Working Draft, October 2002
Comments welcome!

On the Phonetic Reality of Spanish /ɾ/ in Complex Onsets

Travis G. Bradley and Ben Schmeiser
University of California, Davis

Contact information:
Department of Spanish and Classics
622 Sproul Hall
University of California, Davis
One Shields Avenue
Davis, CA 95616
Email: tgbradley@ucdavis.edu, benschmeiser@ucdavis.edu

On the Phonetic Reality of Spanish /ɾ/ in Complex Onsets[*]

[*] The present study is the ongoing extension of ideas initially developed by Bradley (1999, 2002, to appear) and has benefited indirectly from the comments and feedback of several individuals on the latter works: Fanny Argüello, Gary Baker, Eric Bakovic, Jill Beckman, Lev Blumenfeld, Paul Boersma, Barbara Bullock, Patrik Bye, John Dalbor, James Fidelholtz, Adamantios Gafos, Jorge Guitart, Nancy Hall, Barbara Hancin-Bhatt, David Heap, Susan Inouye, Haike Jacobs, Patricia Keating, Robert Kirchner, John McCarthy, Armin Mester, Jean-Pierre Montreuil, Holly Nibert, Richard Page, Joe Pater, Leendert Plug, Jason Riggle, Donna Rogers, Hanne Gram Simonsen, Joe Stemberger, Donca Steriade, Bart van der Veer, Caroline Wiltshire, and members of the WCCFL 21, LSRL 28, and LSRL 32 audiences. Thanks are also due to members of the audience of the Sixth Hispanic Lingustics Symposium held at the University of Iowa, October 18-20, 2002, especially Jill Beckman, Joan Mascaró, and Carlos Eduardo Piñeros. We assume responsibility for any shortcomings in the present study, however.

1. Introduction

A common trend among contemporary generative studies of Spanish rhotics is that of glossing over what are deemed to be irrelevant, low-level details of phonetic implementation. Consequently, much of the variation underlying the phonetic reality of these segments is ignored. Such a move is taken, for instance, by Harris (1983:62), who reduces the "astonishing variety of *r*-quality phones … to just two, [r] and [r̄], which will be understood to jointly exhaust the rich phonetic variety […] I will say little more about phonetic detail…" Recent investigations have begun to redress the lack of attention given to phonetic detail (e.g., Colantoni 2001, Hammond 1999, 2000, to appear-a,b, and Willis and Pedrosa 1998). The present work contributes to this line of research by investigating the phonetic reality of Spanish /ɾ/ in complex onsets.

This paper is organized as follows. In Section 2, we identify some prosodic, segmental, and stylistic influences on the realization of /Cɾ/ clusters. In Section 3, we develop a formal analysis, couched within a phonetically-based version of Optimality Theory (OT; Prince and Smolensky 1993, McCarthy and Prince 1995), in which the articulatory drive to coarticulate adjacent consonantal gestures in the output conflicts with the perceptual requirement that input clusters be recoverable. Section 4 shows how the analysis captures the attested influences on /Cɾ/ realization. In Section 5, we discuss the role of phonetic detail in phonological analysis and suggest some areas for further empirical investigation. Section 6 concludes.

2. Phonetic realizations of Spanish /Cɾ/

2.1 Svarabhakti

It has long been noted that Spanish /Cɾ/ exhibits an intervening *svarabhakti* vowel fragment of variable duration (Gili Gaya 1921, Lenz 1892, Navarro Tomás 1918), as illustrated in (1). Although represented here simply as [ᵊ] in narrow phonetic transcription, this fragment typically has formant structure similar to that of the nuclear vowel appearing on the opposite side of the tap constriction (Quilis 1993:337-42).[1]

(1)	*pronto*	[pᵊɾ]	'soon'
	fresco	[fᵊɾ]	'cool, fresh'
	otro	[t̪ᵊɾ]	'other'
	negro	[ɣᵊɾ]	'black'

[1] See the Appendix for a spectrographic illustration of svarabhakti.

In an early study on /ɾ/ in Peninsular Spanish, Gili Gaya (1921) measured the duration of svarabhakti vowel fragments in different types of /Cɾ/ clusters. Overall, the duration of the vowel fragment was found to be highly variable, even within the same word repeated several times by the same individual. However, when clusters are grouped according to several prosodic and segmental variables, as in Table 1, a trend emerges whereby longer svarabhakti is favored in certain contexts:

Factor	Mean duration of svarabhakti (cs) by cluster type			
Position within the word	Word-initial	**5.3**	Word-internal	3.7
Stress	Stressed syllable	**6.5**	Unstressed syllable	5.2
Order of constriction location	Back-to-front	**6.3**	Front-to-back	5.5

Table 1: *The influence of prosodic and segmental factors on the duration of svarabhakti in* /Cɾ/ *clusters (adapted from Gili Gaya 1921:277-8)*

The means in Table 1 show that vowel fragments tend to be longer in word-initial and stressed /CɾV/ demisyllables than in non-initial or unstressed ones, respectively. Longer svarabhakti is also favored in clusters that exhibit a back-to-front order of constriction location (i.e., dorsal+/ɾ/) than in clusters with the opposite order (i.e., labial+/ɾ/). These results are summarized in (2), where [v] and [^{ə}] denote longer and shorter vowel fragments, respectively:

(2) Relative duration of svarabhakti in /CɾV/ demisyllables according to prosodic and segmental factors shown in Table 1

a. Position within the word: #CvɾV > C^{ə}ɾV
b. Stress: CvɾV́ > C^{ə}ɾV̆
c. Order of constriction location: { kvɾV, gvɾV } > { p^{ə}ɾV, b^{ə}ɾV, f^{ə}ɾV }

We intend for the symbols [v] and [^{ə}] to represent two distinct durational categories, abstracting away from the inherently variable duration of svarabhakti fragments. They represent *relative* durational differences that obtain when two cluster types are compared with respect to each prosodic or segmental factor. Following the results of Gili Gaya's phonetic study, we assume that the clusters shown in (2), although idealized, constitute a more accurate phonetic description of /Cɾ/ clusters than the forms in (1).

