2006. In Optimality-Theoretic Studies in Spanish Phonology, ed. by Fernando Martínez-Gil and Sonia Colina, 15-38. Amsterdam: John Benjamins.

Spanish complex onsets and the phonetics–phonology interface*

Travis G. Bradley
University of California, Davis

This study analyzes cross-dialectal phonetic variation in Spanish complex onsets in light of recent work on the phonetics–phonology interface. Two basic patterns of obstruent-rhotic cluster realization, vowel intrusion and coarticulation-induced rhotic assibilation, receive a phonetically-motivated explanation in terms of the temporal coordination of consonantal gestures, within the framework of Articulatory Phonology (Browman and Goldstein 1989, 1990, 1991, et seq.). Drawing upon recent developments in gestural Optimality Theory (Davidson 2003, Gafos 2002, Hall 2003), I propose an account in which the interaction among gestural alignment constraints generates the range of attested patterns. On the basis of stress restrictions, non-concatenative morphology, the universal non-contrastiveness of intersegmental gestural coordination, and sonority conditions on complex onsets, I show that vowel intrusion and rhotic assibilation are invisible to phonological processes that operate over segments and syllables. In contrast to theories which relegate gestural timing to a low-level phonetic implementation component, this study argues for a unified model in which gestural and non-gestural constraints are present in the same level of the phonology (Hall 2003). Once a *representational* distinction is made between segments and gestures in the phonological representation, a *derivational* mapping between phonology and phonetics becomes unnecessary to account for the invisibility of gestural percepts.

* Ealier versions of this work were presented at the *XXXII Linguistic Symposium on Romance Languages* (University of Toronto, April 2002) and, with Benjamin Schmeiser, at the *6th Hispanic Linguistics Symposium* (University of Iowa, October 2002). For feedback and discussion, I am grateful to Gary Baker, Eric Bakovic, Jill Beckman, Barbara Bullock, Adamantios Gafos, Nancy Hall, David Heap, Haike Jacobs, Joan Mascaró, John McCarthy, Jean-Pierre Montreuil, Carlos Eduardo Piñeros, Donna Rogers, Benjamin Schmeiser, Bart van der Veer, and Caroline Wiltshire. Thanks also to two anonymous reviewers for their constructive comments. I alone am responsible for any remaining shortcomings in the present work. This research was partially supported by funding from the Division of Humanities, Arts, and Cultural Studies (College of Letters and Science, UC Davis) and by a Faculty Research Grant, "Phonetic Variation in The Cross-dialectal Realizations of Spanish Tap and Trill" (Academic Senate, UC Davis).

Keywords: gestural coordination, phonetics–phonology interface, rhotics, vowel intrusion, assibilation

0. Introduction

In Articulatory Phonology (Browman and Goldstein 1989, 1990, 1991, et seq.), the grammar is assumed to operate on articulatory gestures, which are dynamically defined along both spatial and temporal dimensions and produce a constriction in the vocal tract. Many alternations that have previously been explained in discrete, phonological terms can be analyzed in terms of gestural overlap and/or reduction in casual speech. However, the status of gestural representations in the synchronic grammar remains controversial. Should gestures be phonological primitives as well as units of articulation, or is Articulatory Phonology better viewed as a model of phonetic implementation? If gestures are primitives, should they supplant segments or coexist with them? Should the temporal coordination of gestures be specified in underlying representation, or should it be determined by the grammar?

Phonetic variation in Spanish complex onsets — in particular, /Cɾ/ — provides a fertile empirical terrain in which to explore these theoretical issues. Rhotics are known for the considerable phonetic variety they exhibit across languages, dialects, and speech styles. Nonetheless, a common trend among contemporary generative studies of Spanish has been to gloss over what are deemed to be irrelevant, low-level details of phonetic implementation. Such a move is taken, for instance, by Harris (1983: 62), who reduces the "astonishing variety of *r*-quality phones … to just two … which will be understood to jointly exhaust the rich phonetic variety […] I will say little more about phonetic detail…" However, an investigation of phonetic detail is important to uncovering systematic aspects in the patterning of rhotics, which ultimately bears on the status of gestural representations and constraints in the grammar. Recent studies on Spanish rhotics have begun to redress the lack of attention given to phonetic detail (e.g., Blecua 2001, Bradley 1999, 2001a, 2004, Bradley and Schmeiser 2003, Colantoni 2001, Colantoni and Steele 2005, Hammond 1999, 2000, Hualde 2004, Lewis 2004, Schmeiser 2004, Willis 2005, and Willis and Pedrosa 1999). The present work contributes to this line of research by investigating the role of gestural coordination in Spanish complex onsets.

There are two basic patterns of /Cɾ/ cluster realization across Spanish dialects. Vowel intrusion involves the appearance of a vowel-like fragment between the two consonants. Coarticulation entails the absence of such fragments, with concomitant assibilation of the rhotic and gradient laryngeal and/or place accommodation of the cluster. Drawing upon recent developments in gestural Optimality Theory (Davidson 2003, Gafos 2002, Hall 2003), I propose an account of these patterns in terms of differences in the temporal coordination of gestures, which are controlled by gestural alignment constraints. On the basis of stress restrictions, non-concatenative morphology, the universal non-contrastiveness of intersegmental gestural coordination, and sonority conditions on complex onsets, I show that vowel intrusion and rhotic assibilation

are invisible to phonological processes that operate over segments and syllables. Invisibility suggests that gestural coordination is a low-level phenomena of phonetic implementation, which can often produce mismatches with the phonological representation. However, I show that a *derivational* relationship between phonology and phonetic implementation is unnecessary once a *representational* distinction is made between segments and gestures (Hall 2003). In a unified model, gestural alignment constraints determine intersegmental gestural timing, while non-gestural constraints make reference to segments, syllables, and other aspects of prosodic structure. The universal non-contrastiveness of gestural coordination suggests that UG contains no faithfulness constraints on input timing.

This paper is organized as follows. Section 1 documents patterns of phonetic variation in the realization of complex onsets in Spanish. Section 2 develops an account of vowel intrusion in terms of competing gestural coordination constraints, and Section 3 extends the analysis to coarticulation. Section 4 discusses implications for the phonetics–phonology interface. Section 5 compares the proposed analysis with previous ones, and Section 6 summarizes and concludes.

1. Phonetic patterns of Spanish complex onsets

In Spanish, complex onsets consist of an obstruent /p, t̪, k, b, d̪, g, f/ followed by a liquid /l/ or /ɾ/, except for /d̪l/ and, depending on the dialect, /t̪l/ (Harris 1983: 13–14, 20–22, 31–35, Harris and Kaisse 1999: 125, Hualde 1991: 481–483, 1999: 171–172). It has long been noted in the Spanish phonetic literature that the apicoalveolar tap /ɾ/ is usually accompanied by a vocalic element appearing between the rhotic and a preceding or following consonant (Gili Gaya 1921, Lenz 1892, Malmberg 1965, Navarro Tomás 1918, Quilis 1970). Researchers have given a variety of descriptive labels to the vowel fragment, including *svarabhakti, transitional, parasitic, epenthetic,* etc. Adopting terminology from Hall's (2003) cross-linguistic study, I henceforth refer to this phenomenon as *vowel intrusion* and to the fragments themselves as *intrusive vowels.* The reason for this terminological choice is to distinguish vowel intrusion from true epenthesis of a phonological vowel, a distinction that will be motivated in Section 4. The examples in (1) illustrate vowel intrusion in complex onsets.

(1)

pronto	[pᵊɾ]	'soon'
fresco	[fᵊɾ]	'fresh'
tres	[t̪ᵊɾ]	'three'
gracias	[gᵊɾ]	'thanks'

Although represented here in narrow phonetic transcription as a superscript schwa [ᵊ], the intrusive vowel typically has formant structure similar, but not identical, to that of the nuclear vowel appearing on the opposite side of the tap constriction (Quilis 1993: 337–342).

One of the striking characteristics of intrusive vowels in Spanish is their variable duration, as was pointed out early on in phonetic studies based on kymographic inscriptions.

For example, Gili Gaya (1921) made the following observation based on measurements of /Cɾ/ tokens from words pronounced by speakers of Peninsular Spanish:

> The duration of the intervening vocalic element is highly variable even in the same word repeated several times by the same individual. This variability probably stems from rate of speech and from the fact that speakers are unaware of the existence of this vowel fragment, even though in most cases it attains a duration greater than that of the *r* (Gili Gaya 1921: 278–279).

In a later study of Spanish /ɾ/, Malmberg (1965: 10, 35) observed that the duration of the intrusive vowel often approximates that of an unstressed vowel. In fact, the intrusive vowel has occasionally given rise to a lexicalized anaptyctic vowel which copies the nucleus that is tautosyllabic with the complex onset, as shown by the diachronic examples in (2) (Gili Gaya 1921: 280, Quilis 1988: 300).

