

Implications of vowel intrusion for a gestural grammar *

Nancy Hall

February 4, 2004

Abstract

Intrusive vowels result when a vowel gesture overlaps a consonant cluster, and a piece of the vowel gesture is audible during a period of release between the consonants. Such vowel sounds are not phonological segments. Data from Dutch, Scots Gaelic and Hocank (Winnebago) show that intrusive vowels behave as non-syllabic. Intrusive vowels are distinct from epenthetic vowels, which involve insertion of a segment. The proposed Optimality Theoretic analysis uses Articulatory Phonology representations. A typological survey of intrusive vowels provides evidence for constraints on gestural phasing. The environments in which vowel intrusion is typically found suggest that vowels overlap better with sonorants than with obstruents, and that sonorants in some languages have less overlap with adjacent consonants than obstruents do. Vowel intrusion is not contrastive, indicating that gestural alignment is not subject to faithfulness constraints.

1 The vowel intrusion syndrome

About 2500 years ago, the *praticakhyas*, ancient Indian commentaries on Pāṇini, described a type of vocalic interval that the writers call not a vowel, but a ‘vowel fragment’—*svarabhakti* in Sanskrit. This vowel fragment appeared between an [r] and a following consonant, had one-eighth to one-half the length of a normal vowel, and was not considered to add a syllable to the word (Allen 1953:73–80).

This is the earliest description of a phenomenon that I will call ‘vowel intrusion’. In many languages, a sonorant is separated from an adjacent consonant by either a short schwa or a vowel whose quality is identical to that of the vowel next to the sonorant. Examples are given in Figure (1) with the intrusive vowel underlined. I use the term sonorant to include gutturals, following Halle (1995).

Intrusive vowels are often described in the linguistic literature as epenthetic. A typological survey reveals, however, that they have a group of characteristics in

*For useful comments and discussion, thanks to Lisa Selkirk, John McCarthy, John Kingston, Joe Pater, Rex Wallace, Travis Bradley, Adamantios Gafos, Louis Goldstein, Ron Artstein, and members of the Rutgers Optimality Research Group and the UMass Phonology Group.

Arbore	/leh-t-atto/	→	lefi _e 'tatto	'that ewe'
Bulgarian	/garbav/	→	garəbav	'hunchbacked'
Chamicuro	/tuʔlu/	→	'tuʔulu	'chest'
Dutch	/kalm/	→	kaləm	'quiet'
English (dialects)	/arm/	→	arəm	'arm'
Finnish	/kalvo/	→	kalavo	'transparency'
German (S. Hamburg)	/bʁatən/	→	bəʁatən	'to fry'
Hausa	/kurkutu/	→	k ^w ur ^u k ^w u:tu	'small drum'
Hocank	I. /sni/	→	sīnī	'cold'
	II. /ho:tʃāk-ra/	→	ho:tʃəg ^ə rə	'the Hocank'
Hua	/okrumaʔ/	→	ok ^u rumaʔ	'sky'
Irish Gaelic	I. /agla/	→	'agələ	'fear'
	II. /gorm/	→	'gorəm	'blue'
Kekchi	/paʔt/	→	paʔat	'twins'
Lakhota	/gla/	→	g ^a la	'no gloss'
Late Latin	<i>scriptum</i>	→	<i>sc̄i</i> riptum	'a writing'
Mamainde	/mih-takʔu/	→	mihitakʔu	'it is cloudy'
Mono	/gàfrū/	→	gàfūrū	'mortar'
Negev Bedouin Arabic	/gahwa/	→	ga'hawa	'coffee'
Oscan	I. <i>Mulcius</i>	→	<i>Múlúkiis</i>	name
	II. <i>patri</i>	→	<i>patereí</i>	'father'
Popoluca	/itʔa/	→	it ^a ʔa	'your father'
Saami	/skuol:fi:/	→	skuol:əfi:	'owl'
Sanskrit	/darʃata/	→	darəʃata	'?'
Scots Gaelic	/ʃalʲk/	→	ʃalʲak	'hunting'
Spanish (Chilean)	/kronika/	→	k ^ə ronika	'chronicle'
Tiberian Hebrew	/ʃalah-t/	→	ʃalahat	'you (fs) sent'

Sources: Arbore (East Cushitic): Hayward (1984:72); Bulgarian: Jetchev (1995); Chamicuro (Arawakan, Peru): Parker (1994); Dutch: Booij (1995:127); English: Wright (1905:206); Finnish: Harrikari (1999); German: Jannedy (1994); Hausa: Hodge (1947:12); Hocank (aka Winnebago): Miner (1979, 1992); Hua (Papuan, New Guinea) Haiman (1980); Irish: Ó Siadhail (1989); Kekchi (Mayan, Guatamala): Campbell (1974); Lakhota: Albright (1999); Latin: Schuchardt (1867:vol.3:421); Mamaindé (Nambiquára, Brazil): Eberhard (1995); Mono (Niger-Congo; Congo): Olson (2003); Arabic: Blanc (1970); Oscan: Buck (1904); Popoluca (San Juan Atzingo Puebla dialect): Krumholz et al. (1995:282); Saami: Bye (2001); Sanskrit: Allen (1953); Scots Gaelic: Borgstrøm (1940); Spanish: Garcia-Bellido (1999); Tiberian Hebrew: McCarthy (1979).

Figure 1: Examples of vowel intrusion

common that set them apart from epenthetic vowels. The following collection of properties can be called the vowel intrusion syndrome.

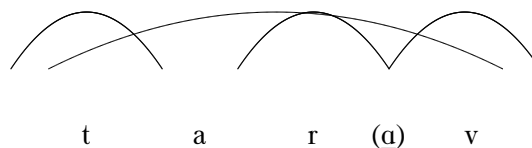
(1) Vowel intrusion diagnostics

- a. Vowel intrusion occurs within a consonant cluster that includes a sonorant. The clusters that have vowel intrusion are often less marked in terms of sonority sequencing than clusters without vowel intrusion.
- b. The intrusive vowel has the quality of the vowel adjacent to the sonorant. This is usually true even when [ə] is transcribed.
- c. Vowel intrusion is usually restricted to heterorganic clusters.
- d. The intrusive vowel behaves as if it does not add a syllable to the word.
- e. In many languages, the intrusive vowel is optional, has a highly variable duration, or may disappear at fast speech rates.

I propose that an intrusive vowel is not an independent segment. Rather, it is the acoustic result of a particular way of timing the articulatory gestures involved in a CCV or VCC sequence. The distribution and behavior of intrusive vowels can best be analyzed using Articulatory Phonology representations, in which abstract articulatory gestures are phonological units (Browman & Goldstein 1986 et seq).

The theory presented here builds on Steriade (1990)'s proposal that copy vowels may result from overlap of the gestures of a consonant and vowel. When two consonant gestures are produced with a low degree of overlap, there is an acoustic release between them. If a neighboring vowel articulation overlaps this period of release, the vowel can be heard briefly between the consonants, sounding like [ə] if it is short or like a copy vowel if it is longer. Diagram (2) shows the proposed gestural representation of a word (Scots Gaelic [tarav], 'bull') with vowel intrusion. Each curve represents the dynamic cycle of one gesture. The intrusive vowel is underlined.

(2) A gestural score of vowel intrusion: Scots Gaelic [tarav] 'bull'



Steriade treats this gestural timing as a method of epenthesis, assuming that moving a consonant gesture over a vowel gesture 'automatically turns a monosyllable into a disyllable.' I will argue that no vowel segment is actually added, and that the surface form [tarav] is monosyllabic, with the segmental representation t-a-r-v. The underlined vowel sound, though acoustically prominent, is not a segment, and does not form the nucleus of a syllable. Vowel intrusion reflects a scheme of organizing

articulatory gestures within a syllable, rather than the addition of a syllable. While vowel intrusion can enhance perceptibility of segments within the syllable (especially the consonants), it cannot repair a marked syllable structure because it does not change the syllable structure. Therefore, vowel intrusion does not particularly target marked clusters, the way structure-changing epenthesis does. I propose that the reason vowel intrusion happens particular next to sonorants, and the reason that it is the vowel adjacent to the sonorant that extends over the release, is that gestural overlap between a vowel and sonorant is less marked than overlap between a vowel and obstruent.

Section 2 reviews some of the evidence that intrusive vowels are non-syllabic. Section 3 demonstrates the difference between vowel intrusion and vowel epenthesis. An Optimality Theoretic analysis of vowel intrusion is laid out in section 4, and some wider claims about the place of gestural representations in the grammar are discussed in section 5. Section 6 responds to Warner et al. (2001)'s recent claim that intrusive vowels behave as syllabic in conditioning allophony in Dutch.

2 Non-syllabicity

'Inserted' vowels with the conditioning environment described in (1) tend, cross-linguistically, to behave as non-syllabic, in a way that is not true of vowels inserted in other environments.

The claim that these vowels are non-syllabic has been made for many individual languages before. The evidence for non-syllabicity is robust, including phonological diagnostics, phonetic characteristics, and native speaker intuitions. In this section I review a sampling of this evidence from Dutch, Scots Gaelic, and Hocank.

2.1 Dutch

Some dialects of Dutch have intrusive [ə] between [l] or [r] and heterorganic consonants, as in [hɛləp] 'help', [hɛrəfst] 'autumn', and [kələm] 'quiet' (Booij 1995). These vowels act non-syllabic in several ways.

Donselaar et al. (1999) (who mention the possibility of a gestural analysis) conducted an experiment in which listeners were asked to reverse monosyllables segment by segment, changing [tap] to [pat], but reverse disyllables syllable by syllable, changing [hotɛl] to [tɛlho]. Subjects preferred to treat words with vowel intrusion like monosyllables, changing [tyləp] 'tulip' to [plyt] rather than [ləpty]. The authors conclude that 'the realizations of real words with schwa epenthesis are represented by listeners as monosyllabic.'

Speaker intuitions could be influenced by spelling, in which intrusive vowels are absent, but phonological patterning supports the conclusion that the vowels are really non-syllabic. For example, intrusive [ə] fails to trigger an optional process of [n]-deletion that occurs after regular [ə]s. In the following words, which have normal, underlying, segmental [ə]s, the coda [n]s are optionally omitted.

(3)	Dutch [n]-deletion		Booij (1995:139)
	<i>regen</i>	reɣən / reɣə	‘rain’
	<i>gouden</i>	ɣoudən / ɣoudə	‘golden’
	<i>bloemen</i>	blumən / blumə	‘flowers’
	<i>lopen</i>	lopən / lopə	‘to walk (pres. pl./inf.)’
	<i>openlijk</i>	opənlək / opələk	‘openly’
	<i>regentje</i>	reɣəntjə / reɣətjə	‘rain (diminutive)’

The deletion of [n] appears to be driven by a constraint against [ən] rhymes. In general, Dutch [ə] is more restricted than other short vowels in what types of coda may follow it.

If the purpose of [n] deletion is to avoid [ən] rhymes, then it should not occur after a non-segmental [ə]. This is in fact the case. For example, *hoorn* ‘horn’ can be pronounced [horən], with an intrusive [ə]. *Horen* ‘to hear’ is also pronounced [horən], but the [ə] is underlying and segmental. I have found that consultants accept the pronunciation [horə] for *horen*, but not for *hoorn*. Since the intrusive [ə] is not a segment, there is no [ən] rhyme to remove, and [n]-deletion is unmotivated.

Words with vowel intrusion also pattern with monosyllables in licensing lexical tone contrasts, in those dialects that have such contrasts. In Venlo Dutch, stressed syllables with two sonorant moras can host a lexical high tone on the second mora (Gussenhoven & van der Vliet 1999:101). Monomoraic syllables do not contrast for tone. Lexical tone does occur on syllables with vowel intrusion, like [ɛrəm] (‘arm’ or ‘arms’, depending on tone). If the [ə] were syllabic, then the stressed syllable would consist only of [ɛ]. Being monomoraic, this syllable would not be expected to license tone. But if [ɛrəm] is monosyllabic, it is heavy and its ability to license tone is normal.