Malmberg's (1965:10, 35) phonetic measurements show that the duration of svarabhakti often approximates that of an unstressed vowel. As shown by the diachronic examples in (3), svarabhakti has occasionally given rise to a lexicalized copy vowel whose quality matches that of the underlying nucleus tautosyllabic with the complex onset (Gili Gaya 1921:280, Quilis 1988:300):

(3)	*pereces*	< *preces*	'prayers'
	tarabilla	< *trabilla*	'stirrup'
	corónica	< *crónica*	'chronicle'
	chácara	< *chacra*	'farm'
	gurupa	< *grupa*	'hindquarters'
	tíguere	< *tigre*	'tiger'

These examples show that in theory, any /CɾV/ demisyllable may be reanalyzed over time as /CVɾV/ regardless of the durational trends observed by Gili Gaya (1921) in (2) above. That is, lexicalized copy vowels may emerge non-initially (*chácara* < *cha<u>cra</u>, tíguere* < *ti<u>gre</u>*), in unstressed syllables (*tarabilla* < *<u>tra</u>billa, chácara* < *cha<u>cra</u>, tíguere* < *ti<u>gre</u>*), as well as in clusters with a front-to-back order of constriction location (*pereces* < *<u>pr</u>eces*).

2.2 Coarticulation

As documented by Alonso (1925), /Cɾ/ clusters in some Peninsular Spanish varieties are often coarticulated in casual speech, with concomitant voicing and/or place assimilation:

(4)	*apretar*	[pɹ̥]	'to squeeze'
	hombre	[bɹ]	'man'
	otro	[tɹ̥]	'other'
	vendrá	[dɹ]	's/he will come'
	padre	[ðɹ]	'father'
	escribir	[kɹ̥]	'to write'
	magras	[ɣɹ]	'lean (fem. pl.)'

Coarticulation entails some frication of the rhotic and the loss of both the intervening svarabhakti fragment and the extra-short constriction period of apicoalveolar [ɾ].[2] As shown in (4), rhotics are progressively devoiced after voiceless consonants (e.g., [kɹ̥] versus [ɣɹ]), and dental /t̪/ and /d̪/ assimilate regressively to the rhotic, yielding an alveolar quasi-affricate realization (e.g., [tɹ̥] versus [t̪ᵊɾ]). The articulatory descriptions provided

[2] See the spectrographic illustration of coarticulation in the Appendix.

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 1925:186-7).

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. That the presence of svarabhakti favors the perceptual recovery of underlying /Cɾ/ is reaffirmed in a more recent cross-linguistic study by N. Hall (in progress): "Inserting a vowel [fragment] gives the first consonant a stronger release gesture and the second consonant a stronger approach phase, enhancing their perceptibility."

While the data in (4) reflect Alonso's (1925) observations of Peninsular Spanish, coarticulation of /Cɾ/ clusters is also attested in contemporary American Spanish varieties. Lipski (1994) points out that in Highland Peru, "pronunciation of the groups /tr/, /pr/, /kr/ is partly determined by ethnolinguistic background. Among bilingual speakers, the /r/ in these combinations is a fricative or retroflex approximant, and in the case of /tr/ may fuse with the preceding consonant to produce a quasi-affricate" (320). 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 from Argüello's (1978) study of Highland Ecuadorian Spanish are given in (5):[3]

(5)	a.	*tres*	[tɹ̥]	'three'
		cuatro	[tɹ̥]	'three'
	b.	*vendrá*	[n^{d}ɹ]	's/he will come'
		saldrá	[l^{d}ɹ]	's/he will leave'
	c.	*padre*	[að^{ə}ɾ]	'father'
	d.	*premio*	[p^{ə}ɾ]	'prize'
		cruz	[k^{ə}ɾ]	'cross'

In casual speech, coarticulation affects clusters such as those in (5a,b), where the preceding coronal is realized as non-continuant. In (5c), however, the voiced coronal surfaces as a continuant after a preceding vowel, and the underlying cluster surfaces intact. Coarticulation also fails to affect heterorganic clusters, as in (5d). 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. 209-10), Central Costa Rica (p. 222), Guatemala (p. 265), Honduras (p. 272), Mexico (p. 279), and Paraguay (p. 308). Available phonetic descriptions suggest widespread coarticulation between /ɾ/ and a preceding homorganic stop, but fail to indicate similar behavior involving heterorganic C_1.

The empirical generalization emerging here suggests an implicational relationship between two types of /Cɾ/ coarticulation 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 only those clusters in which C_1 is a coronal non-continuant. For a given dialect, coarticulation of heterorganic clusters entails coarticulation of homorganic ones (with non-continuant C_1), but the opposite does not hold, as evidenced in by the data (5).

3. Gestural timing in phonetically-based Optimality Theory

In this section, we argue that the phonetic realizations of Spanish /Cɾ/ derive from the relative timing of the articulatory gestures associated with each member of the cluster. First, we motivate a particular approach to

[3] In accordance with the convention of Hispanic linguistics, Argüello employs [ř] and [ř̥] to represent voiced and voiceless variants, respectively, of the *r asibilada* (assibilated/fricative *r*). For consistency, we continue to use Alonso's transcription of the coarticulated rhotic as [ɹ] and [ɹ̥] in (5a,b), and we also indicate the lack of coarticulation with the [^{ə}] fragment in (5c,d).

gestural coordination, and then in Section 4, we show how the prosodic, segmental, and stylistic influences may be accounted for.

3.1 The role of gestural timing in /Cɾ/ *cluster realization*

Articulatory Phonology (Browman and Goldstein 1989, 1990, 1991, 1992) provides a framework within which to examine issues of gestural timing. In this model, gestures are dynamically defined articulatory movements that produce a constriction in the vocal tract. Three aspects of the gestural model are relevant for an analysis of Spanish /Cɾ/ realizations. First, articulatory gestures have internal duration, a property represented abstractly in terms of a 360° cycle. Phonetic timing is thus intrinsic to the phonological representation, and gestures are phonological primitives as well as units of articulation. This property sets Articulatory Phonology apart from most theories of phonology which relegate phonetic timing to an implementation component derivationally ordered after the phonology proper. Second, adjacent gestures are temporally coordinated with respect to each other and may exhibit varying degrees of overlap. Finally, consonantal articulations are superimposed on vocalic gestures, which are themselves articulatorily adjacent (Gafos 1999).

Following Cho (1998b:35), we assume that throughout the course of first language acquisition, learners construct a permissible range of overlap between adjacent gestures and that this range is encoded in lexical entries in terms of a Phase Window (Byrd 1994, 1996). Figure 1 illustrates three hypothesized patterns of gestural overlap between adjacent /C/ and /ɾ/, where the dotted lines delineate the lexically specified Phase Window:

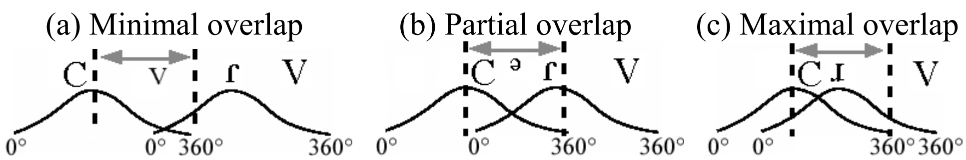


Figure 1: *Three patterns of gestural overlap*

While minimal overlap in (a) permits a greater recovery of the overlapping vowel gesture (not shown in the diagrams), partial overlap in (b) yields only a reduced vowel fragment. On the other hand, maximal overlap in (c) shifts the 0° onset of the /ɾ/-gesture outside the Phase Window, which precludes the svarabhakti fragment and results in coarticulation of the /Cɾ/ sequence.

