Pre-Liquid Excrescent Schwa:
What Happens when Vocalic Targets Conflict

Bryan Gick and Ian Wilson

Department of Linguistics
University of British Columbia, Canada
gick@interchange.ubc.ca

Abstract

Sequences of high tense vowel + liquid in English often result in the percept of an intervening schwa, as in, e.g., *heel, hail, hire*. We argue in this paper that this apparent schwa is simply the incidental acoustic result of the tongue moving through "schwa-space" (a schwa-like position) during the transition between conflicting tongue root targets. This conflict bears on both articulatory timing relationships in syllable codas and tongue root specification for tense vowels. We present two experiments: Experiment 1 shows that excrescent schwa does not correspond with greater duration of syllable rimes; Experiment 2 shows that the tongue moves through schwa space along its trajectory in the excrescent schwa cases. Our results support a timing model whereby coda timing is determined by the relationship between syllable peak and consonant closure, but where timing is unaffected by the number of intervening vocalic events.

1. Introduction

It has been observed that sequences of high tense vowel + liquid in English result in the percept of an intervening schwa (e.g., *hee*[ə]*l, hai*[ə]*l, hi*[ə]*re*). Previous studies have given this phenomenon special phonological status, either by specifying a process of epenthesis [1], or by licensing trimoraic syllables in the phonology [2, 3].

We will argue in this paper that the percept of schwa in these cases is simply the incidental acoustic result of the tongue moving through "schwa-space" (a schwa-like articulatory and acoustic manifestation) during the transition that occurs between conflicting tongue root targets (from an advanced to a retracted position, as in, e.g., the movement from /j/ to /l/ in a words like *feel* and *file*). This conflict, we suggest, is particularly interesting in that it bears on issues concerning both articulatory timing relationships in syllable codas and tongue root specification for tense vowels.

In order to test this hypothesis, we will use acoustic, magnetometric and ultrasound data to evaluate durations, targets, and acoustic outputs in these cases. Our first experiment supports the automaticity of this effect by showing that the presence of excrescent schwa does not correspond with greater duration of syllable rimes (contrary to [1, 2, 3]). Our second experiment shows that the tongue moves through schwa space along its trajectory in excrescent schwa cases.

1.1. Previous analyses

1.1.1. Phonological analyses

Two recent analyses have been proposed to account for the excrescent schwa phenomenon in English. McCarthy [1] claims (p. 198) that "a glide + liquid sequence presents too small a sonority cline" (see also [4]). Consequently, the liquid "cannot be syllabified with the preceding diphthong and schwa epenthesis applies instead." This account predicts that other forms with equally small sonority clines should elicit epenthesis. However, words such as *barn* and *bust* do not surface as [báɹən] or [bʌsət], nor does epenthesis occur even in codas with no sonority cline (e.g., *act*) or a negative cline (e.g., *adze*). This apparent epenthesis thus appears to result from qualities specific to the liquids and tense vowels.

Lavoie & Cohn [2] propose an alternative analysis, whereby tense vowel + liquid sequences constitute trimoraic syllables (which they refer to as "sesquisyllables"). While this approach holds for the high vowels, it does not explain why the low vowel /a/ does not elicit the schwa (e.g., *hall*). Their analysis of the low vowels, proposed in [3], simply stipulates the constraint that "r/l can bear a mora after [-low] vowels but not [+low] vowels."

Another problematic case for Lavoie & Cohn is that they find that the vowel /o/ groups with the low/lax vowels in not eliciting the apparent schwa (e.g., *hole*; see [2], p. 111). In his study of tongue root positions in English vowels, MacKay [5] identifies another case where /o/ patterns with the low/lax vowels: Here, /o/ was the only non-low "tense" vowel with a non-advanced tongue root position. Both of these findings, though exceptional in the context of their previous studies, converge to support the analysis we propose in this paper: As there is no conflict in tongue root position (i.e., as both /o/ and the liquids are retracted), we do not expect the transition to move through schwa space. We will pursue this connection between tongue root position and excrescent schwa throughout the rest of this paper.

