Vowel perception in Victoria: variability, confusability and listener expectation.

Deborah Loakes, Simone Graetzer, John Hajek and Janet Fletcher

School of Languages and Linguistics, The University of Melbourne
dloakes; simone.graetzer; j.hajek; j.fletcher@unimelb.edu.au

Abstract

In the Australian state of Victoria, sound changes involving /el/-/æl/ are reported in production and perception. Previous pilot work has shown that /el/-/æl/ words are confusable for listeners from this region, and interestingly, that words also containing a nasal are typically recognised more poorly than words with no nasals (i.e. Ellen-Allen vs. Ellie-Allie). This study confirms that the sound changes in question affect vowel perception, and that Australian English listeners are actually aided by coarticulation to some degree.

1. Introduction

1.1. /el/->[æl] in Victoria.

In southern Victoria, including Melbourne, a sound change is in progress where /el/-/>[æl] occurs for some Victorians in production [1,2,3] and in perception (see e.g. [4]). Production work has shown that /æ/ lowers due to coarticulatory effects from the following velarised /l/, and for some speakers this lowering is substantial. Following Ohala’s theory of listener motivated sound changes [e.g. 5], it has been shown that listeners do in fact fail to compensate for coarticulation in many cases - commonly confusing /el/-/[æl], performing better when vowels are closer to their non-prelateral counterparts, and when /l/ is less dark [3,4,6].

An example of the coarticulatory effects of lateral /l/ compared to a non-lateral /d/ on preceding /e/ is shown in Fig. 1. This figure shows the F1/F2 formant dynamics across the vowels in the words hell and head. The image shows time normalised formants averaged for three young adult female speakers of Australian English who do not merge /el/-->[æl].

![Image](75x162 to 263x332)

Figure 1. F1 and F2 formant dynamics for the vowels in hell vs. head. Reproduced with permission from [6].

We see a rapidly falling F2 for the prelateral vowel in hell, in contrast with the primarily stable F2 in head. This equates to a rapidly retracting vowel in the former. In hell, F2 is approximately 300 Hz lower at onset, and F1 is around 100 Hz higher, resulting in a more open and retracted vowel in /el/.

Aside from coarticulatory processes involved in /el/ realisation, other observations from preliminary perception work are that Victorian as opposed to non-Victorian Australians make more mistakes, and report more difficulty, in distinguishing /el/-/[æl] (see e.g. [4]). Additionally, lexical frequency has also been shown to play a role in how well these listeners can distinguish such stimuli [3,4].

1.2. Another issue: /æl/->[/el]

A confounding factor as far as /el/ perception is concerned is a lesser-known raising phenomenon where /æl/->[æl]. The relationship between this, and /el/-/[æl], either as a whole, or for individual language users, is not yet clear.

It has been suggested [4] that while /el/-->[æl] - /el/->[æl] relate to the same phonemes and contexts, they may not necessarily be part of the same sound change. The other possibility is that these are two sound changes involving differing directionality: 1) /el/-->[æl] is a merger involving lowering/retraction; and 2) /æl/-->[æl] is a competing tendency for vowels to raise, and especially so in the presence of nasals.

Also relevant is that two variants of /æ/ compete globally [7]. The first is its lowering towards the /a/ vowel (which, e.g. separates northern and southern British). This has also been observed in Australian English [8], as well as for some young adult speakers in Melbourne [6]. In this case, /æ/ can have the percept of the STRUT vowel, and this is also true for some (merged) /el/ tokens for young Melbourne speakers. The second variant noted in world Engishes is /æ/ raising, consistent with /æl/-->[æl]. This raising occurs in New Zealand English, as well as some British and American varieties. Raising of /æ/ also occurs before nasals in Australian English, and this may be phonologised for some speakers [9].

In Australian English, greater confusion of /el/-/[æl] has been observed in words containing nasals, for example in Ellen-Alan as opposed to Ellie-Allie (see [4]). More generally, and as has been reported (and stereotyped) since 1844, Australian English is often considered hypernasal compared to other varieties (see e.g. the summary in [3]), and elevated word-level nasality or nasal spreading across syllables in the presence of a nasal consonant, is normal. Along this line, nasal coarticulation, from either direction, may be responsible for raising effects seen on /æl/ words containing nasals, and perhaps this effect has spread by analogy and/or through the phonetic effects of hypernasality to other /æl/ contexts. Given that the /el/-/[æl] distinction is variable and apparently unstable in Melbourne, and that prenasal /æ/ is subject to raising, it seems likely that /æl/ could move into /el/ space because of, or partly due to, these reasons.
2. Aims.