The timing-based account in Figure 1 provides a phonetic explanation for both the existence of svarabhakti vowels and the nature of coarticulation. The fact that consonantal gestures are superimposed on vocalic gestures explains why svarabhakti is always a continuation of the formant structure present on the opposite side of the tap constriction. Both the nuclear vowel and the svarabhakti fragment stem from the same vocalic gesture, and the superimposed tapping gesture produces a brief interruption separating the two.[4] The assimilatory behavior observed in coarticulated clusters receives a straightforward explanation as the effects of gestural overlap. According to Browman and Goldstein (1990:360), gestures in casual speech are expected "to show decreased magnitudes (in both space and time) and to show increasing temporal overlap." Many types of casual speech alternations, such as deletions, assimilations, and weakenings, can be seen as the consequences of gestural reduction and overlap. On this view, the assimilations observed under coarticulation plausibly result from maximal overlap, as illustrated in Figure 2.

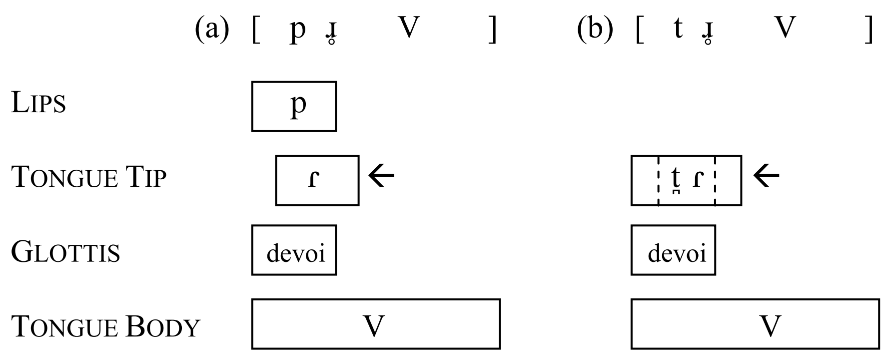


Figure 2: *Maximal overlap in* [pɹ̥] *(a) and* [tɹ̥] *(b) clusters*

In the above gestural representation, the activity of each relevant articulator is depicted on a separate tier, whose labels appear on the left. Boxes represent gestures, and the length of a box denotes the period of time during which the articulator is under active control. The arrow indicates that the tongue tip gesture for /ɾ/ has shifted leftward such that it coincides temporally with preceding gestures. Dotted lines denote overlap on the same tier.

Consider first the progressive devoicing of the rhotic. In Figure 2, coordination of the glottal devoicing gesture with the bilabial and dental closure gestures is responsible for the voicelessness of [p] and [t̪], respectively. Rhotic devoicing stems from greater overlap between the tongue tip gesture for /ɾ/ and the glottal devoicing gesture of the preceding consonant. The fact that overlap-induced devoicing is limited to casual speech lends support to the gestural explanation, given that casual speech

---

[4] A similar gestural explanation is proposed by Steriade (1990) and more recently Bradley (1999, 2001, 2002, to appear) and Hall (in progress). Such an account concords with the definition of Spanish /ɾ/ proposed by Gili Gaya (1921:279): "[E]s un sonido vocálico interrumpido por una oclusión alveolar, sonora, más or menos intensa [It is a vocalic sound interrupted by an alveolar contact that is voiced and more or less tense]."

is characterized by greater overlap. With respect to clusters in which the initial consonant is an underlying dental stop, coarticulation with the following rhotic results in an alveolar quasi-affricate [tɹ̥]. In the gestural model, overlap between adjacent gestures engaging the same articulator will produce *blending* of the characteristics of the two gestures, which "shows itself in spatial changes in one or both of the overlapping gestures" (Browman and Goldstein 1990:362). The retraction of dental stops when overlapped with a following apicoalveolar /ɾ/ plausibly reflects a compromise between the lexically specified constriction locations of the adjacent tongue tip gestures, as shown in Figure 2b.

3.2 Proposed constraints

Cho's (1998a,b) constraint-based analysis of Korean palatalization offers a means of evaluating gestural overlap in the Correspondence-theoretic version of OT (McCarthy and Prince 1995). In the present account, we assume that intergestural timing relevant to /Cɾ/ clusters is governed by the constraints in (6):[5]

(6) a. IDENT(timing)
The relative timing of gestures in the output must fall within the lexically specified Phase Window, which determines a permissible range of gestural overlap.

b. OVERLAP
Adjacent consonantal gestures must be maximally overlapped.

As shown in Tableau 1 below, the faithfulness constraint IDENT(timing) in (6a) bans any output timing relationship in which the 0° onset of the /ɾ/-gesture falls outside the lexically specified Phase Window. Given that phonetic timing is a continuous dimension, the optimal candidates in Tableau 1 should be interpreted as abstractions denoting a range of intermediate degrees of gestural overlap. Therefore, IDENT(timing) will permit a certain amount of variability as long as the timing relation falls

[5] Adamantios Gafos (personal communication) suggests that the "timing" predicate in (6a) and the notion of overlap in (6b) should be formally related in terms of representational primitives (cf. the gestural coordination constraints of Gafos 2002, which refer to specific temporal *landmarks* within gestures such as ONSET, TARGET, C-CENTER, etc.). In contrast, the constraints in (6) assume a lexically-specified Phase Window in the sense of Byrd (1994, 1996), which defines a permissible, gradient *target range* within which the relative timing of adjacent gestures must fall (see the discussion surrounding Figure 1). The phonetic alignment constraints of Zsiga (2000) also specify gestural timing in terms of temporal ranges, although her model assumes that timing is assigned in a post-phonological phonetic implementation component and not specified in the lexical representation. For more discussion, see Section 5.

within the lexically specified Phase Window. (On the influence of prosodic and segmental context, see Section 4.1 below.)

