Excrescent schwa and vowel laxing: Cross-linguistic responses to conflicting articulatory targets*

Bryan Gick and Ian Wilson

Physical conflicts between articulatory goals are inevitable in speech. One such conflict is investigated in this paper, namely that between posterior tongue targets of consonants and vowels. Multiple strategies are identified across dialects and languages for dealing with these conflicts, and described in terms of language-specific rankings of phonological constraints on the organization of articulatory gestures. In English, sequences of high tense vowel + liquid have been observed to result in the percept of an intervening schwa in some dialects (as in, e.g., heel*,* hail*,* hire*, etc.), giving many listeners the impression of an "extra" syllable or half-syllable (cf.* higher*). While previous studies have given this phenomenon special phonological status, the present paper argues that this schwa is the phonetic result of one possible strategy for reconciling an intrinsic conflict between articulatory (tongue dorsum/root) targets. This same conflict, however, results in vowel laxing/lowering in other dialects, implying a different strategy. Preliminary laboratory data are presented, first supporting the notion that the schwa percept is the incidental result of the tongue passing through a schwa-like configuration or "schwa space" during the transition between opposing tongue root targets. Second, data are presented for several additional languages supporting the notion that similar cases of articulatory conflict occur in other languages, often resulting in similar effects. The constraint-based proposal is evaluated in light of these cross-language findings.*

1. Introduction

It has been observed that sequences of high tense vowel + liquid in many dialects of English may elicit the percept of an intervening schwa (as in, e.g., *hee*[ə]*l*, *hai*[ə]*l*, *hi*[ə]*re; cf. higher*), giving many listeners the impression of an "extra" syllable or half-syllable. Previous studies have attributed special phonological status to this phenomenon, either by stipulating a process of phonological epenthesis (McCarthy 1991), proposing segment-specific constraints (Halle and Idsardi 1997; Orgun 2001), or licensing trimoraic syllables in the phonology (Lavoie and Cohn 1999; Cohn and Lavoie 2000).

It is argued in this paper that this impression of an epenthesized schwa is not an isolated phonological process restricted to English, nor is is a schwa percept itself systematically regulated by the grammar (thus this paper will not include a perception study, nor a complete explanation of why a certain string may or may not be perceived as a vowel in general). Rather, this vowel is argued to be an example of "excrescent" schwa (as described by Levin 1987) – in the present case, a phonetic by-product of one strategy for reconciling an intrinsic conflict between articulatory targets. The specific conflict under investigation appears, for example, in words like *feel* and *file* in many English dialects between an advanced tongue root/dorsum target for the palatal vowel or glide, and a retracted target for the following uvular/upper pharyngeal constriction for /l/ [see Narayanan, Alwan, and Haker 1997 and Gick, Kang, and Whalen 2002 for discussion of the dorsal constriction for /l/]). A similar conflict obtains for the retracted tongue root position of /r/ in many dialects (Delattre and Freeman 1968). The strategy invoked here is thus simply to allow both of the conflicting goals to be achieved, forcing a serial ordering with an audible transition.

Another logically possible strategy for resolving a conflict between two targets, however, is to reduce or eliminate one of the conflicting targets (Wood 1996). Indeed, this same sequence of high tense vowel + liquid appears elsewhere in the literature on English dialectal phonology, but with a different result: In Utah English (Di Paolo and Faber 1991) and Pittsburgh English[1] (Walsh Dickey 1997), the vowels in words like *feel*, for example, have laxed or lowered to merge with *fill*, *pool* with *pull*, etc.

The presence of these two distinct strategies is taken to indicate that responses to this physical conflict are not automatic, but are controlled language-specifically. Thus, a proposal is made in the present paper expressing these strategies in terms of rankings of constraints on gestural organization. These rankings are argued to result in excrescent schwa in some languages, vowel laxing in others, and a context-dependent combination of excrescent schwa and laxing in still others. Using this English case as a stepping-off point, this paper also investigates similar cases of conflict across a number of languages to test for the occurrence of similar strategies.

The present paper is organized as follows: Section 2 focuses on the English case, evaluating previous analyses of the English excrescent schwa, developing support for the proposal that these phenomena are in fact the result of articulatory conflict, and proposing a constraint-based account for gestural coordination that can accommodate these observations. In Section 3, previous descriptive data and new pilot laboratory data are presented from several languages beyond English to test for similar effects, and to find whether the additional patterns predicted by the constraint-based account exist in the languages of the world; languages include Beijing Chinese, Nuu-chah-nulth (Wakashan), Chilcotin (Athapaskan) and Korean. Section 4 summarizes the cross-linguistic results and evaluates the constraint-based analysis in light of these data.

2. Excrescent schwa and vowel laxing in English

2.1. Previous analyses of excrescent schwa

As discussed above, three recent approaches have been used to account for the excrescent schwa phenomenon in English. McCarthy (1991: 198) claims that "a glide + liquid sequence presents too small a sonority cline" (see Steriade 1982). 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 [bʌ́sət], 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 more specific to the liquids and tense vowels.

Orgun (2001), following an approach similar to that proposed by Halle and Idsardi (1997), uses Optimality Theoretic constraints to deal with schwa epenthesis. This account is the same in spirit as that of McCarthy (1991) above, however, Orgun uses phoneme-specific constraints (e.g., *Coda-*r*) to derive the context for the schwa. Thus, while able to get the schwa effect, this approach is ultimately driven by specifying a "special" relationship between /r/ and a particular set of preceding vowels. In both Orgun's and Halle and Idsardi's cases, this special status is motivated by the vocalization of postvocalic /r/ in the dialect in question (E. Massachusetts). However, it is not clear why /l/ should also elicit schwa epenthesis in this dialect, or why both /r/ and /l/ should elicit the same effect in other dialects without /r/ vocalization.