1.1.2. Rime durations

Previous data on rime durations support the view that the excrescent schwa does not contribute to the duration of the syllable rime. Lavoie & Cohn [2] give duration measurements for various three-segment sequences, including low vowel + liquid + stop (e.g., *-ald*), low vowel + glide + stop (e.g., *-ajd*) and low vowel + glide + liquid (e.g., *-ajl*). If indeed an additional timing unit or syllable is present in the glide + liquid cases, then we should expect to see an increased duration in *-ajl* relative to *-ald* and *-ajd*. On the contrary, the durations cited for the *-ald* and *-ajd* cases are 213ms and 185ms, respectively (average = 199ms), while that of *-ajl* is 201ms. Thus, the schwa perceived in the glide-liquid combinations does not appear to involve additional timing in the syllable.

Sproat & Fujimura [6] propose that English /l/ is composed of two distinct types of gestures: A "consonantal" gesture (tongue tip [TT] raising) and a "vocalic" gesture (tongue dorsum/root [TR] retraction). They base this categorization in part on the timing relationship between the two gestures: While both of these occur more or less simultaneously in prevocalic allophones, TR retraction occurs much earlier than TT closure in postvocalic allophones. While this observation is certainly relevant to the present discussion, it is unknown whether these timing generalizations extend to postvocalic clusters. Duration facts of the kind discussed in the previous subsection have important implications for this extension.

Consider three syllable rimes parallel to the example types cited above in 1.1.2, where all three rimes involve different numbers of tongue retraction/advancement movements, e.g., /-id/, /-ajd/, and /-ajl/. Of these three rimes, /-id/ involves only a single vocalic gesture (tongue advancement for /i/); /-ajd/ involves two contrary vocalic movements (a retraction for /a/ followed by an advancement for /j/); and /-ajl/ involves three conflicting movements (retraction for /a/, advancement for /j/, and retraction again for /l/). If the overall duration is indeed the same across all of these rimes, then we should expect the temporal lag between vocalic and consonantal gestures to decrease as the number of conflicting movements increases. A schematic illustration of this is shown in Figure 1.
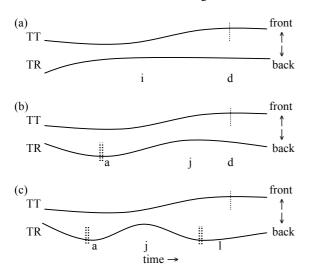

Figure 1: Schematic diagram of tongue tip (TT) vs tongue rear/root (TR) movement in (a) -*id,* (b) -*ajd,* and (c) -*ajl* syllable rimes. Single dotted lines show estimated time of achievement of the tongue tip closure gesture for the final consonant; double dotted lines show time of achievement of TR retraction. Note multiple retraction gestures in (c).

This model suggests that no stable timing is maintained within segments postvocalically. Rather, syllable coda timing is a relationship between the syllable peak and the first consonantal gesture, with vocalic gestures compressed into the available time window. The prediction that the timing between vowel onset and consonant closure is the same regardless of intervening vocalic gestures will be explicitly tested in Section 2 below.

Gick [7] provides magnetometer data consistent with this model, showing that the temporal "lag" between the TR and TT gestures of /l/ are shortest in /-ojl/ and /-ajl/ rimes (those with the largest number of contrary TR targets). Data from [7] are shown in Figure 2.
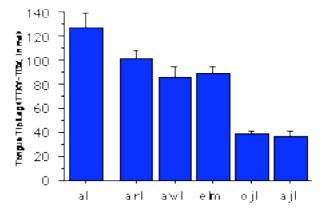

Figure 2: Tongue tip lag in postvocalic /l/ allophones (figure reproduced from Gick [7]). The three sets (al, arl/awl/elm, and ojl/ajl) are all statistically distinct.

2.   Experiment 1: Duration

An experiment was conducted to test the hypothesis that durations from vowel onset to final consonant closure are stable regardless of the number of intervening vowel gestures and regardless of the segment type.

2.1. Methods

2.1.1.   Subjects

Two native speakers of North American English participated in this study, both in their late 20's, and both unaware of the nature of the experiment. W1 (female) was raised in Southwestern Ontario; M1, (male) was raised in Southwestern Manitoba.