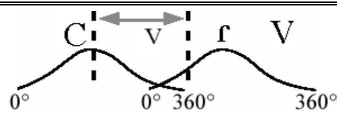
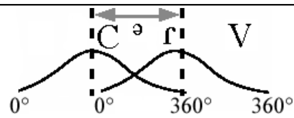
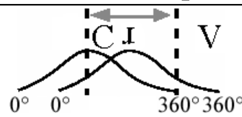
	IDENT(timing)
 ☞ Minimal overlap	
 ☞ Partial overlap	
 Maximal overlap	*!

Tableau 1: *Variable duration of the svarabhakti vowel fragment*

OVERLAP in (6b) is an articulatory markedness constraint that prefers maximal coarticulation between adjacent consonantal gestures. While IDENT(timing) captures the requirements of perceptual recoverability, OVERLAP yields an articulatory advantage in terms of *parallel transmission* (Liberman, Cooper, Shankweiler, and Studdert-Kennedy 1967). That is, coarticulation allows information about several linguistic units to be transmitted simultaneously. In the case of /Cɾ/ clusters, OVERLAP is responsible for the assimilatory effects associated with maximal overlap:

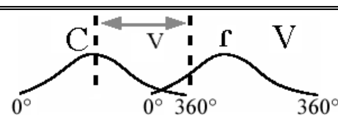
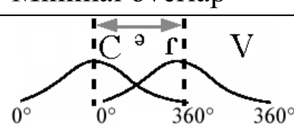
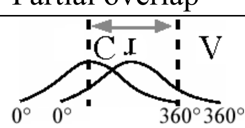
	OVERLAP
 Minimal overlap	*!
 Partial overlap	*!
 ☞ Maximal overlap	

Tableau 2: *Coarticulation of the cluster*

In sum, the constraints proposed above give formal expression to two competing influences on intergestural timing, which Chitoran et al. (to appear) characterize as follows: "the first is the need to ensure recoverability of linguistic units from the signal, and the second is the need to encode and transmit information at a high rate" (26). Because OT is built upon the notion of competition and conflict resolution, it is an appropriate framework for analyzing the intergestural timing patterns of Spanish /Cɾ/ clusters, to which we now turn.

4. Analysis of Spanish /Cɾ/ realizations

In the analysis outlined thus far, faithfulness to input timing in (6a) conflicts with the articulatory imperative in (6b) that adjacent consonantal gestures should be maximally overlapped, such that the higher ranking constraint determines the gestural timing relationship of /Cɾ/ clusters in the output. In this section, we further develop this approach by integrating prosodic, segmental, and stylistic influences on intergestural timing, thereby providing a unified account of the patterns discussed in Section 2.

4.1 Prosodic and segmental effects on svarabhakti

In the Phase Window model, minimal and partial overlap scenarios actually comprise a continuous range of intermediate degrees at which the gestures for /C/ and /ɾ/ may be timed with respect to one another. Variability in the duration of svarabhakti, as observed by Gili Gaya (1921) and Malmberg (1965), stems from the variable timing of gestures during speech production. Recall from Section 2.1, however, that the duration of svarabhakti is influenced by the cluster's prosodic position (word-initial vs. word-internal, stressed vs. unstressed syllable) and segmental composition (back-to-front vs. front-to-back order).

These observations are corroborated by an independent study of gestural timing in Georgian stop-stop sequences. Chitoran et al. (to appear) show that perceptual recoverability conditions place limits on the degree of gestural overlap between adjacent stops in clusters that appear word-initially (vs. word-internally) and that exhibit a back-to-front (vs. front-to-back) order of constriction location. Their explanation for these patterns is as follows. First, word onsets are potential utterance onsets, in which case no preceding vowel is available to provide formant transitions into the first consonant (see Redford and Diehl 1999). Furthermore, word onsets have been shown to be important for lexical access (Marlsen-Wilson 1987). Therefore, it is plausible that minimal overlap is favored word-initially so as to preserve more acoustic information about each consonant of the cluster. Second, gestural overlap in clusters exhibiting a back-to-front order entails that the acoustic release of the first consonant will be perceptually obscured because the second constriction lies ahead of the first constriction in the vocal tract. In contrast, overlap in clusters with a front-to-back order does not obscure the acoustic release of the first consonant because the second constriction lies behind the first. This difference is illustrated in Figure 3 below for two /Cɾ/ clusters. Again, minimal overlap is plausibly favored more so in dorsal+/ɾ/ sequences than in labial+/ɾ/ in order to ensure recoverability of the underlying cluster. (See Byrd 1992 and Surprenant and Goldstein 1998 for further evidence that back-to-front order permits less overlap between consonantal gestures than front-to-back.)

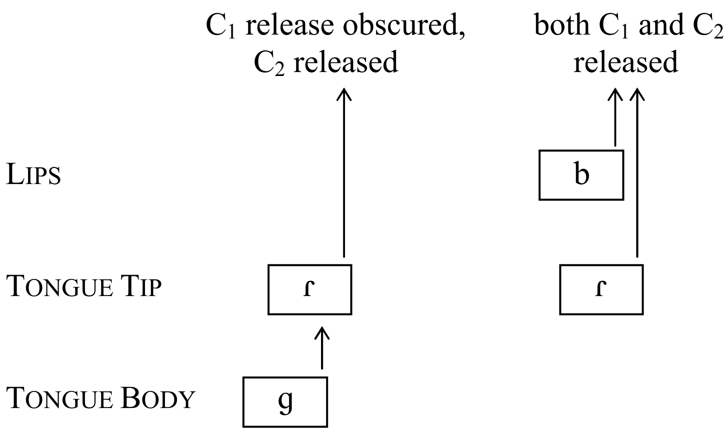


Figure 3: *Gestural overlap obscures the acoustic release of C_1 in the back-to-front cluster* /ɡɾ/ *but not in the front-to-back cluster* /bɾ/

Although Chitoran et al. (to appear) do not examine the possible effects of stress, it seems plausible to expect less overlap in clusters belonging to stressed syllables, again for perceptual reasons. The claim that stressed syllables are perceptually prominent positions vis-à-vis unstressed ones finds support, for example, in American English tapping: intervocalic /t/ and /d/ undergo temporal reduction to [ɾ] only when in onset position of unstressed syllables, whereas the process fails to affect stressed syllables (see Inouye 1995). With respect to overlap in /Cɾ/ clusters, we hypothesize that the prominence of stressed syllables favors the preservation of acoustic information about each consonant. As in the case of word-initial and back-to-front clusters, minimal overlap is the type of gestural coordination that achieves this.