Lavoie and Cohn (1999) 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/ fails to elicit the schwa (e.g., *hall*). Their analysis of the low vowels, proposed in Cohn and Lavoie (2000), offers the constraint that "r/l can bear a mora after [–low] vowels but not [+low] vowels." However, this constraint is essentially descriptive. Another problematic case for the Lavoie and Cohn analyses is that they find that the vowel /o/ groups with the low/lax vowels in not eliciting the excrescent schwa effect (e.g., *hole*; see Lavoie and Cohn 1999: 111). In his study of tongue root positions in English vowels, MacKay (1977) 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 regarding non-high back vowels, though seemingly exceptional, converge to support the analysis we propose in this paper: It is only where there is an opposition in tongue root/dorsum position that we expect a transition to move through "schwa space". Thus, in the case of a vowel such as /o/ or /a/ followed by a liquid, there is no conflict in tongue root/dorsum target (as all of these have a retracted tongue position); therefore no transition occurs, and the tongue does not pass through a schwa-like configuration. We will pursue this connection between tongue root position and excrescent schwa throughout the rest of this paper.

2.1.1. Previous experiments: Syllable rime durations

Some laboratory tests such as measures of duration may be applied in these cases to clarify the status of the schwa. If this apparent schwa in English is in fact an inserted vowel, or if this context corresponds with an additional mora, we should expect the schwa to contribute to the duration of the syllable rime. Previous data on rime durations, however, show that these syllables in English are no longer than other syllables ending in voiced consonants.

Lavoie and Cohn (1999) give acoustic 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*). The presence of an additional timing unit or syllable in the glide + liquid cases should cause a greater duration in *-ajl* relative to *-ald* and *-ajd*. We are unable to make statistical comparisons without the original data, however, the durations cited for the *-ald* and *-ajd* cases are 213ms and 185ms, respectively (averaging 199ms), while that of *-ajl* is 201ms. Thus, these data do not appear to indicate a substantial additional duration contributed by the excrescent schwa in glide-liquid combinations.

As the proposal in the present paper rests upon a conflict between articulatory gestures, verification of the acoustic results presented in 2.1.1 in terms of articulatory timing is perhaps more relevant. Gick and Wilson (2001) report an ultrasound experiment 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 segment type. Their results show no significant differences in duration between rimes ending in the three different final consonants measured (pVn vs. pVl vs. pVd; e.g., pine vs. pile vs. pied, etc.), lending further support to the prediction that the presence of the excrescent schwa does not contribute to the duration of the syllable.

2.2. Gestural modeling of the problem

Articulatory Phonology (Browman and Goldstein 1992) treats articulatory events as basic units of phonology, thus lending phonological meaning to the notions of individual gestural targets and thereby the conflicts between them. As units of phonological representation, phonological constraints may of course be applied to them. Thus, the present paper assumes a representation of articulatory gestures as phonological units in the Articulatory Phonology sense. A further advantage of Articulatory Phonology for the present study is that it allows the component gestures of compound (multi-gesture) segments such as English /l/ to be treated independently. This is crucial as presumably it is normally only individual gestures and not whole segments that come into physical conflict. In the case of /l/, the various manifestations of its component gestures in different allophones, etc., have been reasonably well studied (e.g., Sproat and Fujimura 1993; Browman and Goldstein 1992, 1995; Gick in press), allowing us to draw upon this previous data to further understand the English excrescent schwa phenomenon, at least with regard to interactions with /l/. The following section proposes a model of speech production to help illustrate the implications of the above duration data and to link these findings with the data to follow in the remainder of the paper.

2.2.1. A production model for gestural timing in syllable codas

Sproat and Fujimura (1993) propose that English /l/ is composed of two distinct types of gestures: A “consonantal” gesture (tongue tip [TT] raising) and a “vocalic” gesture (tongue rear/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 significantly earlier than TT closure in postvocalic allophones. This observation is very likely to be relevant to the present discussion, as it is the transition from the preceding vowel toward the TR gesture for /l/ that is proposed here to be the locus for the schwa percept.

Consider a set of three syllable rimes parallel to the example types cited above in 2.1.1, but where all three rimes involve different numbers of retraction/advancement movements of the tongue posterior, e.g., /-id/, /-ajd/, and /-ajl/. Thus, /-id/ involves only a single vocalic gesture (TR 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, as was established in Section 2.1.1, then as more conflicting gestures occur within that duration, they must be more tightly packed. Thus, it must be the case that the duration of the temporal lag between vocalic and consonantal gestures decreases as the number of conflicting movements increases. A schematic illustration of this hypothesis is shown in Figure 1.

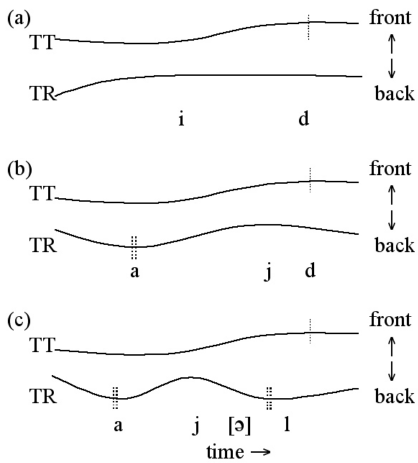

Figure 1. Schematic diagram of tongue tip (TT) vs tongue rear/root (TR) movement in (a) -id, (b) -ajd, and (c) -ajl syllable rimes. Single dotted lines show estimated time of achievement of the tongue tip closure gesture for the final consonant; double dotted lines show time of achievement of TR retraction.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

This model suggests that stable timing is not maintained within segments postvocalically as clusters increase, at least in cases that do not allow for gestural overlap (i.e., cases of direct articulatory conflict). Rather, in this model, syllable coda timing is a relationship between the syllable peak and the first consonantal gesture, with vocalic gestures compressed into the available time window. Gick (1999) 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 opposing TR targets). Data from Gick (1999) are shown in Figure 2. These results again support the notion that /j/ and /l/ have conflicting TR targets, and thus that the excrescent schwa appears during the resulting vocalic transition.