(2) *peréces* < *preces* 'prayers'
tarabilla < *trabilla* 'stirrup'
corónica < *crónica* 'chronicle'
chácara < *chacra* 'farm'
gurúpa < *grupa* 'hindquarters'
tíguere < *tigre* 'tiger'

Quilis (1970) reports similar variability in the duration of the intrusive vowel appearing in /Cɾ/ clusters, which ranges from 8 ms to 56 ms. The mean duration for intrusive vowels is 29 ms versus 20 ms for the tap constriction. In an acoustic study of Castilian Spanish, Blecua (2001) finds that the mean duration of the intrusive vowel in /Cɾ/ is significantly longer than that of the tap constriction itself (27.9 ms versus 20.5 ms). The standard deviation is larger for the mean duration of the intrusive vowel than for that of the tap constriction (9 versus 5.4), which indicates greater variability in the former. Both of these results agree with the findings of Gili Gaya (1921) and Quilis (1970). In a more recent experimental study, Colantoni and Steele (2005) report variable duration of intrusive vowels in Buenos Aires Spanish, ranging from 20 ms to 47 ms. Therefore, vowel intrusion in clusters containing /ɾ/ is not limited to Peninsular Spanish varieties. The phenomenon seems to be quite pervasive across dialects and is also found with /ɾ/ and other sonorants in other languages (see the discussion in Section 2).[1]

1. There are reports, often conflicting, of other aspects of phonetic detail in the realization of Spanish /Cɾ/ onset clusters. The measurements of Gili Gaya (1921) suggest that the intrusive vowel is longer when the /CɾV/ demisyllable is word-initial or stressed and when the initial consonant of the cluster is dorsal. Blecua (2001) finds longer intrusive vowels after voiced consonants and after dorsals. Based on data from Madrid Spanish, Schmeiser (2004) corroborates Blecua's findings, but fails to find any significant effects for word position or stress. For Buenos Aires Spanish, Colantoni and Steele (2005) document significantly longer intrusive vowels after voiced consonants, after dorsals, and in stressed demisyllables. However, they report that word-medial clusters show longer vowels than word-initial ones. Given the variable nature of the phenomenon, discrepancies among reported findings are not unexpected and can be plausibly attributed to differences among speakers and/or dialects. To attempt a comprehensive analysis of all of these patterns is beyond the goals of this paper, but see Colantoni and Steele (2005) for an in-depth discussion of the phonetic factors involved in obstruent-liquid cluster realization.

In contrast to /Cɾ/, onset clusters containing an obstruent followed by a lateral typically do not exhibit vowel intrusion in Spanish. Previous phonetic studies provide ample evidence and description of the intrusive vowel in /Cɾ/, but /Cl/ clusters are remarkably absent from such discussions. Some researchers have proposed that the intrusive vowel is part of the definition of Spanish /ɾ/, which underscores its failure to appear with /l/. For instance, Gili Gaya (1921: 279) states that the rhotic "is a vocalic sound interrupted by an alveolar contact that is voiced and more or less tense." Furthermore, there is an asymmetry in the number of historical examples showing the development of anaptyctic vowels in Spanish onset clusters. The majority of examples cited by Malmberg (1965) and Quilis (1988) involve /Cɾ/, as shown in (2). In contrast, the only given example containing /Cl/ is *Ingalaterra* < *Inglaterra* 'England'. With respect to Buenos Aires Spanish, Colantoni and Steele (2005) demonstrate that intrusive vowels are almost categorically absent in /Cl/ clusters, occurring at a rate of less than 2% in their sample.

The waveforms and spectrograms in Figure 1 provide a visual illustration of obstruent-liquid cluster realizations, with durational measurements indicated in milliseconds. The acoustic examples presented in this paper are based on data from a single female speaker from La Paz, Bolivia, and are thus intended to provide a representative

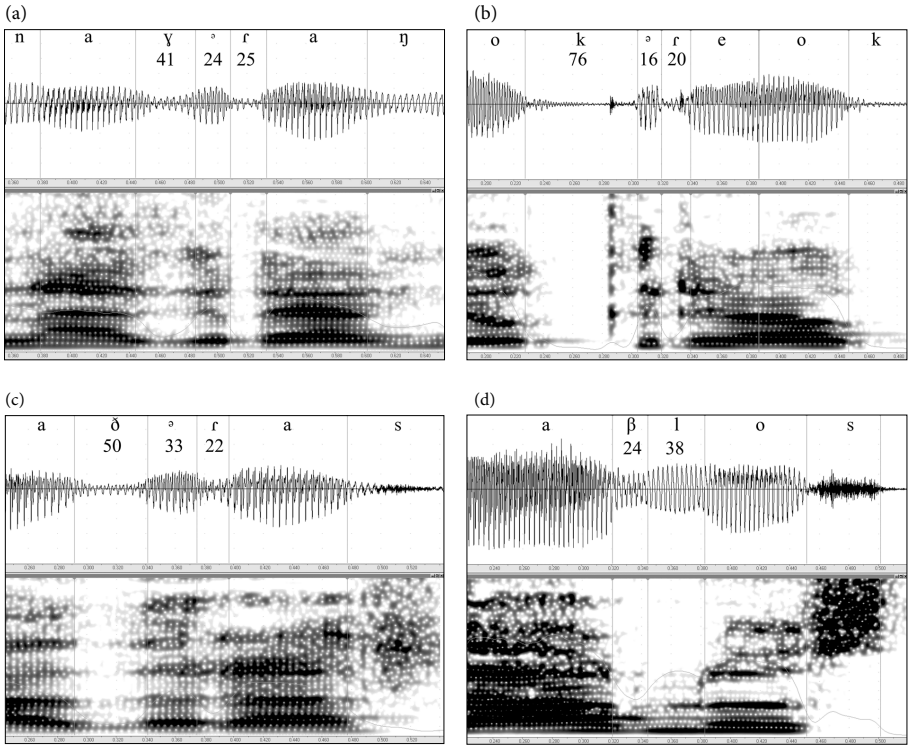


Figure 1. Vowel intrusion in *una gran cantidad* 'a great quantity' (a), *yo creo que* 'I believe that' (b), and *cuadras* 'stables' (c) versus no vowel intrusion in *establos* 'cow sheds' (d)

description of the phenomena under discussion.[2] In (a-c), [ɾ] is phonetically separated from the preceding obstruent by an intrusive vowel of variable duration, in which some trace of formant structure is recoverable from the nuclear vowel opposite the tap constriction.[3] On the other hand, the [βl] cluster in (d) is realized as a contiguous sequence with no intervening vowel fragment.

The discussion of vowel intrusion thus far suggests that /Cɾ/ clusters are realized along an intersegmental duration continuum as a function of the relative phonetic separation between the consonants of the cluster. At the other end of this continuum are cases of what I shall refer to as *coarticulation,* exemplified in (3). These examples come from Alonso's (1925) descriptive study of Peninsular Spanish varieties spoken near Álava, Navarra, Rioja, and Aragón.

(3)	*apretar*	[pɻ̥]	'to squeeze'
	hombre	[bɹ]	'man'
	otro	[tɻ̥]	'other'
	vendrá	[dɹ]	's/he will come'
	padre	[ðɹ]	'father'
	escribir	[kɻ̥]	'to write'
	magras	[ɣɹ]	'lean (FEM PL)'

As suggested by these narrow transcriptions, coarticulation entails some frication, or assibilation, of the rhotic and the loss of both the intervening intrusive vowel and the extra-short constriction period of apicoalveolar /ɾ/. Rhotics are subject to partial devoicing when coarticulated with a preceding voiceless consonant (e.g., [pɻ̥] versus [pʰɾ]). In addition, dental /t̪/ and /d̪/ accommodate the constriction location of the rhotic, yielding an alveolar quasi-affricate realization (e.g., [tɻ̥] versus [t̪ʰɾ]).[4] The

2. The Bolivian data come from a larger corpus of Spanish fieldwork material consisting of recorded interviews and readings of literary texts of varying lengths, constructed through consultation with native speaker informants from over 25 different regions throughout the Spanish-speaking world. Informants were initially recorded on reel-to-reel tape, and the recordings were later digitized and stored on CD-ROM in MPEG format at 22,050 Hz and 16-bit. The fieldwork corpus was made available by John Dalbor at the Pennsylvania State University and subsequently digitized under the supervision of Eric Bakovic at the University of California, San Diego. The tokens presented in this paper were extracted from the corpus recordings, converted to WAV format, and analyzed with version 2.6 of the Summer Institute of Linguistics Speech Analyzer software package.

3. Spanish exhibits a surface alternation in voiced obstruents between stops and continuants. In most dialects, the stops [b, d̪, g] surface after nasal, pause, and in the case of [d̪], after /l/, while the continuants [β, ð, ɣ] appear elsewhere (see Martínez-Gil 2001 and the references cited therein).