We propose to integrate the role of perceptual recoverability into the Phase Window model of intergestural timing presented in Section 3. Recall that the very function of Phase Windows is to define a permissible range of overlap for adjacent gestures. Since what counts as a permissible range is constrained by the requirements of perceptual recoverability, it is plausible that Phase Windows are constructed for individual /Cɾ/ clusters according to the degree of perceptibility of the phonetic contexts in which they appear. Specifically, those clusters appearing word-initially or in stressed syllables, as well as those with a back-to-front order of constriction location, are assumed to have a delayed Phase Window, as in Figure 4b, whereas clusters in other contexts have an earlier Phase Window, as in Figure 4a:

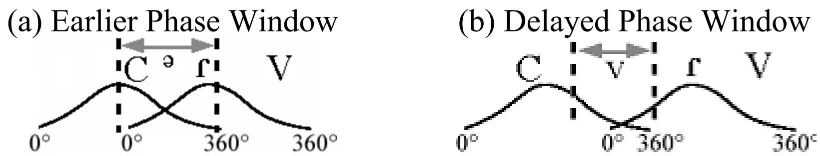


Figure 4: *A delayed Phase Window favors minimal overlap and the perceptual recoverability of* /Cɾ/

The effect of IDENT(timing), then, will be to rule out partial overlap in clusters with a delayed Phase Window, while allowing the usual variability between partial and maximal overlap in other clusters. In this way, greater perceptibility—i.e., longer svarabhakti—ensures that input /Cɾ/ will be recovered in the acoustic representation.

As Tableau 3 demonstrates, the ranking IDENT(timing) » OVERLAP permits a range of gestural overlap in accordance with the lexically specified Phase Window for the /Cɾ/ cluster in question. Since partial overlap is disfavored by the delayed Phase Windows of clusters appearing word-initially or in stressed syllables, high-ranking IDENT(timing) rules out candidates (e) and (k), respectively. The coarticulated candidates (c), (f), (i), and (l) are also ruled out by faithfulness because maximal overlap places the /ɾ/-gesture outside any Phase Window, whether early and delayed. As a result, variability between shorter and longer svarabhakti is optimal in non-initial position (a,b) and in unstressed syllables (g,h), while longer ones are favored word-initially (d) and under stress (j). For reasons of space, we omit clusters differing in the order of constriction location since their evaluation is identical to those shown in Tableau 3.

	IDENT(timing)	OVERLAP
☞ a. /CɾV/ → CvɾV		*
☞ b. CᵊɾV		*
c. CɹV	*!	
☞ d. /#CɾV/ → #CvɾV		*
e. #CᵊɾV	*!	*
f. #CɹV	*!	
☞ g. /CɾV̆/ → CvɾV̆		*
☞ h. CᵊɾV̆		*
i. CɹV̆	*!	
☞ j. /CɾV́/ → CvɾV́		*
k. CᵊɾV́	*!	*
l. CɹV́	*!	

Tableau 3: *Longer svarabhakti in word-initial and stressed* /CɾV/ *demisyllables than in word-internal and unstressed ones*

4.2 Stylistic and segmental effects on coarticulation

Recall Alonso's (1927:186-9) observation, discussed in Section 2, that coarticulation of Spanish /Cɾ/ is characteristic of casual speech, while in careful speech the perceptual integrity of the cluster is preserved. Furthermore, two major patterns of coarticulation were identified among varieties of Spanish: coarticulation of any /Cɾ/ cluster (Peninsular and Peruvian varieties) versus coarticulation of /ɾ/ with only a preceding homorganic stop (Highland Ecuadorian and other Latin American varieties). We propose to capture this segmental effect in terms of an additional articulatory markedness constraint targeting the latter type of cluster:

(7) *FAST/SAME (adapted from Bradley 2001; cf. also Steriade 1995)
Avoid faster-than-usual transitions between adjacent periods of greater stricture involving the same articulator.

The claim that articulatory transitions are more marked between homorganic consonants than heterorganic ones is supported by consonantal transition phenomena in Sierra Popoluca, a Zoquean language spoken in Mexico (Elson 1947, 1956; Foster and Foster 1948). Consonant clusters in this language are realized with an intervening open transition if the consonants are heterorganic, while homorganic clusters lack such a transition, as shown in (8):

(8)	a.	kɛk^h.paʔ	'it flies'	miɲ^ə.paʔ	'he comes'
	b.	kɛk.gak^h.paʔ	'it flies again'	ʔaŋ.kiʔ	'yard'

In (8a), the open transition is realized as aspiration after the voiceless velar stop and as a short schwa-like vowel after the palatal nasal. The homorganic stop sequences [k.g] and [ŋ.k] in (8b) lack an open transition.

While OVERLAP is a general constraint favoring maximal overlap of adjacent consonantal gestures, *FAST/SAME is more specific, targeting only those homorganic clusters in which an open transition intervenes between two periods of maximal oral constriction. The additional articulatory markedness of the latter type of cluster is responsible, we claim, for both the absence of open transitions in Sierra Popoluca, seen in (8b), and the coarticulation of homorganic /Cɾ/ in Highland Ecuadorian Spanish and other Latin American varieties, as in (5a,b).

The stylistic variation effects on coarticulation follow straightforwardly in the constraint-based account proposed here. As demonstrated in Section 4.1, when faithfulness to input Phase Windows is highly ranked in careful speech, perceptually optimal timing is enforced. In casual speech, however, IDENT(timing) is subordinate to articulatory markedness. As shown in Tableau 4 below, high-ranking OVERLAP yields coarticulation of any /Cɾ/ cluster. Maximal overlap of /ɾ/ is optimal after heterorganic consonants in (c) and homorganic continuants in (f), as well as homorganic stops in (i,l). When only *FAST/SAME dominates IDENT(timing), on the other hand, coarticulation is restricted to only those clusters in which C_1 is a homorganic stop. As shown in Tableau 5, candidates (i,l) exhibit coarticulation, while variable svarabhakti obtains after heterorganic C_1 in (a,b) and homorganic non-continuants in (d,e).

