

GESTURES AND SEGMENTS: VOWEL INTRUSION AS OVERLAP

A Dissertation Presented

by

NANCY ELIZABETH HALL

Submitted to the Graduate School of the  
University of Massachusetts in partial fulfillment  
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 2003

Department of Linguistics

© Copyright by Nancy Elizabeth Hall 2003

All Rights Reserved

GESTURES AND SEGMENTS: VOWEL INTRUSION AS OVERLAP

A Dissertation Presented

by

NANCY ELIZABETH HALL

Approved as to style and content by:

---

Elisabeth O. Selkirk, Chair

---

John J. McCarthy, Member

---

John Kingston, Member

---

Joseph V. Pater, Member

---

Rex E. Wallace, Member

---

Elisabeth O. Selkirk, Department Head  
Department of Linguistics

## ACKNOWLEDGMENTS

This dissertation has benefited immeasurably from the input of many linguists, especially the members of my committee. I was lucky to have Lisa Selkirk as an advisor, and thank her for her support, encouragement, and expertise. John McCarthy brought an eagle eye to the formulation of the OT analysis, John Kingston shared his instant recall of all phonetics literature, Rex Wallace combed Oscan inscriptions for relevant data, and Joe Pater pointed out interesting implications for child phonology.

Louis Goldstein, Travis Bradley and Diamandis Gafos provided helpful discussions of gestural grammar. Many linguists shared their expertise on particular languages, especially John Koontz (Hocank), Patrick Bye (Saami), Shigeto Kawahara (Japanese), and Roger Higgins (English dialects, Irish).

I received much useful feedback from the participants in my seminar at Rutgers University: Jill Heather Flegg, Slavica Kochovska, Hyunjoo Kim, and Seunghun Lee, and from Akin Akinlabi, Alan Prince, and Bruce Tesar. Thanks to Markus Hiller and Graham Horwood for late-night phonology discussions and tea (bubble and other). Thanks also to my colleagues at UMass, particularly Meredith Landman, Marcin Morzycki, and the phonology group: Maria Gouskova, Adam Werle, Ania Łubowicz, Jen Smith, Elliott Moreton, Paul de Lacy, Ji-Yung Kim, Anne-Michelle Tessier, Ben Gelbart and Andries Coetzee.

For three years of graduate study, I was supported by a National Science Foundation Graduate Research Fellowship.

Thanks to my family- Ron Artstein, Jim, Jane, Linda, and Michael Hall, and Reba Ray, for their support and love.

## ABSTRACT

### GESTURES AND SEGMENTS: VOWEL INTRUSION AS OVERLAP

NANCY ELIZABETH HALL, B.A., JOHNS HOPKINS UNIVERSITY

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Elisabeth O. Selkirk

This dissertation focuses on a phenomenon that I call *vowel intrusion*. There are cases where a vowel can be heard between two consonants, yet the phonology behaves as if no vowel is present. These “intrusive vowels” are non-syllabic, and native speakers are often unaware of their existence.

I argue that intrusive vowels are a percept resulting from the organization of articulatory gestures. When two consonant gestures have little overlap with one another, there is an acoustic release between them; vowel gestures typically overlap neighboring consonants considerably, and it is possible for an overlapping vowel gesture to be heard in this period of release.

Intrusive vowels are not segments. They behave unlike true epenthetic vowels. A typological survey reveals that vowel intrusion happens in consonant clusters that contain a sonorant or a guttural, and that it is always the vowel adjacent to the sonorant or guttural that is heard during the release. Intrusive vowels occur primarily in heterorganic clusters, especially next to geminates; they often disappear at fast speech rates, and in some languages, they occur only within or only between syllables. I argue that these characteristics are best explained in a theory that uses Articulatory Phonology representations (Browman & Goldstein 1986 et seq.).

I develop a theory called Timing-Augmented Surface Phonology (TASP), cast within the framework of Optimality Theory. TASP contains constraints on the alignment of neighboring gestures (Gafos 2002) and on the permitted degree of overlap between different gestures. The theory requires a segmental representation as well as a gestural representation. Syllables organize segments rather than gestures, and that inter-segmental gestural alignment is universally non-contrastive.

The same gestural framework describes both the short, schwa-like intrusive vowels often described as “excrement”, and also a longer type found in Scots Gaelic and Hocank (Winnebago), in which the vowel is heard in two long parts on either side of the sonorant. In the latter cases the sonorant and vowel together behave like a bimoraic nucleus, and are adjoined in a structure similar to vocalic diphthongs. The theory also has implications for the analysis of Hocank accent.

## TABLE OF CONTENTS

ACKNOWLEDGMENTS .....	iv
ABSTRACT .....	v
Chapter 1. GESTURES AND SEGMENTS: THE STRUCTURE OF NON-SEGMENTAL VOWEL SOUNDS .....	1
1.1. Introduction .....	1
1.1.1. The intrusive vowel syndrome .....	1
1.1.2. Vowel intrusion as gestural overlap.....	3
1.2. Articulatory phonology representations .....	4
1.2.1. Structure of gestures .....	4
1.2.2. Phasing of gestures .....	6
1.2.3. From abstract gestures to articulatory trajectories .....	7
1.2.4. Gestural representation and the vowel intrusion syndrome.....	9
1.2.4.1. The heterorganicity requirement.....	9
1.2.4.2. Speech rate effects .....	10
1.3. Timing-Augmented Surface Phonology (TASP) .....	11
1.3.1. Gestures and segments.....	12
1.3.2. Gestural coordination in Optimality Theory .....	13
1.3.3. Gestures, underlying representations, and faithfulness .....	14
1.3.4. A TASP representation of vowel intrusion .....	16
1.4. Gestural coordination constraints.....	18
1.4.1. C-C and V-C phasing constraints .....	18
1.4.2. Typology of coordination effects .....	20
1.4.2.1. Vowel intrusion .....	20
1.4.2.2. Release without vowel intrusion .....	21
1.4.2.3. No release .....	21
1.4.3. The special behavior of sonorants .....	22
1.4.3.1. Vowel intrusion by sonorants; release elsewhere.....	22
1.4.3.2. Release by sonorants, no release elsewhere.....	24
1.4.4. Hierarchy of sonorant / vowel overlap.....	26
1.4.4.1. Constraints on V / C overlap .....	28
1.4.4.2. Additional evidence for *C in V: articulatory trough data .....	30
1.4.4.3. VC homorganicity.....	31
1.4.4.4. The relevance of vowel intrusion .....	32
1.5. Gestures and features .....	32
1.5.1. Gesture sharing.....	34
1.5.2. Homorganic CR clusters.....	37
1.5.3. Gestures and levels of grammar .....	39
1.5.4. Geminates .....	40
1.6. Conclusion.....	40
Chapter 2. GESTURAL PHASING AND THE SYLLABLE.....	43
2.1. Syllabicity and segmenthood .....	43
2.1.1. The non-syllabicity hypothesis.....	43
2.1.2. Alternative accounts of monosyllabicity.....	45
2.2. Dutch.....	47
2.2.1. Conditioning environment .....	47
2.2.2. The role of syllable structure.....	49
2.2.2.1. Heterosyllabic clusters .....	49
2.2.2.2. Appendices .....	51
2.2.3. Monosyllabic behavior .....	53
2.2.3.1. Syllable duration.....	53
2.2.3.2. Syllable count judgments .....	54
2.2.3.3. [n]-deletion .....	55

2.2.3.4.	Allomorph selection .....	57
2.2.3.5.	Lexical tone .....	57
2.2.3.6.	Distribution of long vowels .....	58
2.2.3.7.	The unmarkedness of RC clusters .....	58
2.2.3.8.	Against ordering .....	59
2.2.4.	Discussion: how can a vowel not be a syllable? .....	59
2.2.5.	Intrusion and [l]-allophony .....	61
2.2.6.	Other allophony .....	62
2.2.7.	Summary .....	63
2.3.	Finnish .....	63
2.3.1.	Conditioning environment .....	64
2.3.2.	Phonetic realization .....	65
2.3.3.	Related phasing effects .....	66
2.3.4.	Syllable boundary effects .....	67
2.3.5.	Geminates .....	69
2.3.6.	Vowel intrusion as contrast enhancement .....	70
2.3.7.	Syllabicity in speaker intuitions and phonology .....	71
2.3.8.	Against an epenthesis analysis .....	71
2.3.9.	Creeping segmentization? .....	72
2.4.	Conclusion .....	73
Chapter 3.	INTRUSION VS. EPENTHESIS .....	74
3.1.	Copy vowel epenthesis .....	74
3.2.	Diagnostics of non-gestural copying .....	76
3.2.1.	Welsh: true epenthesis .....	76
3.2.2.	Stress and epenthesis interaction .....	79
3.2.3.	Typological predictions .....	80
3.3.	Languages with both epenthetic and intrusive copy vowels .....	80
3.3.1.	Kekchi .....	80
3.3.2.	Mono .....	84
3.4.	Vowel intrusion as precursor to epenthesis in loanwords .....	85
3.4.1.	Cook Islands Maori .....	86
3.4.2.	Mawu .....	89
3.4.3.	Japanese .....	89
3.5.	Conclusion .....	90
Chapter 4.	SCOTS GAELIC .....	92
4.1.	Introduction .....	92
4.1.1.	Background: dialects, transcription conventions .....	92
4.2.	Evidence for monosyllabicity .....	93
4.2.1.	Speaker intuitions .....	93
4.2.2.	Duration and pitch .....	96
4.2.3.	Positional licensing .....	97
4.2.4.	Glottal stop epenthesis .....	99
4.2.5.	Mutations .....	100
4.2.6.	Syncope .....	102
4.3.	Evidence for subsyllabic structure .....	104
4.3.1.	Lexical specification .....	105
4.3.2.	Interaction of vowel intrusion and vowel length .....	108
4.4.	Structure $\alpha$ and timing .....	111
4.4.1.	Subsyllabic structure and symmetrical timing: the C-center effect .....	111
4.4.2.	Symmetrical vowel intrusion as vowel / sonorant centering .....	112
4.4.3.	Timing within $\alpha$ : duration .....	114
4.4.4.	Timing within $\alpha$ : symmetry .....	115
4.4.5.	Motivating vowel intrusion synchronically .....	116
4.5.	What is $\alpha$ ? Parallels to vocalic structures .....	119
4.5.1.	Lexical specification of diphthongs .....	120
4.5.2.	Typology of $\alpha$ : coalescence, “VR diphthongs” .....	121

4.5.3.	Licensing of gestures within $\alpha$ : the case of nasality.....	124
4.5.4.	Cooccurrence restrictions.....	126
4.6.	Intrusive vowel quality: sonorant-vowel blending.....	126
4.6.1.	Vowel backing after [r], [n] in Barra.....	127
4.6.2.	Vowel fronting: mutations, verb forms in Barra.....	130
4.6.3.	Exceptions.....	133
4.6.4.	Coarticulation or feature spreading?.....	135
4.6.5.	An alternate strategy: removal of secondary articulations.....	137
4.6.6.	Intrusive [ə].....	138
4.7.	Conclusion.....	139
4.8.	Appendix: consonant cluster charts.....	139
Chapter 5.	HOCANK (WINNEBAGO).....	143
5.1.	Introduction.....	143
5.1.1.	Conditioning environment for vowel intrusion.....	143
5.2.	Evidence for monosyllabicity, monosegmentality.....	144
5.2.1.	Reduplication.....	145
5.2.2.	Root Size.....	146
5.2.3.	Nasalization.....	146
5.2.4.	Ablaut.....	148
5.2.5.	Phonetics.....	149
5.3.	Vowel intrusion as [RV] $_{\alpha}$ .....	149
5.3.1.	Bimoraicity.....	149
5.3.2.	Nasality restrictions.....	150
5.3.3.	Motivating vowel intrusion.....	150
5.3.4.	Syllable structure effects, and a second type of vowel intrusion.....	153
5.3.5.	Historical source of moraic sonorants.....	156
5.4.	Accent.....	157
5.4.1.	Late realization of tone.....	158
5.4.2.	Constraints on accent.....	159
5.4.3.	Initial feet.....	162
5.4.4.	Heavy syllables (without intrusion).....	163
5.4.5.	Intrusive vowel syllables.....	164
5.4.5.1.	Initial feet.....	165
5.4.5.2.	Non-initial feet.....	165
5.4.5.3.	First mora tone; second mora tone; no tone.....	167
5.4.5.4.	Weak local parsing.....	170
5.4.6.	Physical realization of tone on intrusive vowels.....	172
5.4.7.	Summary of accent.....	174
5.5.	Conclusion.....	174
APPENDIX:	LIST OF CONSTRAINTS.....	175
BIBLIOGRAPHY	.....	179

## CHAPTER 1. GESTURES AND SEGMENTS: THE STRUCTURE OF NON-SEGMENTAL VOWEL SOUNDS

### 1.1. Introduction

#### 1.1.1. The intrusive vowel syndrome

This dissertation concerns a phenomenon that was first pointed out by Sanskrit grammarians about 2500 years ago. The *praticakhyas*, ancient Indian commentaries on Pānini, contain descriptions of a type of vocalic interval that the writers call not a vowel, but a ‘vowel-fragment’—*svarabhakti* in Sanskrit. This vowel-fragment is described as appearing between an [r] and a following consonant, or according to one commentator an [r] or [l] and a following consonant, and as having one-eighth to one-half the length of a normal vowel. This vowel-fragment is not described as adding 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 or guttural 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 or guttural. Some examples are given in Figure 1, with the intrusive vowel underlined. Throughout, I will use the term sonorant to include gutturals as well (following Halle 1995, where gutturals are [+sonorant]).