4. An anonymous reviewer points out that since coarticulation processes involve some degree of *assimilation* to neighboring segments, the term *coarticulation* is not completely felicitous as a descriptive label for all of the cluster realizations shown in (3). While progressive rhotic devoicing and regressive dental stop retraction do constitute assimilatory effects, the same cannot be said for [bɹ] or [ɣɹ], which show loss of the intrusive vowel but no devoicing or place accommodation. As I argue in Section 3, however, all of these realizations receive a unified explanation in terms of greater overlap of articulatory gestures. Therefore, I continue to employ the term *coarticulation* as a unifying label in order to emphasize the articulatory basis of the phenomenon.

articulatory descriptions provided by Alonso (1925) and Malmberg (1965) are particularly revealing on all of these points:

> The *r* combines with the consonants with which it groups, without any epenthetic vocalic element (Alonso 1925: 185).
>
> The *r* tends to be formed during the articulation of the preceding voiceless stop, invading its release, letting itself in turn be invaded by the voicelessness of the release … I have heard in speakers from diverse regions of the Peninsula the same fusion in moments of physical fatigue, when speaking casually or in a low voice (Alonso 1925: 186, 189).
>
> This tendency of the consonant *r* to combine with a dental to form a new consonant, which is generally a compromise between the two, is not unknown in other languages (Malmberg 1965: 39).
>
> Careful speech allows the identity of the sounds to be recovered … Careful speech is sufficient to ensure greater intelligibility by isolating the elements of the consonant group (Alonso 1925: 186–187).

Furthermore, realizations of /Cɾ/ clusters are dependent upon speech style, as per Alonso's observations that casual speech favors coarticulation while careful speech enhances recoverability.

Lipski (1994: 320) points out that in highland Peru, "pronunciation of the groups /t̪ɾ/, /pɾ/, /kɾ/ is partly determined by ethnolinguistic background. Among bilingual speakers, the /ɾ/ in these combinations is a fricative or retroflex approximant, and in the case of /t̪ɾ/ may fuse with the preceding consonant to produce a quasi-affricate." Lipski's description of the Peruvian pattern mirrors that of Alonso (1925) for Peninsular Spanish in that coarticulation may affect /Cɾ/ clusters regardless of the place specification of C_1. However, other Latin American varieties appear to limit coarticulation specifically to homorganic clusters in which C_1 is a coronal stop. Representative data in (4) are based on Argüello's (1978) study of highland Ecuadorian Spanish.[5]

(4)	a.	*tres*	[t̪ɻ̪]	'three'
		cuatro	[t̪ɻ̪]	'four'
	b.	*vendrá*	[nᵈɹ]	's/he will come'
		saldrá	[lᵈɹ]	's/he will leave'
	c.	*padre*	[ðᵊɾ]	'father'
	d.	*premio*	[pᵊɾ]	'prize'
		cruz	[kᵊɾ]	'cross'

In casual speech, coarticulation affects clusters such as those in (4a,b), where the preceding coronal is realized as noncontinuant. In (4c), however, the voiced coronal surfaces as a continuant after a preceding vowel, and the underlying cluster surfaces intact

5. In accordance with the convention of Hispanic linguistics, Argüello employs [ɾ̆] and [ɾ̥̆] to represent voiced and voiceless variants, respectively, of the *r asibilada* (assibilated/fricative *r*). For consistency, I continue to employ Alonso's transcription of the coarticulated rhotic as [ɹ] and [ɻ̥] in (4a,b), and I also indicate the lack of coarticulation by transcribing the intrusive vowel [ᵊ] in (4c,d).

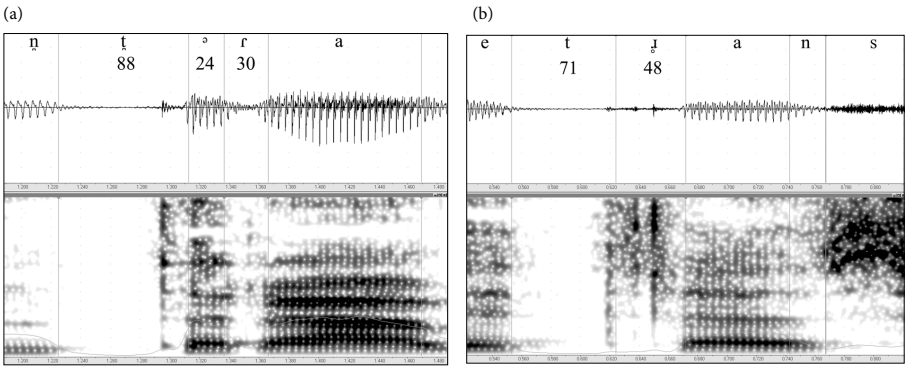


Figure 2. Vowel intrusion in *encontraba* 'found' (a) versus coarticulation in *se transforma* 'changes' (b)

(cf. example (c) of Figure 1). Coarticulation also fails to affect heterorganic clusters, as in (4d). Furthermore, Lipski (1994) documents a similar pattern for other geographic zones, namely Northern interior Argentina (p. 172), highland Bolivia (p. 189), Chile (p. 200), Colombia (pp. 209–210), Central Costa Rica (p. 222), Guatemala (p. 265), Honduras (p. 272), Mexico (p. 279), and Paraguay (p. 308). Available phonetic descriptions suggest that coarticulation between /ɾ/ and a preceding homorganic stop is widespread across these varieties, but similar behavior involving heterorganic C_1 is presumably unattested.

The empirical generalization emerging here suggests an implicational relationship between two types of /Cɾ/ coarticulation in casual speech across Spanish dialects. In Peninsular and Peruvian varieties, coarticulation affects potentially any /Cɾ/ cluster in casual speech, while in other Latin American varieties, it is restricted to clusters in which C_1 is a coronal noncontinuant. For a given dialect, coarticulation of heterorganic clusters entails coarticulation of homorganic ones (with noncontinuant C_1), but the opposite does not hold, as evidenced by the data in (4).

The data in Figure 2 exemplify the variable coarticulation of /t̪ɾ/ clusters, taken from the same highland Bolivian Spanish informant who produced the tokens in Figure 1. In (a), glottal tone and some formant structure are present during the tap constriction following the intrusive vowel. In contrast, vowel intrusion is absent under coarticulation in (b). The rhotic corresponds to a 48 ms period of strident frication, whose turbulence is indicated by the presence of aperiodic energy in the upper spectra. Taken together with (a-c) of Figure 1, these representative data provide some empirical support for the implicational relationship between coarticulation in homorganic and heterorganic clusters. For the Bolivian informant, coarticulation variably affects /t̪ɾ/, whereas other clusters typically exhibit vowel intrusion.

The review of previous phonetic studies yields several generalizations regarding phonetic variation in onset clusters across Spanish dialects:

(5) a. An intrusive vowel of variable duration typically occurs in /CɾV/ but not in /ClV/.
b. The formant structure of the intrusive vowel in /CɾV/ demisyllables is similar, but not identical, to that of the tautosyllabic nuclear vowel.

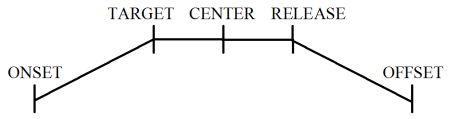
c. Anaptyctic vowels that arise historically from /CɾV/ copy the quality of the tautosyllabic nuclear vowel.
d. Coarticulation in casual speech of heterorganic /Cɾ/ in a given dialect entails coarticulation of homorganic /Cɾ/ (where C_1 is noncontinuant) but not vice-versa.

In the following sections, I propose an account of the generalizations in (5) in which the coordination of articulatory gestures is determined by the interaction of Optimality-theoretic alignment constraints.

2. Vowel intrusion and historical copy vowels

Articulatory Phonology provides an explanatory phonetic account of vowel intrusion in terms of the temporal coordination of consonant and vowel gestures. According to Steriade (1990), vowel intrusion results when an overlapping vowel gesture is heard during the open transition between two consonants. Languages vary systematically in the classes of consonants triggering vowel intrusion (Hall 2003). Such variation can be captured in the framework developed by Gafos (2002), which incorporates the gestural representations of Articulatory Phonology within a constraint-based, Optimality-theoretic grammar (Prince and Smolensky 1993, McCarthy and Prince 1993). Gafos proposes that gestural coordination is determined by alignment constraints of the form (6a), which make reference to temporal landmarks during the activation period of a gesture, shown in (6b):

(6) a. ALIGN(G_1, LANDMARK$_1$, G_2, LANDMARK$_2$)
Align landmark$_1$ of gesture$_1$ with landmark$_2$ of gesture$_2$.
b.



Researchers adopting this framework have posited coordination relations for CV, VC, CC, and VV sequences (Davidson 2003, Gafos 2002, Hall 2003). I propose that an analysis of Spanish onset clusters requires several CC alignment constraints, the first of which is given in (7a). This constraint specifies an OFFSET = ONSET coordination relation in /Cɾ/ sequences, which ensures an open articulatory transition between /ɾ/ and the preceding consonant, as shown in (7b).