	OVERLAP	IDENT(timing)	*FAST/SAME
a. /pɾV/ → pvɾV	*!		
b. p^{ə}ɾV	*!		
☞ c. pɹ̥V		*	
d. /Vd̪ɾV/ → VðvɾV	*!		
e. Vð^{ə}ɾV	*!		
☞ f. VðɹV		*	
g. /t̪ɾV/ → t̪vɾV	*!		*
h. t̪^{ə}ɾV	*!		*
☞ i. tɹ̥V		*	
j. /nd̪ɾV/ → n̪d̪vɾV	*!		*
k. n̪d̪^{ə}ɾV	*!		*
☞ l. n^{d}ɹV		*	

Tableau 4: *Coarticulation of* /ɾ/ *after any consonant*

	*FAST/SAME	IDENT(timing)	OVERLAP
☞ a. /pɾV/ → pvɾV			*
☞ b. p^{ə}ɾV			*
c. pɹ̥V		*!	
☞ d. /Vd̪ɾV/ → VðvɾV			*
☞ e. Vð^{ə}ɾV			*
f. VðɹV		*!	
g. /t̪ɾV/ → t̪vɾV	*!		*
h. t̪^{ə}ɾV	*!		*
☞ i. tɹ̥V		*	
j. /nd̪ɾV/ → n̪d̪vɾV	*!		*
k. n̪d̪^{ə}ɾV	*!		*
☞ l. n^{d}ɹV		*	

Tableau 5: *Coarticulation of* /ɾ/ *only after homorganic stops*

The constraint-based account effectively captures the implicational relationship observed in the coarticulation of different types of /Cɾ/ clusters. Candidates incurring a violation of *FAST/SAME are always a subset of the candidates violating the more general OVERLAP constraint. Given this subset relation, no ranking of the constraints can produce coarticulation of heterorganic /Cɾ/ without also producing coarticulation of /ɾ/ after homorganic stops. In an earlier, non-constraint-based analysis of Highland Ecuadorian Spanish, Bradley (1999) argues that /ɾ/ is overlapped by adjacent consonantal gestures, with overlap resulting in gestural blending next to coronals. Such a general statement of gestural overlap turns out to be descriptively inadequate, since it also predicts coarticulation of /ɾ/ after both homorganic continuants and heterorganic consonants—contrary to the facts in (5c) and (5d), respectively. The analysis developed here is superior, however, because gestural overlap is governed by specific articulatory markedness constraints that are subject to ranking permutation. *FAST/SAME targets [tɹ̥] and [n^{d}ɹ] sequences independently of other types of cluster, while OVERLAP affects all /Cɾ/ clusters equally. As a result, the interleaving of IDENT(timing) between *FAST/SAME and OVERLAP in Tableau 5 successfully predicts coarticulation of the former type of cluster versus svarabhakti elsewhere.

4.3 The development of lexicalized copy vowels

Although minimal overlap between the adjacent gestures of a /Cɾ/ cluster creates the appearance of two vowels, we assume that the longer vowel fragment does not create a new syllable. In an extensive cross-linguistic survey of svarabhakti phenomena, Hall (in progress) observes that svarabhakti vowels are *metrically cohesive* with the adjacent full vowel whose quality they copy. That is, languages tend to count svarabhakti and the original vowel as one for stress purposes. An example from Spanish suffices to illustrate this point:

(9) *hidrómetro* [i.ðɾó.me.t̪oɾo] 'hydrometer'
*[i.ðɾó.me.t̪o.ɾo]

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 the svarabhakti fragment surfacing in the final /t̪ɾ/ cluster in (9) were to create a new syllable, then stress would fall outside the three-syllable window yielding ungrammatical results. This evidence suggests that svarabhakti is not the result of a synchronic process of vowel epenthesis. On this view, the copy vowels in (3) are more appropriately analyzed as occasional historical developments whereby longer svarabhakti fragments are phonologically reinterpreted over time as full lexical vowels, as illustrated in Figure 5 for the hypothetical sequence

/t̪ɾa/. Once phonological restructuring has taken place as in (b), the Phase Window of permissible gestural overlap for /t̪/ and /ɾ/ is dissolved because the epenthetic vowel has broken up the onset cluster and the associated consonantal gestures are no longer adjacent.

(a) Speaker produces svarabhakti fragment of variable duration

(b) Listener reinterprets longer fragment as lexical vowel

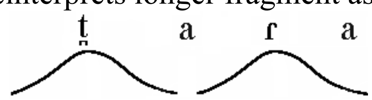


Figure 5: *Before (a) and after (b) stages in* /t̪ɾa/ > /t̪aɾa/

Crucial to the above 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 (1991:331-3) 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 5, the svarabhakti fragments 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.

5. Discussion

Here we discuss the present account with respect to the broader theoretical issue of phonetic detail in phonology and point out some avenues for future research.

5.1 Phonetic detail in phonology

In the analysis developed here, input morphemes already have their gestural timing relations fully and reliably specified, so that faithfulness, i.e., the IDENT(timing), can depend on them. This assumption is fully consistent with Browman and Goldstein's model of Articulatory Phonology, in which gestures are both units of articulation and primitives of phonological organization, and timing relationships are specified directly in the gestural score. On this view, a predictable non-contrastive property of phonetic detail—intergestural timing—is incorporated directly into the phonological representation. As John McCarthy (personal communication) points out, however, assuming fully-specified inputs runs counter to the Richness of The Base hypothesis of OT which forbids placing restrictions directly on input representations. Moreover, the direct appeal to phonetic detail goes against conventional models of phonology in which underlying representation is assumed to be devoid of non-contrastive properties altogether.

In a possible alternative to the account developed here, faithfulness to Phase Windows in the input might be supplanted by phonetic constraints that determine gestural timing relations directly in the output. For example, the alignment constraints of Zsiga (2000) specify gestural coordination in a separate phonetic implementation component, derivationally ordered after the phonology proper. On this view, phonological representations remain abstract, categorical, and timeless, while implementation constraints supply quantitative, non-contrastive temporal specifications to yield a fully-specified phonetic representation. In order to decide between the two competing approaches, future investigation must ultimately evaluate both on the basis of a wider range of empirical test cases. Preliminary evidence supporting the existence of Phase Windows is found in recent analyses of derived environment effects in Korean palatalization (Cho 1998a,b) and in the Norwegian "retroflex rule" (Bradley, to appear). Both of these accounts show that faithfulness to the timing specifications of underived input morphemes explains why overlap-induced morphophonological alternations systematically appear only in derived (heteromorphemic) environments: Phase Windows are specified for gestures within a single morpheme but not across two separate ones. IDENT(timing) is active in the former case, while OVERLAP is free to induce the relevant alternation in the latter.[6]

Finally, Jill Beckman (personal communication) suggests a potential problem with respect to the Phase Window approach advocated in this paper. In Section 4.1, stressed /CɾV/ demisyllables were claimed to exhibit a delayed Phase Window, as shown in Figure 4b, which favors svarabhakti fragments of longer duration. Now, imagine the hypothetical case of a morphologically derived form in which stress is shifted to a different syllable. Such a scenario seems to make the wrong prediction, namely that IDENT(timing) would force adherence to the original delayed Phase Window even though the /CɾV/ demisyllable in question is no longer stressed. Although we are not aware of any examples of stress shift in Spanish that could be brought to bear on the issue, investigation of /Cɾ/ clusters in other languages might be able to provide some insight.