The inserted vowels in these processes are often described in the linguistic literature as epenthetic, but they have a group of characteristics in common that set them apart from other epenthetic vowels. The following list of properties can be called the “vowel intrusion syndrome”.

- (1) Vowel intrusion diagnostics
  - a. Vowel intrusion is triggered by a consonant cluster that includes a sonorant. Gutturals are considered sonorants.
  - b. The intrusive vowel has the same quality as the vowel that is adjacent over 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 may disappear at fast speech rates.

I propose in this thesis that this is not a chance collection of properties. What underlies the vowel intrusion syndrome is a representation in which these vowels are not segments. Rather, they are percepts resulting from a particular timing of articulatory gestures within the syllable, and constraints on articulatory organization cause the trends seen in environment and vowel quality.

### Languages with intrusive vowels in RC or CR clusters (R = sonorant)

Arbore		/leh-t-atto/	→	lefi <sub>e</sub> 'tatto	'that ewe'
Bedouin Arabic dialects		qahwa	>	ga'hawa	'coffee'
Chamicuro		/tuʔlu/	→	'tuʔulu	'chest'
Dutch		/kalm/	→	'kalə <sub>m</sub>	'quiet'
English (various dialects)		/arm/	→	'arə <sub>m</sub>	'arm'
Finnish		/kalvo/	→	'kaləvo	'transparency'
German (S. Hamburg)		/br̩atən/	→	bə <sub>r̩</sub> katən	'to fry'
Hausa		/kurkutu/	→	k <sup>w</sup> ur <sup>u</sup> k <sup>w</sup> u:tu	'small drum'
Hocank	I.	/sni/	→	s <sub>i</sub> n <sup>i</sup>	'cold'
	II.	/ho:tʃ̃ak+ra/	→	ho:tʃ̃ə <sup>g</sup> rə	'the Hocank'
Irish Gaelic	I.	/agla/	→	'agə <sub>l</sub>	'fear'
	II.	gorm	→	'gorə <sub>m</sub>	'blue'
Kekchi		/paʔt/	→	paʔat	'twins'
Lakhota		/gla/	→	g <sup>a</sup> la	no gloss
Late Latin		<i>scriptum</i>	>	<i>sc̣ri<sub>i</sub>ptum</i>	'a writing'
Mamaindé		/mih-takʔu/	→	mih <sub>i</sub> takʔu	'it is cloudy'
Mono		/gàfrū/	→	gà <sub>f</sub> urū	'mortar'
Oscan	I.	<i>Mulcius</i>	>	<i>Mul<sub>u</sub>kiis</i>	name
	II.	<i>patri</i>	>	<i>patere<sub>i</sub></i>	'patri'
Popoluca		/itʔa/	→	it <sup>a</sup> ʔa	'your father'
Saami		/skuol:fi:/	→	skuol:ə <sub>f</sub> i:	'owl, nom. sg.'
Sanskrit		/darʃata/	→	darə <sub>ʃ</sub> ata	no gloss
Scots Gaelic		/ʃaLk /	→	'ʃa <sub>L</sub> ak	'hunting'
Spanish (Chilean)		/kronika/	→	ko <sub>r</sub> onika	'chronicle'
Tiberian Hebrew		/yaʃmod/	→	yaʃ <sub>a</sub> 'mod	'he stands'

**Notes:** '→' indicates synchronic processes; '>' indicates diachronic processes. When a vowel is written in superscript, this follows a convention used by the source to emphasize the vowel's short duration. When two examples are given for a single language, they illustrate distinct types of vowel intrusion.

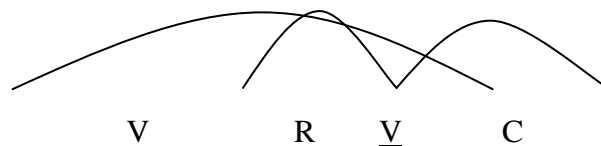
**Sources:** Dutch- Booij 1995:127; English- Wright 1905:206; Finnish- Harrikari 1999:6; German-Jannedy 1994:113; Hausa- Hodge 1947:12; Hocank (a.k.a. Winnebago)- Miner 1992:30, Miner 1979:27; Irish- Ó Siadhail 1989:21,22; Lakhota- Albright 1999:1; Late Latin- Schuchardt 1867 vol 3:421; Mono (Niger-Congo, Congo)- Olson 2003; Oscan- Buck 1904:51-3; Saami- Bye 2001:164; Sanskrit- Allen 1953; Scots Gaelic- Borgström 1940:153; Spanish- Garcia-Bellido 1999:8; Arbore (East Cushitic)- Hayward 1984:72; Chamicuro (Arawakan, Peru)- Parker 1994:266; Kekchi (Mayan, Guatamala)- Campbell 1974:270; Mamaindé (Nambiquára, Brazil) - Eberhard 1995:17; Popoluca (San Juan Atzingo Puebla)- Krumholz, Dolson, & Ayuso 1995:282; Tiberian Hebrew- McCarthy 1979:34.

**Figure 1: Vowel Intrusion**

### 1.1.2. Vowel intrusion<sup>1</sup> as gestural overlap

Working in an Articulatory Phonology framework, Steriade 1990 proposes that copy vowels in Winnebago (Hocank), Late Latin, and Sardinian are formed through overlapping the gestures of a sonorant and an adjacent vowel. (Sardinian is omitted from the list in Figure 1 because the process there seems historical and sporadic.) When two consonant gestures become separated in time, an overlapping vowel gesture can be heard briefly in the interval between them, sounding like [ə] if it is short or like a copy vowel if it is longer. Below, each curve represents the dynamic cycle of one gesture. The intrusive vowel is underlined.

#### (2) Gestural score of vowel intrusion



Steriade points out that this analysis explains the copied quality of the vowel: the inserted vowel has the same quality as the adjacent vowel because they are a single gesture. I argue that the gestural approach explains the rest of the cluster of properties in (1) as well.

In other respects, the gestural analysis of intrusive vowels presented here differs considerably from Steriade's. Steriade views these vowels as segmental and syllabic, but evidence presented in chapters 2–5 will show that this is not the case. Intrusive vowels, despite their phonetic prominence, are completely invisible to a wide range of phonological patterns that count syllables, such as stress, templatic reduplication, syncope, licensing of segmental contrasts, ablaut, and language games.

Incidentally, the idea that one gesture can surround another goes back at least to Grammont 1933:244–9, who uses it to describe a type of metathesis that he calls 'inversion by penetration'. Grammont proposes that while some metatheses are a simple flipping of segments, other metatheses, especially those involving liquids, can result from one segment 'penetrating' another.

In Indo-European *wr* became *rw* or *ru* between consonants and even before vowels when initial ... This is not transposition pure and simple; the *w* doesn't pass over the *r* but through it. First the *r* assimilates to the *w*, without this requiring that the *r* take the *w*'s place of articulation; it can keep pretty much its usual place, while taking the timbre of the *w*, meaning that it is articulated with the resonator of the *w*, raising of the dorsum towards the velum, protrusion and

---

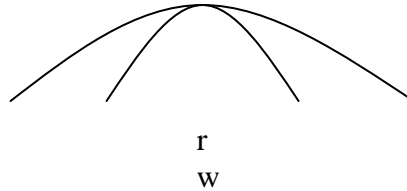
<sup>1</sup>Intrusive vowels include many that have been described as "excrecent" (Levin 1987), "transitional", or "svarabhakti". I have chosen the term "intrusive" (Harms 1976) because I wish to argue for the existence of a class of vowels that is not precisely captured by the commonly understood definition of any of those terms. This class includes some vowels that are fairly long and hence not usually called excrecent or transitional, and it does not necessarily include all vowels that have been called svarabhakti.

rounding of the lips. Once the *r* is thus impregnated with *w*, the *w* springs up again in the form of *w* or *u* (as the case may be) on the side where its appearance will make a better syllable:

Zend *rvāta* ‘dogma’, cf. Sanskrit *vratām* ‘precept’; Zend *rvīnat* ‘compressing’, cf. Sanskrit *vlīnāti, vlināti* ‘he compresses’. These Zend forms seem to represent a phase where the *r* is still submerged in the *w*: it has *w* before and after it and is full of *w* itself. The *urva* of the texts, of which *rva* is an interpretation that in other respects seems correct, is only one syllable. (244; my translation.)

In this passage, Grammont essentially proposes that [rw] metathesis involves a historical stage where the segments have a gestural coordination like the following.

(3) Grammontian “penetration”



Grammont’s claim that a phonetic [ʷrva] is monosyllabic agrees with a claim that I will make about vowel intrusion: that a vocalic-sounding interval produced by gestural overlap, as opposed to insertion of a segment, is not the nucleus of a syllable.

Articulatory phonology, which treats gestures as objects manipulated by the grammar, provides a way to formalize Grammont’s idea and model the role of gestural timing in the grammar. Section 2 introduces this approach to phonology, including its implementation in Optimality Theory, and a basic gestural analysis of vowel intrusion. In section 3 I develop an Optimality-Theoretic Articulatory Phonology framework called Timing-Augmented Surface Phonology (TASP). Section 4 discusses particular gestural constraints needed to produce the vowel intrusion typology. Section 5 compares gesture-based theories to feature-based theories.

## 1.2. Articulatory phonology representations

### 1.2.1. Structure of gestures

Articulatory phonology (Browman & Goldstein 1986 et seq., Byrd 1996, Gafos 2002) is a model where the phonological grammar directly regulates the organization of abstract gestures. A gesture is essentially an instruction for an articulator to achieve a particular constriction in the vocal tract: for example, an opening of the velum, a closing of the glottis, or a raising of the tongue body. At the same time, a gesture is a unit of contrast that plays a role in representations similar to that of the feature.

Formally, a gesture consists of settings for one of the vocal tract variable sets. Examples of variables and settings they can take are listed below.

(4)	ARTICULATOR SETS	VOCAL TRACT VARIABLES	
	tongue body	tongue body constriction location	(TBCL)
		tongue body constriction degree	(TBCD)
	tongue tip	tongue tip constriction location	(TTCL)
		tongue tip constriction degree	(TTCD)
	lips	lip aperture	(LA)
		lip protrusion	(LP)
	glottis	glottal aperture	(GLO)
	velum	velic aperture	(VEL)

CONstriction DEGREE VALUES (Gafos 2002:275)

[closed]	stops
[critical]	fricatives
[narrow]	some vowels
[mid]	approximants and mid vowels
[wide]	some vowels

CONstriction LOCATION VALUES: [labial], [dental], [alveolar], [post-alveolar], [palatal], [velar], [uvular], [pharyngeal]<sup>2</sup>.

Segments in traditional phonological analysis are considered by Articulatory Phonology to be a combination of gestures. For example, an [n] is associated with the following tract variable settings:

(5)	[n]	Tongue tip gesture:	tongue tip constriction location:	alveolar
			tongue tip constriction degree:	closed
		Velum gesture:	velic aperture:	open
		Glottis gesture:	glottal aperture:	critical <sup>3</sup>

For the other tract variables, an [n] specifies no target, since no constriction of the tongue body or lips is involved. The position these articulators actually take during production of an [n] is affected by at least two factors. The first is the influence of other active articulations. If a vowel gesture overlaps the production of [n], it will affect the position of the tongue body. The other factor is a tendency of articulators to return to a neutral, resting state when not activated. In the absence of any other active gestures or influences, non-activated articulators will assume their resting positions while [n] is articulated.

Gestures are spatiotemporal units. Each gesture has a duration in time and an internal cycle. This cycle begins with the onset of movement, progresses to the point when the target is reached, then to the release, where movement away from the

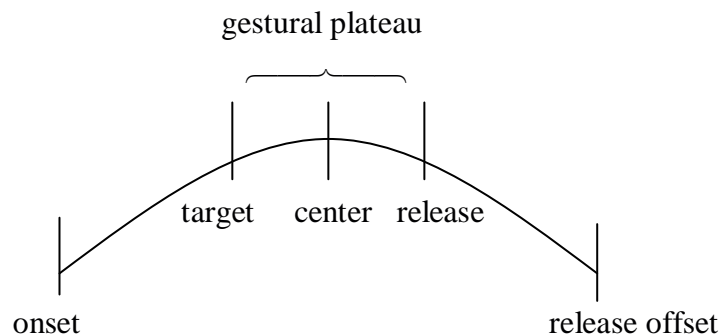
<sup>2</sup> The [pharyngeal] location characterizes not only pharyngeal consonants but also low vowels. A low back vowel involves a [narrow pharyngeal] gesture; a low front vowel a [wide pharyngeal] gesture.

<sup>3</sup> Browman & Goldstein, however, do not show any glottal gesture for voiced segments.

constriction begins, and finally to the offset, the point where the articulator ceases to be under active control of the gesture. Between release and offset, there is a period when there is active control of movement away from the constriction (Browman 1994). After the offset, the articulators either continue moving back to their resting positions, or begin moving towards a new target specified by another gesture. A cycle can be shown by a curve, as below.

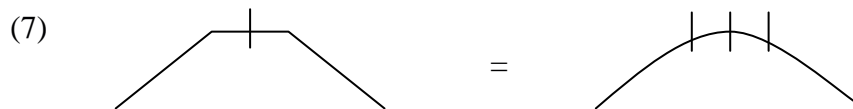
(6) Landmarks in a gestural cycle

Gafos 2002:271



The vertical axis represents the movement of an articulator in space; the horizontal axis represents time. Between the target and release is a period when the constriction is actively held, called the 'gestural plateau'. The middle of the gestural plateau is called the c-center.

For visual clarity, the landmarks can be shown with angles rather than marks on a smooth curve. The two representations below should be understood to be equivalent.

