A gestural-based analysis of /e/ prosthesis in word-initial /sC/ loanwords in Spanish

Mark Gibson
Universidad de Navarra
mgibson@unav.es

Abstract

A theoretical analysis of synchronic /e/ prosthesis in non-native word-initial /sC/ sequences in Spanish is provided. In light of new evidence, it is argued that the appearance of the word-initial /e/ to the left margin of the non-native /sC/ onset is a perception-based phenomenon, and not triggered directly as a productive operation of the grammar. This assertion is robustly supported by the results of a series of perception tests involving 50 Spanish-speaking children in which the subjects routinely reported hearing an illusory vowel to the left of the non-native complex onset. These data are in line with past accounts of perceptual epenthesis and suggest that the auditory signal in some instances may be distorted by a high-level bias for native sound patterns. At the time of production, then this means that the dominant grammatical constraint governing the appearance of /e/ is based on principles of input/output correspondence, and not productive constraints referencing structural phonotactics.

Keywords: L1 biases, relative timing patterns, perceptual epenthesis, Spanish.
Received: 13.IX.2012 – **Accepted:** 18.XII.2012

Table of Contents

1 A grammatical argument for /e/ alignment in word-initial /sC/ sequences
2 A perceptual argument for /e/ prosthesis
3 Expressing vowel prosthesis in a gestural-based grammar
4 A faithfulness account for vowel prosthesis in word-initial /sC/ loan words
5 Conclusion

Ianua. Revista Philologica Romanica
Vol. 12 (2012): 35–56
ISSN 1616-413X
http://www.romaniaminor.net/ianua/

© Romania Minor

BDD-A5302 © 2012 Romania Minor
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 17:58:36 UTC)

1 A grammatical argument for /e/ alignment in word-initial /sC/ sequences

It has gone widely undisputed that the emergence of a prosthetic [e] to the left margin of non-native word-initial [sC] sequences is correlated to stipulations governing the structure of Spanish syllables:

(1) Illicit word-initial /sC/ clusters in foreign loan words

[e]stop	'stop'	/s/ + coronal obstruent
[e]strés	'stress'	/s/ + coronal obstruent/liquid cluster
[e]spray	'spray'	/s/ + labial obstruent/liquid cluster
[e]sprin(t)	'sprint'	/s/ + labial obstruent/liquid cluster
[e]smart	'Smart'	/s/ + labial obstruent
[e]scáner	'scanner'	/s/ + dorsal obstruent
[e]spich	'speech'	/s/ + labial obstruent
[e]slogan	'slogan'	/s/ + lateral sonorant
[e]smoquin	'smoking'	/s/ + nasal sonorant

In these cases, [e] emerges systematically in all word-initial [sC] sequences. The fact that no native Spanish word begins with the [sC] sequence supports the general argument that [e] alignment is governed by principles which regulate the shape of Spanish onsets.

Top-down models invariably explain vowel prosthesis as a grammatically induced repair process triggered by a phonological stipulation which references some aspect of complex onset formation. The pertinent phonological information motivating such an operation has historically been based on an autosegmental approach to the phonology, and the means by which this information interacts with the grammar has been expressed theoretically by way of (i) linearly emerging structure building rules, (ii) the fulfillment of language-specific syllabic templates, and/or (iii) highly ranked universal markedness laws governing structural well-formedness. The vast majority, if not all grammatical justifications of vowel insertion hinge on the supposition that the [sC] structures are banned in Spanish complex onsets, yet detrimentally, none offers a coherent description of the phonological conditions on which such an assertion is determined.

Segment-based approaches to the phonology express the structural well-formedness of complex onset sequences by way of the linear arrangement of timing-independent spatial units. Most often, the correct configuration of segments involves some indirect reference to vocal tract constriction via constraints governing abstract sonority distances between the individual units and their subsequent ranking in a hierarchical-based production algorithm.[1] In this conception, there is no linear path which correlates categorical sonority values with

[1] Patterns of vocal tract constriction in complex onsets are projected onto the grammar as restrictions that stipulate a minimal sonority distance between the segments. First, segments are collocated on a sonority hierarchy on the basis of their respective manner specifications: Stops < Fricatives < Nasal < Liquid < Glide < Vowel. Later, each hierarchical position is allocated an abstract numerical value: (1) Stops < (2) Fricatives < (3) Nasal < (4) Liquid < (5) Glide < (6) Vowel.

© Romania Minor
http://www.romaniaminor.net/ianua/

BDD-A5302 © 2012 Romania Minor
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 17:58:36 UTC)

low-level patterns of vocal tract constriction, nor any programmed explanation as to how low-level spatio-temporal relationships may affect optimal outputs.