(7) a. ALIGN(C, OFFSET, /ɾ/, ONSET)
In /Cɾ/, align the offset of C with the onset of /ɾ/.
b. *Coordination:* C OFFSET = /ɾ/ ONSET

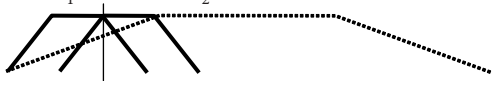
Percept: [C ə ɾ V]

Following Browman and Goldstein (1990) and Steriade (1990), I assume that within a syllable, consonantal articulations are superimposed on the tongue body gesture of the vowel, which is represented by the dotted line in (7b). Open transition allows the initial portion of the full vowel to be perceived between the constriction periods of the initial consonant and /ɾ/. Vowel intrusion is the acoustic result of this articulatory configuration.

The constraint favoring open transition in (7a) competes with other constraints on CC coordination. The alignment constraint in (8a) favors a RELEASE = TARGET relation in which the initial consonant of a /CC/ cluster is unreleased. As illustrated in (8b), close transition between adjacent consonants prevents vowel intrusion in the cluster.

(8) a. ALIGN(C_1, RELEASE, C_2, TARGET)
In /C_1C_2/, align the release of C_1 with the target of C_2.

b. *Coordination:* C_1 RELEASE = C_2 TARGET



Percept: [C ˥ C V]

As Hall (2003: 18) argues, conflicting gestural alignment constraints such as (7a) and (8a) have a functional grounding in terms of perceptibility and effort minimization, respectively. While open transition and vowel intrusion ensure clearer perceptual cues for the adjacent consonants (e.g., consonant release and formant transitions), a greater degree of overlap yields a relatively faster, more efficient overall articulation of the cluster.

The first generalization about Spanish complex onsets given in (5a) is that /Cɾ/ clusters typically have open transition and vowel intrusion, while /Cl/ clusters do not. This pattern actually reflects a broader typology of vowel intrusion behavior involving sonorants. Hall's (2003) cross-linguistic survey shows that vowel intrusion happens more with liquids than with other sonorants, and more with rhotics than laterals, except the alveolar trill. Consider the following implicational hierarchy:

(9) Vowel intrusion triggers (Hall 2003: 28)
obstruents, if ever > other approximants, nasals > [ɾ] > [l] > [ɾ], [ʁ] > gutturals
Among nasals: m > n

In a given language, if a particular class of consonants in (9) triggers vowel intrusion in clusters, then so do all consonant classes further down the hierarchy (modulo phonotactic restrictions and accidental gaps). One way to capture this typology is to posit a universal ranking of CC coordination constraints favoring open transition, each relativized to a different consonant class. The ranking of overlap-inducing constraints such as (8a) with respect to this hierarchy distinguishes consonants that trigger vowel intrusion from consonants that favor close transition.[6]

6. This proposal diverges slightly from that of Hall (2003: 28–30), who posits a hierarchy of *C IN V constraints penalizing the complete overlap of different types of consonant gestures by a tautosyllabic vowel gesture.

Spanish makes the cutoff point between [l] and [ɾ] in (9). Tableau (10) illustrates the difference in gestural alignment between /Cɾ/ and /Cl/ clusters. The ranking of (7a) above (8a) favors candidate (10a) with open transition over candidate (10b) with close transition, thereby ensuring vowel intrusion in [C^{ə}ɾV] demisyllables. In turn, (8a) outranks the alignment constraint favoring open transition in /Cl/, which accounts for the close transition and the lack of vowel intrusion in [ClV] (10d). In sum, generalization (5a) is explained by the Spanish-specific ranking of (8a) along the universal constraint hierarchy projected from the scale in (9).[7]

(10) Open transition in /Cɾ/ versus close transition in /Cl/

	Align (C, offset, /ɾ/, onset)	Align (C_1, release, C_2, target)	Align (C, offset, /l/, onset)
a. ☞ [C^{ə}ɾV]		*	
b. [CɹV]	*!		
c. [C^{ə}lV]		*!	
d. ☞ [ClV]			*

As discussed in Section 1, the duration of the intrusive vowel is highly variable in Spanish /Cɾ/ clusters. Following a proposal by Davidson (2003: 168–174), I suggest that a possible account of such variation might involve specifying a range of landmarks in the initial consonant gesture with which /ɾ/ may be aligned. If the constraint in (7a) is redefined as Align(C, {release ↔ offset}, /ɾ/, onset), then the onset of the /ɾ/ gesture could be aligned with any point between the release and the offset of the preceding consonant, thereby allowing for variability in the duration of the open transition and the intrusive vowel.

7. Although not the central focus of this paper, it is interesting to consider the implications of the analysis for the apicoalveolar trill [r] that appears in many Ibero-Romance varieties. Given the cutoff made between [l] and [ɾ] in the hierarchy in (9), the prediction is that [r] should fail to exhibit vowel intrusion when adjacent to consonants. Blecua (2001, §3.2.1) finds that the overwhelming majority of preconsonantal trills produced by her Castilian Spanish informants do not exhibit any vocalic element between the final closure phase of [r] and the following consonant. Solé (2002a,b) shows that syllable-initial trills in Catalan typically assimilate preceding lingual fricatives across minor prosodic boundaries due to gestural overlap between [r] and the fricative (also see Bradley 2006 on /sr/ clusters in Latin American Spanish). The aerodynamic characteristics of apical trills, together with the gestural coordination predicted in (9), may also explain the failure of [r] to pattern as the second member of complex onsets: "[c]oproduction of trills with tautosyllabic obstruents would affect the narrowly constrained lingual and aerodynamic requirements for tongue-tip trilling" (Solé 2002a: 685).

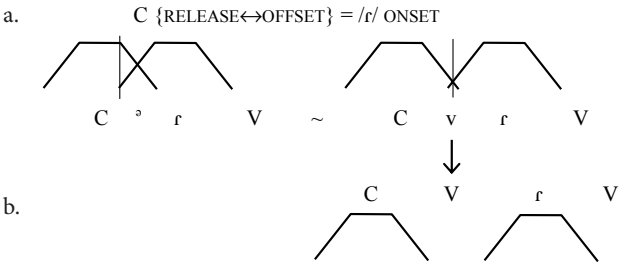


Figure 3. Speaker produces intrusive vowel of variable duration (a), listener reinterprets longer intrusive vowel as lexically-specified (b)

This approach to variable intersegmental timing explains the generalizations in (5b,c) regarding the acoustic quality of intrusive vowels and of the historical copy vowels they occasionally give rise to. As shown in (7b), the intrusive vowel stems from the same tongue body gesture as the tautosyllabic nuclear vowel. During the initial portion of the gesture, the tongue body articulator has not yet attained its target for the steady-state portion of the vowel. Since the pre-target posture of the tongue body is what colors the open transition between the consonantal gestures, complete identity does not obtain between the intrusive vowel and the nuclear vowel. However, as the temporal gap between the consonantal gestures increases, the vowel quality revealed during the open transition becomes more identifiable (Steriade 1990: 393). Over time, longer intrusive vowels may be phonologically reinterpreted and lexicalized as full nuclear vowels, as illustrated in Figure 3. Overlapping tautosyllabic vowel gestures are omitted here, and [ᵊ] and [v] in (a) are intended to denote shorter and longer intrusive vowels, respectively. Once lexical restructuring has taken place, CC coordination constraints are no longer relevant to the gestures of the original consonant sequence. This is because the consonant gestures are no longer segmentally adjacent due to the presence of the intervening lexically-specified vowel, which has effectively broken up the former cluster.

Crucial to the above diachronic explanation is the notion of *gestural misparsing,* whereby language learners erroneously interpret certain aspects of the acoustic signal to be the result of intentional articulatory gestures on the part of the speaker. Browman and Goldstein (1991: 331–333) observe that changes which arise from misparsing "do not involve adding articulations that were not there to begin with; rather they involve changes in the parameters of gestures that are already present." In Figure 3, the intrusive vowels in (a) and the lexicalized copy vowel in (b) all stem from the same overlapping vowel gesture, and the misparsing that occurs in (b) involves a change in the relative timing of adjacent consonantal gestures.

3. Coarticulation

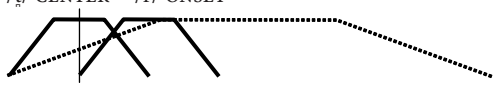
Recall Alonso's (1925: 186–189) observation, discussed in Section 1, that coarticulation of Spanish /Cɾ/ is characteristic of casual speech, while in careful speech the

perceptual integrity of the cluster is preserved. Open transition and vowel intrusion in /Cɾ/ enhances the perceptibility of the cluster, as guaranteed by the careful-speech ranking in which ALIGN(C, OFFSET, /ɾ/, ONSET) is dominant (see tableau (10)). In this section, I show how different patterns of coarticulation emerge in casual speech when this constraint is dominated by constraints favoring greater degrees of overlap.