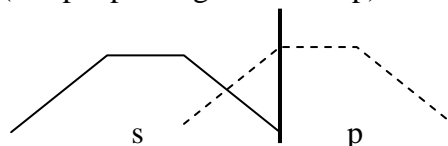


These particular five landmarks are a hypothesis, chosen in order to allow simulations; it is an open question how many landmarks are actually needed in the theory.

**1.2.2. Phasing of gestures**

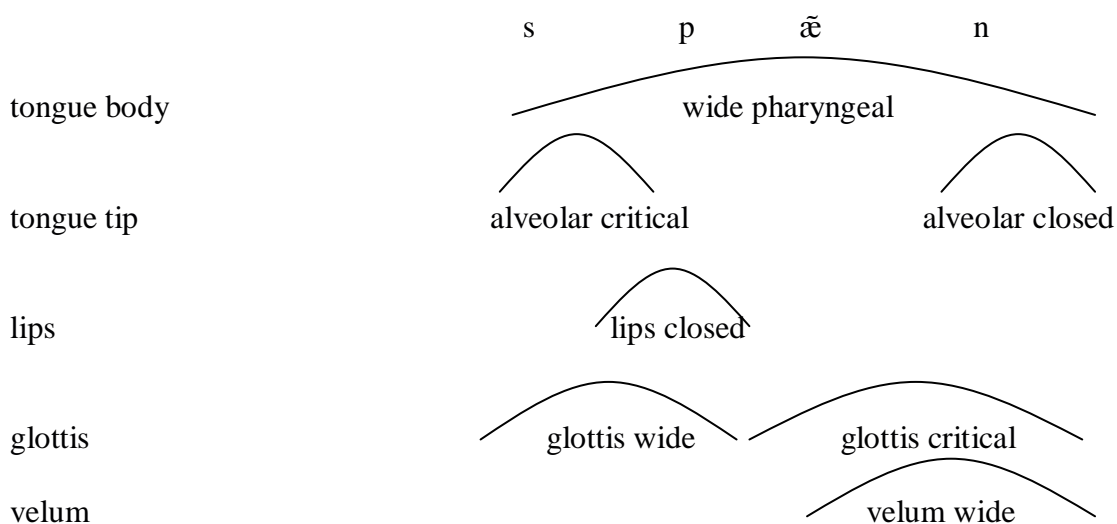
Gestures do not simply fall in a linear order like segments or features; they overlap one another greatly. Even as speech rates change, gestures tend to bear certain consistent relations to one another: for example, one gesture may consistently begin at the point where another gesture reaches its target or release. These intergestural timing relations are called phasing. The grammar specifies the phasing of gestures with respect to one another by aligning landmarks of different gestures. For example, the segments in the onset cluster [sp] might have the phasing relationship  $OFFSET = TARGET$ , meaning that the target of the [p] must be reached as the offset of the [s] occurs, as in (8). The bold line indicates the point in time when both these landmarks are reached.

- (8) OFFSET = TARGET (sample phasing relationship)



The organization of gestures over time is represented in a ‘gestural score’, consisting of the settings of the tract variable parameters and the timing relations between them. Below is a rough gestural score for an utterance of *span*.

- (9) Gestural score of English *span* [spæn]  
adapted from Browman & Goldstein 1992b: 158



Gestures can overlap considerably, but with certain limits. Browman & Goldstein 1990 propose, based on X-ray evidence, that “consonant articulations are superimposed on continuous vowel articulations, which themselves minimally overlap.” A topic of current research is which gestures have specified phasing relationships with one another. There is evidence that phasing relationships are dependent on syllable structure: the same sequence of gestures will show different timing relations depending on whether they are tautosyllabic or heterosyllabic. This will be useful in explaining why vowel intrusion appears only in heterosyllabic clusters in some languages, and only in tautosyllabic clusters in others.

### 1.2.3. From abstract gestures to articulatory trajectories

A gestural score is part of a larger model of speech production developed at Haskins Laboratories called GEST (Browman & Goldstein 1990, Saltzman et al. 1988). GEST includes (among others) the following modules and stages of representation.

(10) Components of GEST

Linguistic gestural model	contains language-specific principles of intergestural phasing
Gestural score	representation of gestural units and their organization over time, produced by the linguistic gestural model for each utterance
Task-dynamic model	calculates pattern of articulatory motions based on the gestural score, according to universal (not language-particular) principles of task dynamics
Articulatory trajectories	the pattern of movements generated by the task-dynamic model

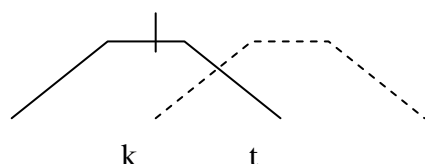
It is important to realize that there is a step between the gestural score and the actual, physical articulatory trajectories. Gestural scores provide input to the “task dynamic” module. Task dynamics is general model of skilled movement control, which can be applied to speech as well as other actions (Saltzman 1986, Saltzman & Munhall 1989). A gestural score may specify, for example, contradictory tract variable settings overlapping in time. The task dynamic model has to deal with such a situation by ‘blending’, which in at least some cases means averaging the parameter values for two gestures in the period where they conflict (Browman & Goldstein 1992a:30). A gestural score is a set of idealized instructions to articulators, which require some interpretation and modification to actually carry out. This interpretation is universal, not language-specific.

Gafos 2002 presents several results of GEST simulations that are useful in representing vowel intrusion. He reports the acoustic effects of different alignments among adjacent consonants. Aligning the c-center of one stop closure gesture with the onset of a heterorganic stop closure gesture results in an audible release between the two (at a certain speech rate). When the c-center of one stop is aligned with the target of a heterorganic stop, there is no audible release.

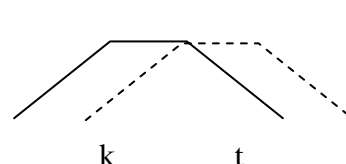
(11) Heterorganic stop clusters

Gafos 2002

a) CENTER (k) = ONSET (t)  
(Audible release)



b) TARGET (k) = ONSET (t)  
(No audible release)



Alignment a) above represents a French-like pronunciation of a [kt] sequence, the [k] having an audible release as in *acteur*. Alignment b) would produce a more English-like

pronunciation, where there is no release of the [k] before the [t] constriction is formed, as in *actor*. Vowel intrusion requires a period of release, consonants in a vowel intrusion structure may have a phasing relationship CENTER = ONSET, or a phasing relationship involving even less overlap.

#### 1.2.4. Gestural representation and the vowel intrusion syndrome

Now that the basic apparatus of gestural phonology has been laid out, it is possible to show how specific properties of the vowel intrusion syndrome are predicted by this model. Vowel intrusion is frequently blocked in homorganic clusters and at fast speech rates. If intrusive vowels have the representation in (4), where vowel intrusion results from acoustic release in a consonant cluster, then both of these characteristics follow from general principles of task dynamics. These characteristics are explained more simply in the gestural approach than in other approaches, and this is the rationale for using a gestural model.

##### 1.2.4.1. The heterorganicity requirement

Intrusive vowels are cross-linguistically rare in homorganic RC clusters. Some examples illustrating this tendency are shown below; a fuller list of conditioning environments for vowel intrusion appears in the appendix.

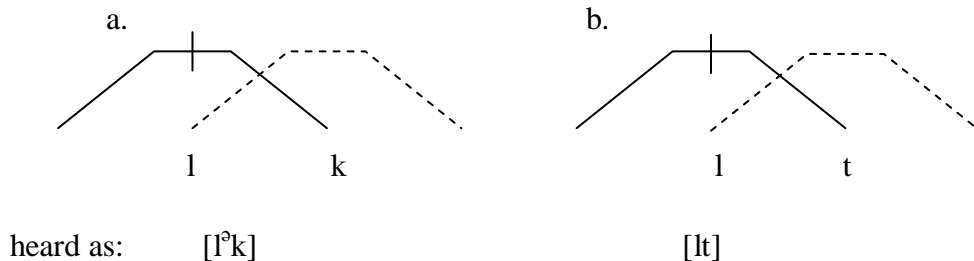
(12) Examples of:	Clusters w/ vowel intrusion	Clusters w/o vowel intrusion
Scots Gaelic <sup>4</sup>	np lb lm Lg rf rx mn mr	Rd Rt r't' Rs Lt L't' mb mp Nd Nt N'd' N't' Ns
Finnish	lh lm lp rp rk rv rj rh rm	rt rs ns mp ls lt
Oscan	lp rf lb rm nf nk rh lk rv	lt rt rn rs nt nd ns

This property turns out to follow directly from task dynamics. Gafos 2002 shows that the alignment CENTER = ONSET produces a release in heterorganic clusters but not in homorganic clusters. The articulatory realization of a particular gestural coordination depends not only on which landmarks that are aligned, but on the types of gestures involved.

---

<sup>4</sup> In the traditional transcription convention for Scots Gaelic, capital letters represent “tense” sonorants and apostrophes represent palatalization. Phonetic equivalents for one dialect are given in chapter 4, section 1.1.

(13) Acoustic effects of alignment CENTER = ONSET Gafos 2002



Intuitively, the reason for non-release in the homorganic cluster is that when the tongue receives instructions to begin moving away from the alveolar ridge, releasing the [l], it is also receiving instructions to begin moving towards the alveolar ridge, for the constriction of the [t]. When an articulator receives contradictory instructions, they must be reconciled through blending, which in this case causes the tongue to stay in place.

In at least the cases of slight, [ə]-like vowel intrusion, there is thus no need to posit any difference between the phasing of homorganic and heterorganic clusters. We can assume that in Finnish, for example, gives the same phasing to all RC clusters, and task dynamics explains why a release is heard in heterorganic [rk] but not homorganic [rt]. The description of output patterns is simplified, and there is no need to propose separate constraints or rules to deal with homorganic and heterorganic clusters.

#### 1.2.4.2. Speech rate effects

Speech rate effects are another piece of the vowel intrusion syndrome that finds a natural explanation under the gestural account.

The duration and appearance of intrusive vowels is often described as inconsistent, often in a way that is dependent on speech rate. For example, Quilis 1981:298 notes that the duration of the intrusive vowel in Spanish Cr clusters is ‘very variable’. The examples in his study range from 8 ms. to 56 ms., with an average value of 29 ms. Jannedy 1994’s experimental work shows that German speakers produce intrusion more at slower speech rates, and that duration of the vowel ranges from 30-60 ms. According to Sanskrit grammarians, “between *r* and a prevocalic fricative, a *svarabhakti* is pronounced having the length of  $\frac{1}{2}$  or  $\frac{1}{4}$  of *a*; before other consonants (than fricatives) its length is  $\frac{1}{4}$  or  $\frac{1}{8}$  of *a*” (Allen 1953:73). In Dutch, vowel intrusion is optional for many speakers even when words are produced in isolation (Donselaar et al. 1999:60-61). For Saami, Bye 2001:139 indicates free variation between clusters like [l:p] and [l:əp]. Holmer 1938:32 describes the intrusive vowel that appears in CR clusters in some Scots Gaelic dialects as ‘an obscure [ə] that may disappear’. For a Finnish dialect that has [ə]-like intrusive vowels, Harms 1976:77, who presents an informal gestural account, describes their appearance as depending on speech rate:

One final support for the view that schwa epenthesis is nonsegmental in nature is the great variation in the quality and duration of the intrusive vocoids. One-word

utterances—e.g., citation forms and exclamations (e.g. *helvetti* ‘hell’, *perkele* ‘devil’)—those with greatest sentence stress, are more likely to contain clearly perceptible schwa-type vowels... on the other hand, in fast speech and in words with weak sentence stress, these vocoids are very short; most frequently they are dropped altogether.

Most of the descriptions of variability concern intrusive vowels that are [ə]-like and short, rather than the longer intrusive vowels that are transcribed with copied qualities. Chamicuro, however, has intrusive vowels that have a distinct quality and yet appear only in emphatic speech.

(14)	Chamicuro	Parker 1994:266	
	Normal speech	Emphatic speech	
a.	túʔlu	túʔulu	‘chest’
b.	uɫáʔlo	uɫáʔalo	‘my wife’
c.	yapléʔti	yapléʔeti	‘lightning’
d.	piʔtótʃi	piʔitótʃi	‘canoe’
e.	maʔnáli	maʔanáli	‘dog’

This variability and optional disappearance is expected in a gestural approach. Gafos 2002 reports from simulations with GEST that the presence of a release between two consonants depends not only on their phasing relation, but also on the rate of speech.

In the dynamic model of gestures, the parameter which determines how fast gestures reach their targets is called stiffness. The higher the value of the stiffness, the faster the gesture. Changing stiffness thus can be used to study the acoustic consequences of fast speech... Maintaining the relational invariance between two consonant gestures—c-center of first to onset of second gesture— I varied the dynamic parameter of gestural stiffness. It was observed... that below a critical value of stiffness an acoustic release is present in the transition between the two consonant gestures; beyond that critical value of stiffness, however, the acoustic release disappears. (286)

The tendency of intrusive vowels to disappear in fast speech and lengthen in emphatic speech is thus a direct prediction of the gestural analysis. This approach dispenses with the need for optional rules of syncope or other devices to explain the vowels’ inconsistent presence.

### 1.3. Timing-Augmented Surface Phonology (TASP)

Gestural representations are clearly useful for describing some sound patterns. But there is controversy over what role they should play in the grammar: whether they should supplant segments and / or features as phonological primitives; whether they are present in underlying representations; and whether gestural phasing is determined in the same level of the grammar where phonological operations like deletion occur or in a separate “phonetic implementation” level. The typology of vowel intrusion bears on many of these questions.