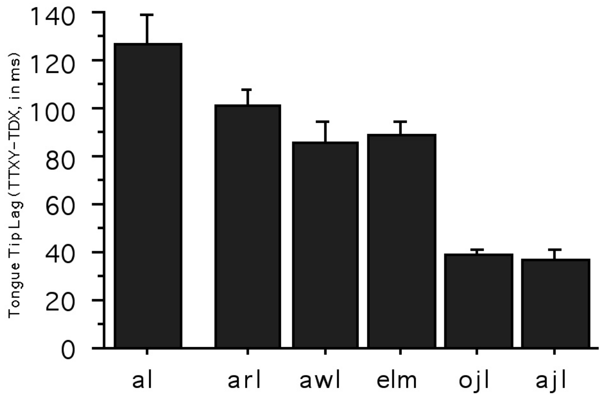

Figure 2. Tongue tip lag in postvocalic /l/ allophones (figure reproduced from Gick 1999). The three sets (al, arl/awl/elm, and ojl/ajl) are all statistically distinct.

2.3. A pilot experiment testing the "schwa space" hypothesis for English

The above sections presented evidence showing that the schwa percept in English syllable rimes does not contribute to rime duration and presented a simple production model wherein such an effect might be feasible. However, while this is consistent with the contention that this schwa is a by-product of articulatory conflict, none of the data thus far offers any explanation as to why this particular conflict should elicit the percept of a schwa, *per se* – nor has the predicted transition actually been observed to occur at all. Thus, before leaving English behind, the results of a pilot experiment clarifying these issues will be presented. An experiment was conducted using direct (ultrasound) and indirect (acoustic) measurement to illustrate: 1) whether the proposed transition from an advanced to a retracted tongue root/dorsum position visibly occurs in English high vowel/glide+/l/ clusters; and 2) whether the tongue and acoustic signal smoothly pass through a schwa-like configuration along their trajectories in the excrescent schwa cases, as proposed above.

2.3.1. Methods

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.

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. In addition to these, the form [pǽpə] was also collected to provide "canonical" schwas for comparison with excrescent schwas.

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. Given the small number of tokens per condition, statistical comparisons were not made in this pilot study.

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. See Gick (2002) for a more in-depth treatment of the ultrasound data collection methods used in this study.

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. Mean tongue posterior positions were calculated using Adobe Photoshop v.6 to measure the distance moved along the trajectory of greatest displacement, starting at the spatial maximum for the vowel immediately preceding /l/ and ending at the maximum point of retraction for /l/. Acoustic signals were analyzed using the freeware Praat v.3.9.13 (http://www.fon.hum.uva.nl/praat/). Formants were calculated at the midpoint of canonical schwas, and compared with the 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.

2.3.2. Results

First, it is apparent from the ultrasound data, as suggested by the acoustic transition, that the successive tongue root/dorsum targets in the excrescent schwa cases are indeed in the predicted conflict relationship. Thus, the tongue moves from a relatively advanced position in the mid pharyngeal region for all of the vowels (except /a/) to a relatively retracted position for the /l/, but not for the other postvocalic consonants. Mean distances of tongue posterior displacement for this transition movement were: (for M1) /il/=9.3mm, /awl/=3.2mm, /ejl/=12.6mm, /ajl/=5.6mm, /ɔjl/=8.5mm; and (for W1) /il/ =17.1mm, /awl/=4.5mm, /ejl/=12.7mm, /ajl/=8.5mm, /ɔjl/=10.0mm.

Second, the mean formant values for the 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 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 3, and formant trajectories are compared in Figure 4.

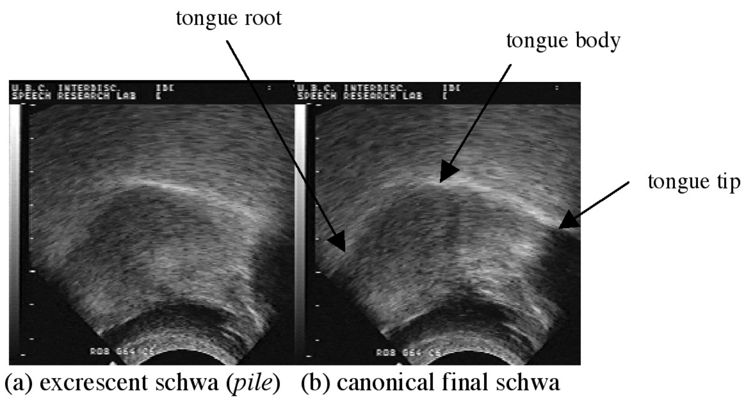

Figure 3. Midsagittal ultrasound images of tongue shape for (a) excrescent and (b) canonical schwa (M1).

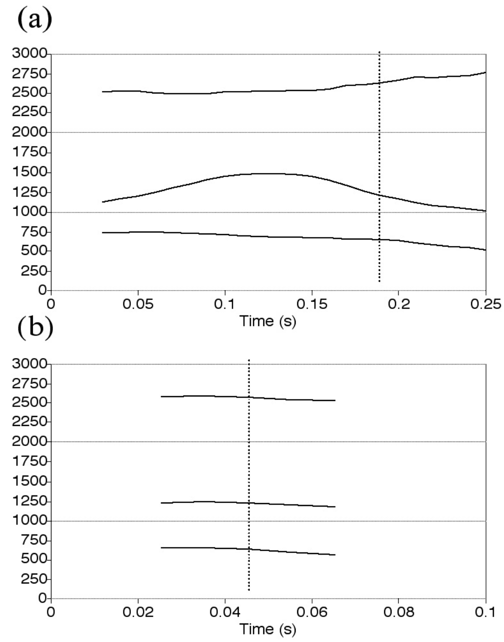

Figure 4. Example formant trajectories for (a) pile and (b) canonical final schwa. The vertical dotted line in (a) shows the point of intersection with canonical schwa F1 and F2; and in (b) shows the midpoint, where measures were made.

2.3.3. Discussion

The results of this experiment illustrate, as predicted, that: 1) the tongue root/dorsum targets for the tense vowels/glides and /l/ appear to be in conflict, forcing a fast transitional movement; and 2) the resulting transition moves the tongue through an articulatory and acoustic space almost identical to that of canonical schwa. Thus, it may be presumed that the percept of schwa in these cases emerges phonetically as a result of more general properties of syllable rime timing, although the actual *percept* of an additional vowel here may be influenced by other factors as well, such as rate of transition or presence of other gestures. This latter point is further suggested by the wide range in the magnitude of the effect, with /il/ and /ejl/ having the largest transitions, and /awl/ having the smallest. There is additionally a good deal of dialectal variation in the strength (impressionistically) of the glide components of certain diphthongs.