As I have pointed out in past articles, a number of conjectural problems arise which dispute the veracity of the purely grammatical explanation of /e/ prosthesis based on abstract sonority values. Principally, these models provide no way to justify the emergence of all optimal output sequences while simultaneously explaining the elimination of sub-optimal sequences.[2] Consider a restriction that establishes a minimal sonority distance of three between complex onset segments (Parker 2002). In this case, it is supposed that a listener of Spanish, upon processing various inputs, is first able to analyze and categorize the sonority properties of each individual segment and later, deduce an abstract rule which is capable of expressing the notion that only certain units may combine to form complex onset sequences based on this criterion. While computationally feasible, the explanatory power of a rule of this nature is put in check by the fact that a minimal sonority distance of three eliminates all [fricative/liquid] sequences, meaning [fl] and [fɾ], which are wholly admissible complex onset sequences in Spanish, must be explained away as exceptions to the general rule.[3] This last pitfall presents serious functional and theoretical difficulties for acquisition and learnability algorithms. Harris (1983) avoids this shortcoming by proposing a minimal sonority distance of two for Spanish complex onsets. The hitch with this explanation, however, is that *all* [fricative/liquid] sequences would surface optimally, of which only [fɾ] and [fl] constitute well-formed complex onsets. Neither minimal sonority distance addresses the elimination of [dl] onsets.[4]

Let us assume that a minimal sonority distance of three were a genuine grammatical principle governing the shape of Spanish syllables. Its high ranking in an immature grammar could, hypothetically, trigger vowel prosthesis in all /sC/ clusters. There is, after all, no possible combination of [sibilant/C] clusters which could satisfy this criterion. Nevertheless, this explanation cannot account for the fact that no repair process surfaces in loanwords containing a word-initial [θɾ] cluster: [ˈθɾi.leɾ], and *[eθ.ˈɾi.leɾ] *thriller*. Notice that according

This topic has been vigorously debated in the Autosegmental literature and is discussed specifically for Spanish in Harris (1983), Clements (1990), Colina (1995),Martínez-Gil (1996, 1997), Parker (2002) and Wright (2004) among others.

[2] The critique presented here should not be misconstrued to be a wholesale criticism of hierarchical-based grammatical models. The present analysis is merely a critical assessment of the ad hoc nature of individual constraints which have gained much currency in the contemporary phonological debate.

[3] Wright (2004) proposes that stops and /f/ be grouped in the same hierarchical position with regard to sonority, whereby a minimal sonority distance of three would justify, at least theoretically, the emergence of all optimal outputs. The consequence of such a separation, though, is a blanket proscription against all fricative/liquid sequences which makes no predictions as to why individual sequences are discarded. In this conception /s// liquid clusters are treated in the same way as /θ// liquid clusters. The drawback is that loanwords actually do exist in which the sequence [θɾ] appears with no surface-level modification, thriller [ˈθɾi.leɾ] as well as the phonetic adaptation of *Heathrow* realized as [hi.θɾow], implying that the absence of this sequence in the data record is not a corollary of the grammar but rather a convention of the lexicon.

[4] In most analyses, /dl/ is prevented by a ban preventing contiguous segments to share the certain [place] specifications (see *Obligatory Contour Principle*, Leben 1973).

BDD-A5302 © 2012 Romania Minor
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 17:58:36 UTC)

to a minimal sonority distance of three, [θɾ] would be just as ill-formed as any [sC] sequence.

The fact is that top-down analyses of complex onset formation based solely on the linear arrangement of spatial units cannot coherently justify the output data in Spanish complex onsets. Undeniably, grammatical constraints referencing sonority distances do reflect output patterns of vocal tract constriction. The problem is that there is no evidence that such abstractions can actually command the low-level articulatory procedures observed in complex onset formation. Because of these incongruities, the idea that injunctions governing sonority sequencing may play a role in triggering the emergence of a post-lexical repair process is theoretically tenuous.

Despite the theoretical shortcomings of the previous grammar-based approaches, the claim that prosthetic outputs are derived from a non-prosthetic lexical form has prevailed in the recent literature. In point of fact, Shepherd (2003, 24) writes «the epenthetic status of the aforementioned vowel is quite uncontroversial, actually, and is most strongly supported by the consistent appearance of the [e] in loanwords with initial s-clusters, as in the following examples, as well as in the interlanguage of native Spanish speakers learning English». It is assumed that the lexical form of word-initial [sC] clusters matches the non-native output form, yet such an assertion ignores strong empirical evidence that the perception, and subsequent lexical storage, of continuous non-native speech streams can be heavily influenced by the phonology (see Best 1995; Hallé *et al.* 1998; Dupoux *et al.* 1999; Kabak & Idsardi 2003; Andersen 2004; Hallé & Best 2007). Dupoux *et al.* (1999), among many others, show that distortion of the auditory input prompted by the phonology may indeed render an innovative lexical form even if acoustic cues to support the novel representation are nonexistent.