I argue for a model which I call Timing Augmented Surface Phonology (TASP), in which gestural phasing is determined by ranked, violable constraints. Constraints referring to gestural and non-gestural phenomena are present in the same level of the grammar. Gestures supplant features but the notion of the segment remains. Intersegmental gestural timing is hypothesized to be universally non-contrastive. Gestures are underlyingly present, but their timing is not. This greatly limits the possibility of gestural contrasts, in a way that is consistent with the typology of vowel intrusion.

### 1.3.1. Gestures and segments

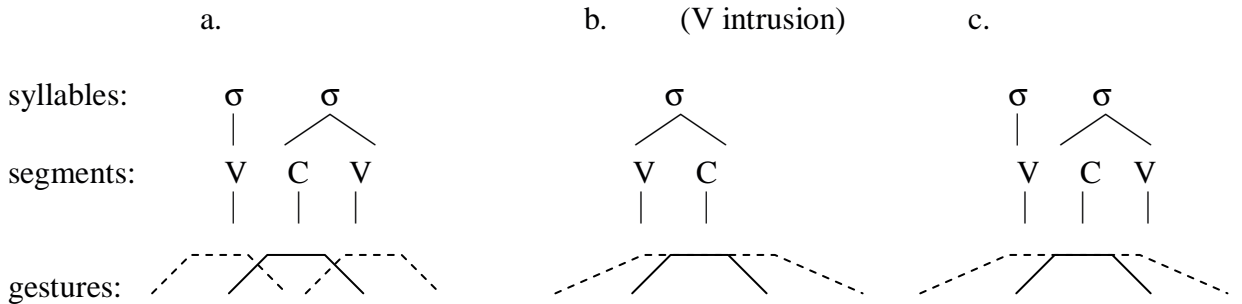
There is controversy about the relation between segments and gestures. In Browman & Goldstein's approach, the concept of the segment as an abstract unit visible to the phonology is abandoned. Gestures are the phonological primes, and segments are viewed at best as epiphenomenal. Byrd 1996:150-160, as well as Nittrouer et al. 1988, Saltzman & Munhall 1989:365 and Lofqvist 1991:346, suggest that the groups of gestures that constitute what are traditionally considered segments bear a more stable timing relationship to one another than do gestures that are parts of different segments: for example, that there is less variability in the alignment of the [k] and [p] gestures in a doubly articulated stop [kp̄] than in a [kp] sequence. The question is whether these groups of gestures have a stable timing relationship because they are associated with a single segmental unit, or whether segmenthood is a percept that results from a stable timing relationship. Byrd assumes the latter. She proposes that some gestures have lexically specified timing relationships, which causes them to have a more stable output timing than other gestures, and that "the percept and functionality of the segmental unit, to whatever extent it exists, results from its characteristic pattern of coordination... it is not the case that the quality of being a segment causes stable timing, but rather than stable timing causes the quality of being a segment." (159–60)

However, the 'segmenthood' of a group of gestures is not only a matter of consistent timing relations. Groups of gestures that are considered segmental act indivisible in certain ways in the phonology. For example, reduplication splits words and syllables but never segments. Language games are similar: they may separate a segment from its tone, but do not split up a segment. Speech errors also tend to work on whole segments.

The study of vowel intrusion can also contribute something to this debate. For example, vowel intrusion shows that gestures are not the units organized by syllables. As shown in chapters 2–5, intrusive vowels are non-syllabic: the vowel gesture corresponds to only one syllable despite the fact that it is heard in two pieces. However, there are other phenomena that plausibly involve a similar gestural score— a vowel completely surrounding a consonant— that are disyllabic. As discussed in chapter 3, there are languages where V.hV or V.?V sequences appear to involve a single vocalic gesture, yet clearly act like two syllables. In this way, syllabic structure is not predictable from the gestural score. Syllables do not organize gestures directly, but rather organize some other unit. This unit is the segment.

In TASP, an output representation contains both segments and gestures. Syllables organize segments rather than organizing gestures directly.

(15) Three possible timing-augmented segmental representations of [V<sub>i</sub>CV<sub>i</sub>].



In this my approach is close to that of Gafos 2002, who also incorporates segments into articulatory phonology and presents a formal theory of the relation between segments and gestures. He treats each segment as a set of gestures, which are organized with respect to one another in a way characteristic of that segment. Each segment has a ‘head gesture’, which is referred to by constraints on intersegmental coordination and serves as an anchor point for the other gestures associated with the segment. For consonants, the head gesture is the oral articulation. Segments are phased with respect to one another via their head gestures, as stated in (16), and the non-head gestures associated with each segment are in turn phased with respect to the head gesture.

(16) Definition: Inter-segmental coordination

Two segments  $S^1, S^2$  are coordinated with some coordination relation  $\lambda$  if the head gestures of these segments are coordinated as in  $\lambda$ . (Gafos 2002:284)

TASP will include the same assumption. Incidentally, Zsiga 1997 also argues for the existence of both segments and gestures, but treats them as belonging to different stages in a derivation: segments cease to be relevant after features are mapped to gestures. TASP assumes that segments and gestures are relevant at the same stage of evaluation.

### 1.3.2. Gestural coordination in Optimality Theory

The gestural representations defended here could be used in various phonological frameworks, both rule- and constraint-based, but the present analysis will be shown in Optimality Theoretic terms.

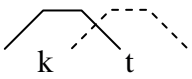
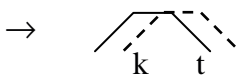
Gafos 2002 proposes that markedness constraints specify various alignments of gestural landmarks. Alignment constraints (McCarthy & Prince 1994) generally align edges of constituents with edges of other constituents. Gestural alignment constraints refer not only to edges of gestures, but to landmarks within gestures. The general form of such constraints is as follows:

(17) ALIGN ( $G^1$ , LANDMARK<sup>1</sup>,  $G^2$ , LANDMARK<sup>2</sup>)      Gafos 2002:278, 292  
Align landmark<sup>1</sup> of gesture<sup>1</sup> to landmark<sup>2</sup> of gesture<sup>2</sup>.

Landmark<sup>1</sup> takes values from the set {onset, target, c-center, release, offset}

The tableau below shows how the ranking of two such constraints would choose between two ways of coordinating the velar and alveolar closures in a [kt] sequence. In tableaux, I will use certain conventions for representing candidates. A candidate consists of a segmental representation, its associated gestures, and timing relations between these gestures. Only one curve, representing the head gesture, is shown for each segment (unless intrasegmental timing is relevant). In the case of this [kt] sequence, the head gestures are the velar closure and the alveolar closure. The symbol under each gestural curve indicates the segment that it is associated with. Below this is a phonetic transcription in brackets that tells how the sequence would sound. The bracketed transcription is *not* part of the representation being evaluated by the constraints; it is parenthetical, to remind the reader how that gestural score would be heard. In this case, one acoustic difference predicted by GEST is that only candidate a) would have an audible release, represented as a superscript [ə] (see section 2).

(18) C<sub>1</sub>C<sub>2</sub> coordination

/kt/	ALIGN (C <sub>1</sub> , RELEASE, C <sub>2</sub> , TARGET)	ALIGN (C <sub>1</sub> , CENTER, C <sub>2</sub> , ONSET)
a.  [k <sup>ə</sup> t]	*!	
b. →  [kt]		*

The ranking of ALIGN (C<sub>1</sub>, RELEASE, C<sub>2</sub>, TARGET) and ALIGN (C<sub>1</sub>, CENTER, C<sub>2</sub>, ONSET) determines whether there will be acoustic release, which is part of what produces vowel intrusion.

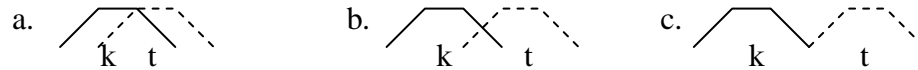
### 1.3.3. Gestures, underlying representations, and faithfulness

In TASP, only markedness constraints govern inter-segmental gestural coordination, not faithfulness constraints. This is necessary in order to avoid overpredicting the types of structures that can potentially contrast. If faithfulness constraints preserved underlying timing relations between segments, it would be predicted that these timing relations would contrast.

In Optimality Theory, positing faithfulness constraints to underlying distinctions expands the range of possible grammars. According to the principle of ‘richness of the base’, the grammar of a language must be able to take any structure as an input and transform it to an output that is a possible utterance of that language. There are no language-particular restrictions on possible inputs. If timing relations are part of the

structure allowed in inputs, then an input in any language could contain any of the following structures.

(19) Possible inputs



Markedness constraints want to change inputs so that they all have the same gestural coordination in the output. Faithfulness constraints want there to be no change, so that the surface timing is identical to the underlying timing.

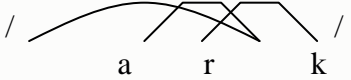
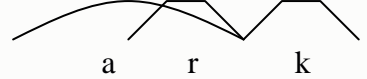
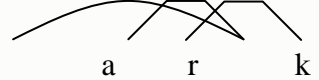
Another basic premise of Optimality Theory is that constraints can have any ranking, and that each ranking must produce a possible grammar. If there are faithfulness constraints on gestural coordination, we must assume that in some languages they are ranked above the markedness constraints on gestural coordination. Therefore, gestural coordination will be contrastive in some languages.

The tableaux below show what happens in a grammar that ranks IDENT-TIMING (“preserve underlying gestural timing relations”) above ALIGN ( $C_1$ , center,  $C_2$ , onset). Such a grammar will preserve the underlying timing of any input. If a CC cluster has the timing OFFSET = ONSET in the input, as in (20), then the output will also have the timing OFFSET = ONSET; if a CC cluster has the timing CENTER = ONSET, as in (21), then the output will have the timing CENTER = ONSET. The markedness constraint ALIGN ( $C_1$ , CENTER,  $C_2$ , ONSET) would prefer all outputs to have the timing CENTER = ONSET, but because this constraint is ranked below faithfulness, it is overruled.

(20) Language with contrastive gestural phrasing

input:	IDENT-TIMING	ALIGN ( $C_1$ , CENTER, $C_2$ , ONSET)
a. → [arək]		*
b. [ark]	*!	

(21) Language with contrastive gestural phasing

input: 	IDENT-TIMING	ALIGN (C <sub>1</sub> , CENTER, C <sub>2</sub> , ONSET)
a.  [arək]	*!	*
b. →  [ark]		

In such a language, degrees of gestural overlap will be contrastive. The timing of an [r] and [k] might be all that distinguishes two morphemes. Hence, phenomena related to gestural timing, such as release and vowel intrusion, should be contrastive as well.

Vowel intrusion seems to be universally non-contrastive. In each language, it either always happens or never happens in a given environment. This strongly argues that if the gestural overlap analysis of vowel intrusion is to be maintained, there cannot be faithfulness to underlying gestural phasing between segments.

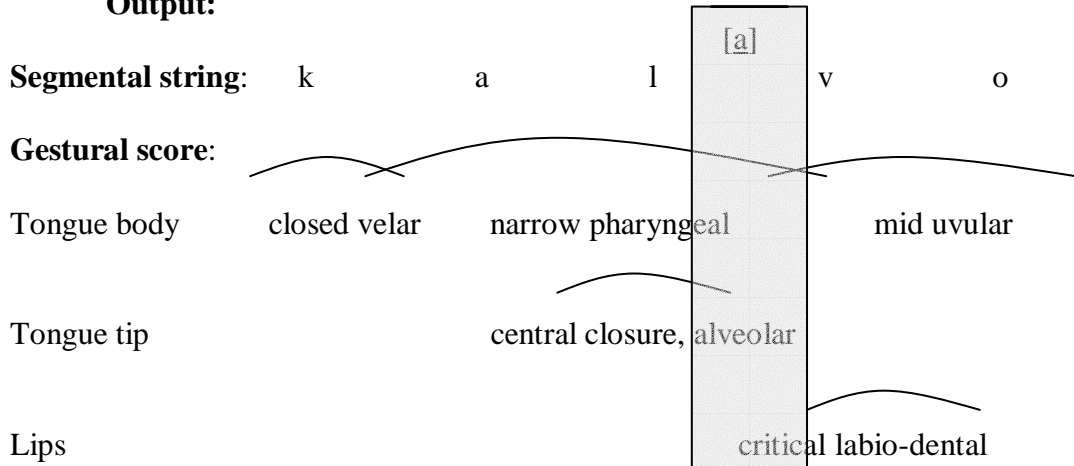
#### 1.3.4. A TASP representation of vowel intrusion

A sample input – output pairing in TASP for a word like Finnish [kalavo] ‘transparency’, looks as in (22). The input consists only of a string of segments. (I argue below that a segment actually consists of a bundle of gestures associated with a root node, so the representation below will need to be expanded slightly). The output consists of a string of segments, syllabified, and their associated gestures, with the timing relations specified between the gestures.

The shaded area in the diagram below indicates the period when the intrusive vowel is perceived, namely the interval after the release of the [l]’s alveolar constriction and before the labio-dental constriction of the [v]. During this period, only the [a]’s articulation is active, and this is what is heard.

(22) **Input:** /kalvo/

**Output:**



The shaded area and the bracketed [a] are not formally part of the representation. The intrusive vowel, despite its acoustic prominence, does not correspond to any independent segment, gesture or other entity in the score, and this is important for explaining its invisibility to some phonological patterns.

The distinctive characteristics of TASP are summarized below.

(23) Characteristics of TASP

1. The input contains segments, but no gestural phasing relations.
2. Only markedness constraints determine gestural alignment, not faithfulness constraints. Hence, gestural alignment cannot be contrastive.
3. Gestural phasing is present in output candidates.
4. Gestural and non-gestural constraints are in the same level of the grammar and can potentially interact.
5. Higher prosodic structures, such as syllables and feet, are composed of segments, not of gestures.

