak a wide range of languages, with the most frequently occurring being Mandarin, Urdu and Arabic. All of these languages include a high back vowel in their phonologies which may result in backer GOOSE vowels in the speech of caregivers who moved to Australia in adulthood [8-10]. In the analysis below, we examine the potential influence of speaking a LOTE at home, as well as exploring SCHOOL-backing as a change in progress.

2. METHODS

2.1. Participants and data collection methods

Children engaged in a self-recorded picture naming task delivered via the Gorilla online platform [11], 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 also 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. The project was approved by the Macquarie University Human Research Ethics Committee. Participants received compensation for their involvement.

This analysis includes recordings from 149 children (143 born in Australia and six who were born overseas but moved to Australia by 18 months of age). Almost all children are from New South Wales, with a large proportion from Sydney, and a cluster from the regional area of the Mid North Coast of NSW. Our original intention was to recruit children at three timepoints: before starting school, soon after starting school, and towards the end of the first year of schooling. However, due to limitations caused by the COVID-19 pandemic, including moving the task to open recruitment online, there was a high level of variation in the ages of the children. We present here mean age values at each timepoint to give an idea of ages and attrition rates. Our analysis in this paper however will use age in months, rather than timepoint, as the key independent variable for modelling real-time change. At their first recording timepoint (T1), the mean age was 59.1 months (n=148), at T2 (n=108) it was 66.5, and at T3 (n=97) it was 76.3.

Parents were invited to complete an in-depth questionnaire to provide details of the child's language and ethnic background. This revealed that 92 participants speak only English at home while 57 speak at least one language other than English (LOTE) at home. The survey also asked for the percentage of time each language is used in the home. This measure is used as a continuous predictor in the current analysis. Where this value was missing, proxy values were used. For those who had a LOTE listed as either their L1, or as their 'main language spoken at home' (or both), a value of 53.4% was assigned. For those with a LOTE listed as an 'other language spoken at home', a value of 24.7% was assigned. These values were chosen as being half a standard deviation either side of the mean % LOTE-usage for those who speak a LOTE at home (39.1).

2.2. Data processing and formant estimation

Data were orthographically transcribed by research assistants, and then processed through webMAUS [12] to provide phoneme level annotation. Boundaries were then hand-checked by phonetically trained research assistants who were instructed to use the automatic boundary unless it was obviously incorrect, in which case to place the boundary at the midpoint of the vowel-lateral sequence.

From an initial dataset of 7540 tokens of /ʉː/, we removed all prevocalic tokens, which can involve a linking glide and consequent coarticulation issues (n=1447), all instances of the lexical items *you* and *to*, which are often subject to reduction (n=864), and tokens that were found to be mislabelled (n=34).

Formants were estimated for the remaining tokens using FastTrack [13], choosing the best fit across a range of settings for the maximum formant value, ranging from 6000Hz–8000Hz. FastTrack estimated formants at seven time bins across each vowel, and the best fit was determined by smoothness of the formant trajectories.

Our analysis in this paper is entirely based on the midpoint measurement for each vowel. While an examination of the dynamics of these vowels would enrich this analysis, it was deemed to be outside the scope of the present paper. Given the automated nature of formant estimation, a stringent series of data pruning steps were then taken. The following 900 tokens were removed:

- Formant estimates missing (n=40).
- Midpoint F1 value > 800Hz (n=243).
- Midpoint measurement of F2 differed from either adjacent measurement by > 400Hz (n=441).
- Vowels shorter than 50ms (n=59).
- F2 measures > 2.5 standard deviations above, or < 2.5 standard deviations below

the mean, calculated separately for GOOSE (n=94) and SCHOOL (n=23).

For each of the above criteria, some tokens were listened to (by the first author) to determine a reasonable compromise between rejecting correctly tracked tokens and retaining mis-tracked tokens. After these exclusions, all tokens produced by five children who moved to Australia after the age of 18 months were also removed from the dataset (n=132). The final dataset used for the analyses below consisted of 4163 tokens (GOOSE n=3716; SCHOOL n=447) produced by 149 children.