In the remainder of this article an alternative account to justify vowel prosthesis is proposed. The focal hypotheses underscoring the following article are that (i) vowel prosthesis is a perception-based repair phenomenon, and (ii) conditioned by a language-specific bias based on principles of relative timing between gestures in complex onsets. In the next section, evidence is reviewed which supports the ongoing hypothesis that vowel prosthesis is a perception-based process. Following, the basic components of a gesture-based grammar are explained and the groundwork is laid by which to illustrate how a perceptual bias of this nature can be derived from such a grammar. Finally, an analysis of vowel prosthesis based on input/output correspondence is presented.

2 A perceptual argument for /e/ prosthesis

The previous grammar-based proposals assume that the non-native input is stored lexically as it emerges in the source language, ignoring strong evidence suggesting that the perception of continuous speech streams is often susceptible to the effects of perceptual biases (Best 1995; Hallé *et al.* 1998; Dupoux *et al.* 1999; Kabak & Idsardi 2003; Andersen 2004; Hallé & Best 2007).

© Romania Minor
http://www.romaniaminor.net/ianua/

BDD-A5302 © 2012 Romania Minor
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 17:58:36 UTC)

The idea that speakers form biases for certain aspects of their native language is robustly supported in the literature. Research in experimental psychology and first language acquisition reveals that biases for native sound patterns emerge during a human infant's first year of life. Werker & Tees' (1984) study of infants under a year old from English speaking homes, for example, found that by 10–12 months old, infants were unable to distinguish non-English contrasts. The authors illustrated that the decline in discrimination of non-English contrasts was due to the lack of experience with these sequences, suggesting a manifest bias for native structures.

With regard to syllable structure, Bertoncini & Mehler (1981) discovered that two month old infants were more apt to distinguish pairs of stimuli that form syllabic patterns. Further studies by Hillenbrand (1983) show that by six months, infants already have some intuitive knowledge of the internal organization of syllables in their native language. In another study, Jusczyk *et al.* (1993) discovered that infants are more apt to listen to lists of native syllables which share common first consonants, or common first consonants and vowels. Nevertheless, when the syllables only share a common vowel, or common vowel and final consonants, listening declines significantly, suggesting an astonishing attentiveness to onsets. By 18 months, Goodman *et al.* (2000) found that rimes and codas were, in general, given almost no consideration.

The claim that the perception of a continuous acoustic stream can be distorted by the phonology of a listener/perceiver was first introduced in Dupoux *et al.* (1999) and elaborated on in Kabak & Idsardi (2003). Specifically, the data show that an illusory vowel can be perceived regardless of whether the acoustic characteristics of the speech stream involve a vocalic unit. Dupoux *et al.* (1999) and Dehaene-Lambertz *et al.* (2000) neurolinguistic study corroborate the general findings outlined in Dupoux *et al.* (1999) concerning perceptual epenthesis in Japanese informants, while Kabak & Idsardi's (2003) study of the perception of vocalic intrusion by Korean informants suggests that the illusory vowel is connected in a programmed way to conventions of L1 syllable structure.

Perceptual epenthesis is based on the idea that phonetic and phonological factors interact in a bottom-up system. The acoustic input is perceived by the listener and later filtered through a system of high-level biases referencing the systematic patterns of L1 syllable structure before lexical storage occurs. Hallé *et al.*'s (1998) study of perceptual epenthesis in Spanish corroborates such a framework, but does not address the specific phonological conditions on which such a bias is based. It is not clear whether the emergence of an illusory [e] is motivated by segmental organization exclusively or if the perception of [e] is conditioned by more microscopic patterns of coordination relations between the units.

Following Hallé *et al.* (1998), Gibson (2012) performed a series of identification and discrimination tests with fifty native Spanish-speaking subjects aged 10–11. In the first test, the subjects were played nonce stimuli containing both licit and illicit word-initial onset clusters. The subjects were asked to respond with the number of syllables they heard. The results indicate that the perception of an illusory syllable (i) is prevalent in all target sequences, and (ii) is

BDD-A5302 © 2012 Romania Minor
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 17:58:36 UTC)

unsupported in non-target stimuli. A set of follow-up tasks revealed that the perception of an extra syllable resulted from an erroneous interpretation of the word-initial /sC/ sequence.