The aim of this study is to extend earlier experimental work which has investigated vowel perception in Victoria, so that the sound change(s) reported earlier may be further understood. The current work is also the last stage of a series of pilot studies informing a highly-controlled perception experiment (see e.g. [10]). Specific research questions for this study are: Q1) How do listeners react to various /el/ and /æl/ stimuli? Q2) Is the ability to distinguish /el/ and /æl/ words made more difficult when a nasal is present in the word? Q3) How well can listeners distinguish /e æ/ in non-nasal and non-lateral environments? Q4) How do listeners perform when asked to identify other vowels? Q5) Does perception of prelateral vowels change when /l/ coarticulation is manipulated?

Regarding Q3 (vowels not in prelateral and prenasal conditions), we need to be certain that listeners can reliably distinguish other vowel contrasts, independent of lateral/nasal effects. Therefore we focus on vowel contrasts in sound change environments, as well as more generally. Regarding Q5, we expect less confusion when /l/ is made clearer (less F2 movement), and more confusion in contexts where /l/ is made especially dark (more F2 movement).

3. Method

3.1. Participants

219 Australian English speaking listeners took part in this experiment. These were high-school students in their final year of study, aged primarily between 16-17 years of age. Participants were equally balanced for gender (108 M, 108 F, 3 unknown), and the majority (89%) were aged 16-18. The remaining 11% were adult teachers. The study was carried out when listeners attended an English Language workshop in 2011 at The University of Melbourne. Participants listened to linguistics-based lectures, and then participated in questionnaire-based research. During the lectures, they heard about language variation and change in Australia, but their attention was not drawn to the sound changes in question at any point.

In this study, we treat listeners who have lived outside Victoria for a year or more separately from those who have lived in the state their whole lives, under the assumption that listeners who have spent time elsewhere (where /el/-/æl/ is distinct) would be at an advantage when distinguishing /el/-/æl/ contrasts. This has been observed in earlier studies (see e.g. [4]), where such listeners performed better on such tasks. However, in those studies, a person was classified as “Victorian” if they had received all of their education in that state.

3.2. Stimuli and Experimental task

Stimuli were naturally produced words in various vowel contexts (Table 1). A small set was manipulated using AKUSTYK for Praat.

The experiment was questionnaire-based. Listeners read instructions, heard a male voice producing a word, and made a forced-choice decision about which word they had heard. The presentation mode has limitations in that distance from the speakers was variable and we are uncertain if all listeners had normal hearing. On the other hand, benefits include the large participant numbers and the fact that this method has also worked well in previous research (e.g. [4] as well as Peterson and Barney’s famous vowel study [11]).

Table 1 shows the contrasts tested. For ease of reference, contrasts have been grouped into categories. For example the code VL is used to refer to all vowels preceding /l/, aside from /e æ/. In this case, we played a word such as belt and gave listeners the options belt-built. While all items played to the listeners are shown in the final column, the actual order that the words were shown on the page differed. Where an item was shown twice, such as the pan-pen contrast, in one instance listeners saw pan on the left, and vice-versa for the other.

Table 1. Codes used.

<table>
<thead>
<tr>
<th>Code</th>
<th>Context</th>
<th>Pairs shown</th>
</tr>
</thead>
<tbody>
<tr>
<td>EAL</td>
<td>prelateral /e æ/</td>
<td>hell-Hal pell-pal (x2) Allie-Ellie Ellen-Alan Mel-Mal</td>
</tr>
<tr>
<td>EALC</td>
<td>prelateral /e æ/ preceding (manipulated) clear /l/</td>
<td>Mal-Mel hell-Hal Hal-hell</td>
</tr>
<tr>
<td>EALD</td>
<td>prelateral /e æ/ preceding (manipulated) dark /l/</td>
<td>Ellen-Alan</td>
</tr>
<tr>
<td>EAN</td>
<td>prenasal /e æ/</td>
<td>pan-pen (x2)</td>
</tr>
<tr>
<td>EAO</td>
<td>/e æ/ preceding non-nasal and non-lateral consonants</td>
<td>peck-pack (x2) head-had</td>
</tr>
<tr>
<td>VL</td>
<td>Vowels preceding /l/, aside from /e æ/</td>
<td>mole-maul built-belt (x2) pull-pool foil-foillillyfeely</td>
</tr>
<tr>
<td>VLC</td>
<td>Vowels preceding clear /l/ (manipulated), aside from /e æ/</td>
<td>pull-pool</td>
</tr>
<tr>
<td>VN</td>
<td>Vowels preceding nasals</td>
<td>meant-mint dent-dint (x2) sum-psalm mint-meannt lint-lent him-hem penned-pinned pin-pen corn-corn</td>
</tr>
<tr>
<td>VO</td>
<td>Vowels preceding non-nasal and non-lateral consonants, aside from /e æ/</td>
<td>hood-horde pick-peck</td>
</tr>
</tbody>
</table>