2.2 Scots Gaelic

In accounts that acknowledge the non-syllabicity of short, transitional schwas, one sometimes finds the assumption that if the transitional vowel becomes longer diachronically, it automatically becomes syllabic. The survey of intrusive vowels does not bear this assumption out. While physical length may increase the probability that an intrusive vowel will eventually be reanalyzed as a regular vowel segment, this is by no means a necessary development. Non-syllabic behavior may persist even when the intrusive vowel becomes phonetically quite long. The Outer Hebrides dialects of Scots Gaelic have intrusive vowels that are as long and distinct in quality as regular vowels, yet still act non-syllabic.

The vowels’ curious patterning was commented on at length by the early field-workers Borgstrøm and Oftedal, based on conversations with speakers. Borgstrøm (1940), who uses square brackets to indicate svarabhakti groups, reports of one speaker:¹

¹In Borgstrøm’s transcriptions, *N* indicates a velarised [n], *L* a velarised [l].

Comparing the two words *fäNak* “a crow” (feannag, type *aran*) and *f₁aL₁a₁k* “hunting” (sealg, type *m₁(ara₁v)*) he said: In *fäNak* there is a “space” between the two syllables, so that he could pronounce *fäN — ak*. In *f₁aL₁a₁k* the *L* and the following *k* are so “close together” that such a separation is impossible; the word is “nearly monosyllabic, but not quite monosyllabic”. (p. 153)

It is generally easier to pause between syllables than within a syllable, so the speaker’s difficulty with the task supports the idea that [*faLak*] is one syllable. While Mr. Sinclair balked at calling it monosyllabic, Borgstrøm reports that when asked to divide a long word into syllables, Mr. Sinclair and other speakers treated intrusive vowels as belonging to the same syllable as the preceding vowel. Oftedal (1956:29) makes similar observations:

[S]varabhakti groups are recognized as monosyllabic by educated native speakers. This may be partly due to the spelling, where the second vowel of a svarabhakti group is left out (orm, falbh); but it is significant that in songs, even local òrain that have never been written down, a svarabhakti group is sung on one note.

Ladefoged et al. (1998) show that syllables with vowel intrusion have the same pitch pattern as other monosyllables. Pitch rises during the first syllable of a Scots Gaelic word and falls during the second, so that a disyllabic word like [du.an] ‘hook’ contains a rise and fall, while monosyllabic [duan] ‘song’ has only a rise. In [bal^vak] ‘skull’, which does not have vowel intrusion, pitch rises during the first vowel and falls during the second. In [bal^vak] ‘belly’, which has vowel intrusion, pitch rises throughout, as expected in monosyllables. Bosch & de Jong 1997 present similar data from natural speech. Bosch & de Jong 1998 show that an intrusive vowel coarticulates with the preceding sonorant more than a non-intrusive vowel does: R has more effect on V₂ in CV₁RV₂C than in CV₁RV₂C. Borgstrøm also indicates this in transcriptions. Within Articulatory Phonology, the degree of consonant-vowel coarticulation has been analyzed as corresponding with the degree of gestural overlap (Zsiga 1995; Cho 1998). The coarticulation facts are consistent with the claim that vowel intrusion involves unusually heavy overlap between a sonorant and vowel.

Syllables with vowel intrusion pattern as monosyllables in several ways in the phonology, as pointed out by Smith (1999) and Bosch (1995). In the Argyllshire dialects, short stressed open syllables (which are normally initial) are followed by an epenthetic glottal stop, unless the following consonant is an obstruent, as shown in (4a). The function of this epenthetic coda is apparently to make the stressed syllable heavy, as required by the constraint STRESS TO WEIGHT (Kager 1999). This is shown by the fact that epenthesis does not occur after long vowels or diphthongs, or in already closed syllables: all of these syllable types are heavy already. Epenthesis also does not happen in syllables with an intrusive vowel, which in this dialect is [ə], as shown in (4b).

- (4) Argyllshire Gaelic [ʔ] epenthesis Holmer (1938)
- | | | | | |
|----|-----------|---|-------------------------|---------------|
| a. | /palax/ | → | 'paʔlax | 'boy' |
| | /haraxəy/ | → | 'k ^h aʔraxəy | 'move, stir' |
| | /u/ | → | 'uʔ | 'egg' |
| b. | /menv/ | → | 'menəv | 'fine, small' |
| | /marv/ | → | 'marəv | 'dead' |

If intrusive vowels are non-syllabic, the lack of [ʔ] epenthesis is expected: ['menəv] is a single closed syllable, and is already heavy without an epenthetic coda.

Intrusive vowels license a range of vowel qualities that is not found in non-initial syllables (Bosch & de Jong 1997). Scots Gaelic permits nine short vowels in stressed syllables, and a reduced inventory elsewhere. For example, [u] occurs in unstressed syllables only as a result of optional vowel harmony, as in [turuus] / [turəs] 'journey' (Oftedal 1956:147). Intrusive vowels can be of any of the nine qualities; they do not undergo the neutralisation that is expected in non-initial syllables. Many words have intrusive [u], such as [uruxər] 'a shot', and these do not have an optional [ə] pronunciation. For positional licensing of contrasts, the intrusive vowel is treated as part of the initial, stressed syllable.

Intrusive vowels also pattern as monosyllables in the pattern of mutations. Many words pluralise, or realise other inflections, by palatalizing their final coda as well as raising and/or fronting the preceding vowel. The words in (5a) show that the palatalisation mutation affects only the final rhyme of the word; it doesn't change medial consonants. When a word with vowel intrusion undergoes mutation, as in (5b), both vowels change, and both the sonorant and following consonant palatalise.

- (5) Bernera Gaelic palatalisation Borgstrøm (1940)
- | | Singular | Plural | | |
|----|-----------------------------------|------------------------------------|--|-------------|
| a. | bal ^v əx | bal ^v iç | | 'boy' |
| | fa ^h kə ^v l | fa ^h kil ^v i | | 'word' |
| | sə ^v ləs | sə ^v lɪf | | 'light' |
| | al ^v t | ul ⁱ tʲ | | 'knuckle' |
| b. | bal ^v ag | bul ⁱ ug ⁱ | | 'bellow' |
| | skarəv | skv ⁱ ð ⁱ ɣv | | 'cormorant' |

This pattern suggests that the whole sequence [arəv] is considered one rhyme.

Finally, many disyllabic verb stems undergo syncope before vowel-initial suffixes, as in a) below. This syncope can be analyzed as a strategy for avoiding sequences of unstressed syllables: it creates a 'σσ sequence instead of a 'σσσ one. When syncope brings together two consonants of the type that trigger vowel intrusion, an intrusive vowel occurs in the cluster, as shown in b).

- (6) Syncope Oftedal (1956:239)
- | | | | | |
|----|---------------------------|---|-----------------------|-----------------|
| a. | /obəð + əx/ | → | 'obrəx | 'work (gen.sg)' |
| b. | /bal ^v əx + u/ | → | 'val ^v əxu | 'boy (voc.pl)' |

If the intrusive vowel were syllabic, it would in effect undo the syncope, rendering syncope pointless. But if vowel intrusion is only an arrangement of articulations within a syllable, then the goal of the syncope is met: the word does not contain a sequence of unstressed syllables. The output [$val^v\grave{a}xu$], like [$obr\grave{e}x$], is disyllabic.

2.3 Hocank

The Siouan language Hocank (also known as Winnebago) has relatively long intrusive vowels in tautosyllabic CR clusters. Alderete (1995) and Clements (1991) have analyzed these vowels as non-syllabic; Clements also gives a gestural analysis.

The evidence for non-syllabicity comes primarily from templatic morphology. Reduplication in Hocank normally copies only the final syllable of a stem, as shown in (7a). If the final syllable has vowel intrusion, the whole $CVCV$ sequence reduplicates, as shown in (7b). This is expected, if it is one syllable.

(7)	Hocank reduplication		Miner (1992), Susman (1943)	
a	gihú	‘swing’	gihuhú	‘wag tail’
	wajú	‘dance’	wajifí	‘dance, stop, dance again’
b	fará	‘bald, bare’	farafára	‘bald in spots’
	parás	‘flat’	parapáras	‘wide’

Similarly, roots in Hocank can be CVC , CVV , or $CVCV$, but not $CVCV$ (Miner). If intrusive vowels are non-syllabic, we can maintain the generalisation that roots are limited to one syllable.

Many Hocank stems have ablaut of [a] to [e] before certain suffixes. This ablaut affects only the nucleus of the final syllable. But when the final syllable has vowel intrusion, both the intrusive vowel and the following vowel undergo ablaut.

(8)	Hocank ablaut		Miner (1992)
	hīxe	hīxawi	‘he buries me, he buries us’
	kere	karaire	‘leave returning, 3pl.’

This is expected, if the two vowels are both part of the final syllable. In any case, it is impossible for only one of the vowels to ablaut if they are halves of a single gesture.

Although instrumental studies are not available, fieldworkers report that $CVCV$ sounds different than ordinary, disyllabic $CVCV$. According to Miner (1979), ‘the sequences are spoken (and apparently, sung) faster than other $CVCV$ sequences.’ Susman (1943) comments that ‘in most surroundings, $CVCV$ is intermediate in length between one long and two short syllables,’ and that ‘secondary stress seems to attach equally to both syllables’ of $CVCV$. The treatment of intrusive vowels in the stress system is too complex to cover here, but Hayes (1995:362) notes that his account is compatible with Clements (1991)’s proposal that intrusive vowel sequences are monosyllabic.

2.4 Summary

The three languages discussed above illustrate how the [ə]s or copy vowels that appear between sonorants and other consonants behave as non-syllabic for phonology, phonetics and speaker intuitions. I have concluded from these facts that the relevant CVCV or VCVC sequences are in fact monosyllabic on the surface. Similar arguments can be adduced from many of the languages listed in Figure 1. For others, I do not have evidence that bears either way on syllabicity, but include them as plausible cases based on the vowels' quality and conditioning environment.

An alternate analysis is that intrusive vowel sequences are surface disyllables that pattern with monosyllables because they were monosyllabic at an earlier stage of the derivation. In a rule ordering framework, vowel intrusion could be universally ordered late, perhaps as a postlexical rule. (This is not possible in Optimality Theory, where the markedness constraints that drive syncope, determine reduplicant size, etc. can refer only to surface representations.)

However, simply analyzing vowel intrusion as a process that applies late in the derivation does not answer the important question: why is there a correlation between vowel quality, conditioning environment, and invisibility to other processes? Metrical invisibility is not a general characteristic of epenthetic vowels, or even epenthetic copy vowels; there are plenty of cases where epenthetic vowels are visible to other patterns and hence epenthesis would have to precede other rules.

Under the gestural analysis, the reason for the correlation between conditioning environment, vowel quality, and phonological behavior is that a) the kind of gestural overlap that creates vowel intrusion is limited by general constraints on gestural overlap, among them a constraint against heavily overlapping vowels with obstruents (see section (4)); b) when two vowel sounds are pieces of a single vowel gesture, they necessarily have the same quality (if that quality can be heard at all in the shorter piece); and c) gestural overlap is simply not the same thing as segmental insertion. They are two unrelated processes. The segmental string is not something inferred from the gestural score, but an independent layer of phonological representation. Shifting levels of overlap among gestures does not change the segmental representation. The syllable organises segments, and anything that does not change the number of segments cannot change the number of syllables. Hence, any syllable-counting phonological pattern will be blind to the presence of intrusive vowels.

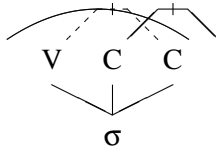
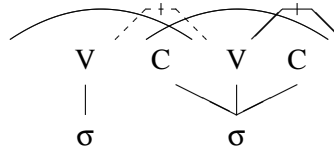
Section 3 reviews several cases of true segmental epenthesis, showing that epenthetic vowels have a different purpose and behavior than intrusive vowels.