[6] However, neither of these studies provides an explicit comparison with the alternative approach in which gestural coordination constraints supplant IDENT(timing).

5.2 Future research on Spanish /Cɾ/ *cluster realization*

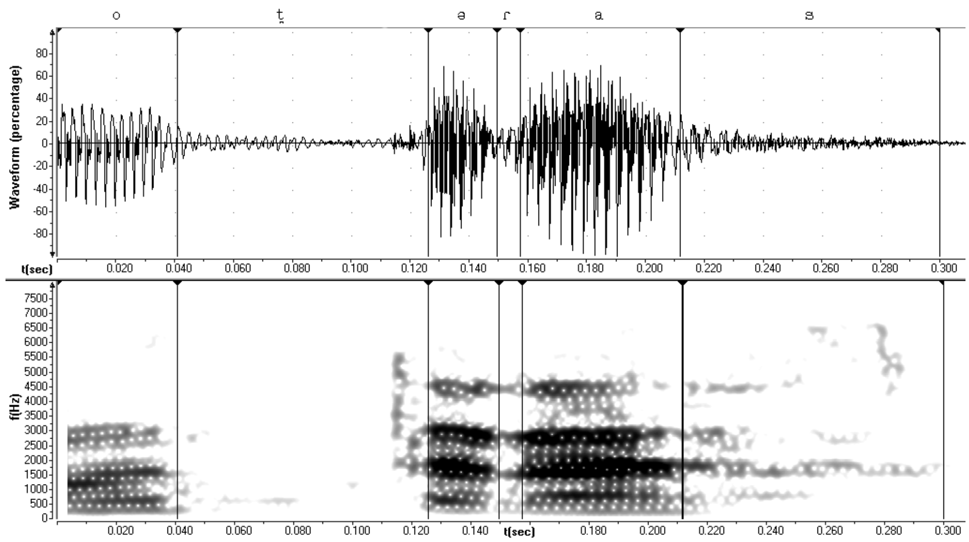
The analysis presented in Section 4 captures the influences on /Cɾ/ realization observed by Gili Gaya (1921) in terms of different Phase Window specifications according to the prosodic position and segmental make-up of the cluster. However, Gili Gaya's early study needs to be replicated on an expanded set of empirical data with appropriate statistical procedures in order to determine the significance of the purported prosodic and segmental effects. Given the extreme variability of the duration of svarabhakti vowel fragments, a much larger corpus of /Cɾ/ tokens is likely to be necessary for any statistically significant effects to emerge. In addition, further research is required to verify the patterns of /Cɾ/ coarticulation across different Spanish varieties, as discussed in Section 2.2. An empirical study of these issues is currently being carried out by the present authors.

6. Conclusion

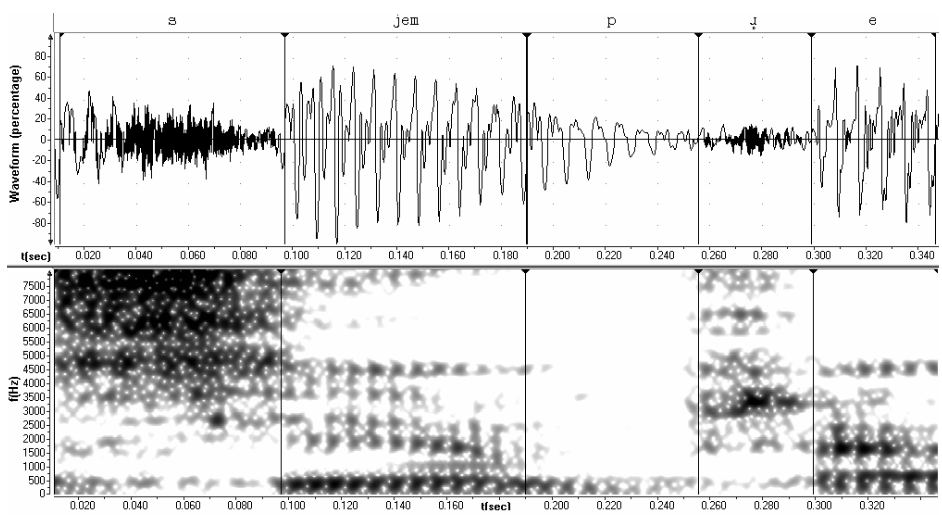
In this paper, we have explained the behavior of Spanish /Cɾ/ by incorporating functional phonetic factors such as intergestural timing into a formal OT analysis. The main advantage of such a direct approach is that it exposes the phonetic motivation underlying the possible realizations of such clusters. Furthermore, the proposed account captures the effects of prosodic and segmental context in terms of perceptual recoverability requirements and explains stylistic and dialectal variation through the re-ranking of a small set of conflicting, universal constraints.

Appendix

Costa Rican Spanish *otras* 'others' illustrating svarabhakti vowel fragment in non-coarticulated [t̪^{ə}ɾ] cluster:



Costa Rican Spanish *siempre* 'always' illustrating coarticulation of [pɹ̥] cluster:



References

Alonso, Amado. 1925. "El grupo *tr* en España y América". *Homenaje a Menéndez Pidal* II.167-191. Madrid.

Argüello, Fanny. 1978. *The žeísta Dialect of Spanish Spoken in Ecuador: A phonetic and phonological study.* Ph.D. dissertation, The Pennsylvania State University.

Beckman, Mary, and John Kingston, eds. 1990. *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech.* Cambridge: Cambridge University Press.

Bradley, Travis G. 1999. "Assibilation in Ecuadorian Spanish: A phonology-phonetics account". *Formal Perspectives on Romance Linguistics* ed. by J.-Marc Authier, Barbara E. Bullock, and Lisa A. Reed, 57-71. Amsterdam: John Benjamins.

----------. 2001. *The Phonetics and Phonology of Rhotic Duration Contrast and Neutralization.* Ph.D. dissertation, The Pennsylvania State University.

----------. 2002. "Gestural Timing and the Resolution of /Cɾ/ Clusters in Romance". Paper presented at the *Linguistic Symposium on Romance Languages 32,* University of Toronto, April 19-21, 2002. [ROA #529]

----------. To appear. "Gestural Timing and Derived Environment Effects in Norwegian Clusters". *WCCFL 21 Proceedings* ed. by L. Mikkelsen and C. Potts, 101-114. Somerville, MA: Cascadilla Press.

Browman, Catherine, and Louis Goldstein. 1989. "Articulatory Gestures as Phonological Units". *Phonology* 6.201-252.