4. Results

4.1. Broad listener responses.

The percentage of correct responses for each category of contrasts is shown in boxplots in Figure 2. Overall, the figure shows that no item was identified correctly by 100% of listeners (the best perceived item had a 94% accuracy rate - see below). By category, the range of results was relatively wide, and some categories contained outliers which were particularly divergent (as such, results are also discussed according to particular lexical items).

As a group, the best perceived categories were EAO and VN. EAO is the group which consists of /e æ/ words preceding non-nasal and non-lateral consonants, and this result
confirms that listeners were able to perceive the /e æ/ contrast in pre-/d/ and pre-/k/ environments relatively well. VN is the vowel contrast preceding nasals. While one of these items was particularly well perceived, this category also contained the item that was most poorly identified. The item best identified in the VN category was corn-cone (best overall with 94%). A good accuracy rate is unsurprising for this pair given it is a monophthong-diphthong contrast. The item identified least correctly, at 36%, was dint-dent, although when presented the other way around, dent-dint, accuracy was 93%.

listeners appreciated having the dramatically falling F2 in this context, with 85% of listeners correctly choosing Ellen here. Exactly the same item but with a naturally occurring (relatively dark, but not as dark) /l/ and presented differently as Alan-Ellen had only a 62% correct identification score. It appears then that enhanced velarization can counteract perceptual confusion effects triggered by nasalization.

Analysis of the list of correctly identified tokens shows that /æ/-pool contrasts where a nasal is present in the word (Mel-Mal, Ellen-Alle) do tend to be more poorly identified, clustering around 62-70% correct identifications.

Figure 2. Boxplots showing percentage of correctly identified items by context.

Other categories, VL and VO, were also identified well. These groups are the vowel contrasts other than /e æ/ before laterals, and vowel contrasts other than /e æ/ before other consonants. The group VL had an outlier which was particularly poorly identified (at a rate of 55% accuracy), and this was for pull-pool, a contrast which is known to be diminishing through another ongoing sound change in Australian English and maintained by length only [1].

Items relatively poorly identified tended to be the cases we might predict based on the sound changes in progress in Australian English. These categories are the ones containing /e æ/ before laterals and nasals (EAL, EALC, EAN) as well as pull-pool (which is the item in the VLC category). Nevertheless, there were two outliers in the EAL category which were particularly well identified; pal-pell and hell-Hal (see below).

EALC and VLC contain the manipulated clear lateral. We had included these categories to try and replicate a situation in which listeners experience a less coarticulated vowel in these prelateral contexts, and we had thought that this would allow better identification rates for /e æ/ as has been seen for naturally occurring speech (where less coarticulated vowels have better recognition rates) [e.g. 4]. However, the reverse was true here, with listeners making more errors in EALC than EAL, and in VLC compared to VL. This result indicates that listeners are most likely expecting a dramatically lowering F2, and are prone to making errors when the stimulus deviates too much from this (i.e. /l/ was evidently made too clear). Likewise, we had thought that creating a particularly dark /l/ token in EALD, for the intervocalic Ellen-Alan, would result in poorer identification. However, it appears that

As well as analysing the percentage of correctly identified items, we also present (Fig. 3), the items not answered at all by listeners as these also give insight into perceptual confusions. Results here correlate, for the most part, with the correctly identified items seen in Fig. 2. The category that listeners failed to answer most was EALC, while the item that was most often unidentified by listeners (at a rate of 45%) was the dint-dent pair which also had the most incorrect responses. While Fig. 3 is useful for giving overall percentages of unanswered items, and showing which categories they belong to in general, a report of some of the other most commonly unanswered items is also illustrative of patterns in the data. These are Hal-hell (EALC, 35% unanswered), pin-pen (23% unanswered), Mel-Mal (21%), pinned-penned (20%), hell-Hal (EALC, 20%), hem-him (20%), peck-pack (20%), had-head (19%).

With respect to these results, the fact that peck-pack and had-head were not answered by one fifth of participants means that we need to qualify the “good results” observed for these contrasts earlier. Evidently, identification rates of /e æ/ were good amongst those who answered the question, but some listeners still had trouble with these categories. The other items that listeners failed to respond to most were all contrasts including /i e/ prenasally, as well as /e æ/ prelaterally, and /e æ/ when a nasal precedes the contrast.