2.1.2.   Stimuli

Stimuli were presented in the carrier phrase "Pop is a __ ." Tokens included real and nonsense words of the form pV(G)C, where V(G) consisted of the set /a, i, aw, ej, aj, ɔj/, and where C consisted of the set /d, n, l/, giving 18 combinations.

2.1.3.   Data collection and analysis

Six repetitions of each token were collected as follows: All tokens were presented in writing to the subjects, who read them aloud; the entire list was repeated six times. The first reading of the list was discarded to ensure that subjects were accustomed to the procedure. Tokens within the list were presented in blocks of six, of which the sixth member was discarded.

Articulatory data were recorded to VHS from an Aloka SSD-900 portable ultrasound machine using a 3.5MHz electronic convex intercostal probe UST-9102 with a 90-degree field of view. The probe was held by the subject against his or her own neck, just above the larynx, so that a midsagittal section of the tongue was visible from the tongue root to the tongue tip. A constant probe position was maintained using a laser pointer attached to the probe.

Subjects were seated facing a wall at a distance of 2 meters. The laser pointer projected an image of crosshairs onto the wall, where a 10cm square was drawn. Subjects were instructed to keep the crosshairs upright and within the square; their accuracy was monitored by the investigators during the experiment.
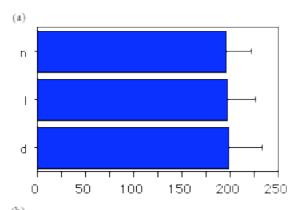
The acoustic signal was simultaneously recorded to VHS to ensure synchronization with the video signal, using a Pro-Sound YU-34 unidirectional dynamic microphone amplified through a built-in amplifier in a Tascam cassette recorder.

After collection, videos were digitized to a Macintosh G4 from the VHS tape using an XLR8 video card with Final Cut Pro v.1.2 video editing software. Images were edited and analyzed using Final Cut Pro.

Acoustic signals were analyzed using the freeware Praat v.3.9.13 (http://www.fon.hum.uva.nl/praat/). Durations were measured from vowel onset to consonant closure. Consonant closure was identified using a combination of audio cues from the waveform and articulatory cues from the ultrasound signal. Several tokens were omitted from the study in cases where the location of closure was not clearly identifiable.

2.2. Results

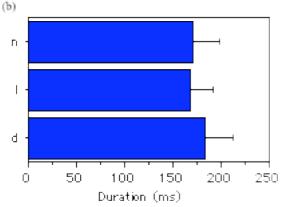The results of the duration experiment are shown in Figure 3.

Figure 3: Rime durations from vowel onset to coda consonant closure for subjects (a) W1 and (b) M1.

ANOVA results indicated that the differences shown in Figure 3 were not significant for either W1 ($F_{(2, 81)} = .053$; $p = .9486$) or M1 ($F_{(2, 71)} = 2.025$; $p = .1396$); post-hoc tests (Fischer's PLSD) were also not significant ($p > .05$) between separate conditions. In other words, across vowels, there were no significant differences in duration between rimes ending in the three different final consonants measured (pVn vs. pVl vs. pVd).

Within-vowel comparisons were also made. For W1, no differences were significant within any vowel; for M1, /d/ was significantly longer than /l/ following /a, ej/, and longer than /n/ following /ɔj/; /n/ was longer than /l/ following /a/; and /l/ was longer than /d, n/ following /i/.

2.3. Discussion

These results support our prediction that the presence of the excrescent schwa does not contribute to the duration of the syllable. Further, the results also support the more general model of coda timing presented in Figure 1, wherein coda timing is not systematically affected by the number of intervening vocalic gestures.

3. Experiment 2: Schwa space

Experiment 2 tests whether the tongue moves through articulatory and acoustic schwa space along its trajectory in the excrescent schwa cases.