----------. 1990. "Tiers in Articulatory Phonology, with Some Implications for Casual Speech". Beckman and Kingston 1990. 341-376.

----------. 1991. "Gestural Structures: Distinctiveness, Phonological Processes, and Historical Change". *Modularity and the Motor Theory of Speech Perception* ed. by Ignatius G. Mattingly and Michael Studdert-Kennedy, 313-338. Hillsdale, New Jersey: Lawrence Erlbaum.

----------. 1992. "Articulatory Phonology: An overview". *Phonetica* 49.155-180.

Byrd, Dani. 1992. "Perception of Assimilation in Consonant Clusters: A gestural model". *Phonetica* 49.1-24.

----------. 1994. *Articulatory Timing in English Consonant Sequences.* Ph.D. dissertation, University of California, Los Angeles.

----------. 1996. "A Phase Window Framework for Articulatory Timing". *Phonology* 13.139-169.

Chitoran, Ioana, Louis Goldstein, and Dani Byrd. To appear. "Gestural Overlap and Recoverability: Articulatory evidence from Georgian". *Papers in Laboratory Phonology 7* ed. by C. Gussenhoven, T. Rietveld, and N. Warner. Cambridge: Cambridge University Press.

Cho, Taehong. 1998a. "Intergestural Timing and Overlap in Korean Palatalization: An Optimality-Theoretic approach". *Japanese/Korean Linguistics* 8 ed. by David Silva, 261-276. Stanford: CSLI Publications.

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

Colantoni, Laura. 2001. *Mergers, Chain Shifts, and Dissimilatory Processes: Palatals and rhotics in Argentine Spanish.* Ph.D. Dissertation, University of Minnesota.

Elson, B. 1947. "Sierra Popoluca Syllable Structure". *International Journal of American Linguistics* 13:13-17.

----------. 1956. *Sierra Popoluca Morphology.* Ph.D. Dissertation, Cornell University.

Foster, M. L., and G. M. Foster. 1948. *Sierra Popoluca Speech.* Washington, D.C.: U.S. Government Printing Office, Smithsonian Institution, Institute of Social Anthropology.

Gafos, Adamantios. 1999. *The Articulatory Basis of Locality in Phonology.* New York: Garland.

----------. 2002. "A Grammar of Gestural Coordination". *Natural Language and Linguistic Theory* 20.269-337.

Gili Gaya, Samuel. 1921. "La *r* simple en la pronunciación española". *Revista de Filología Española* VIII.271-280.

Hall, Nancy. In progress. *Svarabhakti as Segmental Overlap*. Ph.D. dissertation, University of Massachusetts, Amherst.

Hammond, Robert. 1999. "On the Non-Occurrence of the Phone [r̃] in the Spanish Sound System". *Advances in Hispanic Linguistics* ed. by Javier Gutiérrez-Rexach and Fernando Martínez-Gil, 135-151. Somerville, Massachusetts: Cascadilla Press.

----------. 2000. "The Phonetic Realizations of /rr/ in Spanish: A Psychoacoustic Analysis". *Hispanic Linguistics at the Turn of the Millennium* ed. by Héctor Campos, Elena Herburger, Alfonso Morales-Front and Thomas J. Walsh, 80-100. Somerville, Massachusetts: Cascadilla Press.

----------. To appear-a. "The Phantom of the Spanish Sound System - The Voiced Alveolar Multiple Vibrant". *Probus.*

----------. To appear-b. "A Psychoacoustic Analysis of The Phonetic Realizations of Spanish /rr/". *Lingua.*

Harris, James. 1983. *Syllable Structure and Stress in Spanish: A nonlinear analysis.* Cambridge, Massachusetts: MIT Press.

----------. 1995. "Projection and Edge Marking in the Computation of Stress in Spanish". *The Handbook of Phonological Theory* ed. by John Goldsmith, 867-887. Cambridge, Massachusetts: Blackwell.

Inouye, Susan. 1995. *Trills, Taps and Stops in Contrast and Variation.* Ph.D. Dissertation, University of Los Angeles, California.

Lenz, Rodolfo. 1892. "Chilenishe Studien". *Phonetische Studien* V.272-293.

Liberman, A., Cooper, F.S., Shankweiler, D., and Studdert-Kennedy, M. 1967. "Perception of the Speech Code". *Psychological Review* 74, 431-461.

Lipski, John. 1994. *Latin American Spanish*. London: Longman.

Malmberg, Bertil. 1965. *Estudios de fonética hispánica*. (= *Collectanea Phonetica*, I.) Madrid: Consejo Superior de Investigaciones Científicas.

Marslen-Wilson, W.D. 1987. "Functional Parallelism in Spoken Word-Recognition". *Cognition* 25.71-102.

McCarthy, John, and Alan Prince. 1995. "Faithfulness and Reduplicative Identity". *University of Massachusetts Occasional Papers 18: Papers in Optimality Theory* ed. by Jill Beckman, Laura Walsh Dickey and Suzanne Urbanczyk, 249-384. Amherst, Massachusetts: Graduate Linguistic Student Association.

Navarro Tomás, Tomás. 1918. "Diferencias de duración entre las consonantes españolas". *Revista de Filología Española* V.367-393.

Prince, Alan, and Paul Smolensky. 1993. "Optimality Theory: Constraint interaction in generative grammar". Ms., Rutgers University, New Brunswick, New Jersey and University of Colorado, Boulder, Colorado.

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

----------. 1993. *Tratado de fonología y fonética españolas*. Madrid: Editorial Gredos.

Redford, M.A. and Diehl, R.L. 1999. "The Relative Perceptual Distinctiveness of Initial and Final Consonants in CVC Syllables". *Journal of the Acoustical Society of America* 106.1555-65.

Steriade, Donca. 1990. "Gestures and Autosegments: Comments on Browman and Goldstein's paper". Beckman and Kingston 1990. 382-397.

----------. 1995. "Licensing Retroflexion". Ms., University of California, Los Angeles.

Surprenant, A.M. and Goldstein, L. 1998. "The Perception of Speech Gestures". *Journal of the Acoustical Society of America* 104.518-529.

Willis, Erik and M. Begoña Pedrosa. 1998. "An Acoustic Analysis of the Rhotic System of Spanish: Isolated words, read texts and spontaneous speech". Paper presented at the *Second Hispanic Linguistics Symposium*, Ohio State University, October 9-11, 1998.

Zsiga, Elizabeth. 2000. "Phonetic Alignment Constraints: Consonant overlap and palatalization in English and Russian". *Journal of Phonetics* 28.69-102.