With these basic assumptions about the grammar in place, we can explore the particular constraints that produce vowel intrusion. Vowel intrusion must be shaped entirely by markedness constraints on gestural phasing. In the following section, I show which constraints are needed to capture the typical conditioning environments of intrusive vowels.

## 1.4. Gestural coordination constraints

### 1.4.1. C-C and V-C phasing constraints

For an intrusive vowel to be heard in a CCV or VCC sequence, at least two things must be true: the consonants must be phased so that there is a release between them, and the vowel articulation must overlap this period of release.

The phasing of the consonants can be produced by ALIGN (C<sub>1</sub>, CENTER, C<sub>2</sub>, ONSET).

(24) ALIGN (C<sub>1</sub>, CENTER, C<sub>2</sub>, ONSET)

In a C<sub>1</sub> C<sub>2</sub> sequence, the center of C<sub>1</sub> is aligned with the onset of C<sub>2</sub>.

According to GEST simulations (Gafos 2002:271-272), a phasing relation of CENTER = ONSET produces a schwa-like vocalic element between heterorganic consonants. For an intrusive [ə] to be heard, then, it is necessary for the consonants to be phased at least this far apart. In languages where the intrusive vowel is described as having a clear and audible quality, the consonants are probably even more spread apart, perhaps at RELEASE = ONSET or OFFSET = ONSET. Not knowing exactly which phasing relationship is appropriate in each language, I will generally use CENTER = ONSET to produce vowel intrusion in tableaux.

ALIGN (C<sub>1</sub>, CENTER, C<sub>2</sub>, ONSET) competes with another CC phasing constraint, ALIGN (C<sub>1</sub>, RELEASE, C<sub>2</sub>, TARGET). This constraint favors a phasing that does not produce release.

(25) ALIGN (C<sub>1</sub>, RELEASE, C<sub>2</sub>, TARGET)

In a C<sub>1</sub> C<sub>2</sub> sequence, the release of C<sub>1</sub> is aligned with the target of C<sub>2</sub>.

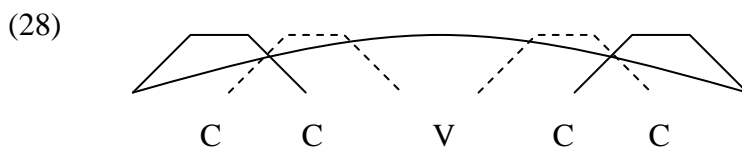
The ranking of these two determines whether there will be release or not in a consonant cluster.

Why should there be multiple constraints on the phasing of CC clusters? The competing C-C coordination constraints reflect the competing priorities of ease of articulation and ease of perception. Constraints favoring less overlap have a functional grounding in perceptibility: a consonant has clearer formant offsets when it is not heavily overlapped by another consonant. Donselaar et al. 1999 show that in Dutch, where vowel intrusion is optional, reaction times to lexical decision tasks and sonorant identification tasks are quicker when a word like *tulp* ‘tulip’ is pronounced with vowel intrusion ([tuləp]) than without ([tulp]). The alignment CENTER = ONSET helps listeners to recognize sounds more easily. The alignment RELEASE = TARGET, on the other hand, allows a faster and in that sense more efficient articulation.

A second type of constraint must ensure that the vowel overlaps the release period between the consonants. I propose that there are constraints demanding that a vowel articulation span its entire syllable. These can be instantiated as constraints aligning the onset and offset of a vowel to the left and right edges of the syllable, respectively.

- (26) ALIGN (V, OFFSET, SYLL, OFFSET)  
The offset of every vowel is aligned with the offset of the rightmost segment that belongs to the same syllable as that vowel.
- (27) ALIGN (V, ONSET, SYLL, ONSET)  
The onset of every vowel is aligned with the onset of the rightmost segment that belongs to the same syllable as that vowel.

These constraints prefer gestural alignments producing vowel intrusion, like the following:



Since vowel intrusion does not happen univervally, these constraints must be countered by opposing ones. I propose that there is a type of constraint that penalizes too much overlap of one gesture by another. Such constraints take the form below.

- (29) \*GESTURE<sub>x</sub> IN GESTURE<sub>y</sub>  
A gesture of type *x* does not fully surround a gesture of type *y* (extending on both sides of it).

An example of such a constraint is \*C IN V.

- (30) \*C IN V  
A vowel articulation does not fully surround a consonant articulation.

This constraint will favor the vowel's articulation beginning no earlier than the last prevocalic consonant, and ending no later than the first postvocalic consonant, as below. This degree of overlap presumably is not sufficient to produce a vocalic-sounding release.



These four basic constraints determine whether there is release in a consonant cluster, and whether a vowel overlaps this period of release.

### 1.4.2. Typology of coordination effects



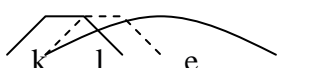
This section will demonstrate how different rankings of the constraints proposed above produce three different patterns of gestural organization within a /CCV/ or /VCC/ syllable: vowel intrusion, release without vowel intrusion, and no release. In this chapter, I discuss only phasing within the syllable. Vowel intrusion between syllables, and the interaction between syllable structure and gestural organization, are treated in chapter 2.

#### 1.4.2.1. Vowel intrusion

Vowel intrusion occurs in a /CCV/ string when  $\text{ALIGN}(\text{V}, \text{ONSET}, \text{SYLL}, \text{ONSET})$  is ranked above  $*\text{C IN V}$ , and  $\text{ALIGN}(\text{C}_1, \text{CENTER}, \text{C}_2, \text{ONSET})$  is ranked above  $\text{ALIGN}(\text{C}_1, \text{RELEASE}, \text{C}_2, \text{TARGET})$ . These two rankings ensure that the vowel will fully overlap the consonant cluster, and that there will be a period of release during which the vowel can be audible.

In tableau (32), candidate a) has the phasing  $\text{CENTER} = \text{ONSET}$  in the [kl] cluster, producing a release. But the vowel gesture does not fully overlap the cluster, so the release is not heard as vocalic. The vowel's phasing violates  $\text{ALIGN}(\text{V}, \text{ONSET}, \text{SYLL}, \text{ONSET})$ , so a) is eliminated. Candidate b) also has the  $\text{CENTER} = \text{ONSET}$  phasing for [kl], and the vowel does overlap the whole cluster. Candidate c) has the phasing  $\text{RELEASE} = \text{TARGET}$ , which produces no release in [kl]. Since  $\text{ALIGN}(\text{C}_1, \text{CENTER}, \text{C}_2, \text{ONSET})$  is high-ranked, c) is eliminated and b) wins.

(32) Vowel intrusion: Oscan *pukele* 'son'

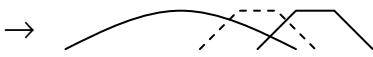
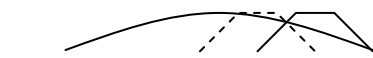
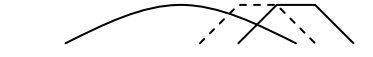
/pukle/	$\text{ALIGN}(\text{C}_1, \text{CENTER}, \text{C}_2, \text{ONSET})$	$\text{ALIGN}(\text{V}, \text{ONSET}, \text{SYLL}, \text{ONSET})$	$*\text{C IN V}$	$\text{ALIGN}(\text{C}_1, \text{RELEASE}, \text{C}_2, \text{TARGET})$
a.  [pukəle]		*!		*
b. →  [pukele]			*	*
c.  [pukle]	*!	*		

For vowel intrusion in a /VCC/ string, the analysis is the same except that  $\text{ALIGN}(\text{V}, \text{ONSET}, \text{SYLL}, \text{ONSET})$  must be ranked above  $*\text{C IN V}$ .

### 1.4.2.2. Release without vowel intrusion

If a language ranks ALIGN (C<sub>1</sub>, CENTER, C<sub>2</sub>, ONSET) above ALIGN (C<sub>1</sub>, RELEASE, C<sub>2</sub>, TARGET), but ALIGN (V, OFFSET, SYLL, OFFSET) below \*C IN V, it has releases between all consonants, but no intrusive vowels. In French, for example, the word *acte* is generally pronounced with a release between the [k] and [t]. Candidates a)–c) and their violations are the same as in tableau (32), but since the constraint ranking is different, candidate a) wins.

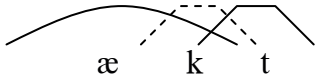
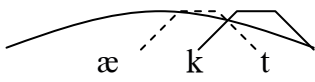
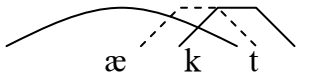
(33) French: release without vowel intrusion

/akt/	ALIGN (C <sub>1</sub> , CENTER, C <sub>2</sub> , ONSET)	*C IN V	ALIGN (V, OFFSET, SYLL, OFFSET)	ALIGN (C <sub>1</sub> , RELEASE, C <sub>2</sub> , TARGET)
a. →  a k t [ak <sup>a</sup> t]			*	*
b.  a k t [ak <sup>a</sup> t]		*!		*
c.  a k t [akt]	*!		*	

### 1.4.2.3. No release

In English, consonants generally are not released in a cluster. This pattern results from ranking ALIGN (C<sub>1</sub>, RELEASE, C<sub>2</sub>, TARGET) above (C<sub>1</sub>, CENTER, C<sub>2</sub>, ONSET). The candidates have the same violation marks as in (32) and (33), but under this ranking, c) wins.

(34) English: no release

/ækt/	ALIGN (C <sub>1</sub> , RELEASE, C <sub>2</sub> , TARGET)	*C IN V	ALIGN (V, OFFSET, SYLL, OFFSET)	ALIGN (C <sub>1</sub> , CENTER, C <sub>2</sub> , ONSET)
a.  [ækʰt]	*!		*	
b.  [ækʰt]	*!	*		
c. →  [ækt]			*	*

In clusters without release, it is hard to tell from transcriptions whether the vowel fully overlaps the cluster or not. This probably can be determined only instrumentally, although perhaps the degree of C-V coarticulation might provide a rough diagnostic.

### 1.4.3. The special behavior of sonorants

Vowel intrusion happens mostly, and perhaps 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. This suggests that special phasing constraints apply to sonorants.

There is evidence from other phenomena that sonorants have special phasing relations both with vowels and with other consonants. In some languages, like Upper Chehalis, there is always a release before sonorants without regard to vowel overlap, which shows there is a special C-C coordination constraint referring to these sound classes.

Also, vowels seem to overlap better with sonorants than with other consonants. In Hua, all consonants clusters have release, but a vocalic quality is heard in the release only if a sonorant is involved.

#### 1.4.3.1. Vowel intrusion by sonorants; release elsewhere

Hua has releases, transcribed as [ə], between all consonants in careful speech. The release has the quality of the following vowel only in a C[r] or C[ɣ] cluster. I show the vowels in superscript following Haiman's transcription.

## (35) Plain release in Hua

Haiman 1980:26-7

- |    |  |  |
|----|--|--|
| a. | f <sup>a</sup> tú                                | ‘smell’  |
| b. | k <sup>a</sup> t <sup>a</sup> r <sup>a</sup> gúʔ | ‘kind of mushroom’                               |
| c. | t <sup>a</sup> v <sup>a</sup> r <sup>a</sup> gíe | ‘he sharpened it’                                |
| d. | m <sup>a</sup> ni                                | ‘water, river, beverage other than milk or beer’ |
| e. | r <sup>a</sup> m <sup>a</sup> zoe                | ‘I laid an offering (at a girl’s door)’          |
| f. | k <sup>a</sup> s <sup>a</sup> sípaie             | ‘he wrapped it in leaves with a vine string’     |

## Vowel intrusion

- |    |  |                               |
|----|--|-------------------------------|
| g. | pot <sup>a</sup> yaie                  | ‘it glanced off (the target)’ |
| h. | ok <sup>u</sup> rumaʔ                  | ‘sky’                         |
| i. | k <sup>a</sup> ra                      | ‘dog’                         |
| j. | f <sup>i</sup> rie                     | ‘he died’                     |
| k. | t <sup>e</sup> re t <sup>e</sup> refie | ‘it flowed’                   |
| l. | d <sup>a</sup> yai                     | ‘I’                           |
| m. | b <sup>o</sup> ʔot <sup>a</sup> va     | ‘one’                         |

The consistent presence of release shows that all CC clusters have a wide phasing. It may be as wide as OFFSET = ONSET, since release happens even in homorganic clusters. But evidently, the vowel gesture overlaps this release period only if the intervening consonant is [r] or [ʎ].<sup>5</sup> To capture this, we need a more specific version of \*C IN V, which refers to particular classes of consonants.

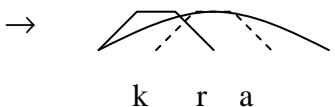
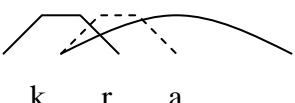
## (36) \*OBSTRUENT IN V

A vowel gesture does not completely surround the gesture of an obstruent consonant.

Under the ranking \*OBSTRUENT IN V >> ALIGN (V, ONSET, SYLL, ONSET) >> \*C IN V, a vowel will fully overlap a consonant cluster only if the closest consonant is a non-obstruent. This allows vowel intrusion to occur in a C<sub>1</sub>C<sub>2</sub>V input where C<sub>2</sub> is non-obstruent. In the tableau below, both candidates have a CR cluster with a release. In candidate a), the vowel overlaps this release and produces an intrusive vowel.