3 Intrusion versus epenthesis

While vowel intrusion is the extension of a vowel gesture over a neighboring consonant rather than addition of a segment (and gesture), there are also cases of true epenthesis where a segment and gesture are inserted into the word and create a new syllable. The representations of intrusive and epenthetic copy vowels are given in

(9). I assume (see section 5) that inputs consist of a segmental representation, without any gestural phasing. In the output, each segment is associated with a (group of) gestures, and segments are organised into syllables. In the case of vowel intrusion, there is no change to the segmental representation between input and output. The gestural phasing chosen produces a vocalic-sounding period between the two consonants, but this period does not correspond to a segment and hence is not relevant to syllabification. The output VCC is one syllable. Epenthesis, on the other hand, involves the insertion of a new segment that was not present in the input, along with an associated gesture. This new segment becomes the nucleus of a syllable, so that the output is a disyllable VCVC. Epenthetic copy vowels do not seem to actually share the gesture of the vowel they copy; they do not show the restrictions on intervening consonants that I analyze as resulting from gestural overlap constraints. Rather, the copied quality must be enforced by correspondence constraints (Kitto & de Lacy 2000) or some other means.

(9)

	Vowel Intrusion	Epenthesis
Input:	/VCC/	/VCC/
Output representation:		
Acoustic realisation:	[VCVC] / [VC ^V C]	[VCVC]

It is predicted that the acoustic realisations of vowel intrusion and vowel epenthesis should not be exactly the same. Assuming that the lengths of individual gestures remain constant, the intrusive vowel sequence should be shorter than the disyllabic output of epenthesis. The two outputs might also show different patterns of coarticulation (i.e., different allophonic variation of the consonants or vowels involved), because the gestures overlap one another to different degrees.

In contrast to the vowel intrusion syndrome, the following characteristics usually hold of true epenthetic vowels:

- (10)
- a. The epenthetic vowel removes a structure that is marked, in the sense of being cross-linguistically rare or avoided by other processes within the same language.
 - b. If the vowel's quality is copied, there are no restrictions as to which consonants may be copied over.
 - c. The epenthetic vowel behaves as syllabic.

The last criterion may provoke objections: aren't there cases where epenthetic vowels are skipped by phonological processes? Yes, but this skipping is of a different kind. True epenthetic segments sometimes receive special treatment, but this does not mean that their presence is altogether ignored.

This difference is illustrated by the interaction of stress and epenthesis in Selayarese (Mithun & Basri 1986) Broselow 1999. Selayarese epenthesises copy vowels to remove certain codas in loanwords. Restrictions on what segments can appear in codas are cross-linguistically well-attested, so epenthesis removes a marked structure in accordance with (10a). As stated in (102), copying can happen over non-sonorants, such as [p], [k], and [s] (as in [tapasere] from [tapsir] 'interpretation', [bakari] from the name [bakri], and [baruasa] from [baruas] 'cookie', Broselow 1999). The third criterion, that the vowel should behave syllabically, is superficially contradicted by the fact that, as happens in many languages (Broselow 1982), the stress count skips over epenthetic vowels. Normally stress is penultimate, but it shifts to the antepenult if the final vowel is epenthetic.

- (11) Selayarese stress Broselow (1999)
 Normal: sam'pulo 'ten'
 Final epenthesis: /sahal/ → 'sahala 'profit'

This skipping does not, however, result from the vowels' not being syllabic, or not being phonologically present. Epenthetic vowels are not skipped if this would cause stress to fall outside of the final three syllables. When epenthetic vowels appear in both the final and antepenultimate syllables, stress occurs on the penult as usual.

- (12) Selayarese stress with multiple epenthesis Broselow (1999)
 /solder/ → solo'dere 'weld'
 /kartis/ → kara'tisi 'ticket'

This pattern shows that stress sees the epenthetic vowels, and treats them as syllabic, but simply prefers to avoid them. Broselow (1999) analyzes the pattern with constraints against the presence of an intrusive vowel in a foot (Alderete 1999), combined with constraints on the alignment of the main stress foot with the word edge.

A supporting argument that Selayarese epenthetic vowels are syllabic is simply that they could not fulfill their apparent purpose, resyllabifying codas into onsets, without creating a new syllable. The assumption that the vowels are syllabic is implicit in the claim that they repair syllable structure.

A particularly strong demonstration of the correlation between non-syllabic behavior and conditioning environment comes from languages that have both vowel intrusion and epenthesis. Kekchi and Mono each have two types of inserted copy vowels, one of which copies only over sonorants and one of which copies over any consonant. The one that copies only over sonorants behaves as non-syllabic, while the other behaves as syllabic.

3.1 Kekchi

In the Cobán dialect of Kekchi, a Mayan language of Guatemala, a copy vowel appears within final [ʔ]C clusters (Campbell 1974). These clusters may be tautomorphemic or else result from attachment of the intransitive infinitive suffix $-[k]$ to a [ʔ]-final root.

- (13) Kekchi vowel intrusion Campbell (1974)
- | | |
|-------------------------|-------------------|
| poʔ <u>o</u> t | ‘huipil (blouse)’ |
| kaqtuʔ <u>u</u> j | ‘red ant’ |
| ij <i>i</i> ʔi <u>k</u> | ‘(finger)nail’ |
| seʔ-ek | ‘to laugh’ |
| kwaʔ- <u>a</u> k | ‘to eat’ |

This conditioning environment is typical of vowel intrusion: the copy vowel occurs after a sonorant and copies over that sonorant.

Kekchi also inserts copy vowels between C-final roots and certain C-initial verbal suffixes, as below.² These vowels copy the quality of the vowel to the left, regardless of the identity of the intervening consonant.

- (14) Kekchi vowel epenthesis Campbell (1974)
- | | |
|---------------|------------------|
| ninkwiq’-i-b’ | ‘I bend it’ |
| ninhup-u-b’ | ‘I turn it over’ |
| k’ox-o-b’a:nk | ‘to begin’ |
| atʃ’-a-b’a:nk | ‘to loosen’ |

CC clusters are avoided in many languages, so the epenthesis removes a marked structure.

These two types of copy vowels are treated differently in a language game called Jerigonza. The game consists of inserting after every vowel a sequence [pV], where V is a copy of the preceding vowel. For example, the name of the game [xerigonsa] is rendered as [xeperipigoponsapa]. Since the game causes an alteration in each syllable, it is an indirect test for speakers’ intuitions about syllabicity and the number of vowels in a word.

Words that contain true epenthetic copy vowels have two possible outputs in the game. Either [pV] is inserted after the epenthetic vowel, or the epenthetic vowel deletes.

- (15) kw*i*qib’ank ‘to bend it’ → kw*i*q*i*q*i*q*i*b’apank
or kw*i*qib’apank

²Campbell describes the vowels as appearing in C[m] or C[b’] clusters in verbal forms, but this generalisation is contradicted by examples like [molb’ek] ‘to lay eggs’ and [kwaxb’ak] ‘to play’ (270). Judging from the data in the paper, the process appears to be regular and predictable only before the suffixes [-b’] and [-b’a:nk]. I assume that the basic purpose of the vowel is to avoid consonant clusters, and that the process is morphologically restricted in a way that happens to result in the second consonant usually being [b’].

Campbell suggests that the optional omission of the epenthetic vowel indicates that the game can access the underlying representation, in which these vowels are absent. In any case, this special treatment of epenthetic vowels shows that speakers distinguish them from underlying vowels.

When Jerigonza applies to the intrusive copy vowels that break [ʔ]C clusters, there are also two possible outcomes. As with epenthetic vowels, it is possible to insert a [p]V after each vocalic period, suggesting that these V[ʔ]V sequences are optionally treated as disyllabic. But it is also possible to insert a [p]V only into the intrusive portion of the vowel, while leaving the other portion alone.

- (16) tʃaʔax ‘difficult’ → tʃapaʔapax
or tʃaʔapax

As Campbell points out, ‘The very fact that jerigonza can skip over the first vowel demonstrates that the complex vocalic nucleus (V₁ʔV₁) is perceived in some sense as a single unit.’ The game acts as if such sequences are monosyllabic and the two vowels are one segment, which only needs to be copied once.

3.2 Mono

Mono, a Niger-Congo language of Congo, also has two types of inserted copy vowels (Olson 2001, 2003). As in Kekchi, the vowels that copy over only sonorants behave as non-syllabic, while the vowels that copy over any consonant behave as syllabic. In this case, the diagnostic of syllabicity is a constraint on minimal word size.

All words in Mono must be at least disyllabic. When a monosyllabic root appears unaffixed, a vowel is added to the beginning of it, as shown in (17). The evidence that the vowels are epenthetic is that they disappear in compounds, and also are absent when the same roots occurs in verbal forms, which have more than one syllable due to affixes. The vowels copy the quality and tone of the first vowel in the root, regardless of the identity of the intervening consonant.

- (17) Mono epenthesis Olson (2003)
- | | | | |
|-------|---|------|---------|
| /ɔ̄/ | → | īɔ̄ | ‘tooth’ |
| /bè/ | → | èbè | ‘liver’ |
| /mà/ | → | àmà | ‘mouth’ |
| /ngú/ | → | úngú | ‘water’ |

These vowels show all the properties of true epenthetic copy vowels listed in (10). They copy over any consonant. They appear for the purpose of avoiding monomoraic lexical words, which are cross-linguistically marked (McCarthy & Prince 1993), and they cannot accomplish this goal if they are not syllabic.

Mono also inserts vowels in CR clusters, as shown in (18). These vowels have the same quality as the vowel adjacent to the sonorant.

- (18) Mono vowel intrusion Olson (2003)
- | | | | |
|--------------|---|--------------------|------------|
| /gàfrū/ | → | gàfū <u>rū</u> | ‘mortar’ |
| /plézū/ | → | pél <u>ézū</u> | ‘bat’ |
| /jābrū/ | → | jāb <u>urū</u> | ‘goat’ |
| /dókólóngbā/ | → | dókól <u>óngbā</u> | ‘scorpion’ |

These vowels show the properties of intrusive vowels. They appear only by sonorants, they copy only over sonorants, and they are optionally absent in casual speech. Most importantly, they do not behave as syllabic. Vowel intrusion is not sufficient to bring a word up the disyllabic minimum: even if a root like /gré/ ‘big’ is pronounced with an intrusive copy vowel, an epenthetic copy vowel must also be added.

- (19) Mono intrusion & epenthesis Olson (2003)
- | | | | |
|--------|---|----------------|---------|
| /gré/ | → | ég <u>éré</u> | ‘big’ |
| /kplú/ | → | úk <u>púlú</u> | ‘heap’ |
| /prō/ | → | ōp <u>ōrō</u> | ‘egg’ |
| /krō/ | → | ōk <u>ōrō</u> | ‘skull’ |

Mono and Kekchi show that the correlation between conditioning environment and syllabic behavior holds even of two types of copy vowels within the same language. The type of vowel that copies over only sonorants does not create a new syllable, while the type of vowel that copies over any consonant does behave as syllabic.

3.3 What is a syllable?

An important assumption of the theory presented here is that the syllable is not an acoustic object. A syllable is an abstract unit that organises segments. Segments are not the same thing as sounds: a word may contain a very audible vocalic-sounding period that does not correspond to any independent segment. Such a vocalic period cannot be the nucleus of a syllable.

For this reason (and others), speakers of different languages may interpret the same acoustic signal as containing different numbers of syllables. For example, Harms (1976), who gives an informal gestural account of intrusive [ə] in Finnish, reports that Finnish and English speakers interpret acoustically similar [ə]s differently.

[mɛləkein] (*melkein*) ‘almost’ has essentially the same vowel qualities ([ɛ, ə, ei]) and relative durations as the English verb *delegate*—[dɛləgeɪt]. From a descriptive phonetic point of view, the Finnish epenthetic schwa and the English reduced-vowel schwa represent very nearly identical classes of vowel sounds; i.e., they vary over a wide central area, with their range of variation conditioned by the preceding and following segments. But here the similarity ends. The schwa in the above Finnish forms is purely transitional in nature. Speakers perceive these forms as containing only two syllables, not three.