3. RESULTS

3.1. Raw results

The mean F2 for children who speak only English at home was 2251Hz for GOOSE and 1028Hz for SCHOOL, compared to 2050Hz and 1077Hz, relpectively for children who speak a LOTE at home at least some of the time. Amongst these LOTE-speaking children, the mean F2 of GOOSE for speakers of Mandarin, Urdu and Arabic was 2019Hz, 1955Hz and 2074Hz, respectively.

3.2. Methods for statistical modelling

Linear mixed effects regression models were fit to predict the midpoint F2 of the /ʉː/ vowels, using *lme4* [14], considering the following four predictors:

- 'preL', that is, whether the following environment was /l/ or not, (a binary factor, non-prelateral GOOSE vs. prelateral SCHOOL);
- percentage of the time a LOTE is spoken at home (as a continuous predictor, centred and scaled), referred to as LOTE-usage;
- age in months (as a continuous predictor, centred and scaled);
- whether or not the child has an older sibling (a binary factor).

Random intercepts were included for participant and word. A slope for age on participant was also included since most children are represented at more than one age. This allows the model to predict the effect of age above and beyond idiosyncratic changes over timepoints for individual children. The slope for following environment on participant was also attempted but was abandoned as it led to singular fit errors. We followed a backward model selection process, beginning with the four-way interaction and sequentially removing least significant terms, using model comparison to confirm that each step in model simplification was justified.

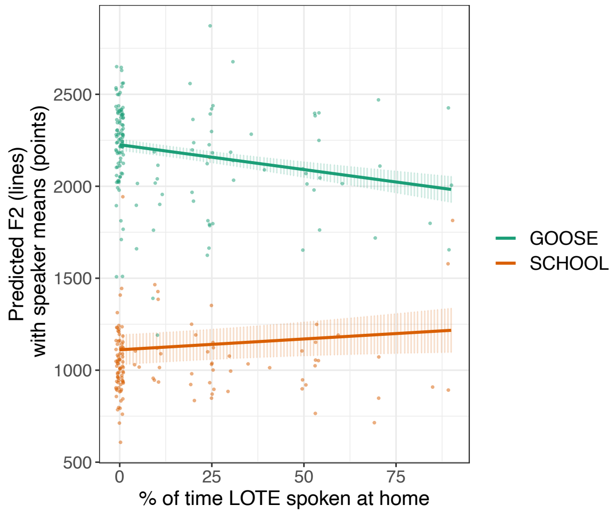

Figure 1: Predicted F2 (in Hertz) of GOOSE and SCHOOL according to percentage of time a given child speaks a language other than English (LOTE) at home.

3.3. Final statistical model

No four-way or three-way interactions were significant at an alpha of 0.05. The final model included three significant two-way interactions (all p<0.01), and the following syntax:

```
F2 ~ preL*OlderSibling + preL*LOTE-usage +
preL*Age + (1 + Age | Speaker) + (1 | Word)
```

Effect sizes from the model for each of these interactions are represented in Figures 1–3. Overall, a following /l/ robustly decreases the F2 of /ʉː/, that is, SCHOOL is consistently backer than GOOSE. The F2 of these two allophones varies as a function of LOTE-usage, age and having an older sibling.

Figure 1 shows that high levels of LOTE-usage are associated with backer tokens of GOOSE, and fronter tokens of SCHOOL. The figure shows the model predictions for this interaction with lines, along with points for the speaker means according to LOTE-usage, calculated separately for the prelateral and non-prelateral tokens.

Figure 2 shows the interaction of following environment with age. GOOSE does not change in any robust way with increasing age, but there is a strong tendency for SCHOOL to be more and more backed as the children get older.

Whether or not the child has an older sibling was found to significantly interact with following environment, as shown in Figure 3. Children with an older sibling have fronter GOOSE and more retracted SCHOOL. Though not included in the final model, there was a trend (p=0.077) for a three-way interaction between following environment, older sibling and age. This possible three-way interaction suggested that while SCHOOL retracts over real-time for all children, the slope of this change was steeper for children that have an older sibling.