2.4. A constraint-based proposal

When one talks of different phonologies employing different strategies for resolving conflicts, it is natural to think of Optimality Theory (McCarthy and Prince 1993; Prince and Smolensky 1993) because, as Kager (1999: xi) states, "the central idea of Optimality Theory (OT) is that surface forms of language reflect resolutions of conflicts between competing demands or constraints." If the resolution of gestural conflicts is phonological, then we expect it to be possible for different languages to have different winning candidates. From the English examples of excrescent schwa and vowel laxing, it is apparent that different strategies are available for resolving cases of articulatory conflict. The two strategies that we will focus on most in the present paper are: 1) achieving all specified targets, but not achieving them simultaneously, thus allowing excrescent schwa to occur in some cases, and 2) compromising one of the component targets. As there are different strategies available for different languages, or different contexts within a language, it is a reasonable next step to propose a formalization that will allow us to identify the full set of logically possible responses to gestural conflict.

There appear in these data to be two fundamental tendencies coming into opposition. The first of these is that all gestures want to reach their targets; the second is that gestures want to occur as close as possible to the syllable peak. In cases of physical conflict, both of these constraints cannot be satisfied. In OT there are two general types of constraints, faithfulness constraints and markedness constraints (Kager 1999), each corresponding to one of these two tendencies. Thus, as the excrescent schwa scenario discussed above is the result of two segments both having their specified gestural targets achieved, this may be easily expressed in terms of a general faithfulness constraint on gestures. The constraint we propose is IDENT(target), defined as follows:

IDENT(target): The gestural target of an input segment must be preserved in its output correspondent.

As for the markedness constraints, we use two Align constraints that are independently motivated in Gafos (2002: 281). These constraints are CV-COORD(INATION) and VC-COORD(INATION) and they are defined as in Gafos (p.281) as follows:

CV-COORD: the c-centre of the C gesture must be synchronous with the onset of the V gesture.

VC-COORD: the target of the C gesture must be synchronous with the release of the V gesture.

As we are only concerned with the TR targets, only these gestures will be discussed in the present paper. However, it should be noted that previous studies, as well as the production model presented in Section 2.2 above, have found that the component gestures of postvocalic compound segments are asynchronous, suggesting that VC-COORD does not apply equally to all C gestures in English. In fact, the posterior gestures in question are specifically the ones termed “vocalic” by Sproat and Fujimura (1993), and “V-gestures” by Gick (in press). It is perhaps natural that they should be more closely coordinated with the syllable peak.

The three constraints listed above can be ranked with respect to each other in numerous ways. Table 1 outlines the possible rankings. The following section investigates similar cases of conflict in several additional languages that are included in Table 1 to test whether these constraints adequately account for cross-linguistic patterns in gestural conflict resolution. The proposed constraint-based analysis is evaluated with respect to these data in Section 4.

Table 1.

	CV (onset) result	VC (coda) result	Constraint Ranking	Languages
A	—	ə	IDENT >> VC-COORD (CV-COORD doesn't apply because no conflict in onset)	Many dialects of English; Beijing Chinese
B	laxing/ lowering	ə	CV-COORD >> IDENT >> VC-COORD	Nuu-chah-nulth
C	ə	ə	IDENT >> CV-COORD, VC-COORD	?
D	ə	laxing/ lowering	VC-COORD >> IDENT >> CV-COORD	Chilcotin
E	laxing/ lowering	laxing/ lowering	CV-COORD, VC-COORD >> IDENT	Skye Scots Gaelic
F	—	laxing/ lowering	VC-COORD >> IDENT (CV-COORD doesn't apply because no conflict in onset)	Pittsburgh/Utah English; Korean
G	ə	—	IDENT >> CV-COORD (VC-COORD doesn't apply because no conflict in coda)	No known potential cases (i.e., where a back gesture is present in onset but not coda allophone)
H	laxing/ lowering	—	CV-COORD >> IDENT (VC-COORD doesn't apply because no conflict in coda)	No known potential cases (i.e., where a back gesture is present in onset but not coda allophone)
I	—	—	Not ranked because there is no conflict	All languages lacking or avoiding conflicting sequences, e.g., Yoruba, Cook Island Maori, Bernera Scots Gaelic

3. Cross-linguistic cases of gestural conflict

This section draws on field description and pilot ultrasound imaging and acoustic data to observe cross-linguistic responses to tongue posterior conflicts similar to that discussed above for English. A number of languages have been identified where similar articulatory conflicts are expected to arise, and conflict resolution strategies used in different languages are surveyed.

3.1. Excrescent schwa

In this section we identify several cross-linguistic examples of one strategy languages use to resolve conflicting TR targets, namely that all targets are achieved in sequence, and that this strategy may at least sometimes result in the percept of an intervening excrescent schwa. This shows first that English is not an isolated case, and second, allows for a wider range of manifestations of this strategy to be observed.

3.1.1. Beijing Chinese

A similar pattern to that of English excrescent schwa has been reported in Beijing Chinese. Beijing Chinese does not normally have non-nasal postvocalic consonants at all, with the one exception of the diminutivizing suffix *-r*. This postvocalic *-r* impressionistically has a pharyngeal retroflex quality surprisingly similar to that of the rather unusual postvocalic /r/ of English.