Harms' comment runs counter to a common belief that where there is a vocalic-sounding period of a certain duration or perceptual prominence, there must be a syllable—i.e., that the syllable is at least partially an acoustic object. This assumption crops up frequently in the literature. For example, Donselaar et al. (1999:64) conclude from their experiment on Dutch words like [tʏləp] that 'the realizations of real words with schwa epenthesis are represented by listeners as monosyllabic', and suggest that 'from the speakers' point of view, schwa epenthesis may not arise via insertion of a segment as such, but simply via realization of the gestures corresponding to articulation of the consonant cluster'—exactly the analysis advocated here. Yet they also comment that 'obviously, adding a vowel between two consonants adds an extra syllable to the word; the optional form with epenthesis has one more syllable than the underlying form without.' The unstated assumption is that segments and syllables have some type of reality outside of the speaker's representation of them. It is precisely this that I wish to argue against. Syllables exist in the mind, not in the air.

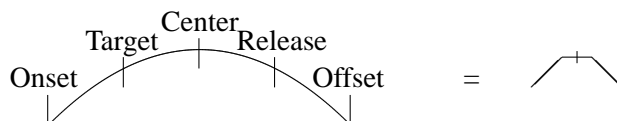
4 Constraints on gestural coordination

4.1 Consonant-consonant coordination

Since intrusive vowels are not segments or syllables, their distribution cannot be explained with the kind of markedness constraints that determine permissible segmental sequences (such as constraints against codas or complex margins, or constraints on sonority sequencing). Rather, vowel intrusion is produced by constraints on gestural coordination, like those in the Optimality-Theoretic gestural grammar proposed by Gafos 2002. These constraints have the purpose of maximizing perceptual cues in consonant clusters, and avoiding types of gestural overlap that are marked, either for perceptual or articulatory reasons. Before laying out the analysis, it is first necessary to introduce the basic apparatus of Articulatory Phonology (Browman & Goldstein 1986 et seq).

A gesture is an abstract specification of a constriction within the vocal tract, essentially a command to an articulator to achieve specified positions at specified time points. Most segments correspond to multiple gestures. A [t], for example, requires the gestures [tongue tip alveolar closure], [glottis wide], and [velum raised]. But in this paper, since intra-segmental gestural coordination is not at issue, each segment will be treated as if it consisted of a single gesture, the oral constriction. Unlike a feature, a gesture has an internal temporal structure. Every gesture has five 'landmarks': the ONSET of movement, the TARGET, or point when maximal constriction is reached, the CENTER of the constriction phase, the RELEASE of the constriction, when the articulator begins a controlled movement away from the target position, and the OFFSET, when the articulator ceases to be under active control. For simplicity, the gestural curve may be drawn with angles to represent the landmarks, as shown below right. I will show all consonant gestures as the same length, although this is of course not true in reality.

(20) Landmarks in gestural life



Gestures overlap one another, and the grammar determines the degree of this overlap by specifying alignments of landmarks in overlapping gestures. Although the duration of a gesture will vary due to speech rate, the relation between certain landmarks in overlapping gestures remains constant. For example, in a sequence of two consonants, the grammar might specify that the center of the first consonant should be simultaneous with the onset of the second. A sequence of gestures with specified phasing relationships is called a gestural score.

The gestural score is part of a computational gestural model developed at Haskins Laboratories, described in Browman & Goldstein (1990). A model of task dynamics, based on a general theory of skilled motion, converts the gestural score to articulatory trajectories. The task dynamic model cannot always perfectly carry out the idealised commands encoded in all of the abstract gestures. If two overlapping gestures want the tongue to be in two different positions at once, for example, the tongue's actual trajectory will be a compromise between these two targets. This is coarticulation. Once the articulatory trajectories are calculated, an articulatory simulator converts these trajectories to an acoustic output. In this way it is possible to test the acoustic result of a given alignment of given gestures at a given speech rate.

The acoustic result of a gestural score depends not only on the phasing of the gestures, but on the characteristics of the gestures involved. Simulations reported by Gafos (2002:271–272) show that a phasing relation of CENTER = ONSET produces a schwa-like sound (an acoustic release) between heterorganic consonants but not between homorganic consonants.

(21) Homorganic and heterorganic clusters, CENTER = ONSET phasing



Intuitively, the reason for the lack of release in [lt] at this level of overlap is that the tongue tip has the same target constriction for both consonants. At the time when the tongue tip receives the order to move away from the first target (RELEASE of [l]), it has already received an order to begin moving towards the same target again (ONSET of [t]). The easiest path is to stay in place. Only if the [t] gesture began after the release of [l] might an acoustic release occur between the two consonants.

If consonant clusters with short intrusive vowels have the phasing CENTER = ONSET, this helps explain why vowel intrusion is typically restricted to heterorganic clusters.³ Finnish, for example, has vowel intrusion in [lv lj lh lm lp nh] but

³However, it may be that this restriction holds, crosslinguistically, only in RC clusters and not

not [lt ls ns]. Non-gestural theories have to include separate rules for the two sets of cases, but in the gestural approach we can say that all the consonant clusters above have the same CENTER = ONSET alignment. The clusters sound different, because this alignment produces a release only between heterorganic gestures, but from the point of view of the grammar they are all treated the same. Harms (1976) gives a description suggesting that this is actually what occurs in Finnish:

Of special relevance for the epenthesis case is the careful control of the syllable division, providing a clear separation between the final consonant of the first syllable and the initial consonant of the following syllable. This ‘control’ process occurs with all nongeminate sonorant plus consonant clusters, both homorganic (as in *valta* ‘power’, *parta* ‘beard’, *kansa* ‘folk’) and nonhomorganic (e.g., *valmis*, *kalja* ‘near beer’, *surma* ‘death’, *korko* ‘interest’)... In all these clusters, to a greater or lesser extent, the energy of the first syllable is ‘spent’ before the onset of the next syllable, and the result is often a very short pulse of energy, vocoid-like in nature, at the end of the syllable. (p. 77)

Attributing vowel intrusion to CC alignment also helps explain why intrusive vowels are often variable in duration, and may disappear at fast speech rates (as noted by Bye (2001:139) for Saami, Holmer (1938:32) for Argyllshire Gaelic, Harms (1976:77) for Finnish, Quilis (1981:298) for Spanish, Jannedy (1994) for German, Parker (1994) for Chamicuro). In Chamicuro, for example, intrusive vowels occur only in slower speech.

(22)			Parker (1994)
	Normal speech	Emphatic speech	
	'tuʔlu	'tuʔ <u>u</u> lu	‘chest’
	jap'leʔti	jap'leʔ <u>e</u> ti	lightning’
	maʔ'nali	maʔ <u>a</u> 'nali	‘dog’

In Articulatory Phonology, speech rate is modeled by adjusting a property of the gesture called stiffness, which is part of the equation that describes the gestural curve. Gafos (2002) shows that increasing the stiffness of gestures with the CENTER = ONSET alignment eventually results in the disappearance of the release even in heterorganic clusters. An utterance may have the same gestural alignments at all speech rates, yet have different acoustic realisations in slow and fast speech.

Thus, two common characteristics of intrusive vowels— their restriction to heterorganic clusters and disappearance at fast speech rates—result in this approach from general principles of task dynamics rather than from the grammar

in CR. Oscan, for example, which developed copy vowels in both RC and CR clusters, has them in [dr] and sometimes [tr] but not in [rt] or [lt] (*húrtúm* ‘grove’ vs. *patereí* ‘father’). Hocank has vowel intrusion in [sn] and [sr] clusters as well.

itself. There is no need for rules that refer specifically to homorganic or heterorganic clusters, nor for rules referring to speech rate. The gestural approach allows a simpler grammar because the model includes a way to incorporate physical factors that are not specifically linguistic.

In an Optimality Theoretic approach, gestural phasing is determined by ranked constraints. Gafos (2002) proposes a basic form for constraints on gestural overlap, as follows:

- (23) ALIGN (G_1 , LANDMARK₁, G_2 , LANDMARK₂) Gafos 2002
Align landmark₁ of gesture₁ with landmark₂ of gesture₂.

In tableaux, I abbreviate such constraints as G_1 LANDMARK₁ = G_2 LANDMARK₂, and abbreviate landmarks as ONS(et), TAR(get), CEN(ter), REL(ease), and OFF(set). The phasing CENTER = ONSET in consonant clusters is produced by the constraint C_1 CEN = C_2 ONS, whose full definition is as follows.

- (24) ALIGN (C_1 , CENTER, C_2 , ONSET)
In a C_1 C_2 sequence, the center of C_1 is aligned with the onset of C_2 .

A competing constraint favors the alignment RELEASE = TARGET in CC clusters. This alignment produces no release even in heterorganic clusters (Gafos 2002).

- (25) ALIGN (C_1 , RELEASE, C_2 , TARGET)
In a C_1 C_2 sequence, the release of C_1 is aligned with the target of C_2 .

The ranking of these two constraints determines the whether consonant clusters will have an alignment that produces release in at least some cases (CENTER = ONSET) or an alignment that never produces release (RELEASE = TARGET).

These competing C-C coordination constraints likely reflect the competing priorities of ease of articulation and ease of perception. The alignment RELEASE = TARGET allows a faster and in that sense more efficient articulation. Constraints favoring less overlap, like CENTER = ONSET, have a functional grounding in perceptibility: a consonant has clearer perceptual cues (such as a release burst, or formant transitions) when it is not heavily overlapped by another consonant. Donselaar et al. (1999) show that in Dutch, where vowel intrusion is optional (at any speech rate), reaction times to lexical decision tasks and phoneme identification tasks are quicker when a word like *tulp* ‘tulip’ is pronounced with vowel intrusion ([tʏlɒp]) than without ([tʏlp]).

Enhancing perceptibility of the consonants is likely the main motivation for the gestural phasing that produces intrusive vowels in consonant clusters. This helps explain a peculiarity of the vowels’ distribution noted in (1): the fact that the clusters they occur in are rather unmarked clusters, at least in terms of sonority sequencing. The same languages that have vowel intrusion in rising sonority onsets or falling sonority codas often leave obstruent-obstruent clusters alone.

(26) Examples of clusters:

	Triggering vowel intrusion	Not triggering vowel intrusion
Hocank	pr- kr- kw- sn-	ps- kf- xg- st-
Scots Gaelic	-rb -lv -rf -nx -mj	-g'f -xg -sb -vh

If vowel intrusion were epenthesis, with the goal of resyllabifying marked margins, then this distribution would be curious. Cross-linguistically, rising sonority onsets and falling sonority codas are the most favored cluster types; if a language allows only a few clusters they are usually these. Generally processes like deletion and epenthesis remove the worst structures, not the best ones. But if vowel intrusion does not, in fact, change the structure of the syllable by addition of a segment, but only reflects a particular scheme of phasing gestures within the syllable, then there is no reason to expect that vowel intrusion would be related to the markedness of the clusters involved.

4.2 Consonants / vowel phasing

The second requirement for copy vowel intrusion, in addition to the presence of an acoustic release in a consonant cluster, is that a vowel articulation overlap this period of release. Without this, the release may sound [ə]-like, but will not be transcribed as a copy vowel. To achieve the needed alignment, constraints must cause a vowel to heavily overlap adjacent consonant clusters.

Browman & Goldstein (2001) report that in English, a vowel bears a consistent timing relation with the center of an onset consonant. This suggests that there exists a constraint aligning the onset of each vowel to the center of the leftmost consonant in its syllable. I add a symmetrical constraint aligning the offset of each vowel to the center of the rightmost consonant in the syllable. (In English the phasing of onsets and codas is actually not symmetrical, and the treatment of onsets with multiple segments is more complex than this. But these constraints suffice for modeling vowel intrusion.)

(27) ALIGN (V, OFFSET, C]_σ, CENTER)

The offset of every vowel is aligned with the center of the rightmost consonant that belongs to the same syllable as that vowel.