4.2. Statistical analysis of results using generalized linear mixed models.

A generalised linear regression model (LMM) was fit to the binary data (using the ‘glmer’ function in the lme4 package in
the R programming language) in a sample of 219 hearers. The binary response “answer” (T = correct, F = incorrect) was associated with five potential explanatory variables: segmental context (nine levels: EAL, EALC, EALD, EAN, EAO, VL, VLC, VN, VO), hearer sex/gender (two levels: male, female), Victorian residence (two levels: year or more outside of Victoria, less than a year outside Victoria), father’s language (two levels: native speaker of Australian English or otherwise) and mother’s language (two levels: native speaker of Australian English or otherwise). Hearer was included as a random factor. Alpha was set at 0.05. Of the fixed factors, only segmental context had a significant effect on the response: EAN, EAP, VL, VLN and VN coefficients were significant at p<0.0001 and VO at p<0.05, but the EALC coefficient was not at p=0.15 (EAL was treated as the intercept).

Pairwise post-hoc comparisons were performed for the significant factor of segmental context with Tukey’s HSD method. These showed: 1) EAL responses were less often correct than EAN, EAO and VL responses. 2) EALC responses were less often correct than EAN and VL responses. 3) EAO responses were less often correct than EALC, EAN, VL, VLC, VN and VO responses. 4) EALD responses were less often correct than EAN and VL responses. 5) EAN responses were more often correct than VLC, VN and VO responses. 6) VL responses were more often correct than VLC, VN and VO responses.

5. Discussion

Results confirm that Australian English listeners have the most trouble discriminating /el/-/æl/ contrasts (Q1 and Q3), /el/-/æl/ contrasts when a nasal is present in the word (Q2), prenasal front vowels /i æ/ (Q2), pool-pull contrasts (Q4), and words where an exaggerated clear lateral has been falsely added to a usually dark context (Q5). In other words, all environments where sound changes are in progress are harder for listeners to discriminate than vowel contrasts in general. However, it must be said that overall the /æ/ vowels contrast is not especially robust (being acoustically similar and typologically rare, as well as likely to change in English [8]), and that while listeners could discriminate peck-pack and head-had quite well, a relatively large number failed to respond to these stimuli at all, indicating some level of confusion, but not to the extent seen for prelateral and prenasal vowels. Finally, atypical clear /l/ in syllable final contexts goes against listener expectation, creating perceptual confusion, while a falsely dark /l/ actually aids perception. This finding about coarticulation aiding perception has been seen elsewhere, for example [12]. As far as Australian English /l/ is concerned, this suggests a complex relationship between coarticulation that drives sound change in the first place, and its place in discrimination of contrasts generally.

While both /el/-/æl/ were confusable to varying degrees for many listeners, some words were better identified than others. Words without nasals were better identified than those containing nasals. Second, the stimuli pal-pell and hell-Hal were recognised well overall, which may be due to lexical frequency and category (real word v. non-word or name) effects.1 Other problematic contrasts include those which vary in the F1 dimension, especially the prenasal and prelateral /i æ/

1 Pell could also have been recognised as a surname (cf. Australian Catholic primate, Cardinal George Pell).

Australian English listeners are good at perceiving vowel contrasts in environments containing contrasts that are thought to be relatively invariable (e.g., corn-cone, etc.). However, in contexts where sound changes are reported (esp. lax front vowels in prelateral and prenasal contexts), perceptual confusion arises. Additionally, when a nasal is present in an /el/-/æl/ word, a greater level of misperception is observed amongst listeners, although this effect can be reduced if counterbalanced by greater lateral velarization. All of this points to significant interactions between different, apparently unrelated, phenomena in the language, which requires further investigation and analysis.

6. Conclusions

Australian English listeners are good at perceiving vowel contrasts in environments containing contrasts that are thought to be relatively invariable (e.g., corn-cone, etc.). However, in contexts where sound changes are reported (esp. lax front vowels in prelateral and prenasal contexts), perceptual confusion arises. Additionally, when a nasal is present in an /el/-/æl/ word, a greater level of misperception is observed amongst listeners, although this effect can be reduced if counterbalanced by greater lateral velarization. All of this points to significant interactions between different, apparently unrelated, phenomena in the language, which requires further investigation and analysis.

7. References

Author/s:
Loakes, DE; GRAETZER, N; Hajek, JT; Fletcher, J

Title:
Vowel perception in Victoria: variability, confusability and listener expectation

Date:
2012

Citation:

Persistent Link:
http://hdl.handle.net/11343/129769

File Description:
Published version