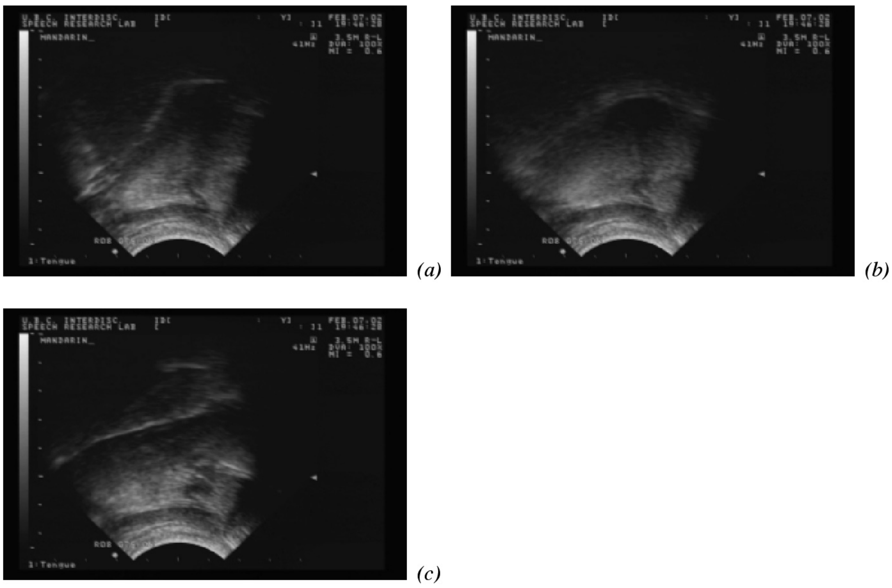

Figure 5. Beijing Mandarin /pi-r/ 'skin-DIM'. (a) 0 ms: 1st part of /i/. (b) 66 ms: 2nd part of /i/ ([ə]); note the tongue root retraction. (c) 210 ms: end of retroflex approximant; note the tongue root retraction and tongue body lowering.

According to Pulleyblank (2003), adding this suffix to a word ending in /i/, such as *jī* [tɕi] 'chicken' results not in *jīr* [tɕiɻ], but rather it "requires the insertion of a schwa vowel before the suffix, giving [tɕjər][2]."

Pilot ultrasound data was collected to test whether this phenomenon is parallel to the English excrescent schwa, by verifying that the Beijing Chinese postvocalic /r/ involves a tongue root or dorsum retraction opposing the advanced position for /i/.

The speech of three adult speakers of Beijing Chinese was measured (see 2.3.1 above for data collection and processing methods). All speakers exhibited tongue root advancement for /i/, a tongue root retraction for /r/, and the predicted transition through an apparent schwa-like configuration. As there is no tongue root retraction in the prevocalic liquid (see Gick, Campbell, Oh and Tamburri-Watt in press), and hence no conflict, this appears to be a language of type A in Table 1. Results for one subject are shown in Figure 5 for illustration.

3.1.2. Nuu-chah-nulth

Nuu-chah-nulth (formerly known as Nootka), a Wakashan language spoken on the west coast of Vancouver Island, British Columbia, is a language that includes a number of consonants articulated in the uvular and pharyngeal areas of the vocal tract. In order to produce constrictions at these areas of the vocal tract, it is necessary for the tongue dorsum/root to retract. There are primarily three vowels in the language, /i/, /u/, and /a/ (/e/ is also used, but only occurs in very specific styles of speech examined in the Ahousaht dialect). Given that the language has uvular and pharyngeal consonants as well as a high front vowel /i/, an articulatory conflict is predicted (assuming that /i/ specifies tongue root advancement in this language). It will be shown that Nuu-chah-nulth employs two different methods for resolving this conflict. When /i/ precedes a retracted consonant such as /q/, tongue root targets are not compromised, and an intervening excrescent schwa offglide results: [iᵊq]. However, when /i/ follows /q/, the tongue root target of /i/ is compromised (becomes retracted) and the tongue body (TB) commensurately lowers, resulting in a lowered and/or laxed vowel [qe] ~ [qɛ]. The vowel laxing/lowering case will be discussed in Section 3.2.2 below.

In words that have an /i/ followed by the uvular stop /q/, a very distinct schwa is apparent in the transition from /i/ to /q/. For example, in the word /ci:qci:qa/ 'talking' two perceived excrescent schwas are reminiscent of the

English and Beijing Chinese examples discussed above. On the other hand, no excrescent schwas are observed in a word with velar, as opposed to uvular stops: /ci:kci:ka/ '(a vessel) listing/tilting back and forth'. If the excrescent schwa in Nuu-chah-nulth is, as in English and Beijing Chinese, a purely phonetic by-product of two opposing tongue root targets (i.e. advanced for /i/ and retracted for /q/), then this conflict will be evident from the tongue images. Furthermore, as the tongue retracts from /i/ to /q/ position, the tongue is expected to pass through "schwa space" in the course of its transition. The TB lowering suggested in the transcription to occur simultaneously with TR retraction is likewise not surprising, given the volume-preserving nature of the hydrostatic tongue.

Using the same methods as described in 2.3.1, ultrasound data was collected from two native speakers of Nuu-chah-nulth to test this hypothesis. The movement from advancement to retraction can been seen in Figure 6, with TR retraction and TB lowering occurring simultaneously in the transition from the /i/ in (a) to the /q/ in (d).

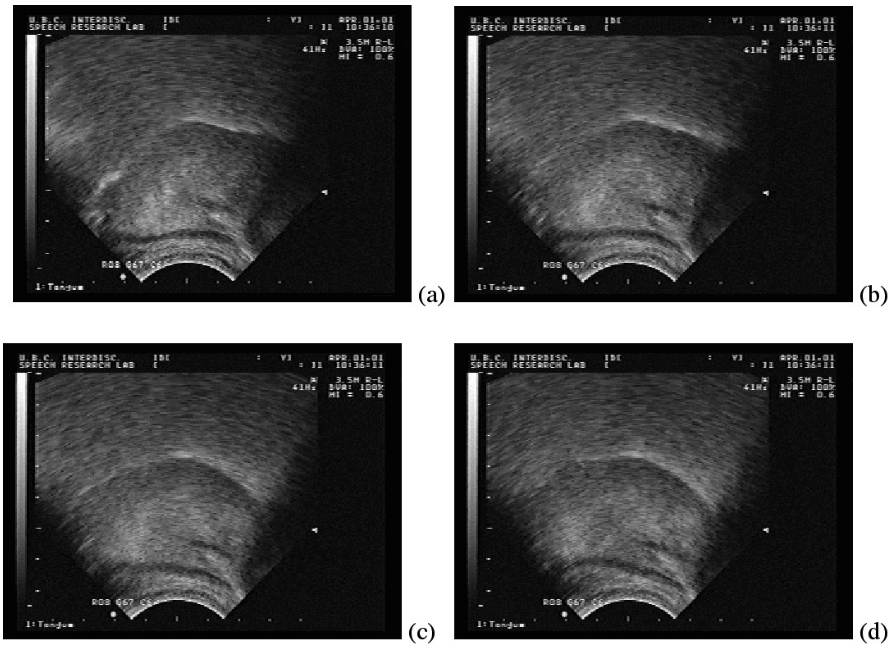

Figure 6. Nuu-chah-nulth /ci:qci:qa/ 'talking'. (a) 0 ms: 1st part of 1st /i:/. (b) 33 ms: 2nd part of 1st /i:/; note the tongue root retraction and body lowering. (c) 66 ms: end of [ə]; lead-in to /q/. (d) 99 ms: /q/ (uvular stop).