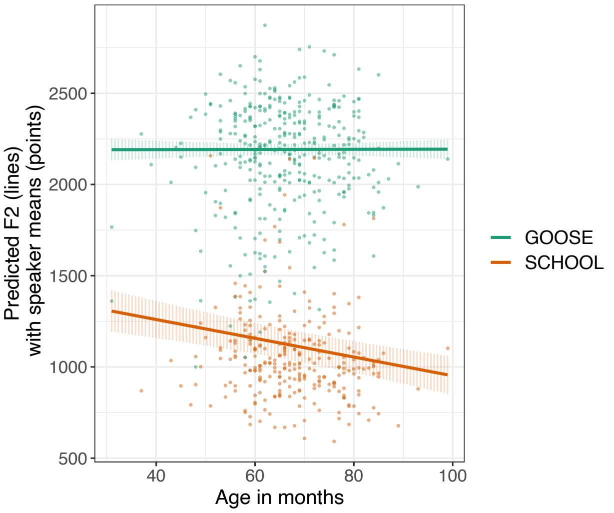

Figure 2: Predicted F2 (Hz) of GOOSE and SCHOOL according to the child's age in months.

4. DISCUSSION

4.1. Effect of LOTE-usage on GOOSE-fronting

Across the sample as a whole, we found a significant effect of the amount of LOTE-usage at home on the F2 of /ʉː/. Those who use a LOTE at home more often tend to have backer GOOSE vowels, supporting our first prediction. While this result may seem to depart from those of [6] discussed earlier, they are not directly comparable since we have focused on the extent to which children actually speak a language other than English in the home. Our results may thus capture transfer effects happening within the speech of these bilingual/multilingual children.

We were surprised at the non-significance of the three-way interaction of following environment, LOTE-usage, and age. We might expect that children who begin with a retracted goose would have begun to undergo rapid GOOSE-fronting to converge to their peers in the time scope of our recordings. The continuous LOTE measure used here may not have been the right measure to explore any fronting that may have occurred over real-time for the subset of LOTE-speaking children that had the most backed GOOSE vowels at the time of their first recording.

4.2. SCHOOL-backing as a change in progress

The second and third interactions reported revealed signs of incrementation of a change in progress, supporting our second and third predictions. While GOOSE showed no signs of change with increasing age, there was a significant backing of SCHOOL as children got older in real-time.

SCHOOL backs with increasing age, suggesting the real-time incrementation of an ongoing change in progress. These young children are acquiring and then incrementing the allophonic split between GOOSE and SCHOOL.

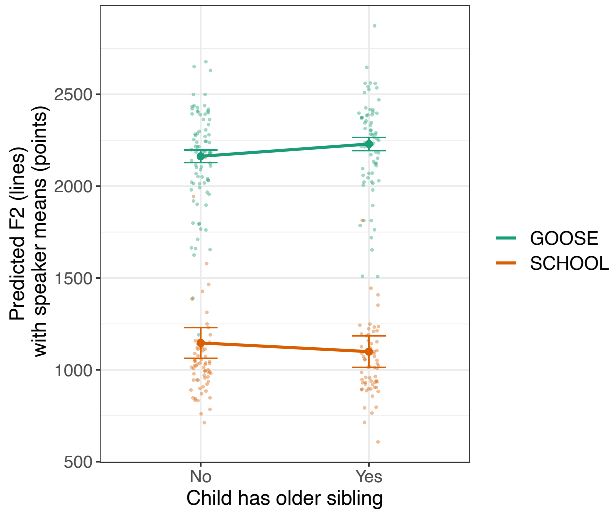

Figure 3: Predicted F2 (Hz) of GOOSE and SCHOOL according to whether the child has an older sibling or not.