Importantly, the data from the tests reveal that the fundamental conditions of the bias motivating the prosthetic vowel go beyond the mere linear organization of spatial characteristics. There was no significant difference between the perception of prosthetic [e] in [s+stop], [s+nasal] or [sl] sequences, meaning that spatial characteristics of individual units alone cannot be supposed to be responsible for the perception of prosthetic [e] in all the cases. A bias which includes the temporal relationships between those spatial characteristics offers a refined account which coincides with the output patterns of intergestural timing in Spanish complex onsets.

Apropos the specific vocalic segment that inserts in all of the examples in (1), it is not clear what motivates the specific insertion of /e/, and not another vowel from the Spanish vocalic repertoire. Harris (1983) asserts that the appearance of /e/ is somehow related to its relative markedness value, although Hume (2002) provides empirical data suggesting problems with claims of segment insertion based on markedness cross-linguistically. Lombardi (2002), to the contrary, presents an alternative universal schema grounded in the relative markedness of major [place] classes. Nonetheless, Lombardi concludes that /a/, and not /e/, is the universally unmarked segment.

Although still in the exploratory phase of investigation, another line of research could offer fruitful insights into an articulatory-based justification for the specific insertion of /e/. As far back as 1964 (in modern speech science), Honikman (1964) introduced the idea that individual languages have an organic basis which aids in the acquisition of language. Honikman posteriorly refers to this organic basis as language-specific 'articulatory settings'. According to Honikman (1964, 73) articulator settings serve as the 'gross oral posture and mechanics' which underlie the production of a language's phonetic inventory and make the speech of a native speaker distinguishable from the speech of non-native speakers.

Although various proposals have been suggested with regard to how speech scientists may gain access to these elusive settings for observation, one of the most fertile hypotheses has been to measure the default speech positions for individual languages during rest. It has been found that when preparing to speak, the various articulators of the vocal tract assume a sort of 'pre-speech' posture from which to execute the sounds of the language (Öhman 1967; Perkell 1969). Barry (1992) went a step further to show that pre-speech posture is in fact discernible from the absolute rest posture of the vocal tract, which Gick (2002) upholds with X-ray data from English.

Now with regard to the perception of /e/ in word-initial /sC/ clusters in Spanish, it is feasible to suspect a relationship between the specific perception of /e/ and the articulatory posture of the vocal tract in pre-speech rest position for Spanish. Consider the non-linguistic pre-speech sound that native Spanish-speakers articulate when preparing to speak: *eh, no lo sé* 'uhm, I don't know':

© Romania Minor
http://www.romaniaminor.net/ianua/

BDD-A5302 © 2012 Romania Minor
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 17:58:36 UTC)

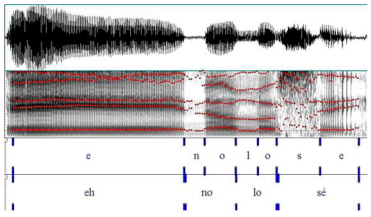


FIGURE 1.

Acoustically, the initial vowel /e/, considered a pre-speech utterance, and the second /e/ from the word *sé*, are nearly identical. Although there are differences in amplitudes between the two sounds, a close reading of the formants reveals that the sounds are quite similar. Interestingly, the vocalic unit which appears as a pre-speech utterance in this context also appears in a myriad of other contexts in Spanish: *no me parece bien, eh* 'I don't think that's right' (just so you know); *¿eh?* (to express befuddlement).

It is viable that the posture assumed by the vocal tract in this instance, which is articulatorily similar to that of /e/, is referenced as the default setting for Spanish when perceiving the illicit syllable-initial onset. Of course this claim would imply a certain phonological condition for language-specific articulatory settings, or at least imply a close relationship with the phonology. Gick *et al.* (2004) data come close to corroborating such an assertion, stating clearly that (i) articulator settings do show tendencies of functioning as speech targets, and (ii) asserting influence on other speech targets in the phonetic environment. Currently, ultrasound data is being examined which seems to corroborate this claim and will be presented in a separate article in the future.

3 Expressing vowel prosthesis in a gestural-based grammar

3.1 Articulatory Phonology

Articulatory Phonology (Browman & Goldstein 1986) (AP, henceforth) provides an empirically-based explanation of speech which falls within a more general model of action and coordination dynamics. In this framework, speech gestures are considered to be «point attractors, depicted as critically damped mass-spring systems which interact in a task space- whose dimensions are delineated as constrictions that can be produced and released by the various constricting effectors of the vocal tract» (Saltzman & Munhall 1989 in Goldstein *et al.* 2006).