The generalization in (5d) reveals two major patterns of coarticulation among varieties of Spanish: coarticulation of any /Cɾ/ cluster versus coarticulation of /ɾ/ with only a preceding homorganic stop. Again, cross-linguistic evidence suggests the pervasiveness of this pattern, given that some languages forbid vowel intrusion specifically in homorganic clusters while allowing it in heterorganic ones (see Hall 2003: 9–10). An account of this difference requires the additional constraint shown in (11a), the effects of which are illustrated in (11b) for homorganic /t̪ɾ/ and in (11c) for heterorganic /pɾ/. The percept [t̪ɹ̥] in the first cluster denotes close transition and the lack of vowel intrusion (but see below for a more detailed discussion of the assimilatory effects observed in coarticulation).

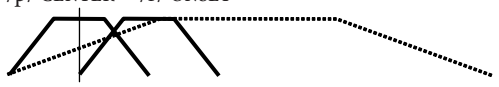
(11) a. ALIGN(C, CENTER, /ɾ/, ONSET)
In /Cɾ/, align the center of C with the onset of /ɾ/.

b. *Coordination:* /t̪/ CENTER = /ɾ/ ONSET



Percept: [t ɹ̥ V]

c. *Coordination:* /p/ CENTER = /ɾ/ ONSET



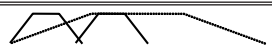
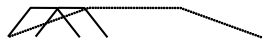
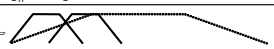
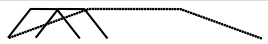
Percept: [p ə ɾ V]

Gafos (2002: 283–287) argues that the coordination relation CENTER = ONSET can produce different acoustic results depending on the consonant gestures involved. Consider first (11b), where the gestures for /t̪/ and /ɾ/ both involve the tongue tip. At the moment when this active articulator receives instructions to begin the release of the first consonant, it is simultaneously receiving instructions to move in the opposite direction toward the constriction target of the second consonant. In the gestural model, significant overlap between adjacent gestures engaging the same articulator results in a *blending* of gestural characteristics, which "shows itself in spatial changes in one or both of the overlapping gestures" (Browman and Goldstein 1990: 362). Since the temporal distance between the /t̪/ release and the /ɾ/ target in (11b) is not enough to allow the tongue tip to return to a neutral position, gestural blending keeps the articulator in place, producing a close transition. On the other hand, blending does not occur when the adjacent consonant gestures involve different active articulators, as in (11c). In heterorganic clusters, therefore, the CENTER = ONSET coordination relation will produce an acoustic release between the two consonants.

Tableau (12) illustrates the casual speech ranking necessary to account for those dialects in which coarticulation affects only homorganic /Cɾ/. Candidates (12a,b,c)

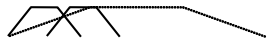
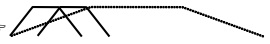
involve different coordination relations for heterorganic /pɾ/, namely OFFSET = ONSET, CENTER = ONSET, and RELEASE = TARGET, respectively. Candidates (12d,e,f) present the same three relations but for homorganic /t̪ɾ/. High-ranking (11a) optimizes the CENTER = ONSET coordination for both clusters, and candidates (12b) and (12e) win. As explained above, heterorganic (12b) is realized with open transition and vowel intrusion, but gestural blending ensures close transition in homorganic (12e).

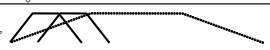
(12) Open transition in heterorganic /pɾ/ versus close transition in homorganic /t̪ɾ/

	ALIGN (C, CENTER, /ɾ/, ONSET)	ALIGN (C, OFFSET, /ɾ/, ONSET)	ALIGN (C_1, RELEASE, C_2, TARGET)
a.  [pvɾV]	*!		*
b. ☞  [pᵊɾV]		*	*
c.  [pɹ̪V]	*!	*	
d.  [t̪vɾV]	*!		*
e. ☞  [t̪ɹ̪V]		*	*
f.  [t̪ɹ̪V]	*!	*	

To account for dialects in which coarticulation affects both heterorganic and homorganic /Cɾ/ clusters, it is necessary to posit a casual speech ranking in which the constraint favoring close transition, (8a), dominates both (7a) and (11a). As seen in tableau (13), this ranking optimizes close transition in candidates (13c) and (13f).

(13) Close transition in heterorganic and homorganic /Cɾ/

	ALIGN (C_1, RELEASE, C_2, TARGET)	ALIGN (C, CENTER, /ɾ/, ONSET)	ALIGN (C, OFFSET, /ɾ/, ONSET)
a.  [pvɾV]	*!	*	
b.  [pᵊɾV]	*!		*
c. ☞  [pɹ̪V]		*	*

d.  [t̪vɾV]	*!	*	
e.  [t̪̊V]	*!		*
f. ☞  [t̪̊V]		*	*

Coupled with the notion of gestural blending, an analysis based on competing gestural alignment constraints effectively captures the implicational relationship observed in the coarticulation of different types of /Cɾ/ clusters in Spanish. No ranking of the constraints can produce close transition in heterorganic /pɾ/ without also producing it in homorganic /t̪ɾ/. On the other hand, the ranking in tableau (12) is capable of producing close transition in /t̪ɾ/ without also producing it in /pɾ/. Ultimately, this difference is due to the fact that the same gestural coordination relation, CENTER = ONSET, can have different acoustic consequences depending on the articulators involved (Gafos 2002).

The gestural alignment account provides a straightforward explanation of the assimilatory effects observed in coarticulation, documented in Section 1. For instance, Alonso's (1925: 186, 189) observation that rhotics may be partially devoiced after voiceless consonants suggests some degree of overlap between the rhotic constriction and the glottal devoicing gesture associated with the preceding consonant. With respect to /Cɾ/ clusters in which C_1 is a dental stop, the result of coarticulation is often described as an alveolar quasi-affricate. The retraction of the dental stop when overlapped with a following apicoalveolar /ɾ/ plausibly reflects a blended compromise between the lexically-specified constriction locations of the adjacent tongue tip gestures. The gestural account also explains cases of perceived consonantal shortening, as when /d̪ɾ/ clusters are preceded by a nasal or lateral in highland Ecuadorian Spanish (4b). Since there is a decrease in the relative timing between the tap gesture and the secondary gesture responsible for nasality or laterality of the initial sonorant, the duration of the intervening [d] is also decreased, yielding the percept of a reduced [^{d}].[8]

4. Gestural representations and constraints in the phonology

In this section, I argue that both gestures and segments are present in the phonological representation but are subject to different constraints interacting at the same level in the phonological grammar. I show how the absence of input-output faithfulness constraints on gestural coordination keeps the grammar from overgenerating unattested contrasts based on intersegmental timing.

8. It remains unclear why in a dialect such as highland Ecuadorian Spanish, coarticulation affects the [^{d}ɹ] clusters in (4b) but not [ðˠɾ] in (4c). In the gestural account proposed here, a CENTER = ONSET coordination relation would predict close transition and no vowel intrusion in both cases. I suggest that the difference is most likely related to the stop-continuant alternation in which the voiced coronal obstruent participates.

Many phonologists assume a division between phonological and phonetic components in the grammar (see Liberman and Pierrehumbert 1984, Keating 1990, Cohn 1990). Underlying forms are devoid of non-contrastive properties such as syllabification or temporal relations between articulatory gestures. The phonological component derives a syllabified surface representation that is categorical, qualitative, and timeless, and phonetic implementation then supplies gradient, quantitative aspects of non-contrastive detail to yield a fully-specified phonetic representation. Evidence that gestural coordination belongs in phonetic implementation comes from the observation that the acoustic consequences of different timing relations are in many ways invisible to the phonology. For instance, the duration of the intrusive vowel in a /C*ɾ*V/ demisyllable often approximates that of an unstressed vowel, which may give the appearance of two vowels, i.e., [Cv*ɾ*V]. However, there is good reason to believe that vowel intrusion does not create a new syllable, unlike true phonological epenthesis of a nuclear vowel (Hall 2003).

Two arguments from Spanish support the phonological invisibility of vowel intrusion. First, intrusive vowels are never counted in stress computation. In Spanish, main stress is confined without exception to a three-syllable window at the right edge of the morphological word (Harris 1995: 869). If vowel intrusion in the /C*ɾ*/ clusters of proparoxytones such as *kilómetro* [ki.ˈlo.me.t̪o*ɾ*o] 'kilometer' and *demócrata* [de.ˈmo.ka*ɾ*a.t̪a] 'democrat' were to create a new syllable, then stress would fall outside the three-syllable window: *[ki.ˈlo.me.t̪o.*ɾ*o], *[de.ˈmo.ka.*ɾ*a.t̪a]. Although theoretically possible, stress shift is unattested as a means of repairing the prosodic ill-formedness of examples such as these. Second, in the Spanish language game Jerigonza, often used by younger speakers as a secret speech code, intrusive vowels again pattern as invisible. In one version of the game, an epenthetic CV syllable is inserted to the right of every syllable boundary in a word. The consonant is from the set /p, t̪, k, ʧ/, and the vowel is a copy of the preceding syllable nucleus (Piñeros 1999). If the intrusive vowel in *libro* [ˈli.βo*ɾ*o] 'book' were syllabic, then CV-insertion should also target this nucleus. The fact that Jerigonza word formation yields *li.pi.bro.po* instead of **li.pi.bo.po.ro.po* shows that speakers treat the [βo*ɾ*] sequence as a complex onset and ignore the intrusive vowel.