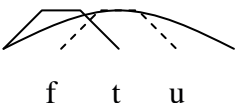
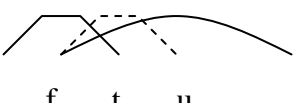
<sup>5</sup> [ʎ] patterns with the gutturals in other cases of vowel intrusion as well, such as Negev Bedouin Arabic (see (51)). It is possible that the sound is often more uvular, even when transcribed as a velar.

(37) Hua vowel intrusion

/kra/	*OBSTRUENT IN V	ALIGN (V, ONSET, SYLL, ONSET)	*C IN V
a. →  k r a [k <sup>a</sup> ra]			*
b.  k r a [k <sup>a</sup> ra]		*!	

In a C<sub>1</sub>C<sub>2</sub>V input where C<sub>2</sub> is an obstruent, however, the high ranked \*OBSTRUENT IN V prevents vowel intrusion. Only plain release can occur, as in candidate b) below.

(38) Hua release

/ftu/	*OBSTRUENT IN V	ALIGN (SYLL, ONS, V, ONS)	*C IN V
a.  f t u [f <sup>u</sup> tu]	*!		*
b. →  f t u [f <sup>ə</sup> tu]		*	

Hua shows that sonorants are more conducive to being overlapped by vowel articulations. This is part of the reason that they are especially conducive to vowel intrusion.

**1.4.3.2. Release by sonorants, no release elsewhere**

There is also evidence that there are special C-C phasing constraints for clusters that include a sonorant. Some languages have releases only next to sonorants.

Upper Chehalis (Q<sup>w</sup>ay'áyíłq') (Kinkade 1963 et seq.,) has excrescent schwas (which are presumably an uncolored release) between a consonant and any one of [m n j l w ?]. Kinkade argues that these schwas are purely transitional, which is equivalent to saying that they result from C-C phasing. In their phonological behavior, they contrast with phonetically similar epenthetic [ə]s that appear in a different set of environments.

- (39) Upper Chehalis excrescent schwas Kinkade 1963:192-3
- |    |                         |                        |
|----|-------------------------|------------------------|
| a. | č́ə́ləlp                | ‘open-weave basket’    |
| b. | č́’op’wə́xənstwən       | ‘squeeze with the arm’ |
| c. | q’ <sup>w</sup> ó təweʔ | ‘maple’                |
| d. | sqəʔúq <sup>w</sup> əls | ‘skull’                |

The Upper Chehalis schwas differ from intrusive vowel in two ways. First, there is no indication that these schwas have any quality that varies depending on the neighboring vowel. Second, in vowel intrusion, the sonorant is always adjacent to a vowel. The Upper Chehalis schwa does not need to be near a vowel at all: many consonants may intervene between the schwa and the nearest full vowel, as in [č́’op’wə́xənstwən], and those intervening consonants can be of any type.

These two facts together suggest that these schwas are a result only of distance between consonants, without any overlapping vowel articulation. There must therefore be special coordination constraints on CC clusters that include sonorants.

- (40) ALIGN (R, CENTER, C, ONSET)  
 In a C<sub>1</sub> C<sub>2</sub> string, C<sub>1</sub> a sonorant, the center of C<sub>1</sub> is aligned with the onset of C<sub>2</sub>.
- (41) ALIGN (C, CENTER, R, ONSET)  
 In a C<sub>1</sub> C<sub>2</sub> string, C<sub>2</sub> a sonorant, the center of C<sub>1</sub> is aligned with the onset of C<sub>2</sub>.

These constraints produce release next to sonorants, without vowel intrusion, as shown below. Candidates a) and b) have a CENTER=ONSET phasing for the [qʔ] cluster, as demanded by ALIGN (C, CENTER, R/H, ONSET). Candidate c) has a RELEASE = TARGET alignment for [qʔ], so it is eliminated. Whether a) or b) wins is determined by the relative ranking of \*C in V and ALIGN (V, ONSET, SYLL, ONSET). When the former is higher-ranked, as below, the vowel gesture is not permitted to extend fully over a consonant in order to span the syllable. Candidate b) wins.

(42) Upper Chehalis [sqəʔúq<sup>w</sup>əls] ‘skull’

/sqʔuq <sup>w</sup> ls/	ALIGN (C, CENTER, R/H, ONSET)	*C in V	ALIGN (C <sub>1</sub> , RELEASE, C <sub>2</sub> , TARGET)	ALIGN (V, ONSET, SYLL, ONSET)
<p>a.</p> <p>s q ʔ u</p> <p>[sqʔúq<sup>w</sup>əls]</p>		**!	*	
<p>b. →</p> <p>s q ə ʔ u ə</p> <p>[sqəʔúq<sup>w</sup>əls]</p>			*	*
<p>c.</p> <p>s q ʔ u</p> <p>[sqʔúq<sup>w</sup>əls]</p>		*!		*

Diamandis Gafos (p.c.) reports similar facts for Moroccan Arabic: there are audible releases after gutturals in clusters.

Thus, there are special constraints both on the phasing of sonorants with respect to other consonants, and on vowels overlapping sonorants. Both of these contribute to restrict vowel intrusion to clusters with sonorants.

#### 1.4.4. Hierarchy of sonorant / vowel overlap

Besides being restricted to sonorants generally, vowel intrusion happens only with a subset of the inventory of sonorants in most languages. Cross-linguistically, vowel intrusion happens more with liquids than with other sonorants, and more with rhotics than laterals. This indicates a need for more specific constraints specific consonant classes, in addition to those in (36), (40), and (41).

The chart below contains the inventory of sonorants in several languages, divided into those that are and aren't found with intrusive vowels. In some cases, a certain segment is never found in a position where it could trigger vowel intrusion and hence is irrelevant to determining whether there is a hierarchy of segments allowing vowel intrusion. Such segments, where known to me, are in boldface. For example, Dutch has vowel intrusion only in codas, and [h] does not appear in codas, so we cannot tell whether in principle [h] would allow vowel intrusion or not.

## (43) Sonorant inventories

	Triggering V intrusion	Not triggering V intrusion
Kekchi	ʔ	r l n m w j h
Mamaidé	h	n m l j w
Tiberian Hebrew	ʔ ʕ h h	r l n m w j
Spanish	r	r l m n ñ ʎ
German	ʁ	l m n j v ŋ
Sanskrit	r (l)	m ŋ n ɲ ŋ y v h
Dutch	r/ʁ l	m n ŋ h
Hausa	r l	r n m w j ʔ
Finnish	(r?) <sup>6</sup> l n h	(r?) m ŋ j
Oscan	r l n	m h
Saami	r l j β ð <sup>7</sup>	m n ɲ h
Hocank	r n w	h j m ʔ
Scots Gaelic	r <sup>v</sup> r r <sup>j</sup> l <sup>v</sup> l <sup>vj</sup> l n <sup>v</sup> n <sup>vj</sup> n m	w j h

This list reveals a partial implicational hierarchy as to which sonorants allow vowel intrusion. First, there are clear-cut cases where gutturals do have vowel intrusion and other sonorants don't, such as Kekchi and Tiberian Hebrew. This suggests the following implication:

(44) Vowel intrusion triggers  
non-guttural sonorants → gutturals

A number of languages have vowel intrusion only over non-guttural sonorants and not over gutturals, but in most or all of these cases, whatever gutturals the language has never appear in positions where vowel intrusion would be expected (as with Dutch [h]). Thus, the lack of vowel intrusion with gutturals can be attributed to other factors that prohibit gutturals appearing in clusters, and does not form a counterexample to the generalization.

A potential counterexample to (44) comes from Saami. Saami transcriptions often appear to contain clusters of [h] and a voiceless stop, in words like [toh:pa] 'sheath', without vowel intrusion (Bye 2001:138). Bye analyzes these as single preaspirated consonants, not as clusters, but if the cluster interpretation were favored, it would be a problem for the proposed hierarchy. Also, Hocank is a counter-example, because it has vowel intrusion in CR but not Cʔ clusters. I will argue in chapter 5, however, that Hocank vowel intrusion happens for different reasons than the other cases here, and there is a different type of constraint that rules out vowel intrusion with [ʔ].

<sup>6</sup> There is a disagreement about this data, as discussed in chapter 2; apparently [r] does not trigger vowel intrusion in modern Finnish although it was reported to do so in the 1920's.

<sup>7</sup> According to Bye, β and ð are approximants.

In languages that have vowel intrusion with gutturals, there are no cases where only a subset of the gutturals trigger it (unless one of the gutturals does not occur in the correct position, like Kekchi [h]). No implicational hierarchy within the guttural class can be established at present.

Within the class of non-guttural sonorants, there is a clear preference for vowel intrusion with liquids, especially rhotics. Intrusion with laterals implies intrusion with rhotics, except the alveolar trill. For example, an experiment by Jannedy 1994 produces intrusion in German [bʁ] but not [gl]. In Sanskrit, some authorities report intrusion after both [r] and [l], but others only after [r]. For Irish Gaelic, which has had vowel intrusion after many sonorants, Greene 1952:217 mentions that there is evidence that the intrusive vowels historically arose first after [r]. The preference for intrusion with rhotics is even revealed in variation. For Dutch, where vowel intrusion is optional, a production experiment by Kiuipers & Donselaar 1997 found that speakers had vowel intrusion 60% of the time after /r/ (which can be alveolar or uvular) and only 40% of the time after [l].

The alveolar trill is the only rhotic that produces vowel intrusion less than [l] does. Hausa and Finnish (according to most descriptions) have vowel intrusion with [l] but not [r]. Intrusion is common with [n] than [m], suggesting a preference for overlap with coronals.

These facts support the following hierarchy:

- (45) Vowel intrusion triggers  
obstruents, if ever → other approximants, nasals → r → l → r , ʁ → gutturals

Among nasals: m → n

Beyond this the pattern becomes less clear, and no implicational hierarchy can be established yet. Saami has vowel intrusion after glides but not nasals; Finnish and Scots Gaelic have it after some nasals but no glides.

#### 1.4.4.1. Constraints on V / C overlap

As argued above, there are two overlapping reasons that sonorants are particularly conducive to vowel intrusion: they are subject to special constraints favoring release; and they are *not* subject to the special constraints that disfavor other types of segments being fully overlapped by a vowel. The implicational hierarchy as to which segments within the sonorant class allow vowel intrusion could be due to either of these factors or both of them. Hua, which has release before all consonants but vowel intrusion only before [r] and [ɣ] (and not other sonorants like [n]), is evidence that vowels overlap some sonorants better than others. I know of no Upper Chehalis-like language that has release only by certain sonorants.

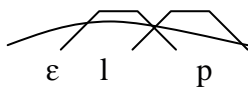
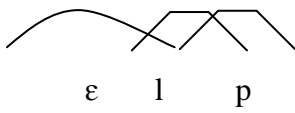
Therefore, I propose that the \*V IN C constraints are relativized to particular classes of sounds within the sonorant group, and that this family of constraints has a universally fixed ranking that correlates with the hierarchy of vowel intrusion triggers. This yields the following set of constraints.

- (46) \*OBSTRUENT IN V >> \*NASAL IN V; \*GLIDE IN V >> \*[r] IN V >> \*[l] IN V >> \*[r] IN V, \*[ɾ] IN V >> \*GUTTURAL IN V

The ranking of ALIGN (V, OFFSET, SYLL, OFFSET) with respect to this scale will determine which sonorants are eligible to trigger vowel intrusion.


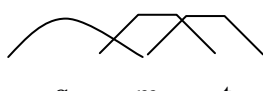
Dutch, for example, has ALIGN (V, OFFSET, SYLL, OFFSET) ranked below \*NASAL IN V but above \*[l] IN V. The tableaux below show how this ranking produces vowel intrusion into clusters like [lp] but not into clusters like [mt]. Both cluster types have a release, but only in the case of [lp] can the vowel fully overlap this period of release, as in candidate a).

(47) Dutch vowel intrusion

/hɛlp/	*NASAL IN V	ALIGN (V, OFFSET, SYLL, OFFSET)	*[l] IN V
a. →  ε l p [hɛlɔp]			*
b.  ε l p [hɛlp]		*!	

In tableau (48), the intrusive vowel candidate a) is eliminated because it involves total overlap of a vowel and [m].

(48) Dutch blocking of vowel intrusion

/kamt/	*NASAL IN V	ALIGN (V, OFFSET, SYLL, OFFSET)	*[l] IN V
a.  ɑ m t [kamət]	*!		
b. →  ɑ m t [kamt]		*	

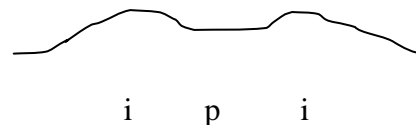
Dutch has, in fact, intrusive *stops* between nasals and heterorganic consonants, so that /kamt/ is pronounced [kampt]. Stop intrusion is another phenomenon that is likely a percept caused by a gestural timing pattern rather than epenthesis of a stop segment: it is a release of the [m] after the raising of the velum and cessation of voicing.

#### 1.4.4.2. Additional evidence for \*C in V: articulatory trough data

The phenomenon of articulatory “troughs” provides an additional source of evidence that some types of consonants are more easily overlapped by vowels than others.

An articulatory trough occurs when articulations associated with the identical vowels in a  $V_iCV_i$  sequence are relaxed during the C (Bell-Berti & Harris 1974, Gay 1975). The trough phenomenon suggests that the two vowels have separate gestures, and that the trough represents the juncture where the first vowel gesture is leaving its target and the second has not yet reached its target.