A gesture coordinates the movements of a limited set of articulators for a linguistically significant goal (Goldstein *et al.* 2006). The set of articulators are described as tongue body, tongue tip, glottis, velum and lips, each one forming a separate articulator tier (Browman & Goldstein 1986). When gestures form to compose words, the activation times of individual gestures in the vocal tract are

BDD-A5302 © 2012 Romania Minor
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 17:58:36 UTC)

coordinated in relation to the activation of other gestures by way of an ensemble of limit-cycle oscillators which are connected to each gesture. When the atomic gestures couple temporally to other units to form syllables, they constitute what Browman and Goldstein call a gestural molecule. The following is an example of a temporal plan of gesture activation (temporal score) for a word *pen* (Browman & Goldstein 1995):

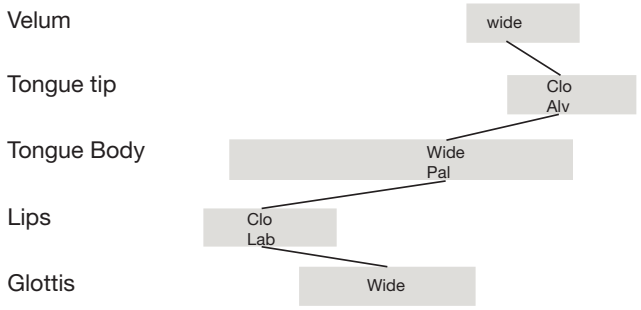


FIGURE 2.

The length of the shaded boxes corresponds to the temporal duration of a given gesture. Notice that gestures overlap with one another in time.

It has been proposed that syllable structure falls out from a self-organizing system of non-linear limit-cycle oscillators which govern intergestural timing (Nam & Saltzman 2003). In this system, timing relationships between gestures are not prescribed by a central command control, common to information processing conceptions of cognition, but rather emerge dynamically from the interaction of oscillator ensembles. Thus the stability, entrainment and alteration of relative phasing modes which result from speech timing can be understood in relation to the interaction of the corresponding oscillator networks governing speech production.

The Coupled Oscillatory Theory conjectures that onset and coda constitute the two possible intrinsic modes of coordinating a consonant and vowel gesture. It is generally accepted that coda gestures are arranged linearly (anti-phase, $\Phi = \pm\pi 180°$) in relation to the preceding vocalic target while onset gestures are phased globally (in-phase, $\Phi = 0°$) in relation to the following nucleus (Browman & Goldstein 1986, 1988, and much subsequent work).

Although both consonant gestures in [CCV] structures are phased globally (in-phase = 0°) in relation to the following vocalic target, each gesture is also coupled in an anti-phase (180°) relationship with the other consonant gesture (Browman & Goldstein 2000), resulting in a competitive coupling graph (Nam & Saltzman 2003), the outcome of which ensures that each gesture of the com-

© Romania Minor
http://www.romaniaminor.net/ianua/

BDD-A5302 © 2012 Romania Minor
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 17:58:36 UTC)

plex sequence is recoverable (Browman & Goldstein 1988). The final coupling phase which emerges at the end of planning is thus a concession between the competing target phases (Nam & Saltzman 2003).

Syllable coordination results when the articulatory system settles into specific attractor states in the perceptual motor landscape. The landscape of attractors must gradually be altered in order to allow the speaker to learn new patterns. The learner must rely on learned patterns in order to make predictions regarding the production of novel patterns. Learning occurs when the new pattern becomes an attractor, a «stable state of the (now modified) pattern dynamics» (Kelso 1995, 163).

Evidence for phasing can be found in the systematic patterns of gestural timing. Hermes *et al.* (2008) analysis of Italian complex onsets illustrate that (i) the c-center in complex onsets remains stable in relation to the vowel, corroborating predictions made by Browman & Goldstein (1988) and Honorof & Browman (1995), and that (ii) the duration between the rightmost consonant and vowel decreases, being forced rightward to accommodate the addition of the extra consonant, verifying predictions made by Goldstein *et al.* (2007) and elaborated on in Shaw *et al.* (2009). Shaw *et al.*'s (2009) study of temporal patterns in Moroccan Arabic show similar results, substantiating the global hypotheses espoused by Browman & Goldstein (1988), Krakow (1989, 1999), Sproat & Fujimura (1993), Byrd (1995, 1996), Honorof & Browman (1995) and Goldstein *et al.* (2007), concerning the temporal arrangement and duration effects of consonant gestures in complex onsets.