The stress and Jerigonza facts demonstrate that vowel intrusion is not a synchronic process of vowel epenthesis, even when the intrusive vowel is as long as a full unstressed vowel. Rather, intrusive vowels attain phonological status only in the diachronic dimension as a result of gestural misparsing (see the discussion surrounding Figure 3). This account predicts that diachronic reanalysis of intrusive vowels cannot yield forms such as the ungrammatical *[ki.ˈlo.me.t̪o.*ɾ*o] and *[de.ˈmo.ka.*ɾ*a.t̪a], which, to the best of my knowledge, is correct.

Perhaps the best evidence for the invisibility of intrusive vowels is that gradient differences in intersegmental timing are universally non-contrastive. Hall's (2003) cross-linguistic survey shows that in each language, vowel intrusion either always happens or never happens in a given environment (modulo variability due to fast/casual speech). This places the intrusive vowel on a par with consonant release, which plays an important role in the perceptual licensing of contour segments although it is never

phonologically contrastive per se (Steriade 1993). Moreover, Spanish speakers are typically unaware of the existence of intrusive vowels in clusters containing /ɾ/. It seems unlikely that any language would have minimal pairs based solely on minute differences in the phonetic timing of adjacent consonant gestures.[9]

Phonological invisibility is also found in the case of rhotic assibilation. As discussed in Section 1, Spanish onset clusters are limited to two consonants, where the first is an obstruent and the second is a liquid. Assuming that the assibilated rhotic is an obstruent since it is clearly fricative, coarticulation of /Cɾ/ clusters yields an apparent violation of the sonority conditions that determine complex onset phonotactics in Spanish. In the same way that intrusive vowels are invisible in stress computation and non-concatenative morphology, the assibilated rhotic is invisible to sonority sequencing, which otherwise disallows onset clusters consisting of two obstruents.

How is the phonological invisibility of vowel intrusion and rhotic assibilation to be accounted for? In the model of Zsiga (2000), the phonology acts upon abstract features and segments, which are then mapped to gestures that are coordinated by language-specific alignment constraints in phonetic implementation. Intrusive vowels and assibilated rhotics arise in the phonetics, where stress and sonority constraints are no longer operative and where segments cease to be relevant after features are mapped to gestures. On this view, vowel intrusion and rhotic assibilation constitute a *phonetics–phonology mismatch* in the sense of Blevins (1995: 232–234). More recently, however, Hall (2003) argues against the necessity of a derivational mapping between segments and features in the phonology and gestures in the phonetics. She argues instead for a unified representational model in which gestures are associated to segments, which in turn group together into higher prosodic constituents such as syllables, feet, prosodic words, and so on. If the constraints responsible for stress computation, Jerigonza word formation, and syllabification refer only to higher segmental and prosodic structure, then it follows that they will be insensitive to any percepts arising from specific gestural coordination relationships.

Consider Figure 4, in which gestures, segments, and syllables exist simultaneously in the same phonological representation. As we have seen in Section 2, gestural coordination constraints interact to produce open transition and vowel intrusion in (a) versus close transition and no vowel intrusion in (b). At the same time, phonological constraints evaluate segmental and prosodic structure. The fact that /t̪ɾ/ is a

9. While minute differences in intersegmental timing never form the sole basis of a phonological contrast, timing can impact the ability of listeners to successfully recover other contrasts. For example, the non-affrication of the /t/-/ʃ/ sequence in the English minimal pairs *white shoes* versus *why choose* and *might shop* versus *my chop* is dependent on intergestural timing. The very fact that intrusive vowels are credited with enhancing the perceptual cues of adjacent consonants in open transition deems them as playing at least some role in phonological contrasts. To these cases we can add the apparent deletion of English word-final /t/ when heavily overlapped and perceptually hidden by a following word-initial consonant, e.g., [mʌsbi] *must be* and [pəfekmemɹi] *perfect memory* (Browman and Goldstein 1990: 361). Also, the failure of trills to contrast with taps in the second position of complex onsets in Ibero-Romance has a plausible basis in aerodynamic factors and gestural coordination, as suggested in Fn. 7.

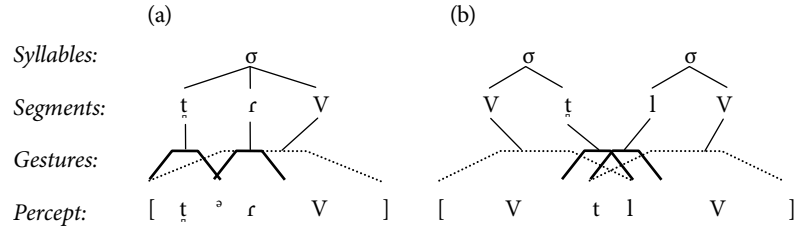


Figure 4. Open transition in tautosyllabic /t̪ɾ/ (a) versus close transition in heterosyllabic /t̪l/ (b)

permissible onset cluster in Spanish is accounted for by a language-specific ranking of universal sonority and syllabification constraints (see Colina 1995, 1997, this volume). Recall that /d̪l/ onsets are not allowed in any Spanish variety and that dialects differ in the syllabification of /t̪l/. Martínez-Gil (2001: 209) suggests that cooccurrence restrictions on dental-lateral clusters can be explained as an effect of the Obligatory Contour Principle (OCP; McCarthy 1986, Odden 1986), whereby adjacent coronal noncontinuant segments are prohibited in the syllable onset. Dialect-specific rankings of the relevant OCP constraints with respect to NoCoda determine whether a dialect prefers heterosyllabic /t̪l/, as shown in (b).[10]

Crucially, the intrusive vowel in (a) of Figure 4 is not part of the segmental representation, which accounts for its invisibility to phonological processes that make reference to syllables. As we have seen, intrusive vowels are the acoustic consequence of non-overlapping consonant gestures and are not part of the formal representation of segments. In short, the mismatch between phonological vowel epenthesis and phonetic vowel intrusion requires not a *derivational* difference between phonological and phonetic components but rather a *representational* difference between segments and gestures in the phonological representation. The same argument holds with respect to rhotic assibilation and sonority sequencing. Under a different ranking of gestural alignment constraints, coarticulation of the /t̪ɾ/ cluster in (a) would yield the percept of an assibilated rhotic, i.e., [tʒ̪]. However, the resulting sequence would still syllabify as a complex onset because sonority and syllabification constraints operate over the segmental string [t̪ɾV] and are not sensitive to percepts that arise from gestural coordination.

Finally, consider the observation that intersegmental gestural timing never forms the sole basis of a phonological contrast. If UG had an input-output faithfulness constraint such as Ident(timing), then some language might rank it above gestural alignment constraints. This would allow input coordination relations between adjacent consonants to surface faithfully, thereby overpredicting a contrast based on gestural coordination, such as the presence versus absence of consonant release. The possibility of contrast overgeneration fails to capture the fact that Spanish speakers treat phonetic forms such as [o.t̪oɾo], [o.t̪ᵊɾo], and [o.tʒ̪o] as phonologically equivalent realizations of the same word *otro* 'other'. What this suggests is that UG contains no faithfulness

10. The gestural account predicts blending of the dental /t̪/ in close transition with the following alveolar lateral, independently of the cluster's heterosyllabic status. Further articulatory investigation is needed to verify this prediction empirically.

to underlying gestural alignment and that phonetic timing relations must result from interacting markedness constraints alone (Hall 2003; also see Kirchner 1997). Underlying gestures may be temporally coordinated in various ways or perhaps not at all, but the absence of faithfulness to input timing ensures that timing will never be contrastive per se.[11] Rather, surface coordination will always be determined by the interaction of gestural alignment constraints, as shown in Sections 2 and 3.

It is important to emphasize that the proposal to include both segments and gestures within the same phonological representation does not entail the existence of two separate evaluation mechanisms in an Optimality-theoretic grammar. In the unified model, "[c]onstraints referring to gestural and non-gestural phenomena are present in the same level of the grammar" (Hall 2003: 12). Tableaux (10), (12), and (13) show only the interaction among conflicting gestural alignment constraints, but this is not meant to imply that such constraints evaluate output candidates independently of the non-gestural constraints that determine stress or syllabic configurations. For a given input, the optimal surface form, such as those exemplified in Figure 4, is accounted for by the interaction of constraints in a single, unified tableau. In many cases, it is likely that gestural and non-gestural constraints will not conflict with each other, simply because they evaluate different aspects of the phonological representation. More intriguing is the possibility that gestural alignment constraints may make reference to higher-level aspects of representation, such as prosodic domains or underlying morphological structure. Models that assume a derivational relationship between phonology and phonetic implementation predict that morphology cannot influence gestural coordination. However, such interaction is predicted by a unified model in which gestural alignment constraints have access to underlying morphological structure (see Bradley 2005 and Cho 2001 for further discussion of the influence of prosodic and morphological structure on gestural coordination).