(28) ALIGN (σ[C, CENTER, V, ONSET)

The onset of every vowel is aligned with the center of the leftmost consonant that belongs to the same syllable as that vowel.

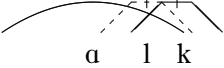
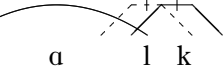
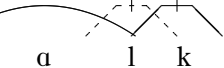
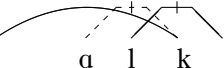
In a VCC or CCV sequence, the vowel necessarily has greater overlap with the nearest consonant than with the peripheral consonant. The phasing constraints in (28) and (27) only care about the alignment of the vowel with respect to the peripheral consonant. But satisfying them requires a heavy overlap between the vowel and non-peripheral consonant. I propose that heavy overlap is dispreferred by constraints of the type below.

- (29) *GESTURE_x LANDMARK₁ PAST GESTURE_y LANDMARK₂
 In a sequence gesture_x gesture_y, landmark₁ of gesture_x is not later than
 landmark₂ of gesture_y

The constraints *C CENTER PAST V ONSET and *V OFFSET PAST C CENTER prevent a vowel from overlapping any consonant more than halfway. This limits the degree to which a vowel can extend across a consonant cluster and hence be heard during the release, if any. When these constraints are high-ranked, release in a consonant cluster is not colored by an adjacent vowel.

At this point, we have four constraints that apply to any given CCV or VCC cluster. Tableau (30) demonstrates how these constraints evaluate output candidates for the input /alk/. The representations being evaluated each consist of a gestural score and a segmental string. The phonetic transcription, given in brackets, is parenthetical, to remind the reader how the phasing would sound. It is not part of the representation evaluated by the constraints. The output candidates differ on two parameters: the degree of overlap between the two consonants, and whether the vowel gesture extends to the center of [k] or not. Candidates a) and b) have a close transition between consonants, with no release. They differ only in how much the vowel gesture overlaps the cluster. In a) the vowel extends to the center of [k], which satisfies V OFFSET = C]_σ, CENTER. This alignment violates *V OFFSET PAST C CENTER, because the offset of V is later than the center of [l]. In b) the vowel extends only to the center of [l], satisfying *V OFFSET PAST C CENTER but violating V OFFSET = C]_σ, CENTER. Candidates a) and b) would sound similar, except that the vowel and consonants would coarticulate more in a) due to their greater degree of overlap. Candidates c) and d) have the phasing CENTER = ONSET between the consonants, which produces a release between them. In c), this release is not (strongly) colored by the vowel, since the vowel gesture extends only to the center of [l]. This would describe a language that has open transitions between consonants, but without copy vowel intrusion. In candidate d), the vowel overlaps the release between the consonants, and a piece of the vowel gesture is heard as an intrusive copy vowel. There are twenty-four possible rankings of the four constraints; each candidate wins under six rankings.

(30) Constraints on gestural phasing

/alk/	C ₁ CEN = C ₂ ONS	C ₁ REL = C ₂ TAR	V OFF = C]σ CEN	*V OFF PAST C CENT
a.  [alk]	*			*
b.  [alk]	*		*	
c.  [alə̯k]		*	*	
d.  [alək]		*		*

These four constraints establish a basic typology of phasing in VCC or CCV sequences: close transition without release, open transition with an uncolored release, and copy vowel intrusion.

4.3 The behavior of sonorants

Vowel intrusion happens only with sonorants. Languages that have intrusion vowels next to some or all sonorants have either plain, non-vocalic sounding releases or no release at all next to other consonant types. Furthermore, intrusive vowels typically have the quality of the vowel that is next to the sonorant. Even in languages where [ə̯] is transcribed, the [ə̯] is often colored by the vowel that is next to the sonorant, as has been instrumentally confirmed for Lakhota (Albright 1999) and Chilean Spanish (Quilis 1981). Capturing the difference between sonorants and obstruents requires constraints referring to these consonant classes. There is evidence that sonorants are subject to special C-C phasing constraints, and also that overlap between a vowel and sonorant is less marked than overlap between a vowel and obstruent.

First, there seems to be a preference for having less gestural overlap in RC or CR clusters than other CC clusters. The Salish language Upper Chehalis has excrescent [ə̯] where one of [m n j l w ʔ] follows another consonant (Kinkade 1963–4). Kinkade argues that these schwas are purely transitional; in gestural terms this means they result from a phasing like CENTER = ONSET between adjacent

consonants. In their phonological behavior, these [ə]s contrast with acoustically similar epenthetic [ə]s that appear in different environments.

- (31) Upper Chehalis excrescent schwas Kinkade (1963–4:192–3)
- | | |
|-------------------|------------------------|
| xə́łə́lp | ‘open-weave basket’ |
| č’óp’waxə́nstwə́n | ‘squeeze with the arm’ |
| q’wółə́weʔ | ‘maple’ |
| sqə́ʔúq’wə́ls | ‘skull’ |

These transitional schwas do not take the quality of a neighboring vowel, and do not need to be near a vowel at all. Many consonants of any type may intervene between the schwa and the nearest full vowel. Thus, Upper Chehalis is a useful example because it allows us to examine CC phasing constraints in isolation from any influence of VC or CV phasing constraints. What Upper Chehalis shows is that sonorants may have a different phasing relation with neighboring consonants than other consonants do. This pattern can be produced by adding a more specific version of the constraint in (24) that refers specifically to sonorants.

- (32) ALIGN (C, CENTER, R, ONSET)

In a C₁ C₂ string, C₂ a sonorant, the center of C₁ is aligned with the onset of C₂.

Upper Chehalis has the constraint ranking ALIGN (C, CENTER, R, ONSET) » ALIGN (C, RELEASE, C, TARGET) » ALIGN (C, CENTER, C, ONSET). Under this ranking, CR clusters have CENTER = ONSET phasing, while all other CC clusters have the phasing RELEASE = TARGET. Hence, release occurs only in CR clusters. In a language that had vowels overlapping consonant clusters, this ranking would produce vowel intrusion in CR clusters but no release in CC clusters.

Even in languages that have releases in all cluster types, however, copy vowel intrusion may happen only by sonorants. The Papuan language Hua (New Guinea) has releases, transcribed as [ə], between all consonants in careful speech, but the release has the quality of the following vowel only in a C[r] or C[ɣ] cluster (Haiman 1980). ([ɣ] often patterns as a guttural, as in Arabic.)

- (33) Plain and colored releases in Hua Haiman (1980:26–27)
- | | |
|--|--------------------|
| f ^ə tu | ‘smell’ |
| k ^ə t ^ə r ^ə guʔ | ‘kind of mushroom’ |
| t ^ə v ^ə r ^ə gie | ‘he sharpened it’ |
| pot ^ə yaie | ‘it glanced off’ |
| ok ^l rumaʔ | ‘sky’ |
| f ⁱ rie | ‘he died’ |
| t ^ə re t ^ə refie | ‘it flowed’ |

As Upper Chehalis allowed us to isolate the role of CC phasing constraints, Hua allows us to isolate the role of CV phasing constraints. Apparently a vowel gesture overlaps a consonant cluster in Hua only if the closest consonant is a sonorant. The

language avoids heavy overlap of vowels with obstruents. This can be captured by positing anti-overlap constraints, following the general form in (34), that refer to different consonant types. One such constraint is *OBSTRUENT CENTER PAST V ONSET.

(34) *OBSTRUENT CENTER PAST V ONSET

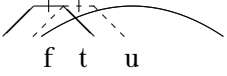
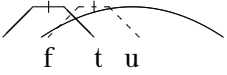
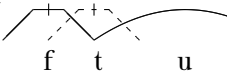
In a sequence CV, where C is an obstruent and V a vowel, the center of C is not later than the onset of V.

When this constraint is higher ranked than ALIGN (σ [C, CENTER, V, ONSET]), a vowel will heavily overlap a CR cluster but not a CC cluster, as in Hua. The tableaux below demonstrate how these two realisations of different clusters are chosen. For the two inputs /kra/ and /ftu/, the alignment of the CC clusters in the output is the same: in each case, a CENTER = ONSET phasing is chosen. Only the degree to which the vowel overlaps the clusters is different. In /kra/, the vowel gesture can extend to the center of [k], as demanded by σ [C CENTER = V ONSET]. This alignment doesn't violate *OBSTRUENT CENTER PAST V ONSET, because [r] is not an obstruent. Since the vowel extends over the period of release, the release sounds like a short copy vowel in [k^ara]. But for /ftu/, *OBSTRUENT CENTER PAST V ONSET blocks the vowel gesture from extending to the center of [f]. Thus, the release in [f_ətu] does not sound like a copy vowel.

(35) Hua

/kra/	C ₁ CEN = C ₂ ONS	C ₁ REL = C ₂ TAR	*OBS CEN PAST V ONS	σ [C CEN = V ONS
a. k r a [kra]	*!			
b. k r a [k ^a ra]		*		
c. k r a [k ^a ra]		*		*!

(36) Hua

/ftu/ 'smell'	C ₁ CEN = C ₂ ONS	C ₁ REL = C ₂ TAR	*OBS CEN PAST V ONS	V OFF = σ OFF
a.  [ftu]	*!			
b.  [fʰtu]		*	*!	
c.  [fʰtu]		*		*

To summarise, two types of constraints may refer specifically to different classes of consonants: constraints on C-C alignment and constraints limiting C-V or V-C overlap. Both of these constraints types contribute to limiting vowel intrusion to CR and RC clusters. As the next section will show, there is also evidence for constraints that refer to more specific classes than obstruents and sonorants.

4.4 Hierarchy of C/V overlap

Besides being restricted to sonorants generally, vowel intrusion happens only with a subset of the inventory of sonorants in most languages. Some examples are given in (37).

(37) Sonorant inventories

Language	Triggering intrusion	Not triggering intrusion
Kekchi	ʔ	r l n m w j h
Mamainde	h	n m (l j w)
Tiberian Hebrew	ʔ ʕ h ħ	r l n m w j
Spanish	r	(r) l m n ɲ ʎ
German	ʁ	l m n j v (ŋ)
Sanskrit	r (l) ^a	m n ṅ ɲ ŋ j h
Dutch	r/ʁ l	m n ŋ (h)
Hausa	r l	r n m w j ʔ
Finnish	l n h	r (m ŋ j)
Oscan	r l n	m h
Saami	r l j β ð ^b	m n ɲ (h)
Hocank	r n w	ʔ (h j m)
Scots Gaelic	r rʲ rʲ ʲ ʲ n nʲ nʲ m	j (w h)

^aSome authorities report svarabhakti after only [r], others after both [r] and [l].

^bAccording to Bye 2001, [β] and [ð] are approximants in Saami.

Sometimes, a particular sonorant simply never appears in an environment where vowel intrusion is possible: for example, Dutch has [h] but it does not occur in codas, which is where vowel intrusion takes place. Such cases, where known to me, are in parentheses. But in other cases, a sonorant does occur in the appropriate position yet lacks vowels intrusion. From the list above, there emerges a partial cross-linguistic implicational hierarchy as to which sonorants trigger vowel intrusion.⁴

(38) Vowel intrusion triggers

(obstruents, if ever) > other approximants, nasals > [r] > [l] > [r], [ʁ] > gutturals
Among nasals: [m] > [n]

This hierarchy suggests that some alignment constraints distinguish not just between sonorants and obstruents, but between different classes of sonorants.