3.2 A gestural-based grammar: Gestural Optimality Theory

Gafos (2002) integrates the spatio-temporal coordination of gestures into a grammar based on constraint conflict (Optimality Theory, OT) in what has subsequently become known as Gestural Optimality Theory. The basic premise on which this grammar is based is that gestural coordination is determined by alignment constraints which reference a set of «temporal landmarks» during the activation period of a gesture: onset, target, center, release, offset. Essentially, there are three basic ways in which consonant sequences may be arranged in relation to the following vocalic target:

It has been proposed that Spanish permits two ways to coordinate [CC] sequences with the following vocalic target. Close transitions occur when the release if C_1 can align directly to the target of C_2, as the configuration in (i) illustrates above. This is accomplished when the phonetic properties of the gestures permit the onset of C_2 to occur in tandem with the realization of the target gesture of C_1, for example when the consonants come from different tiers. Open transitions, on the other hand, arise when the phonetic characteristics of the individual gestures do not permit synchronization between the release of the target gesture of C_1 and the start of the target gesture for C_2. As Bradley & Schmeiser (2003) suggest, this type of temporal pattern most often occurs in complex onsets in which the C_2 is [ɾ]. Both patterns require the center of the sequence to align globally to a following vocalic target.

BDD-A5302 © 2012 Romania Minor
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 17:58:36 UTC)

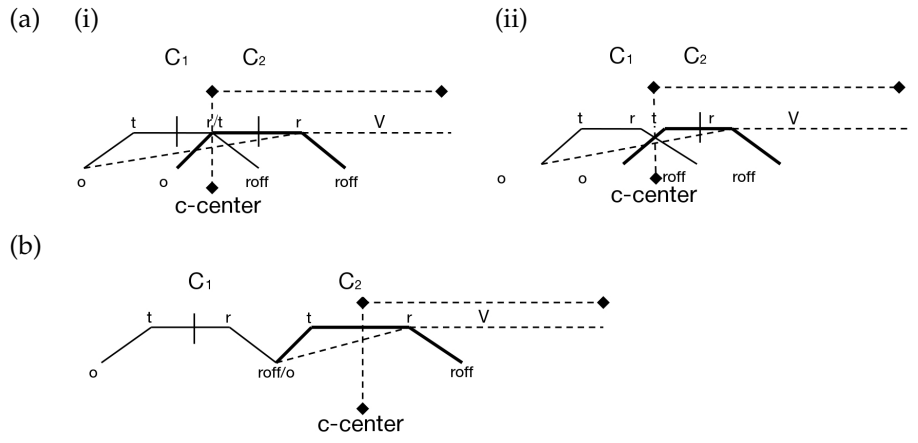


FIGURE 3.

The temporal relations in (a) represent the two permissible patterns of intergestural timing in complex onsets. The timing relation in (b) represents a first gesture which is arranged linearly in relationship to the preceding vowel and a second gesture which is arranged globally in relation to the following vocalic target. In (b), the first consonant constitutes a coda gesture and the second consonant constitutes an onset gesture.

Gafos (2002) proposes that coordination between the consonant gestures is carried out by a function of the grammar which aligns the landmarks of distinct gestures in order to form coordination relations. For complex onsets, a constraint ALIGN(C_1, c-center, C_2 onset) stipulates that the temporal center of C_1 must align to the onset of C_2 (for close transitions):[5]

(2) ALIGN(C_1, c-center, C_2 onset)
Align the temporal center of C_1 with the onset of C_2.

In intervocalic position, there are two options by which to arrange the gestures with respect to the vowels. In the case that the phonetic postures of the gestures permit coordination as a complex onset, then ALIGN will require them to do so. However, in the case that no coordination is possible due to the articulatory requirements of the gestures, there is a natural conflict which must be resolved. In the following section I will illustrate how this conflict can be resolved in a gesture-based grammar and show how the resolution of this conflict can serve as an articulatory basis for the perceptual bias which leads to the misinterpretation of the illusory vowel in word-initial /sC/ sequences.

[5] Further constraints are necessary in order to bond the consonant sequence with the following vocalic target. This point will be developed further in the following section.

© Romania Minor
http://www.romaniaminor.net/ianua/

BDD-A5302 © 2012 Romania Minor
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 17:58:36 UTC)

3.2.1 Gestural coordination in VsCV sequences: a gestural basis for a perceptual bias

To date, no articulatory study addressing the spatio-temporal coordination in /sC/ sequences in Spanish has been performed, although such data would certainly be welcomed. Data from other Romance languages though may provide interesting insight into the coordination relations which govern Spanish syllable formation.