3.1.3. Chilcotin

In Chilcotin, an Athapaskan language spoken in northern British Columbia, there are a number of so-called “flat” consonants produced with a retracted tongue root (Cook 1983, 1993). These consonants include pharyngealized fricatives and affricates, as well as a set of plain and labialized uvulars (stops, fricatives, and approximants). In addition, there are three pairs of vowels described by Cook (1993: 152) as being “tense (long)” versus “lax (short)”: /i/ & /ɪ/, /æ/ & /ɛ/, and /u/ & /ʊ/ (no articulatory data exists as yet with which to verify the tongue root/dorsum position for the tense/lax distinction in these vowels). Thus, the stage is set again for a conflict in TR targets between the retracted consonants and the (presumably) advanced vowels.

Like Nuu-chah-nulth, Chilcotin exhibits the same two strategies as discussed above for resolving this TR conflict. Interestingly, however, the two strategies are employed in exactly the opposite contexts from Nuu-chah-nulth (i.e., vowel laxing/lowering occurs *preceding* back consonants, and excrescent schwa *after* back consonants). Vowel lowering will be described in Section 3.2.3 below.

According to Cook’s transcriptions, sequences of these flat consonants and the high front vowels do result in the percept of an excrescent schwa. However, unlike Nuu-chah-nulth, where the excrescent schwa appears in the transition from /i/ to a following retracted consonant, in Chilcotin the excrescent schwa occurs in the transition from a preceding retracted consonant to a tense vowel. A minimal pair from Cook (1983: 128) is given below:

(1) /sid/ [sit] ‘I’
(2) /ŝid/ [sˤᵊit] ‘kingfisher’

Example (1) shows that when /i/ precedes/follows a non-retracted consonant, it is realized as [i]. Example (2) shows that an excrescent schwa results when /i/ is preceded by a pharyngealized consonant.

3.2. Vowel laxing

In this section, several cases are presented where the same combination of segments (high front tense vowel + retracted consonant) consistently yields a different result: vowel lowering/laxing. It appears that, unlike the languages in the previous section, this set of languages prefers to compromise the achievement of articulatory targets. That is, rather than let the TR go to its fully advanced position for /i/, these languages reduce or eliminate that vowel gesture and produce /i/ without advanced TR, resulting in a laxed or lowered vowel (usually something impressionistically in the range of [ɪ̞e̞ɛ̞]). The implications of this pattern are discussed in more detail in Section 4 below.

3.2.1. Korean

Korean has not been previously described as a laxing language. However, Korean was chosen for this study for two reasons: first, because it gave the auditory impression of some laxing of the vowel /i/ before /l/; and second, our earlier pilot ultrasound work on Korean vowel articulation showed that /i/ appeared to be the only vowel in Korean with tongue root advancement. These two factors conspired to give the impression that this again may be a potential case of articulatory conflict. A small acoustic study was conducted, therefore, to test whether this auditory impression of laxing is accurate.

A single female native speaker of Korean in her thirties participated in this experiment. The subject produced repetitions of the words *p^{h}it*, *p^{h}in*, and *p^{h}il* in a fixed carrier phrase context. Values for F1 and F2 were extracted from the mid-point of each vowel using Praat software, according to the methods described in 2.3.1 above.

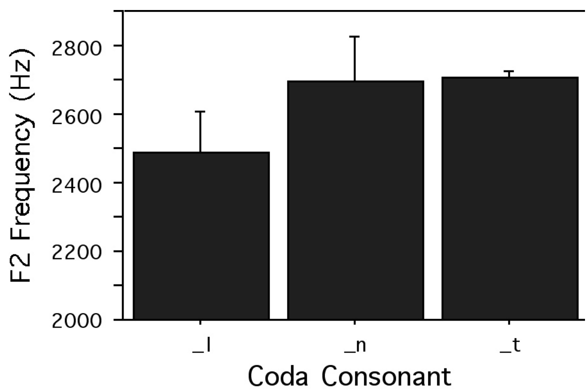

Figure 7. Effect of coda consonant on F2 of preceding /i/.

ANOVA results for this speaker indicate that F1 of /i/ is not significantly affected by coda consonant, but that F2 is ($p < .05$). Post-hoc tests (Fisher's

PLSD) indicate a significant difference ($p < .05$) between the effects on F2 both between final /l/ and /t/ and between /l/ and /n/, but no significant difference between the effect of final /t/ and /n/. These results are shown in Figure 7. This F2 lowering effect supports the auditory impression of laxing of Korean /i/ before /l/. As there is no such evidence of conflict in the initial allophone (see Gick, Campbell, Oh and Tamburri-Watt [in press] for relevant articulatory data), Korean must be classed with Pittsburgh and Utah English as a type F language in Table 1.

3.2.2. Nuu-chah-nulth

In Section 3.1.2, it was mentioned that in addition to the excrescent schwa cases, Nuu-chah-nulth also shows vowel lowering, i.e. a compromise of the [ATR] target of /i/ resulting in lowering of the TB. Two examples from the Ahousaht dialect of Nuu-chah-nulth are given in (3) and (4) below:

(3) /siqiːɬ/ [sɪˀqeːɬ] ‘to cook’
(4) /qitʃin/ [qetʃɪn] ‘louse’

The difference between /i/ in a non-retracted environment and /i/ following a retracted consonant is clearly seen in Figure 8.