In most languages, it is not reported whether the RC clusters without vowel intrusion have release or not. This makes it difficult to tell whether the constraints that distinguish among different sonorants are C-C phasing constraints like (32), anti-overlap constraints like (34), or both. However, a few phenomena outside

⁴Hocank is an exception to this hierarchy, since it does not have vowel intrusion in [Cʔ] clusters. It is possible that these are really glottalised single segments (although Susman (1943:21) argues for a cluster analysis). Finnish is also an exception: while some sources (Harrikari 1999) claim there is vowel intrusion after [r], this seems to be untrue of at least the modern Ostrobothnian dialect (Suomi 2000). However, Harms (1976) remarks that '[r] clusters are basically no different from [l] clusters...although the available energy generally results in a stronger trill instead of a vocoid-like release.' In other words, [rC] clusters may be treated the same way as [nC] and [lC] by the grammar, but have a different acoustic output due to characteristics of the [r] gesture.

vowel intrusion suggest that at least specialised V-C overlap constraints must exist. Occasionally, other copy vowel phenomena show sensitivity to the hierarchy in (38), in situations unconnected with consonant clusters. One such case, discussed in section 4.5, is the development of vowels in foreign loanwords. Another kind of evidence comes from the phenomenon of articulatory troughs. In a V_iCV_i sequence, there is generally an articulatory trough, or relaxation of the vowel articulation, during the consonant. Articulatory troughs have been argued to indicate the juncture between two separate vowel gestures (Lindblom et al. 2002), while the absence of a trough would presumably indicate a single vowel gesture overlapping the consonant. Harris & Bell-Berti (1984) find that troughs are lacking when an English speaker produces [h] or [ʔ] between identical vowels. This suggests that in this case only, a single long vowel gesture surrounds the consonant, rather than each vowel segment corresponding to a separate gesture as usual. Whatever the reason for this gestural realisation, the fact that it happens over gutturals but not other consonants is consistent with the idea that vowels overlap gutturals most easily.

This universal hierarchy of C / V overlap can be achieved by aligning the constraint barring heavy C / V overlap with the hierarchy above. For onsets, this yields the following constraints:

- (39) *OBSTRUENT CENTER PAST V ONSET \gg *GLIDE CENTER PAST V ONSET, *NASAL CENTER PAST V ONSET \gg *[r] CENTER PAST V ONSET \gg *[l] CENTER PAST V ONSET \gg *[r] CENTER PAST V ONSET, *[β] CENTER PAST V ONSET \gg *GUTTURAL CENTER PAST V ONSET

These constraints are in a fixed ranking in every language. Thus, a language cannot allow heavy overlap between vowels and nasals, for example, without allowing the same degree of overlap between vowels and liquids. A fixed ranking produces the type of implicational hierarchy seen in the data (Prince & Smolensky, 1993).

4.5 The V-C overlap hierarchy in loanword adaptation

Another phenomenon in which vowel copy may happen only over certain consonants is the addition of vowels to foreign words in the course of loanword adaptation. In Cook Islands Maori and Mawu, vowel copy in loanwords occurs only over the consonants for which I have proposed that VC overlap is least marked. This is likely because the relevant vowels developed, in some cases, through an intermediary vowel-intrusion-like stage in the interlanguage that precedes full nativisation of foreign words. Although the resulting vowels are sometimes called epenthetic, they must be synchronically represented as underlying because their quality is unpredictable. It is important to distinguish these diachronically added vowels from synchronically epenthetic vowels, because they would otherwise be a counterexample to the claim in (10) that epenthetic copy vowels copy over any consonant.

Cook Islands Maori adds vowels to the end of C-final borrowings (Kitto 1997). The added vowel is usually either a copy of the preceding vowel or else a default [i]. The choice between copy vowel and [i] is unpredictable, but Kitto & de Lacy (2000) note that statistically, the patterns are strongly influenced by the identity of the intervening consonant. The copy vowel option is chosen most consistently when the final consonant is [r], as in [pe:re] ‘bail’. After [n], both copy vowels and [i] are fairly common, as in [ri:pene] ‘ribbon’ vs. [mere:ni] ‘melon’. Here the identity of the preceding vowel plays a role as well; copy vowels are chosen more often after [en] than after [an]. After [t], copy vowels are rare. The added vowel is usually a default [i], as in [ti:keti] ‘ticket’. Thus, the consonants least > most likely to have vowel copy occur over them are [t] > [n] > [r]. This is fully consistent with the hierarchy in (38). If, as I have proposed, this hierarchy is a result of gestural overlap constraints, it suggests that gestural overlap played a role in shaping the Maori pattern. I suggest that the copy vowels began as something similar to an intrusive vowel, created by extending the preceding vowel articulation over the final consonant, and were reinterpreted as separate segments. This reanalysis happened through mishearings and misinterpretations of the gestural phenomenon, and hence happened inconsistently. In cases where speakers did not perceive a copy vowel after the consonant, they added an [i] to satisfy Maori syllable patterns. This happened mostly after obstruents, which had not been overlapped as much by preceding vowels.

Synchronically, the copy vowels are not likely to involve gestural overlap, or indeed to be represented as epenthetic. If words like [ri:pene] and [mere:ni] still had the underlying representations /ripen/ and /meren/, it would be hard to explain why the grammar inserts an [e] after the first but an [i] after the second.

A similar example comes from Mawu, a Manding language of Côte d’Ivoire (Moussa 1996; Kenstowicz 2001) In adapting loanwords, Mawu breaks up French C[ɸ] clusters with a copy vowel, but C[l] with a default high vowel. Copying over [ɸ] but not [l] is consistent with (38), and suggests that the pattern stems from vowel gestures extending over [ɸ] but not [l]. But the striking twist is that French [ɸ] and [l] are both realised as [l] in Mawu.

(40) Mawu loanwords from French

brosse bòlèsí
France fàlàzí
plan pílá
bloque búlókí

If the ranking *[l] CENTER PAST V ONSET ≫ *[ɸ] CENTER PAST V ONSET is indeed responsible for this pattern, the arena where these constraints act cannot possibly be located within the Mawu grammar. There is no distinction there between [l] and [ɸ] for the constraints to refer to. It is more likely that the loanwords reflect the Mawu speakers’ sensitivity to a greater overlap of vowels with [ɸ] in French, or perhaps in an interlanguage used by bilingual speakers that still distinguishes the two liquids. The French encountered by Mawu speakers may have

a slight intrusive vowel in C[ɸ] clusters, which Mawu monolinguals expand to a full vowel. There is less quality to the release (if any) in C[l] clusters, so here speakers insert a default vowel. But the surface unpredictability of the historically epenthetic vowel's quality indicates that it must now be represented underlyingly, and not shaped purely by markedness constraints in the way that a synchronically epenthetic or intrusive vowel is.

4.6 Other factors affecting vowel intrusion

The constraints proposed above capture the most salient facts about the distribution of intrusive vowels. However, I wish to briefly mention a few more factors that appear to be conducive to vowel intrusion in some languages: C-V homorganicity and gemination.

Two languages show a particular preference for vowel intrusion next to geminates. Saami has vowel intrusion after geminate sonorants but not singleton sonorants (Bye 2001), and some dialects of Finnish have vowel intrusion in RC: but not RCC clusters (Harrikari 1999). This suggests that there may be special C-C phasing constraints that apply only to geminates, such as the following.

(41) ALIGN (C, CENTER, C:, ONSET)

In a C₁C₂ sequence, where C₂ is a geminate, the center of C₁ is aligned with the onset of C₂.

Additional support for such a constraint is the fact that other languages prefer release before or after geminates. Imdlawn Tashlhiyt Berber allows releases in C_iC_i or C_i:C_i clusters but not in most C_iC_j clusters (Dell & Elmedlaoui 1996). Plain release in clusters is produced by the same kinds of C-C phasing constraints that are involved in vowel intrusion. Languages also may favor release of geminates even when they are not in clusters: word-final geminate stops in Amharic are obligatorily released, while after singletons release is optional in some environments (Hudson 1995:664). Stefania Marin (p.c.) reports similar facts for final geminates and singletons in Wolof. Word-final release cannot be modeled with constraints on C-C phasing (as far as I am aware, no gestural treatment of final release or non-release has been proposed), but it probably reflects the same perceptual pressures that are the phonetic grounding for CENTER = ONSET phasing. Intuitively, it is clear why release should be especially important in geminates: it helps the listener to detect the length of the closure, and hence to distinguish geminates from singletons.

Another factor that can affect gestural overlap is the homorganicity of V / C pairs. Negev Bedouin Arabic has vowel intrusion⁵ in guttural-initial clusters only when they follow the vowel [a], as shown in (42a) vs. (42b). In Articulatory

⁵Synchronically, these vowels are probably no longer intrusive in this dialect, as shown by the fact that the open-syllable raising process optionally happens in one vowel, as in [ʃahar] / [ʃihar] 'month'.

Phonology, both low vowels and guttural consonants have a [pharyngeal] constriction. In Negev Bedouin Arabic, then, heavy vowel-consonant overlap only happens when the vowel and consonant are homorganic.

- (42) Negev Bedouin Arabic Blanc (1970)
- | | | | | |
|----|------|---|-------|-----------|
| a. | baxt | > | baxat | ‘luck’ |
| | mahl | > | maħal | ‘drought’ |
| b. | | | uxt | ‘sister’ |
| | | | ihfir | ‘dig’ |

This restriction suggests that there is a special avoidance of vowels overlapping consonants that have a different place of articulation. Support for this view comes from cases of diachronic C / V metathesis. Steriade (1990) proposes that metathesis can result from movement of a consonant gesture across a vowel gesture. If correct, this makes metathesis another source of evidence for constraints on C / V overlap. Grammont (1933) also classifies some metatheses as ‘inversion by penetration’, which he describes as the movement of one sound through another. He proposes that a vowel and consonant must be articulated in the same general region in order for ‘penetration’ to occur, noting that in the French dialects Pléchatel and Havre, [r] has metathesised with [e] but not with other vowels, and that in unaccented syllables in Latin, there is historical metathesis of [i] with [r] and [u] with [l], but not of the same liquids with other vowels. Recall also that in Cook Islands Maori, discussed in section 4.5, [e] is more likely to copy over [n] than [a] is. All these facts suggest there are constraints like the following:

- (43) *V OFFSET PAST HETERORGANIC C CENTER
 In a heterorganic VC sequence, the offset of V is not later than the center of C.

This constraint blocks vowel intrusion only in the case that the vowel and consonant are heterorganic.

Vowel intrusion is a good source of evidence for constraints on gestural organisation because it is a gestural phenomenon unusually amenable to typological survey. Since vowel intrusion is clearly audible, it is often mentioned in descriptive works, unlike, for example, release. Furthermore, intrusive vowels are distinguishable from ordinary vowels by their phonological behavior. Since vowel segments form the nuclei of syllables, they influence a large range of syllable-conditioned phonological processes. Intrusive vowels are conspicuous in their failure to do so. (By contrast, intrusive stops, while also audible and often transcribed, offer few clear diagnostics to tell whether speakers treat them as normal segments or not. For example, a [t] is often heard in the final cluster of *mince* [mmʰs]. But since an extra consonant in that position would not be expected to affect the syllable’s phonological behavior, there is little way to tell from phonology whether the [t] is segmental or not.) Vowel intrusion is only one case within wider patterns of release and VC overlap, and its importance ultimately is in the light it sheds on more general constraints on gestural phasing.

5 Timing-Augmented Surface Phonology

Besides providing evidence for particular gestural constraints, vowel intrusion has implications for the overall shape that a gestural grammar should take. The approach I use here can be called Timing-Augmented Surface Phonology (TASP). TASP incorporates gestures into the surface phonological representation, but limits the types of constraints that can refer to them and analyzes many phonological processes as non-gestural in nature.

In TASP, an output representation consists of a segmental score and a gestural score. Segments and gestures are associated with one another; most segments are associated with more than one gesture, and it is possible that in some cases a single gesture is associated with more than one segment. Segments are organised into syllables. Gestures belong to a syllable only indirectly, through their associations with segments; they are not directly organised by syllables themselves. Diagram (9) gives examples of such output structures.

The area where TASP differs most from some other gestural analyses is in the nature of the mapping between inputs and outputs. First, segments (and their associated gestures) can be added and deleted. TASP does not claim that all phonological processes consist of gestural shifting, masking, weakening, etc. The goal of an analysis should be to differentiate the processes that involve changes to the abstract phonological units present, versus processes that may affect the acoustic result without addition or deletion of gestures and segments. I have argued that epenthesis is in the former category, and vowel intrusion in the latter.