(49) Tongue movement: (approximated from Harris & Bell-Berti 1984)

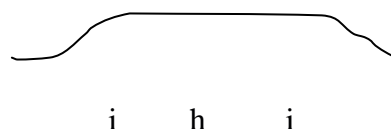


Hypothesized gestures:



Articulatory troughs may be absent for some consonants, however. Harris & Bell-Berti 1984 find that, in nonsense words produced by an English speaker, there is a decrease in lip rounding during the medial consonant in [ɔʔɔ] sequences but not [ɔʔɔ], and a relaxing of the tongue position during the medial consonant in [ipi] but not [ihi] or [iʔi]. This suggests that the sequences involving laryngeals have a single vocalic gesture spanning the two syllables, as below, and that this vowel gesture fully surrounds the consonant gesture.

(50) Tongue movement: (approximated from Harris & Bell-Berti 1984)



Hypothesized V gestures:



The fact that troughs are lacking only with gutturals supports the idea that these are most amenable to vowel overlap.

### 1.4.4.3. VC homorganicity

Besides the widespread restrictions on consonant types triggering vowel intrusion, there is one example of vowel intrusion happening only with a particular vowel. In some dialects of Bedouin Arabic, copy vowels appear [a]HC clusters only (H = guttural). These arose through a historical process known as the *gahawa* syndrome, after the word for ‘coffee’, which is [qahwa] in Classical Arabic ([qahawa] shows the result of the process.) In other Bedouin and Gulf dialects, the gutturals have completely metathesized with the [a], in a process known as the *ghawa* syndrome. As Steriade 1990 points out, metathesis of this type can be seen as the movement of a consonant articulation all the way across a vowel.

(51) Negev Bedouin Arabic reflexes of Classical Arabic words  
(NBA- Blanc 1970:124-5, CA- Wehr 1971)

a.	ʃahr	>	ʃa'har	‘month’
b.	ʃaʕr	>	ʃa'ʕar	‘hair’
d.	qahwa	>	ga'hawah	‘coffee’
e.	bayl	>	ba'yal	‘mule’
f.	wahʃ	>	wa'haʃ	‘beast’
g.	baxt	>	ba'xat	‘luck’

These vowels may not be synchronically represented as intrusive vowels. They seem to have been historically reanalyzed as regular segments for at least some speakers. What indicates this is that [ʃahar] has an alternate pronunciation [ʃihar], reflecting the open syllable raising process of Arabic. Such a process would not affect a vowel without affecting the following intrusive vowel. But given the similarity of the conditioning environment to examples of current vowel intrusion, it is reasonable to hypothesize that the *gahawa* and *ghawa* syndromes are relics of a vowel intrusion process that was productive at some point in the past.

The most likely reason that gutturals in this language overlap heavily only with [a] is that [a] is articulatorily most similar to the guttural consonants, involving a narrow pharyngeal constriction. There are several other cases of historical metathesis that support this conclusion. Grammont 1933 notes that in various French dialects such as Pléchéâtel and Havre, [r] has metathesized with [e] but not with other vowels.

(52)	Pléchâtel	Standard French	Grammont 1933:245
	a. kerve	<i>crever</i>	
	b. gerlô	<i>grelot</i>	
	Havre		
	c. bertô	<i>breton</i>	
	d. kersô	<i>cresson</i>	
	e. gerloté	<i>grelotter</i>	

Grammont analyzes this metathesis as ‘inversion by penetration’, which means, essentially, metathesis by way of vowel intrusion. He points out that [r] is articulated in the same general region as [e], and proposes that this is what allows the ‘penetration’ to occur. He gives the same reason for why in unaccented syllables in Latin, there was historical metathesis of [i] with [r] and [u] with [l], but not the same liquids with other vowels.

These cases are evidence that vowel type as well as sonorant type can affect the potential for VC overlap. Vowels prefer to overlap with consonants that are articulatorily similar to them.

#### 1.4.4.4. The relevance of vowel intrusion

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. This typological survey provides evidence for constraints favoring a low degree of overlap between sonorants and adjacent consonants, a high degree of overlap between sonorants and adjacent vowels, and better overlap between homorganic vowels and consonants.

Among types of gestural coordination, vowel intrusion is ideally suited for a typological survey. Because it is clearly audible, it is often mentioned in descriptive works, unlike, for example, release. Furthermore, intrusive vowels are distinguishable from epenthetic 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 *Nancy* [næntʃi]. But since an extra consonant in that position would not be expected to affect the stress pattern or anything else easily detectable, there is little way to tell from a description whether the [t] is epenthetic or intrusive.

### 1.5. Gestures and features

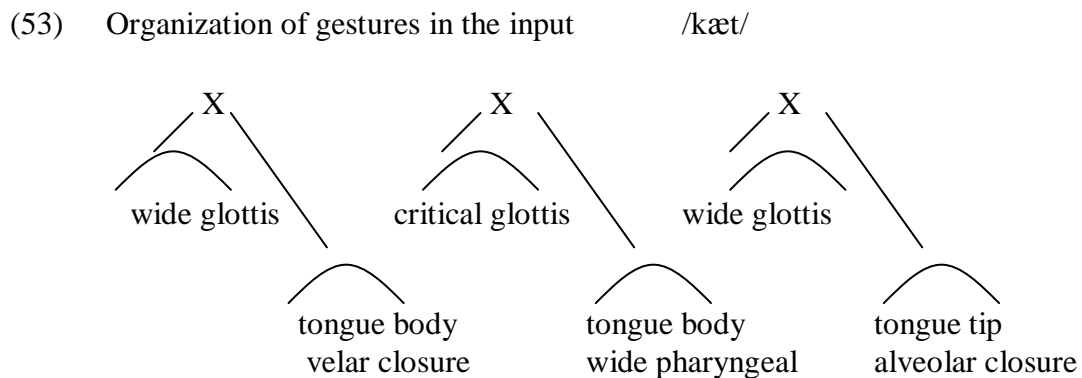
The above analysis has focused on the role of segment type and order in triggering the gestural phasing pattern that is heard as vowel intrusion. But some patterns can only be explained by reference to other types of structure. In particular, phenomena often analyzed as involving feature-sharing, such as geminates and assimilation

structures, affect the distribution of release. I will suggest that for two segments to share a feature is equivalent to sharing a gesture. To extend the theory to account for these cases, it is first necessary to clarify the relation between features and gestures.

Zsiga (1995, 1997) proposes that phonological representations include both features and gestures. She divides phonological processes into those that change features, and those that do not change features but involve gestural overlap. Zsiga 1995 examines cases of [s] → [ʃ] palatalization in English, and concludes that these are of two kinds: a feature-changing assimilation, as in [kənfeʃən] *confession*, and assimilation through gestural overlap as in [preʃju] *press you*. The gestural assimilation is revealed instrumentally to be more variable and incomplete.

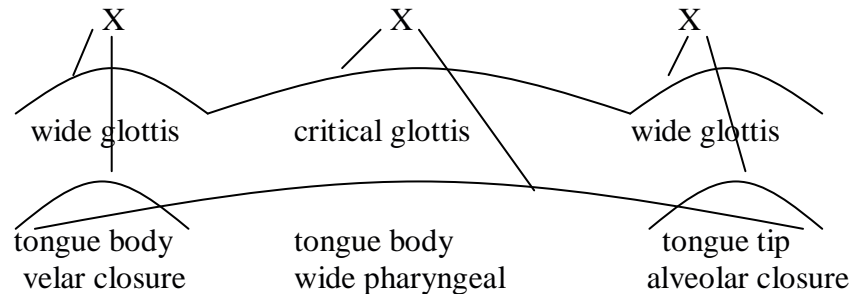
My view is similar in recognizing the need for distinguishing between categorical (“featural”) and gradient (“gestural”) effects. However, it is not necessary to have both a featural and a gestural representation, and a mapping between the two, in order to achieve this. Rather, gestures themselves can be associated with root nodes, in addition to having phasing relations with other gestures. I will assume that gestures are, in a sense, features enhanced with duration.

In TASP, there is an underlyingly association between root nodes and gestures. For example, the underlying representation of *cat* contains three root nodes, each associated with several gestures, shown below. These gestures do not have phasing relationships with one another specified, and their strength or stiffness is not specified either. (To be more precise, there are no faithfulness constraints preserving any underlying phasing, stiffness, etc., so its presence in the underlying representation, while not impossible, would be irrelevant.)



In the output candidates, gestures still are associated with root nodes, and in addition, phasing relationships and other characteristics such as stiffness of gestures are specified. Also, gestures that were present underlyingly may be absent (violating MAX), non-underlying gestures may be added (violating DEP) or gestures may be associated to root nodes they were not underlyingly associated with. The same gesture may be associated with two root nodes, or a single root node may be associated with multiple gestures for a single articulator.

(54) Organization of gestures in the output [kæt]



In the representation above, the gestures are shown in a tier-like organization, mostly for graphical convenience. However, as Browman & Goldstein 1990 point out, gestures do seem to interact in a somewhat tier-like fashion: consonant gestures are overlaid on a sequence of more or less continuous vowel articulations that minimally overlap with one another. I assume that this is because gestures must, for simple physical reasons, interact differently with one another depending on whether they use the same articulators. If two gestures make demands on the same organ, especially as a primary articulator, then there are probably more constraints against them heavily overlapping than if they used different articulators. For this reason it may make sense to think of, for example, a “tongue-body tier”, comprising a sequence of tongue-body articulations whose mutual phasing relations are subject to different considerations than the phasing relationships between tongue-body gestures and other gestures. It is not necessary to include tiers in the formal representation to capture this (special constraints on the phasing of same-articulator gestures will be enough), but the informal separation into tiers is intuitively helpful.

Thus far, TASP simply transfers the role of features to gestures. But there are important differences between the two approaches. Feature theory does not in itself make predictions about the relationship between featural structure and non-contrastive phonetic characteristics like release. But in TASP, predictions about the relation between, for example, categorical assimilation and release, fall out of the representation.

### 1.5.1. Gesture sharing

I assume that many kinds of assimilation result from two segments sharing a single gesture, such as [constricted glottis] or [tongue tip alveolar closure]. If two segments share a constriction gesture, it is not possible to have a release of that constriction within the cluster, so place assimilation should preclude release.

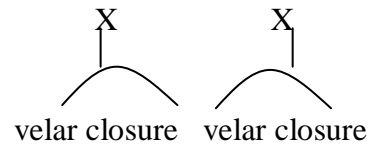
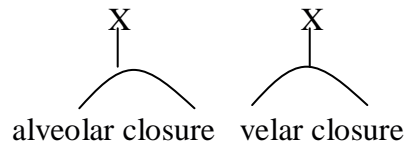
There is evidence that this is true. Dell & Elmedlaoui 1996:386 find that in Imdlawn Tashlhiyt Berber, release cannot happen between stops that are homorganic as a result of assimilation, although it can occur between stops that are underlyingly homorganic. Assimilation is the linking of a segment to a feature of an adjacent segment; so under the theory that features license gestures, assimilation necessarily involves a single gesture, which of course cannot contain a release and reconstruction.

(55) Imdlawn Tashlhiyt Berber

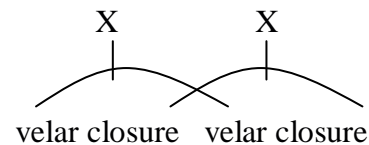
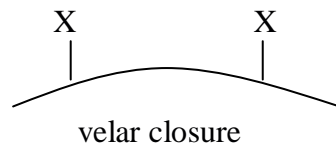
a. Assimilation structure:  
release impossible

b. Non-assimilation structure:  
release possible

Input:



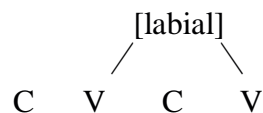
Output:



In general, release and assimilation do not occur in the same clusters in this language. If, in a sequence of stops such as /td/ or /kk<sup>w</sup>/, the first is released, there cannot be assimilation to the second in voicing or in secondary labiality.

Another type of evidence that feature-sharing is equivalent to gesture-sharing comes from Turkish vowel harmony. This is usually analyzed as the sharing of a single feature— such as [labial] for round vowels— by multiple segments.

(56) Autosegmental representation of vowel harmony



In Turkish, there is evidence (Boyce 1990) that a single lip-rounding gesture spans the word. In a word with multiple round vowels, the intervening consonants have some lip rounding as well. In English, where each round vowel bears its own separate [labial] feature, lip-rounding does not continue between two round vowels in a word. This is additional evidence that feature-sharing has phonetic as well as phonological implications.

The idea that adjacent segments can share a gesture helps to explain one of the cross-linguistically recurrent characteristics of vowel intrusion: its failure to occur in homorganic clusters.

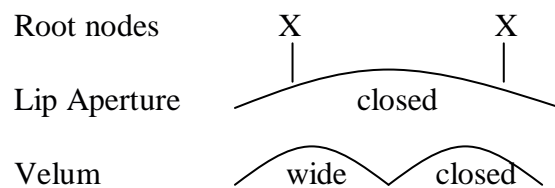
A partial explanation for this fact was already offered: we have seen that an alignment of CENTER = ONSET produces vowel intrusion in heterorganic but not homorganic RC clusters, and this can explain the heterorganicity condition on vowel intrusion in any language that uses this phasing. But this explanation cannot work for all cases, because not all clusters with vowel intrusion have the CENTER = ONSET phasing. In

some languages, intrusive vowels grow historically grow longer until they are as long as ordinary vowels. Sometimes they are eventually reanalyzed as regular vowels, as in Negev Bedouin Arabic, discussed under (51). In other languages, the sonorant apparently moves all the way across the vowel, resulting in metathesis. In either case, we would expect that as the overlap between consonants decreases, intrusive vowels would eventually develop in homorganic clusters as well. According to Gafos 2002, this will happen by the time the alignment  $OFFSET = ONSET$  is reached.