5. Comparison with previous accounts

Steriade (1990) is the first to demonstrate the utility of gestural representations in accounting for vowel intrusion phenomena in Winnebago, Late Latin, and Sardinian. The first application of Articulatory Phonology to Spanish rhotics is found in Bradley (1999; see also 2001a). Specifically, I proposed an account of open and close transition in highland Ecuadorian Spanish clusters containing /ɾ/, as well as an explanation for the assibilation of syllable-initial trills in terms of gestural reduction. In that work, I assumed a division between phonology and phonetic implementation in the grammar, with gestural timing determined in the latter component, as in Zsiga's (2000)

11. In phonological frameworks that take a systemic view of contrast, it is insufficient to ban IDENT(timing), either because no underlying representation is assumed (Flemming 1995) or because generalized systemic faithfulness exists as an independent constraint in the grammar (Padgett 2003a,b,c). See the discussion in Bradley (2005), in which I argue that imperceptible contrasts based on gestural timing must be universally ruled out by inviolable perceptual distinctiveness constraints.

model. However, the mechanisms governing the coordination of gestures were never made explicit. Given the study's relatively limited dialectal focus, no attempt was made to connect the gestural coordination patterns of Spanish clusters to vowel intrusion phenomena in other languages. The present work formalizes the competition among gestural coordination relations in terms of Optimality-theoretic alignment constraints and connects the variation observed in Spanish with other languages, following Hall's (2003) work.

On the basis of derived-environment effects in Korean palatalization, Cho (1998a,b) proposes an Optimality-theoretic analysis in which a markedness constraint favoring gestural overlap competes with a constraint requiring faithfulness to input Phase Windows, which define permissible ranges of overlap between adjacent gestures (see Byrd 1996). In Bradley (2002), I propose a similar approach to derived-environment effects in Norwegian /ɾC/ clusters. The approach is further extended to dialectal variation in Spanish /Cɾ/ clusters by Bradley and Schmeiser (2003) and to coda rhotics in highland Ecuadorian Spanish by Bradley (2004). The assumption underlying all of these works is that input morphemes already have their gestural timing relations fully and reliably specified in terms of Phase Windows so that faithfulness, IDENT(timing), can depend on them. On this view, a predictable non-contrastive property of phonetic detail — intersegmental timing — is incorporated directly into the phonological representation. As McCarthy (personal communication; 2003, Fn. 7) points out, however, the assumption that inputs are fully-specified for intersegmental timing runs counter to the Richness of The Base hypothesis of Optimality Theory, which forbids placing restrictions directly on input representations (Prince and Smolensky 1993). Moreover, the use of IDENT(timing) faces the problem of overgenerating contrast, as discussed in Section 4 above. The analysis developed in the present study avoids these complications because optimal gestural coordination relations are determined by interacting markedness constraints. The problem of having to specify intersegmental timing in underlying forms becomes irrelevant under this approach.

Finally, Blecua (2001) argues against a gestural account of vowel intrusion in Spanish /CɾV/ and /VɾC/ sequences, as outlined in Bradley (1999, 2001a). She maintains that if the intrusive vowel were merely the portion of the nuclear vowel that appears between the tap constriction and the adjacent consonant, then we would expect the formant structure of both vocalic sounds to be identical.

> However, the results obtained in our study indicate that the structures are not identical; although the vowel has an important influence on the formant structure of the vocalic element, the triangle formed by this element in a formant chart is included within that of the vowel […] In this sense, it is difficult to accept that the vocalic element is the part of the vowel that has been separated by the tap constriction (Blecua 2001: §4.1.1).

Rather, Blecua proposes that the intrusive vowels appearing on either side of the /ɾ/ constriction are an inherent part of the rhotic itself. Since the formant structure of these vocalic elements serves no distinctive function, it simply adopts characteristics that are similar to the tautosyllabic vowel. Although not explicitly discussed by Blecua,

a similar representation of the tap as a tripartite contour segment has been independently proposed by Inouye (1995) and further elaborated in Bradley (2001a,b).

While the representation of /ɾ/ as a sequence of approach, closure, and release phases works for Spanish, such an account fails to explain the fact that in other languages, intrusive vowels can be triggered by consonants other than /ɾ/ (Hall 2003). An analysis in terms of interacting gestural coordination constraints effectively situates Spanish vowel intrusion within a broader cross-linguistic typology. Specifically, a universal ranking of alignment constraints favoring open transition captures the implicational relations among vowel intrusion triggers and also explains why /ɾ/ is the only such trigger in Spanish (see tableau (10)). As discussed in Section 2, a gestural account of vowel intrusion also explains the lack of complete identity between intrusive vowels and their tautosyllabic nuclei. During the open transition between the adjacent consonantal gestures, the tongue body has not yet attained its target for the steady-state portion of the nuclear vowel. As we have seen in Figure 3, full identity is achieved diachronically in some cases when longer intrusive vowels become lexicalized as nuclear vowels.

6. Conclusion

In this paper, I have shown that Articulatory Phonology, in conjunction with Optimality Theory, makes possible an explanatory account of the phonetic patterning of complex onsets in Spanish. Furthermore, I have shown that there is no danger in assuming phonetically rich gestural representations along with segments and syllables in the phonology. The fact that vowel intrusion and rhotic assibilation derive from gestural coordination accounts for the invisibility of these phenomena to processes that refer to segmental and prosodic structure. In Optimality Theory, the absence of faithfulness to input timing ensures that no language grammar can generate contrasts based solely on differences in gestural coordination — even if such differences happen to be present in the input. Finally, since the gestural alignment constraints necessary to account for Spanish vowel intrusion are universal and violable, different rankings of these constraints can account for vowel intrusion patterns observed in other languages.

References

Argüello, F. 1978. The Žeísta Dialect of Spanish Spoken in Ecuador: A phonetic and phonological study. PhD dissertation, The Pennsylvania State University.

Alonso, A. 1925. El grupo tr en España y América. *Homenaje a Menéndez Pidal* II: 167–191. (Madrid).

Blecua, B. 2001. Las vibrantes en español: Manifestaciones acústicas y procesos fonéticos. PhD dissertation, Universitat Autònoma de Barcelona.

Blevins, J. 1995. The syllable in phonological theory. In *Handbook of Phonological Theory,* J. Goldsmith (ed.), 206–244. Cambridge MA: Blackwell.

Bradley, T.G. 1999. Assibilation in Ecuadorian Spanish: A phonology–phonetics account. In *Formal Perspectives on Romance Linguistics,* J.-M. Authier, B.E. Bullock, and L.A. Reed (eds), 57–71. Amsterdam: John Benjamins.

Bradley, T.G. 2001a. The Phonetics and Phonology of Rhotic Duration Contrast and Neutralization. PhD dissertation, The Pennsylvania State University.

Bradley, T.G. 2001b. A typology of rhotic duration contrast and neutralization. In *Proceedings of NELS 31,* M.J. Kim and U. Strauss (eds), 79–97. Amherst MA: Graduate Linguistics Student Association.

Bradley, T.G. 2002. Gestural timing and derived environment effects in Norwegian clusters. In *WCCFL 21 Proceedings,* L. Mikkelsen and C. Potts (eds), 101–114. Somerville MA: Cascadilla.

Bradley, T.G. 2004. Gestural timing and rhotic variation in Spanish codas. In *Laboratory Approaches to Spanish Phonology,* T. Face (ed.), 197–224. Berlin: Mouton de Gruyter.

Bradley, T.G. 2005. Systemic markedness and phonetic detail in phonology. In *Experimental and Theoretical Approaches to Romance Linguistics,* R. Gess and E. Rubin (eds), 41–61. Amsterdam: John Benjamins.

Bradley, T.G. 2006. Phonetic realizations of /sr/ clusters in Latin American Spanish. In *Selected Proceedings of the 2nd Conference on Laboratory Approaches to Spanish Phonetics and Phonology* [Proceedings Project]. M. Díaz-Campos (ed.), 1–13. Somerville MA: Cascadilla.

Bradley, T.G. and Schmeiser, B.S. 2003. On the phonetic reality of /ɾ/ in Spanish complex onsets. In *Selected Proceedings of The 6th Hispanic Linguistics Symposium,* P.M. Kempchinsky and C.-E. Piñeros (eds), 1–20. Somerville MA: Cascadilla.

Browman, C. and Goldstein, L. 1989. Articulatory gestures as phonological units. *Phonology* 6: 201–252.

Browman, C. and Goldstein, L. 1990. Tiers in articulatory phonology, with some implications for casual speech. In *Papers in Laboratory Phonology I: Between the grammar and physics of speech,* J. Kingston and M.E. Beckman (eds), 341–376. Cambridge: CUP.

Browman, C. and Goldstein, L. 1991. Gestural structures: Distinctiveness, phonological processes, and historical change. In *Modularity and the Motor Theory of Speech Perception,* I.G. Mattingly and M. Studdert-Kennedy (eds), 313–338. Hillsdale NJ: Lawrence Erlbaum.