Second, gestural phasing is not by itself contrastive in TASP. This claim is based on the observation that vowel intrusion is not contrastive: it is always predictable whether a release can occur between two consonants in a given environment, and whether the release will sound like a copy vowel. I have found no reports of words that are distinguished purely through the presence or absence of intrusive vowels.⁶ If this generalisation is a universal, it means that neither C-C nor V-C or C-V phasing can be contrastive. There are, of course, other types of gestural phasing besides C-C coordination and V-C or C-V coordination, but it is reasonable to hypothesise that if these are non-contrastive, other types of gestural phasing are the same.

In an Optimality Theoretic analysis, the non-contrastive nature of gestural phasing is captured by not letting faithfulness constraints refer to gestural alignment. Faithfulness constraints regulate the correspondence between input and output representations, demanding that representational elements that are present in the input be present in the output or vice versa. By prohibiting changes, faithfulness constraints preserve underlying distinctions and hence allow particular distinctions to be contrastive. Markedness constraints, on the other hand, refer only to output

⁶Scots Gaelic is an apparent exception. However, I argue elsewhere that the dramatic vowel intrusion in this language reflects a type of VR diphthongal structure, and diphthongs are contrastive with sequences of simple segments in Scots Gaelic. In other words, it's not the gestural phasing that is contrastive, but a layer of segmental organisation that conditions the phasing is contrastive.

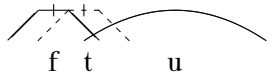
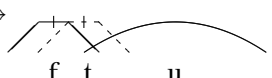
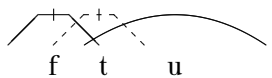
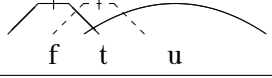
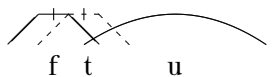
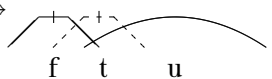
representations, and consist of prohibitions of particular structures. Markedness constraints can only prevent distinctions from being contrastive, never cause them to be so.

Tableau (45) shows how ranking a gestural faithfulness constraint above gestural markedness constraints causes gestural phasing to be contrastive in consonant clusters. The constraint FAITH TIMING, defined in (44), demands that gestural phasing in the input be preserved in the output. In the hypothetical language of the tableau, this constraint is ranked above all markedness constraints that would demand uniform phasing in CC clusters. Thus, for a CC cluster with the underlying phasing CENTER = ONSET, as in the first input, the phasing CENTER = ONSET is chosen in the output. For the second input, which has underlying RELEASE = TARGET phasing, an output with RELEASE = TARGET phasing is chosen. In this language, there can be two words [ftu] and [fətu] that differ only in the presence of release in their clusters.

(44) FAITH TIMING

If landmark_{*i*} of gesture_{*x*} is simultaneous with landmark_{*j*} of gesture_{*y*} in the input, then landmark_{*i*} of gesture_{*x*} is simultaneous with landmark_{*j*} of gesture_{*y*} in the output.

(45) Contrastive gestural timing

Input: 	FAITH TIMING	C ₁ REL = C ₂ TAR	C ₁ CEN = C ₂ ONS
a. →  f t u [ftu]			*
b.  f t u [fətu]	*!	*	
Input: 	FAITH TIMING	C ₁ REL = C ₂ TAR	C ₁ CEN = C ₂ ONS
a.  f t u [ftu]	*!		*
b. →  f t u [fətu]		*	

In Optimality Theory, constraints are freely rerankable and any ranking of constraints may exist in some language. If a constraint like FAITH TIMING exists, then we expect to find some language where release between consonants (with or without copy vowel intrusion) is contrastive. I am not aware of such a language. For this reason, all the constraints that refer to gestural phasing in TASP are markedness constraints. All of the gestural constraints proposed in section 4 specify alignments between various gesture types in the output; they never refer to the input.

This does not mean that gestural phasing cannot exist in the underlying representation (Richness of the Base, a basic Optimality Theory tenet, holds that any representation can be an input) but only that whatever gestural phasing may exist at the underlying level is not relevant to determining the output. For simplicity, I therefore show the input only as a string of segments, without gestures.

6 Intrusive vowels and allophony

One challenge to the gestural account of vowel intrusion has come from Warner et al. (2001)'s recent study of allophony in Dutch. The authors claim that intrusive

vowels cause resyllabification of preceding consonants, based on the patterning of [l] allophones.

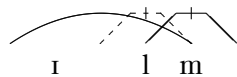
Dutch [l] is hypothesised to have the characteristics that Sproat & Fujimura (1993) have shown for English [l]. English [l] consists of both a tongue tip raising gesture and a tongue body backing gesture. The former is stronger when [l] is in the syllable onset, and the latter when [l] is in coda position. Warner et al. (2001) use this fact as a diagnostic for testing the syllabic structure of vowel intrusion. They compare tongue tip positioning during the [l]s of triplets like [film] / [filəm] ‘film’, and [vɪləm] (name): that is, the same word pronounced with and without intrusive [ə] versus a word with an underlying [ə] in the same position. For all seven subjects, the tongue tip was significantly higher during the [l] of [filəm] than [film], although for three there was also a significant difference between tongue tip positioning in [filəm] and [vɪləm]. The authors conclude that intrusive [ə] causes resyllabification of the preceding [l] into onset position.

In order for the schwa to condition allophonic variation, it must be present as a phonological unit, because the allophonic variation involves timing of the /l/ gestures relative to the vowel. If the schwa were simply a period of time without gestural specifications, which happens to be interpreted perceptually as a schwa-like sound, this targetless schwa would be merely a perceptual epiphenomenon, not a linguistic unit, and could not possibly condition allophonic variation. (395)

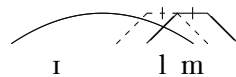
There is another possible explanation for this pattern of allophony, however, which does not require that the intrusive vowel be a phonological unit. This explanation rests on two assumptions. The first assumption is that allophony, in Articulatory Phonology, is not determined by rules of the form ‘use light [l] in onsets and dark [l] in codas.’ Allophones of a phoneme are not categorically different objects, but different realisations of (in most cases) the same group of gestures. These differences result from general rules about phasing within the syllable, and also from gestural overlap. In other words, the realisation of a particular [l] segment depends both on the rules the language has for phasing onset and coda [l], and also on what other segments overlap and perhaps conflict with that [l]. Two coda [l]s won’t necessarily be articulated the same way if they are in codas that involve different segments or different phasing. The second assumption is that the two pronunciations [film] and [filəm] in Dutch are the outputs of different grammars. They are not realisations of the same C-C alignment at different speech rates; rather, they involve different C-C alignments. The reason for this assumption is that the presence of intrusive [ə] in Dutch is not dependent on speech rate. Speakers who have vowel intrusion optionally, like Warner et al. (2001)’s subjects, are able to use two different grammars and hence produce two different gestural scores in free variation. Hypothetical gestural scores for the two pronunciations are given in (46).

(46) Gestural scores

a. [fɪl̩əm]



b. [fɪlm]



There are two differences between these gestural scores that could result in a stronger tongue tip raising gesture in a) than b), in addition to producing vowel intrusion in a). The first is the difference in C-C phasing. When the [lm] cluster is produced without vowel intrusion, the [l] overlaps more heavily with the following consonant. Overlap between consonants typically causes their constrictions to weaken (Browman & Goldstein 1995), and this should translate into less tongue tip raising. It is worth mentioning one study which, although limited, suggests that a following consonant indeed contributes to the weakening of the tongue tip gesture in coda [l] in Dutch. Reenan (1986) asked subjects by questionnaire whether they vocalised coda [l] (that is, lost the tongue tip constriction altogether) in various words. For each word where [l] preceded a consonant, 31% to 64% of respondents reported vocalizing the [l], while for each word where [l] occurred in final position, 26% to 34% did so. Of course, the fact that [l] did not become light and onset-like when it was in final position shows that overlap with a following consonant cannot be the only factor contributing to the lowering of the tongue tip in coda [l], but it is likely one factor.

Another factor that may affect the [l] in [fɪl̩əm] is the degree of overlap between [l] and the preceding vowel. Suppose that the vowel has the same phasing relation with [m] in both [fɪlm] and [fɪl̩əm]: then it must overlap the [l] more in the latter case, as shown in (46a). This is significant because the darkness of coda [l] is related to the tongue body gesture, and a vowel also consists of a tongue body gesture. There is evidence from Scots Gaelic that secondary articulations on sonorants (such as the velarisation in dark [l]s) conflict with overlapping vowel articulations. In the Ross-shire dialect (Borgstrøm 1941), when vowel intrusion occurs with a back vowel and a palatalised sonorant, as in /ar'kət/ 'silver', there are two possible outcomes. Either the sonorant's secondary articulation strongly influences the intrusive vowel, producing [ar'ikət], or the intrusive vowel has the same quality as the preceding vowel but sonorant loses its secondary articulation, producing [arəkət]. In this case the depalatalisation cannot be attributed to resyllabification of /r'/ as an onset, since [r'] is permitted in onset position even before back vowels, as in [r'iu:ms] (Borgstrøm 1941:148). Rather, the overlapping low vowel articulation causes the loss of palatalisation, because palatalisation and back vowels require conflicting tongue body positions. For similar reasons, the gestural phasing that results in vowel intrusion in Dutch may cause a loss of [l] velarisation. If the strength of the tongue body and tongue tip gestures are inversely proportional, then this weakening of the tongue body articulation could also enable a stronger tongue tip raising.

In short, it is not necessarily the case that when the presence of an intrusive

vowel between two consonants correlates with the presence of some allophone of one of the consonants, the intrusive vowel must condition this allophone. Rather, vowel intrusion and consonant allophony are both results of the gestural score, and a single change in the gestural score may affect both. For this reason, I do not consider the fact that Dutch intrusive vowels correlate with a pattern of [l] allophony to be sufficient evidence that the intrusive vowels are syllabic.

In general, consonant allophony is not the clearest test for the syllabicity of intrusive vowels. Although patterns of allophony are often stated in terms of position within the syllable, allophony is also related to gestural overlap. Metrical diagnostics like those mentioned in 2 (stress, minimal word requirements, templatic morphology, syncope, etc.) are more reliable because they seem to have nothing to do with gestural phasing. Looking at cases beyond Dutch, intrusive vowels sometimes do and sometimes do not appear to condition allophonic processes where ordinary vowels would. In Tiberian Hebrew, for example, stops are normally spirantised after a vowel. This is shown in the contrast between (47a) and (47b). The suffix /t/ becomes [θ] after a vowel-final stem, but remains [t] after a consonant. When a root ends in a guttural, as in (47c), an intrusive vowel appears before the suffix, but /t/ does not spirantise after the intrusive vowel. In this case, an intrusive vowel fails to condition an allophonic variation in the way a normal vowel would.

(47) Tiberian Hebrew spirantisation

- a. /ktv + t/ → katavt ‘write 2fs. perfective’
- b. /glj + t/ → galiθ ‘go into exile 2fs. perfective’
- c. /ʃlh/ + t → ʃalahat ‘send 2fs. perfective’

In Saami, intrusive vowels correlate with a special type of consonant allophony not triggered by other vowels. There is optional degemination of a sonorant preceding an intrusive vowel, but not before other vowels. For example, [skuol:əfi:] ‘owl, nom. sg.’ can also be pronounced [skuoləfi:] (Bye 2001). In this case, the allophony says nothing about how the sonorant is syllabified, since degemination is not associated with any position within the syllable. It happens only before intrusive vowels.

Dutch, Tiberian Hebrew, and Saami show that no single generalisation describes how vowel intrusion interacts with allophonic variation in neighboring consonants. Intrusive vowels may condition the same alternations that normal vowels do; they may condition alternations that normal vowels don’t; or they may fail to condition any alternations. The cause of each pattern must be sought in the particular gestural overlap patterns that lead to each type of allophony.