Yet in Scots Gaelic, for example, this has not happened. In Scots Gaelic  $VRVC$ , the intrusive vowel is now as long as or longer than the preceding vowel portion, as shown in Bosch & de Jong's 1997 phonetic study<sup>8</sup>. The sonorant is now around the center of the vowel gesture. Yet vowel intrusion is still triggered only by heterorganic clusters: it doesn't occur in  $[Lt]$ ,  $[mb]$ , etc. This suggests there is another reason that homorganic clusters don't have release.

I suggest that it is because they share a single oral constriction gesture. An  $[mb]$  sequence has just one labial closure gesture. Therefore, C-C phasing constraints cannot possibly cause a release within  $[mb]$ .

(57) Homorganic cluster



There is no reason to believe that the  $/m/$  in Scots Gaelic  $/mp/$  clusters has undergone assimilation ( $[np]$  clusters occur as well, but with vowel intrusion), so there must be a constraint forcing gesture-sharing in clusters that are underlyingly homorganic. This can be accomplished with a gestural version of the Obligatory Contour Principle, similar to that proposed by Gafos 2002:295: identical gestures cannot overlap.

(58) OCP-GESTURE

Overlapping identical oral gestures are prohibited.

When two segments with identical oral gestures are adjacent, one way of eliminating this overlap is for the segments to share a single long oral gesture.

Since gesture-sharing structures do not seem to be chosen universally, OCP-GESTURE must compete with another constraint. I propose that there is a general preference for a one-to-one mapping between features and gestures.

<sup>8</sup> I argue in chapter 4 that Scots Gaelic vowel intrusion synchronically involves a different type of structure than the other languages discussed here: the vowel and sonorant now seem to form a complex nucleus. The following discussion, therefore, may be accurate only for an earlier stage of the development of Scots Gaelic.

(59) \*MULTIPLE LINKING

A single gesture is not associated with more than one segment.

The ranking between these two constraints determine whether a language chooses gesture-sharing structures or sequences of identical gestures. In any language where OCP- GESTURE is ranked above \*MULTIPLE LINKING, homorganic clusters will involve gesture sharing, and release will be impossible. In a language that has the opposite ranking, sequences of identical gestures will be permitted, and homorganic clusters will potentially have releases, if the phasing constraints permit.

At this point we have two overlapping explanations of the lack of vowel intrusion in homorganic RC clusters: task dynamics and gesture sharing. Both seem needed. There are some languages where the gesture-sharing account is the only possible one: task dynamics cannot explain the lack of release in homorganic clusters in languages that ordinarily phase clusters to have little overlap. But there are other languages that demonstrate only a tendency not to have vowel intrusion in homorganic clusters, without the ban being absolute. This is the case in Saami, where [l:s] and [l:n] clusters can have vowel intrusion but not [l:t].

For Finnish, (Harms 1976) reports that homorganic clusters don't have a release but do have some kind of decrease of "energy" between the consonants. This is consistent with the idea that homorganic Finnish clusters involve two separate gestures with a low degree of overlap, and that the nonrelease is due only to task dynamics, not gesture sharing.

### 1.5.2. Homorganic CR clusters

The examples of vowel intrusion being blocked in homorganic clusters above all involve RC clusters; the list yields no clear-cut examples where vowel intrusion in CR clusters is subject to the same restriction. Hocank, for example, breaks up the homorganic CR clusters [sr] and [sn].

A striking illustration of the difference between RC and CR with respect to homorganicity restrictions comes from ancient Oscan, which developed copy vowels in both cluster types. Intrusion does not occur into the homorganic RC clusters [rt], [rn], [rs], and [lt], but it does occur in the homorganic CR clusters [tr] and [dr]. The heterorganicity requirement depends on the direction of consonant contact in Oscan. In Scots Gaelic, also, a few dialects have also developed vowel intrusion in [dl] and [dr] clusters but none in [ld] or [rd] clusters (Borgström 1941:90).

The vowel intrusion in homorganic CR onsets shows that these sequences do not involve gesture-sharing. This is reminiscent of the fact that place assimilation happens more in RC than CR clusters.

There is a functional reason that gesture sharing should be more marked in [tr] cluster than in an [rt] cluster. The perceptibility of [t] depends heavily upon its release burst, since the period of constriction itself is silent. An [r], on the other hand, is audible during its constriction period. Since gesture-sharing eliminates release in a cluster, it damages the perceptibility of [tr] far more than of [rt]. (In a [tr] cluster, it is not strictly correct to say that there is no release: the constriction degree must be relaxed slightly for the [r]. Clusters that involve the same constriction location but different degrees must

have a more complex type of shared gesture than that of an [nt] cluster. But such a slight opening of the constriction will produce less burst than the normal, controlled movement away from the constriction.)

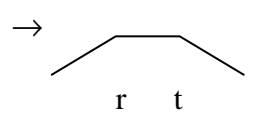
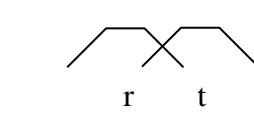
This can be captured with the following constraint:

(60) \*MULTIPLE LINKING: STOP-C

A single head gesture is not associated with a stop and a following consonant.

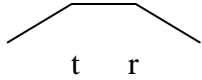
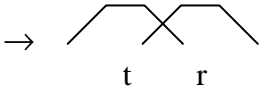
When high-ranked, this constraint will prevent a language such as Oscan from creating a gesture-sharing structure for [tr], as shown below. When an input contains an RC cluster, gesture-sharing is preferred, as in candidate a) below. Candidate b) is eliminated because it involves two alveolar constrictions overlapping.

(61) RC clusters

/rt/	*MULTIPLE LINKING: STOP-C	OCP-GESTURE	*MULTIPLE LINKING
a.  (no release possible)			*
b.  (release possible)		*!	

When the input contains a CR cluster, however, the higher ranked \* MULTIPLE LINKING: STOP-C eliminates candidate a), and candidate b), with overlapping gestures, is permitted to surface, as shown below.

## (62) CR clusters

/tr/	* MULTIPLE LINKING: STOP-C	OCP-GESTURE	*MULTIPLE LINKING
a.  (no release possible)	*!		*
b.  (release possible, depending on phasing)		*	

In this way, \* MULTIPLE LINKING: STOP-C can produce languages where the heterorganicity restriction holds for RC clusters but not CR.

### 1.5.3. Gestures and levels of grammar

A further question about the phonological role of gestures is whether they belong to a separate “phonetic implementation” level of the grammar. It is often assumed that one level of the grammar determines the arrangement of segments, features, syllables, etc., and then passes this output to a separate module that determines the phonetic realization of the structure. Zsiga 1997:270, for example, proposes that “segments (root nodes) are relevant for phonological alternations, but cease to be relevant when features are mapped into gestures.”

In TASP, there are no separate levels of the grammar for constraints on the associations of gestures to root nodes (the equivalent of Zsiga’s “phonological alternations”) and constraints on gestural phasing. This unity allows a more straightforward modeling of the ways in which phonological alternations are guided by a desire to create a good gestural alignment in the output.

For example, the asymmetry between release in RC and CR clusters above could be analyzed as caused by a constraint against gesture-sharing in CR clusters (as argued above), or by a constraint against feature-sharing in CR clusters. The non-shared features would be mapped to separate gestures by the phonetic component, and the end result would be the same.

However, the split-level approach relies on a featural constraint whose functional purpose is precisely to block an undesirable gestural alignment. The set of featural constraints is grounded in a knowledge of what the gestural results of each featural structure will be. This seems like an unnecessary complication: it is simpler to have gestural constraints refer to gestural alignment, and constraints on other structures, such as the branching of association lines, refer strictly to those structures.

#### 1.5.4. Geminates

Another case in which gesture-sharing structures influence vowel intrusion is the special behavior of geminates, which I assume also involve a single gesture associated with two root nodes. Geminates trigger vowel intrusion more than singletons: Saami has vowel intrusion in R:C but not RC clusters, and some dialects of Finnish have vowel intrusion in RC: clusters but not RC or RCC clusters.

There is a general pattern of languages preferring releases before and after geminates. Imdlawn Tashlhiyt Berber allows releases in  $C_iC_i$  or  $C_iC_i$  clusters but not in most  $C_iC_i$  clusters (Dell & Elmedlaoui 1996). Word-final geminate stops in Amharic are always obligatorily released, while in singletons release is optional in some environments (Hudson 1995:664). Stefania Marin reports similar facts for final geminates and singletons in Wolof (p.c.). Wolof is also described as inserting a [ə], which may be simply a release, between a geminate and a consonant-initial suffix (Ka 1994:105). It is obvious why release is especially important perceptually for geminates: without an audible release it would be difficult to perceive the length of closure, and hence to tell the difference between geminates and singletons.

I propose that there are specific phasing constraints for clusters that include geminates, such as that below.

(63) ALIGN ( $C_i$ , CENTER, C, ONSET)

In a  $C_1 C_2$  sequence, where  $C_1$  is a geminate, the center of  $C_1$  is aligned with the onset of  $C_2$ .

(64) ALIGN (C, CENTER,  $C_i$ , ONSET)

In a  $C_1 C_2$  sequence, where  $C_2$  is a geminate, the center of  $C_1$  is aligned with the onset of  $C_2$ .

The existence of a specific constraint on geminate phasing in addition to the general constraints on CC phasing predicts that in some languages geminates will show a greater tendency to allow vowel intrusion.

#### 1.6. Conclusion

This chapter highlights some of the cross-linguistically common properties of intrusive vowels. I have defended Steriade 1990's proposal that vowel intrusion involves gestural overlap, but claimed furthermore that it is unlike epenthesis in that it does not add a segment or syllable to the word. I have also discussed the implications that this phenomenon has for our understanding of gestural overlap.

1. It supports the idea that gestural coordination constraints on onsets and codas are not symmetrical, since intrusive vowels can appear in one without appearing in the other.
2. It shows that, in a language that has vowel intrusion in a  $VC_1C_2$  sequence, there must be phasing relationships between V and  $C_2$ , unlike in Browman & Goldstein's

model of English. This suggests that the organization of gestures within a syllable is not uniform cross-linguistically.

3. It shows that some consonants are more likely to allow or give rise to this configuration than others, and that there is an implicational hierarchy cross-linguistically as to which consonants trigger it.

4. It suggests that V / C overlap is less marked if V and C are homorganic.

5. It provides evidence against underlying gestural phasing relationships, since intrusive vowels, the result of a particular gestural coordination, are non-contrastive.

Chapter 2 contains more detailed case studies of languages with vowel intrusion, concentrating on ways syllable structure affects vowel intrusion and ways in which intrusive vowels act non-syllabic.

## APPENDIX

### Examples of vowel intrusion conditioning environments

	clusters with intrusive vowels	RC or CR clusters surfacing without intrusive vowels	examples of other clusters
Dakota	bl, gl, gn, gm	sl, fl, xl, sn, fn, xn, sm, fm, xm	kp, pt, sk, kf, pc, xt
Dutch	lm rm lp rp rf lf lk rk lg rg rn	rt rs lt nt ls ns	kt
Finnish	r / l + C; h + voiced C hv hj hm hn hl hr lk lv lj lh lm lp rp rk rv rj rh rm nh	hk ht lt ls ll mm mp nn ns nt nk nj rn rr rt rs	kl kr ks ps pl pr tk tr tj sk sl sm sp sv tt ts
Hausa	r / l + C. examples given: rh lb lk rk rm lm	rC mC nC wC yC ?C (many examples of each)	sk pt kw bd fs br y?
Hocank	pn pr kn kr kw sn sr sw fn fr fw tw xn xr xw	C?	st ps kdʒ kf tʃg xg t?
Irish Gaelic I (diachronic)	l / r / n + fricative or voiced stop lb nb r'b' l'v' n'v' r'v' nf g'm' n'm' rm nx rx lg rg lf rf rn	rk lk rp lp mb nd ng nr lr	
Irish Gaelic II	plosive / m / f / s / h + nasal / liquid gl g'n' dr bn m'n' mr f'r' hn' sr (not nec. complete)		
Lakhota	bl gl gm gn mn		
Oscan	RC: lp rf lf rv lv rm rg rk lk nf ng nk rh	lt rt rn rs nd nt ns	

	CR: kr kl tr dr fr tn kn fn mn		
Saami	r:k r:j r:m r:n r:n β:n l:j l:f l:p l:m l:n l:n l:s l:f l:v	lf lk lp rf rk rp l:t l:ts̄ l:tʃ̄ h:p h:t h:ts̄ h:tʃ̄ h:k <sup>9</sup>	s:k s:t s:m s:p ʃ:m ʃ:t
Sanskrit	r + fricative (some authorities) r + any C (some authorities) l / r + fricative (some authorities)		
Scots Gaelic	mr mn mʃ mç Nk Nç Nx np nv nç nx nm Lb Lp Lg Lv Lx Lm lg' lv lm lj L'x Rg' Rk rb rg rv rf rʃ rx rm r'g' r'v r'ç r'x r'm r'j	<u>RR</u> : jh Rh rh r'h Lh lh mr nr Rn Rl <u>RC</u> : jp rb rp Rd Rt r't' r'k' Lp Lt L't' lk' Lk mb mp Nd Nt N'd' N't' N'g' Ng Nk rf Rs rx lf Ns N'ʃ	see full chart in chapter 3
Spanish	all rC Cr some possibly historical examples with [l]	CR: pl bl fl kl gl gn RC: mp ns nt lm lv lt	st sp bs kt

<sup>9</sup> These are analyzed by Bye 2001 as superheavy preaspirated stops rather than clusters.