Children with an older sibling have both fronter GOOSE and backer SCHOOL than children who lack a slightly older role model in the home. The presence of an older sibling may give these children a 'head-start' in extracting the direction of changes in progress. While we did not see evidence of GOOSE fronting in real-time, the older sibling effect for GOOSE suggests there may still be residual levels of GOOSE-fronting, as the change nears completion.

4.3. Limitations of automated methods

This study has used automated formant estimation methods, and it should be kept in mind that some tokens with formant tracking errors are almost certainly included in the results. While the gold-standard would be to check the formant tracking of every token, FastTrack is a useful and efficient tool for automatically achieving reasonable formant estimates from a large number of tokens. Additionally, the boundary between /ʉː/ and /l/ was largely determined by automated methods (though manually checked). Post-hoc analysis ruled out possible biases related to word-length: neither child age nor F2 of /ʉː/ correlated with word duration.

5. CONCLUSION

The allophonic split between GOOSE and SCHOOL appears to be an ongoing sound change for young speakers in NSW. Additionally, we see an effect of LOTE-usage on these vowels. A full understanding of this phenomenon needs to take into consideration the rich ethnolinguistic diversity in urban centres like multicultural Sydney.

6. ACKNOWLEDGEMENTS

This research was supported by Australian Research Council Grant DP190102164 and Australian Research Council Future Fellowship Grant FT180100462 to the second author.

7. REFERENCES

[1] T. Szalay, T. Benders, F. Cox, S. Palethorpe, and M. Proctor, "Spectral contrast reduction in Australian English /l/-final rimes," *J Acoust Soc Am,* vol. 149, no. 2, p. 1183, Feb 2021.

[2] B. Purser, J. Grama, and C. E. Travis, "Australian English over Time: Using Sociolinguistic Analysis to Inform Dialect Coaching," *Voice and Speech Review,* vol. 14, no. 3, pp. 269-291, 2020.

[3] H. Oasa, "Phonology of current Adelaide English," in *Australian English: The language of a new society*, P. Collins and D. Blair Eds. St Lucia: University of Queensland Press, 1989, pp. 271–287.

[4] D. Bradley, "Regional characteristics of Australian English: phonology," in *Varieties of English. Volume 3: The Pacific and Australasia*, K. Burridge and B. Kortmann Eds. Berlin: Mouton de Gruyter, 2008.

[5] T. Szalay, T. Benders, F. Cox, and M. Proctor, "Prelateral vowel change in Australian English POOL–PULL," in *Proceedings of the International Congress of Phonetic Sciences 2023*, Prague, 2023.

[6] J. Cheshire, P. Kerswill, S. Fox, and E. Torgersen, "Contact, the feature pool and the speech community: The emergence of Multicultural London English," *Journal of Sociolinguistics,* vol. 15, no. 2, pp. 151–196, 2011.

[7] W. Labov, *Principles of Linguistic Change, vol. 2: Social Factors*. Oxford: Wiley, 2001.

[8] N. Amir, O. Amir, and J. Rosenhouse, "Colloquial Arabic vowels in Israel: A comparative acoustic study of two dialects," *The Journal of the Acoustical Society of America,* vol. 136, no. 4, pp. 1895-1907, 2014.

[9] I. Rehman, "Urdu vowel system and perception of English vowels by Punjabi-Urdu speakers," University of Kent, 2019.

[10] B. D. Barkana and A. Patel, "Analysis of vowel production in Mandarin/Hindi/American-accented English for accent recognition systems," *Applied Acoustics,* vol. 162, p. 107203, 2020.

[11] A. L. Anwyl-Irvine, J. Massonnié, A. Flitton, N. Kirkham, and J. K. Evershed, "Gorilla in our midst: An online behavioral experiment builder," *Behavior Research Methods,* vol. 52, no. 1, pp. 388-407, 2020/02/01 2020.

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

[13] S. Barreda, "Fast Track: fast (nearly) automatic formant-tracking using Praat," *Linguistics Vanguard,* vol. 7, no. 1, 2020.

[14] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting Linear Mixed-Effects Models Using lme4," *Journal of Statistical Software,* vol. 67, no. 1, 2015.