7 Conclusion

This paper began with the observation that there seems to be a correlation between the conditioning environment, quality, and phonological behavior of ‘inserted’ vowels. Cross-linguistically, copy vowels or schwas that appear predictably

between a sonorant and a heterorganic consonant, and which copy their quality from the vowel adjacent to the sonorant, tend to behave as non-syllabic. These characteristics are tied together through a representation in which the intrusive vowel is part of a neighboring vowel gesture that is heard during the period of release between two consonants. The quality of the intrusive vowel sounds similar to that of a neighboring vowel because the two vowels are a single gesture. A level of consonant-consonant overlap that produces release between heterorganic consonants does not necessarily produce release between homorganic consonants. This gestural phasing is chosen only with sonorants because sonorants are more amenable to heavy gestural overlap with a vowel than obstruents are. Gestural overlap, although it may create a vowel-like percept between consonants, is quite distinct from segmental insertion. When no abstract segment is added to the representation, no syllable is added.

Articulatory Phonology provides a useful way of representing intrusive vowels. Gestural representations allow us to analyze the distribution and typology of non-segmental sounds, without treating them as phonological units. However, the presence of non-gestural phonological units, such as segments and syllables, is crucial for analyzing the differences in behavior between intrusive vowels and segmental vowels (whether underlying or epenthetic). Thus, this study of vowel intrusion does not support the idea that all phonological alternations are gestural in nature.

References

- Albright, Adam (1999). The effect of context on epenthetic vowels in Lakhota. Ms., UCLA.
- Alderete, John (1995). Winnebago accent and Dorsey's Law. In Jill Beckman, Suzanne Urbanczyk & Laura Walsh Dickey (eds.), *University of Massachusetts occasional papers in linguistics 18: Papers in Optimality Theory*. Amherst, MA: GLSA.
- Alderete, John (1999). Head dependence in epenthesis-stress interaction. In Ben Hermans & Marc van Oostendorp (eds.), *The Derivational Residue in Phonology*. Amsterdam: John Benjamins.
- Allen, W.S. (1953). *Phonetics in Ancient India*. London: Oxford University Press.
- Blanc, Haim (1970). The Arabic dialect of the Negev Bedouins. In *Proceedings of the Israel Academy of Sciences and Humanities 4*. 112–50.
- Booij, Geert (1995). *The Phonology of Dutch*. Oxford: Clarendon Press.
- Borgström, C. H. (1940). *A Linguistic Survey of the Gaelic Dialects of Scotland, vol. 1: the Dialects of the Outer Hebrides*. Oslo: Norwegian Universities Press.

- Borgstrøm, C. H. (1941). *The Dialects of Skye and Ross-shire*. Oslo: Norwegian Universities Press.
- Bosch, Anna (1995). A gestural analysis of epenthesis in Scottish Gaelic. Ms., University of Kentucky.
- Bosch, Anna & Kenneth de Jong (1997). The prosody of Barra Gaelic epenthetic vowels. *Studies in the Linguistic Sciences* 27:2–15.
- Broselow, Ellen (1982). On predicting the interaction of stress and epenthesis. *Glossa* 16:115–131.
- Broselow, Ellen (1999). Stress-epenthesis interactions. Paper presented at the Phonology 2K symposium, Harvard/ MIT, April 1999.
- Browman, Catherine & Louis Goldstein (1986). Towards an articulatory phonology. *Phonology* 3:219–252.
- Browman, Catherine & Louis Goldstein (1990). Tiers in articulatory phonology, with some implications for casual speech. In John Kingston & Mary Beckman (eds.), *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*. Cambridge: Cambridge University Press, 341–386.
- Browman, Catherine & Louis Goldstein (1995). Dynamics and articulatory phonology. In Robert Port & Timothy Van Gelder (eds.), *Mind as motion*. Cambridge, MA: MIT Press, 175–94.
- Browman, Catherine & Louis Goldstein (2001). Competing constraints on intergestural coordination and self-organization of phonological structures. *Bulletin de la communication parlée* 5:25–34.
- Buck, Carl Darling (1904). *A grammar of Oscan and Umbrian*. Boston: Ginn and Company.
- Bye, Patrik (2001). Virtual phonology: rule sandwiching and multiple opacity in North Saami. PhD dissertation, Universitetet i Tromsø.
- Campbell, Lyle (1974). Theoretical implications of Kekchi phonology. *International journal of American linguistics* 40:269–278.
- Cho, Taehong (1998). Specification of intergestural timing and gestural overlap: Ema and epg studies. MA thesis, UCLA.
- Clements, George N. (1991). The structure of vowel and consonant gestures. Paper given at the University of Illinois conference on The organization of phonology: features and domains, May 1991.
- Dell, Francois & Mohamed Elmedlaoui (1996). On consonant releases in Imdlawn Tashlhiyt Berber. *Linguistics* 34:357–395.

- Donselaar, Wilma van, Cecile Kuijpers & Anne Cutler (1999). Facilitatory effects of vowel epenthesis on word processing in Dutch. *Journal of memory and language* **41**:59–77.
- Eberhard, David (1995). *Maimande stress: the need for strata*. Summer Institute of Linguistics publication 122.
- Gafos, Adamantios (2002). A grammar of gestural coordination. *Natural language and linguistic theory* **20**:269–337.
- Garcia-Bellido, Paloma (1999). Local movement and the syntax of Spanish sounds. In Raffaella Folli & Roberta Middleton (eds.), *Oxford University working papers* 4.
- Grammont, Maurice (1933). *Traité phonétique*. Paris: Librairie Delagrave.
- Gussenhoven, Carlos & Peter van der Vliet (1999). The phonology of tone and intonation in the Dutch dialect of Venlo. *Journal of linguistics* **35**:99–135.
- Haiman, John (1980). *Hua: a Papuan language of the eastern highlands of New Guinea*. Amsterdam: John Benjamins.
- Halle, Morris (1995). Feature geometry and feature spreading. *Linguistic inquiry* **26**:1–46.
- Harms, Robert T. (1976). The segmentalization of Finnish ‘nonrules’. In *Texas linguistic forum 5 (Papers from the Transatlantic Finnish Conference)*. 73–88.
- Harrikari, Heli (1999). Epenthesis, geminates, and the ocp in Finnish. *Nordic journal of linguistics* **22**:3–26.
- Harris, Katherine S. & Fredericka Bell-Berti (1984). On consonants and syllable boundaries. *Language and cognition* :89–95.
- Hayes, Bruce (1995). *Metrical stress theory: Principles and case studies*. Chicago: University of Chicago Press.
- Hayward, Dick (1984). *The Arbore language: A first investigation*. Hamburg: Helmut Buske Verlag.
- Hodge, Carleton T. (1947). *An outline of Hausa grammar*. Baltimore: Linguistic Society of America. Language dissertation no. 41.
- Holmer, Nils (1938). *Studies on Argyllshire Gaelic*. Uppsala: Almqvist & Wiksells Boktryckeri A.B.
- Hudson, Grover (1995). Consonant release and the syllable. *Linguistics* **33**:1–21.
- Jannedy, Stefanie (1994). Rate effects on German unstressed syllables. In *OSU Working Papers in Linguistics* 44. 105–124.

- Jetchev, Georgi (1995). Rhotics, jers and schwa in the history of bulgarian. Paper presented at the 13th ICPhS, Stockholm, 13-19 August 1995.
- Kager, René (1999). *Optimality Theory*. Cambridge: Cambridge University Press.
- Kenstowicz, Michael (2001). The role of perception in loanword phonology, a review of les emprunts linguistiques d'origine européenne en fon by flavien gbéto, köln: Rüdiger köppe verlag. To appear in *Linguistique Africaine*.
- Kinkade, M. Dale (1963–4). Phonology and morphology of Upper Chehalis. *International journal of American linguistics* **29–30**:181–95, 345–356, 32–61, 251–260.
- Kitto, Catherine (1997). Epenthesis in Polynesian loanwords. MA thesis, University of Auckland.
- Kitto, Catherine & Paul de Lacy (2000). Correspondence and epenthetic quality. In Carolyn Smallwood & Catherine Kitto (eds.), *Proceedings of AFLA VI: The sixth meeting of the Austronesian formal linguistics association*. 181–200.
- Krumholz, Jeanne, Marjorie Dolson & Miguel Hernández Ayuso (1995). *Diccionario Popoluca de San Juan Atzingo Puebla*. Tucson, Arizona: Instituto Lingüístico de Verano.
- Ladefoged, Peter, Jenny Ladefoged, Alice Turk, Kevin Hind & St. John Skilton (1998). Phonetic structures of Scottish Gaelic. *Journal of the international phonetic association* **28**:1–41.
- Lindblom, Björn, Harvey M. Sussman, Golnaz Modarresi & Elizabeth Burlingame (2002). The trough effect: Implications for speech motor programming. *Phonetica* **59**:245–262.
- McCarthy, John (1979). Formal problems in Semitic phonology and morphology. PhD Dissertation, MIT.
- McCarthy, John & Alan Prince (1993). Prosodic morphology I: constraint interaction and satisfaction. In *Technical report 3*. Rutgers University center for cognitive science.
- Miner, Kenneth (????). Winnebago grammar. Ms., University of Kentucky.
- Miner, Kenneth (1979). Dorsey's law in Winnebago-Chiwere and Winnebago accent. *International journal of American linguistics* **45**:25–33.
- Miner, Kenneth (1992). Winnebago accent: the rest of the data. In *Indiana University Linguistics Club Twenty-Fifth Anniversary Volume*. Bloomington, Indiana: Indiana University Linguistics Club, 28–53.

- Mithun, Marianne & Hasan Basri (1986). The phonology of selayarese. *Oceanic Linguistics* **25**:210–254.
- Moussa, Bamba (1996). Mawukakan loanword phonology. Ms., University of Pennsylvania.
- Ó Siadhail, Micheál (1989). *Modern Irish: Grammatical structure and dialectal variation*. Cambridge: Cambridge University Press.
- Oftedal, Magne (1956). *A linguistic survey of the Gaelic dialects of Scotland, vol. III: the Gaelic of Leurbost, Isle of Lewis*. Oslo: H. Aschehoug & Co.
- Olson, Kenneth (2001). The phonology and morphology of Mono. University of Chicago dissertation.
- Olson, Kenneth (2003). Word minimality and non-surface-apparent opacity in Mono. Talk given at WOCAL 4, Rutgers University.
- Parker, Steve (1994). Laryngeal codas in Chamicuro. *International Journal of American Linguistics* **60**:261–71.
- Quilis, Antonio (1981). *Fonética acústica de la lengua española*. Madrid: Editorial Gredos.
- Reenan, Pieter van (1986). The vocalization of /l/ in standard Dutch, a pilot study of an ongoing change. In F. Beukema & A. Hulk (eds.), *Linguistics in the Netherlands*. Amsterdam: John Benjamins, 189–199.
- Schuchardt, Hugo (1867). *Der Vokalismus des Vulgärlateins*. Leipzig: Teubner.
- Smith, Norval (1999). A preliminary account of some aspects of Leurbost Gaelic syllable structure. In Harry van der Hulst & Elizabeth Ritter (eds.), *The Syllable: views and facts*.
- Sproat, Richard & Osamu Fujimura (1993). Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics* **21**:291–311.
- Steriade, Donca (1990). Gestures and autosegments. In John Kingston & Mary Beckman (eds.), *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*. Cambridge: Cambridge University Press, 382–397.
- Suomi, Kari (2000). Nearly a million Ostrobothnians cannot be wrong: comments on an Optimality Theoretic analysis of “Ostrobothnian”. *Nordic Journal of Linguistics* **23**:89–93.
- Susman, A. (1943). The accentual system of Winnebago. PhD Dissertation, Columbia University.
- Warner, Natasha, Allard Jongman, Anne Cutler & Doris Mücke (2001). The phonological status of Dutch epenthetic schwa. *Phonology* **18**:387–420.

Wright, Joseph (1905). *The English dialect grammar*. Oxford: Clarendon Press.

Zsiga, Elizabeth (1995). An acoustic and electropalatographic study of lexical and postlexical palatalization in American English. In Mary Beckman & John Kingston (eds.), *Phonology and phonetic evidence: Papers in laboratory phonology IV*. 282–302.