3.1. Methods

The same methods were employed as described in Experiment 1 above, except as follows: Repetitions of the form [pápə] were also collected. The canonical schwas in these forms were used as controls in comparison with excrescent schwas in the forms collected in the previous section. Formants were calculated at the midpoint of canonical schwas, and compared with their crossover points in words with excrescent schwas. A linear interpolation Burg LPC was used to automatically extract formant trajectories, with a time step of 10ms, window length of 25ms, and pre-emphasis from 50 Hz.

3.2. Results

The mean formant values for the final canonical schwas were: (for M1) F1=658 Hz, F2=1217 Hz, F3=2539 Hz; and (for W1) F1=710 Hz, F2=1550 Hz, F3=2980 Hz. These formants were used to identify the crossover point in the transition from /j/ to /l/ in the word *pile*. Our findings show that in the region at approximately the midpoint of this transition, during the schwa percept, both F1 and F2 cross from above to below the formant values recorded for canonical schwa. Example ultrasound images of the tongue shapes during both excrescent and final schwa are shown in Figure 4, and formant trajectories are compared in Figure 5.

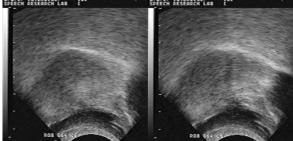(a) excrescent schwa (*pile*)   (b) canonical final schwa

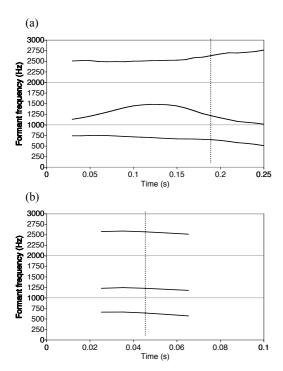*Figure 4*: Ultrasound images of tongue shape for (a) excrescent and (b) canonical schwa (M1).

(a)

(b)

Figure 5: Example formant trajectories for (a) *pile* and (b) canonical final schwa. The vertical dotted line in (a) shows the point of intersection with canonical schwa F1 and F2; and in (b) shows the midpoint, where measures were made.

3.3. Discussion

The results of Experiment 2 show that, as predicted, the transitions in the excrescent schwa cases move through canonical articulatory and acoustic schwa space along its trajectory. Thus, the percept of schwa in these cases emerges phonetically as a result of more general properties of syllable rime timing.

More importantly, we propose that the presence of the conflicting tongue root targets in the excrescent schwa cases indicates that these targets are not incidental, but rather that the TR gestures for both English liquids and non-low tense vowels must be specified.

4. Conclusion

We have argued in this paper that the apparent schwa in English tense vowel + liquid combinations is simply the incidental result of the tongue moving through canonical "schwa-space" during the transition forced by conflicting tongue root targets. More importantly, we suggest that this forced transition indicates that the TR targets of English liquids and non-low tense vowels must be specified. Finally, this analysis supports a timing model whereby coda timing is determined by the relationship between syllable peak and consonant closure, but where timing is unaffected by the number of intervening vocalic events.

5. Acknowledgments

6. References

[1] McCarthy, J. J. "Synchronic rule inversion," *Proceedings of the Berkeley Linguistics Society* 17:192-207, 1991.

[2] Lavoie, L. M. and Cohn, A. C. "Sesquisyllables of English: the structure of vowel-liquid syllables," *Proceedings of the XIVth International Congress of Phonetic Sciences*. San Francisco, 109-112, August, 1999.

[3] Cohn, A. C. and Lavoie, L. M. "English vowel-liquid monosyllables: a case of trimoraic syllables," Paper presented at Universtiy of Massachusetts, Amherst, Linguistics Colloquium Series, March, 2000.

[4] Steriade, D. "Greek prosodies and the nature of syllabification," Doctoral dissertation, Massachusetts Institute of Technology, 1982.

[5] MacKay, I. R. A. "Tenseness in vowels: an ultrasonic study," *Phonetica* 34: 325-351, 1977.

[6] Sproat, R. and Fujimura, O. "Allophonic variation in English /l/ and its implications for phonetic implementation," *Journal of Phonetics* 21: 291-311, 1993.

[7] Gick, B. "The organization of segment-internal gestures," *Proceedings of the XIVth International Congress of Phonetic Sciences*. San Francisco, 1789-1792, August, 1999.