In Hermes *et al.*'s (2008) study of complex onset formation in Italian word-initial consonant clusters it was found that non-sibilant clusters in word-initial position adhere to the predictions regarding syllabic constituencies outlined in Shaw *et al.* (2009) with regard to the latencies from the rightmost consonant to the vocalic target. Importantly, a c-center effect was found for all non-sibilant word-initial clusters, meaning non-sibilant complex clusters indeed form complex onsets.

Interestingly it was found that when /s/ was added to the clusters, the target of the rightmost consonant does not shift towards the vocalic target, as in non-sibilant headed clusters, but remains stable. This, according to Hermes *et al.* (2008), suggests that /sC/ and /sCC/ clusters word-initially are in fact heterosyllabic. The heterosyllabic status of /sC/ clusters has been proposed for French, yet no conclusive data have been provided to corroborate such a claim. These assumptions contrast with data from Spanish, not in the fact that Italian and French may permit sibilant-headed word-initial complex onsets, but rather that these languages permit non-tautosyllabic sequences to emerge in word-initial position. In Spanish, it appears, word-initial consonant clusters may not emerge heterosyllabically.

In intervocalic position, it is not at all clear how speakers of Italian and French coordinate the sibilant in relation to the preceding vocalic targets. The results of Bertinetto's (2004) psycholinguistic study suggest little systematicity concerning the syllabification of the sibilant in Italian. In a language-game task, subjects tended to marginally treat word-internal /sC/ sequences as tautosyllabic, whereas in a follow-up concept formation experiment, the sequences were given a treatment more akin to heterosyllabic sequences. A superficial reading of available electropalatographic data for intervocalic [sibilant+consonant] sequences in French reveals no convincing evidence for the duration effects most commonly found for complex onset formation;[6] that is, there appears to be no reduction of the target of the rightmost consonant due to a shift toward the vocalic target. Patterns of overlap between the intervocalic /sC/ sequences are equally inconsistent across tokens, rendering any categorical declaration concerning the heterosyllabic status of the sequence merely speculative.

Spanish syllabification, conversely, has traditionally been considered somewhat more systematic than French and Italian in the sense that there are no examples of /sC/ clusters word-initially. In this way, the distribution data from modern Spanish support Itô's (1989) syllabic parsing hypothesis which states

[6] The data I refer to here were collected at the Laboratoire de Phonetique et Phonologie of the Université de Paris Nouvelle Sorbonne 3.

BDD-A5302 © 2012 Romania Minor
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 17:58:36 UTC)

that parsing at word-boundaries is subject to stipulations regulating syllable-boundaries. Diachronic data of /e/ prosthesis to the left margin of word-initial /sC/ clusters support this assertion (Gibson 2010):

(3) Prosthetic /e/ in phonologized Latinate forms

Spanish adaptation	Latin root	
esperado	*sperāre*	'hoped' past part.
espejo	*specŭlum*	'mirror'
estado	*status*	'status'
esfera	*sphaera*	'sphere'

As in the synchronic cases discussed in the present article, all word-initial /sC/ sequences in the Latinate items also surface with the prosthetic vowel.

Results from a series of psycholinguistic tests with native Spanish-speaking children may shed light on the articulatory arrangement of output gestures in Spanish. In an independent test whose results will be published in a future article, a group of three-, four-, and five-year old Spanish-speaking children were asked to parse words into syllables in game form. The subjects were given a polysyllabic token containing the word-medial /sC/ sequence and asked to clap out and repeat orally the syllables, each syllable receiving one clap. The results indicate that Spanish-speaking children routinely coordinate the sibilant and following consonant to different vocalic units, forming a heterosyllabic sequence word-medially.

Following Gafos (2002) and Davidson (2003), here it is proposed that the intervocalic coordination of /sC/ sequences can be expressed by the hierarchical organization of three basic families of constraints which govern the spatio-temporal coordination of output gestures. The constraint *Overlap/Fα,Fβ prohibits overlap between contiguous gestures which are specified for a given feature. In the case of /sC/, for now this will be expressed as *Over/F[sibilant], [+consonant]. Due to the lack of articulatory data, however, commentary will be withheld for now as regard the nature of this failure to overlap. Simply, there are not enough data to determine whether this point is regulated by biomechanic principles, or is governed by linguistic factors:[7]

(4) *Overlap/Fα,Fβ
Do not overlap the release of a gesture specified for afeature Fα with the plateau of a following associated gesture specified for a feature Fβ. (Davidson 2003)

Such a constraint presents a natural conflict with any constraint which seeks to coordinate two gestures as a complex onset (Gafos 2002): Coordinate-cc (C_1 c-center, C_2 onset) (henceforth, Coord-cc (C_1 c-center, C_2 onset)).