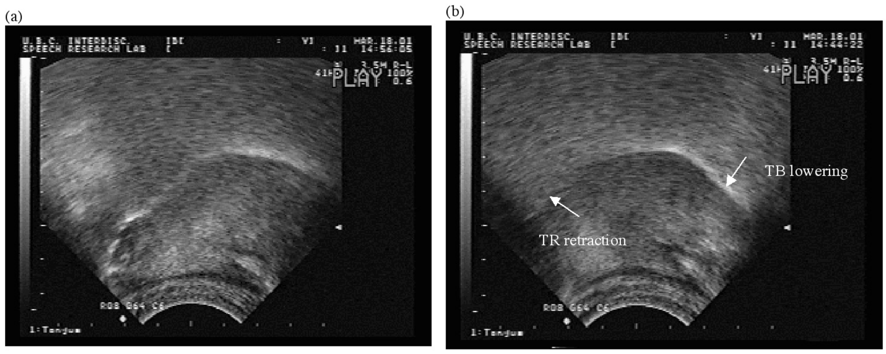

Figure 8. (a) Final /i/ in the word /-a-itʃi/ (eyebrow) – phonetically [i]. (b) First /i/ in the word /qitʃin/ (louse) – phonetically [e].

In Figure 8 (a), the advanced TR and raised TB are key to producing this [i]. However in Figure 8 (b), /i/ is seen to have a retracted TR and a lowered TB, resulting in a vowel transcribed as being close to [e] or [ɛ]. Given that Nuu-chah-nulth shows laxing/lowering in onset sequences, but serial production with an accompanying schwa percept in coda sequences, this language must be classed as type B in our Table 1.

3.2.3. Chilcotin

As discussed above, while vowel lowering takes place after /q/ in Nuu-chah-nulth, in Chilcotin it happens *before* /q/. This mirror image effect was also seen in the behaviour of the excrescent schwa in these two languages in Sections 3.1.2 and 3.1.3. Examples (5) and (6) from Cook (1993: 155) show the lowering that occurs before /q/.

(5) /niqin/ [neqᵊin] 'we paddled'
(6) /tŝ˞iqi/ [tsˤ˞eqᵊi]ˣ [tsˤ˞ᵊiqᵊi] 'woman'

Thus, we must categorize Chilcotin as a type D language in our Table 1.

Note that both vowel lowering preceding /q/ and the excrescent schwa following /q/ occur in the word in (5). Also note that in (6), where a retracted consonant both precedes and follows the (first) /i/, there is no clear precedence: the realization of /i/ varies between [e] and [ᵊi].

3.2.4. Skye Scots Gaelic

According to the description of Borgstrøm & Oftedal (1941: 18), Skye Scots Gaelic lowers and retracts /i/ both before and after /l/. While no primary data on this language is available to the authors at present, if this description is indeed correct, then this language constitutes the only known case of symmetric responses to this conflict (i.e., where the same strategy is employed in both pre- and post-vocalic positions), and must be categorized as a type E language in Table 1.

3.3. Conflict avoidance[3]

In addition to the two main effects discussed in the above sections, it is worth mentioning an additional response to conflict: systematic avoidance (i.e., cases where sequences of ATR vowels + liquids or other retracted consonants are systematically avoided). Examples of this type of language include Yoruba, where /r/ is deleted altogether before /i/ (anonymous reviewer, p.c.); Cook Island Maori, where that language's usual epenthetic vowel, /i/, surfaces as [a] following /r/ (Kitto and de Lacy 1999); and Bernera Scots Gaelic, where tense /l/ does not appear before or after a stressed front vowel (Borgstrøm and Oftedal 1940: 68). These languages thus all fall into category I in Table 1.

4. Summary and Conclusion

The results of the survey in Section 3 are summarized in in the right-most column of Table 1. It can be seen from Table 1 that only patterns of type C (serial, symmetric) remain unattested in our sample. Given the small sample size in the present study, we are optimistic that a more thorough search will reveal languages of this type. If, however, such a language were found not to exist, it would be interesting to consider the possible relationship between this and previous observations relating to gestural positional assymmetries in syllable structure (e.g., Krakow 1999; Redford 1999). A specific question that emerges from this study relates similarly to syllable assymetries: Why are the back/vowel gestures of liquid consonants in syllable-initial allophones of liquids in most languages relatively smaller in magnitude than those in final allophones (Gick, Campbell, Oh and Tamburri-Watt in press)? Further research in these areas will be needed to clarify this relationship between the phonetic and phonological asymmetries in syllable production.

It has been argued in this paper that the apparent schwa in English tense vowel + liquid combinations is the incidental result of the tongue moving through a schwa-like configuration during the transition forced by conflicting tongue root targets. This phenomenon has helped to identify other similar cases of conflicting targets, allowing for a cross-linguistic comparison of strategies employed in resolving such articulatory conflicts. To accommodate these data, a model for gestural coordination in syllable codas has been presented 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. Finally, it has been shown that the different strategies typically employed in cases of gestural conflict may be straightforwardly described in terms of the ranking of phonological constraints on articulatory gestures, and that the predictions of this constraint-based approach are supported in a cross-language survey.

An additional implication not discussed above involves the perception-production mismatch in the number of syllables in the excrescent schwa cases. If the argument in the present paper is correct, then the apparent "extra" syllable in these cases is most clearly explained as being a purely physical by-product of gestural conflict, and itself has no phonological status. Thus, for example, *feel* and *feed* are most efficiently represented as being monosyllabic, contrary to perception. This type of case will need to be addressed explicitly by theories of phonology where perception plays a more prominent role (e.g., Ladd, this volume).

Notes

* The authors wish to thank Shaffiq Rahemtulla, Sunyoung Oh, and Christine Boggs for assistance with data collection and analysis; Lisa Davidson, Abigail Cohn, Lisa Lavoie, and Doug Whalen for helpful discussion and comments; the editors of this volume and two anonymous reviewers for their insightful comments on an earlier version; previous audiences who heard earlier versions of this material and contributed their questions and insights; and our subjects for their participation. This work was funded by a grant from the Natural Science and Engineering Research Council (NSERC) of Canada, an equipment grant from the Canadian Foundation for Innovation (CFI), and NIH Grant DC-02717 to Haskins Laboratories.