Byrd, D. 1996. A phase window framework for articulatory timing. *Phonology* 13: 139–169.

Cho, T. 1998a. Intergestural timing and overlap in Korean palatalization: An optimality-theoretic approach. In *Japanese/Korean Linguistics* 8, D. Silva (ed.), 261–276. Stanford CA: CSLI.

Cho, T. 1998b. The Specification of Intergestural Timing and Overlap: EMA and EPG studies. M.A. thesis, University of California, Los Angeles.

Cho, T. 2001. Effects of morpheme boundaries on intergestural timing: Evidence from Korean. *Phonetica* 58: 129–162.

Cohn, A. 1990. Phonetic and Phonological Rules of Nasalization. PhD dissertation, University of California, Los Angeles.

Colantoni, L. 2001. Mergers, Chain Shifts, and Dissimilatory Processes: Palatals and rhotics in Argentine Spanish. PhD dissertation, University of Minnesota.

Colantoni, L. and Steele, J. 2005. Phonetically-driven epenthesis asymmetries in French and Spanish obstruent-liquid clusters. In *Experimental and Theoretical Approaches to Romance Linguistics,* R. Gess and E. Rubin (eds), 77–96. Amsterdam: John Benjamins.

Colina, S. 1995. A Constraint-Based Analysis of Syllabification in Spanish, Catalan and Galician. PhD dissertation, University of Illinois, Urbana-Champaign.

Colina, S. 1997. Identity constraints and Spanish resyllabification. *Lingua* 103: 1–23.

Colina, S. Forthcoming. Optimality-theoretic advances in our understanding of Spanish syllabic structure. In *Optimality-Theoretical Studies in Spanish Phonology,* S. Colina and F. Martínez-Gil (eds). Amsterdam: John Benjamins.

Davidson, L. 2003. The Atoms of Phonological Representation: Gestures, coordination, and perceptual features in consonant cluster phonotactics. PhD dissertation, Johns Hopkins University.

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

Flemming, E. 1995. Auditory Representations in Phonology. PhD dissertation, University of California, Los Angeles.

Gili Gaya, S. 1921. La r simple en la pronunciación española. *Revista de Filología Española* 8: 271–280.

Hall, N. 2003. Gestures and Segments: Vowel intrusion as overlap. PhD dissertation, University of Massachusetts, Amherst.

Hammond, R. 1999. On the non-occurrence of the phone [r̃] in the Spanish sound system. In *Advances in Hispanic Linguistics,* J. Gutiérrez-Rexach and F. Martínez-Gil (eds), 135–151. Somerville MA: Cascadilla.

Hammond, R. 2000. The phonetic realizations of /rr/ in Spanish: A psychoacoustic analysis. In *Hispanic Linguistics at the Turn of the Millennium,* H. Campos, E. Herburger, A. Morales-Front and T.J. Walsh (eds), 80–100. Somerville MA: Cascadilla.

Harris, J. 1983. *Syllable Structure and Stress in Spanish: A nonlinear analysis.* Cambridge MA: The MIT Press.

Harris, J. 1995. Projection and edge marking in the computation of stress in Spanish. In *The Handbook of Phonological Theory,* J. Goldsmith (ed.), 867–887. Cambridge MA: Blackwell.

Harris, J. and Kaisse, E. 1999. Palatal vowels, glides and obstruents in Argentinian Spanish. *Phonology* 16: 117–190.

Hualde, J.I. 1991. On Spanish syllabification. In *Current studies in Spanish Linguistics,* H. Campos and F. Martínez-Gil (eds), 475–493. Washington DC: Georgetown University Press.

Hualde, J.I. 1999. La silabificación en español. In *Fonología generativa contemporánea de la lengua española,* R.A. Núñez-Cedeño and A. Morales-Front (eds), 170–188. Washington DC: Georgetown University Press.

Hualde, J.I. 2004. Quasi-phonemic contrasts in Spanish. In *WCCFL 23 Proceedings,* B.S. Schmeiser, V. Chand, A. Kelleher, and A. Rodríguez (eds), 374–398. Somerville MA: Cascadilla.

Inouye, S. 1995. Trills, Taps and Stops in Contrast and Variation. PhD dissertation, University of Los Angeles, California.

Keating, P.A. 1990. Phonetic representations in a generative grammar. *Journal of Phonetics* 18: 321–334.

Kirchner, R. 1997. Contrastiveness and faithfulness. *Phonology* 14: 83–111.

Lenz, R. 1892. Chilenishe Studien. *Phonetische Studien* 5: 272–293.

Lewis, A. 2004. Coarticulatory effects on Spanish trill production. In *Proceedings of the 2003 Texas Linguistics Society Conference* [Proceedings Project], A. Agwuele, W. Warren and S.-H. Park (eds), 116–127. Somerville MA: Cascadilla.

Liberman, M. and Pierrehumbert, J. 1984. Intonational invariance under changes in pitch range and length. In *Language Sound Structure,* M. Aronoff and R.T. Oehrle (eds), 157–233. Cambridge MA: The MIT Press.

Lipski, J. 1994. *Latin American Spanish.* London: Longman Group.

Malmberg, B. 1965. *Estudios de fonética hispánica* [Collectanea Phonetica, I.] Madrid: Consejo Superior de Investigaciones Científicas.

Martínez-Gil, F. 2001. Sonority as a primitive phonological feature: Evidence from Spanish complex onset phonotactics. In *Features and Interfaces in Romance,* J. Herschensohn, E. Mallén, and K. Zagona (eds), 203–222. Amsterdam: John Benjamins.

McCarthy, J. 1986. OCP effects: Gemination and antigemination. *Linguistic Inquiry* 17: 207–263.

McCarthy, J. 2003. Comparative markedness. *Theoretical Linguistics* 29: 1–51.

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

Navarro Tomás, T. 1918. Diferencias de duración entre las consonantes españolas. *Revista de Filología Española* 5: 367–393.

Odden, D. 1986. On the role of the obligatory contour principle in phonological theory. *Language* 62: 353–383.

Padgett, Jaye. 2003a. Contrast and post-velar fronting in Russian. *Natural Language and Linguistic Theory* 21: 39–87.

Padgett, Jaye. 2003b. The emergence of contrastive palatalization in Russian. In *Optimality Theory and Language Change,* D.E. Holt (ed.), 307–335. Dordrecht: Kluwer.

Padgett, Jaye. 2003c. Systemic contrast and Catalan rhotics. Ms., University of California, Santa Cruz.

Piñeros, C.-E. 1999. Head dependence in Jerigonza, a Spanish language game. In *Advances in Hispanic Linguistics,* J. Gutiérrez-Rexach and F. Martínez-Gil (eds), 265–277. Somerville MA: Cascadilla.

Prince, A. and Smolensky, P. 1993. Optimality Theory: Constraint interaction in generative grammar. Ms., Rutgers University and University of Colorado.

Quilis, A. 1970. El elemento esvarabático en los grupos [pr, br, tr…]. In *Phonétique et linguistique romanes: Mélanges offerts à M. Georges Straka,* 99–104. Lyon-Strasbourg : Société de Linguistique Romane.

Quilis, A. 1988. *Fonética acústica de la lengua española.* Madrid: Editorial Gredos.

Quilis, A. 1993. *Tratado de fonología y fonética españolas.* Madrid: Editorial Gredos.

Schmeiser, B.S. 2004. On the durational variability of svarabhakti vowels in Spanish complex onsets. Paper presented at the *Western Conference on Linguistics,* Los Angeles, November 2004.

Solé, M.J. 2002a. Aerodynamic characteristics of trills and phonological patterning. *Journal of Phonetics* 30: 655–688.

Solé, M.J. 2002b. Assimilatory processes and aerodynamic factors. In *Papers in Laboratory phonology* 7, C. Gussenhoven and N. Warner (eds), 351–386. Berlin: Mouton de Gruyter.

Steriade, D. 1990. Gestures and autosegments: Comments on Browman and Goldstein's paper. In *Papers in Laboratory Phonology I: Between the grammar and physics of speech,* M. Beckman and J. Kingston (eds), 382–397. Cambridge: CUP.

Steriade, D. 1993. Closure, release, and nasal contours. In *Nasals, Nasalization, and the Velum,* M. Huffman and R. Krakow (eds), 401–470. New York NY: Academic Press.

Willis, E. 2005. An acoustic study of the trill in Dominican Spanish. Paper presented at *The XXXVth Linguistic Symposium on Romance Languages,* University of Texas at Austin, February 24–27, 2005.

Willis, E. and Pedrosa, M.B. 1999. An acoustic analysis of the rhotic system of Spanish: Isolated words, read texts and spontaneous speech. Paper presented at the *3rd Hispanic Linguistics Symposium,* Washington, D.C., October 1999.

Zsiga, E. 2000. Phonetic alignment constraints: Consonant overlap and palatalization in English and Russian. *Journal of Phonetics* 28: 69–102.