1. Pittsburgh is unique in that the non-low tense vowels are laxed not only before /l/ (e.g., *feel* [ɪ], *fool* [ʊ], *fail* [ɛ], etc.) and /r/ (e.g., *here* [ɪ]), but also before /g/ (e.g., *league* [ɪ]), etc. While this case will not be pursued in detail in this paper, it is particularly interesting as it implies the testable prediction that (as we shall see below) the Pittsburgh /g/ is somehow at articulatory odds with the high tense vowels (e.g., perhaps the /g/ is produced at a more posterior place of articulation than in other dialects of English).
2. sic [r] (Pulleyblank 2003) – retroflex approximant [ɻ] in IPA.
3. Many thanks to an anonymous reviewer for suggesting this additional possibility, and for providing the Borgstrøm & Oftedal (1940, 1941) references.

References

Borgstrøm, C. H. and M. Oftedal
1940 *A Linguistic Survey of the Gaelic Dialects of Scotland: Vol 1 - The Dialects of the Outer Hebrides*. Oslo: Oslo Aschehoug.

1941 *A Linguistic Survey of the Gaelic Dialects of Scotland: Vol. 2 - The Dialects of Skye and Ross-shire*. Oslo: Oslo Aschehoug.

Browman, C. P. and L. M. Goldstein
1992 Articulatory phonology: An overview. *Phonetica* 49: 155–180.
1995 Gestural syllable position effects in American English. In: F. Bell-Berti & L. J. Raphael (eds.), *Producing Speech: Contemporary Issues. For Katherine Safford Harris*, 19–34. Woodbury, NY: American Institute of Physics.

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

Cook, E.-D.
1983 Chilcotin flattening. *Canadian Journal of Linguistics* 28: 123–132.
1993 Chilcotin flattening and autosegmental phonology. *Lingua* 91: 149–174.

Delattre, P. C. and D. C. Freeman
1968 A dialect study of American r's by x-ray motion picture. *Linguistics* 44: 29–68.

Di Paolo, M. and A. Faber
1991 Phonation differences and the phonetic content of the tense-lax contrast in Utah English. *Language Variation and Change* 2: 155–204.

Gafos, A. I.
2002 A grammar of gestural coordination. *Natural Language & Linguistic Theory* 20: 269–337.

Gick, B.
1999 The organization of segment-internal gestures. In: *Proceedings of the XIVth International Congress of Phonetic Sciences*, 1789–1792. San Francisco, CA. August, 1999.
2002 The use of ultrasound for linguistic phonetic fieldwork. *Journal of the International Phonetic Association* 32(2): 113–121.
2003 Articulatory correlates of ambisyllabicity in English glides and liquids. In: J. Local, R. Ogden & R. Temple (eds.), *Papers in Laboratory Phonology VI: Constraints on Phonetic Interpretation,* 222–236. Cambridge: Cambridge University Press.

Gick, B., F. Campbell, S. Oh and L. Tamburri-Watt
in press Toward universals in the gestural organization of syllables: A cross-linguistic study of liquids. *Journal of Phonetics*.

Gick, B., A. M. Kang and D. H. Whalen
2002 MRI evidence for commonality in the post-oral articulations of English vowels and liquids. *Journal of Phonetics* 30: 357–371.

Gick, B. and I. Wilson
2001 Pre-liquid excrescent schwa: What happens when vocalic targets conflict. In: P. Dalsgaard, B. Lindberg and H. Benner (eds.), *Proceedings of the 7th European Conference on Speech Communication and Technology (Eurospeech 2001)*, 273–276. Aalborg, Denmark: Center for Personkommunikation. September 2001.

Halle, M. and W. Idsardi
1997 r, hypercorrection, and the Elsewhere Condition. In: I. Roca (ed.), *Derivations and Constraints in Phonology*, 331–348. Oxford: Clarendon Press.

Kager, R.
1999 *Optimality Theory*. Cambridge, U.K.: Cambridge University Press.

Kitto, C. and P. de Lacy
1999 A correspondence theory of epenthetic quality. In: C. Kitto and C. Smallwood (eds.), *Proceedings of the Austronesian Formal Linguistics Association (AFLA) VI*, 181–200. Toronto: Toronto Working Papers in Linguistics.

Krakow, R. A.
1999 Physiological organization of syllables: A review. *Journal of Phonetics* 27: 23–54.

Ladd, D. R.
this vol. 'Contrastive phones' in surface representation.

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

Levin, Juliette
1987 Between epenthetic and excrescent vowels. In: M. Crowhurst (ed.), *Proceedings of the West Coast Conference on Formal Linguistics (WCCFL) 6*.

MacKay, I. R. A.
1977 Tenseness in vowels: An ultrasonic study. *Phonetica* 34: 325–351.

McCarthy, J. J.
1991 Synchronic rule inversion. *Proceedings of the Berkeley Linguistics Society* 17: 192–207.

McCarthy, J. J. and A. S. Prince
1993 Generalized alignment. In: G. E. Booij and J. van Marle (eds.), *Yearbook of Morphology 1993*, 79–153. Dordrecht: Kluwer.

Narayanan, S. S., A. A. Alwan and K. Haker
1997 Toward articulatory-acoustic models for liquid approximants based on MRI and EPG data, Part I: The laterals. *Journal of the Acoustical Society of America* 101(2): 1064–1077.

Orgun, C. O.
2001 English r-insertion in optimality theory. *Natural Language & Linguistic Theory* 19: 737–749.

Prince, A. S. and P. Smolensky
1993 Optimality theory: Constraint interaction in generative grammar. Manuscript, Rutgers University, New Brunswick, NJ, and University of Colorado, Boulder, CO.

Pulleyblank, E.
2003 Pharyngeal glide. Ms., University of British Columbia.

Redford, M. A.
1999 An articulatory basis for the syllable. Ph.D. dissertation, University of Texas, Austin.

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

Steriade, D.
1982 Greek prosodies and the nature of syllabification. Ph.D. dissertation, Massachusetts Institute of Technology.

Walsh Dickey, L.
1997 The phonology of liquids. Ph.D. dissertation, University of Massachusetts, Amherst.

Wood, S. A. J.
1996 Assimilation or coarticulation? Evidence from the temporal co-ordination of tongue gestures for the palatalization of Bulgarian alveolar stops. *Journal of Phonetics* 24: 139–164.