[7] Modifications to this basic constraint may be needed to account for sibilant+glide sequences, which are licensed in Spanish. However, based on the historical evolution of the sibilant+glide sequences in Spanish, a nucleus analysis of glides is warranted. That is, in sequences such as [sw], *sueco* 'Swedish', and [sj] *siete* 'seven', the glides form part of the syllable nucleus.

© Romania Minor
http://www.romaniaminor.net/ianua/

BDD-A5302 © 2012 Romania Minor
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 17:58:36 UTC)

(5) Coord(inate)-CC (C_1, c-center, C_2 onset)
Align the temporal center of C_1 with the onset of C_2.

Finally, as Gafos and Davidson have pointed out, coordination constraints must interact with a series of constraints which define association relations between consonant and vocalic gestures in order to bond the sequence with a following vocalic target. Gestural Association Theory (Davidson 2003) posits that (i) gestural coordination is a grammatical operation, and as such is (ii) subject to the exponential proliferation of possible output coordination patterns by GEN. In this theory, associations are the «natural correlate of syllable structure» (Davidson 2003, 105). Essentially, association constraints formalize relations for a series of gestures which are regulated by Coord constraints, specifically CV, CC, and VC gestures (Davidson 2003):

(6) Assoc(iate)-CV
A consonant gesture must have a coordination relationship with the nearest following vowel gesture.

(7) Assoc(iate)-CC
A consonant gesture must have a coordination relationship with adjacent consonant gestures.

(8) Assoc(iate)-C
A (non-nuclear) consonant gesture must not be unassociated.

(9) *Mult(iple)Assoc(iation)
A consonantal gesture must not be associated with multiple vowels. Consonantal gestures associated with different vowels must not be associated with one another.

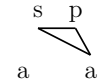
The following hierarchical ranking of these constraints yields correct predictions for intervocalic /sC/ parsing (Davidson 2003):

(10) *MultAssoc, Assoc-C, *OV/[sib],[+cons], CV-Coord, CC-Coord, DEP ≫ Assoc-CV, Assoc-CC, *OV/[sib],[+cons]

Notice how the conflict plays out in the following tableau for an intervocalic /sC/ sequence. Solid lines between gestures represent coordination relations:

BDD-A5302 © 2012 Romania Minor
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 17:58:36 UTC)

(11) Nonce input: /aspa/

s p a a	*Mult Assoc	Assoc-C	*OV/[sib],[+cons]	CV-Coord	CC-Coord	DEP	Assoc-CV	Assoc-CC	*OV/[sib],[+cons]
☞ a.  [as.pa]							*	*	
b.  [a.spa]			*(!)						*
c.  [asp.a]				*(!)					

Here candidate (11b) turns out sub-optimal due to the fact that it coordinates two segments which do not permit phasing as a complex onset. Candidate (11c) is eliminated for coordinating both segments as a complex coda while leaving the following vocalic unit unassociated, fatally violating CV-Coord. Candidate (11a) turns out optimal due to its satisfaction of the higher ranked constraints while committing only a consequential infraction to the inferiorly ranked Assoc-CC, which regulates the association between two contiguous consonant gestures.

Over time a child raised in a Spanish-speaking environment will learn that no native /VsCV/ input structure ever violates the veracity of the previous hierarchy. From a purely action-based account then, two generalizations can be deduced by the language learner. First, the learner can presume that /sC/ are always coordinated to different vocalic targets. Secondly, yet interrelatedly, she will gather that the consonant gestures are flanked on either side by glottal aperture and corresponding vocal fold activation related to some vowel gesture. These are important generalizations from the perspective of statistical learning since the language-learner never has positive output evidence for which to question the hierarchical organization she has learned. Spanish simply does not present alternative possibilities by which to arrange consonants and vowels in /VsCV/ input.

As the learner acquires new inputs from her native language, the production hierarchy becomes reinforced, forming a bias which distorts the novel auditory input. Thus when an innovative structure is perceived which does not coincide

© Romania Minor
http://www.romaniaminor.net/ianua/

BDD-A5302 © 2012 Romania Minor
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 17:58:36 UTC)

with the rules of intergestural coordination provided by Spanish, the language learner will process the acoustic input in accordance with the rules governing spatio-temporal coordination in Spanish syllables. Thus an unlicensed word-initial /sC/ sequence will be interpreted as intervocalic (between two glottal apertures with corresponding vocal fold vibrations) since that is the only permissible way to coordinate these gestures in Spanish. In absence of acoustic evidence specifying the vocalic unit which corresponds to the first vowel, the system inserts the most «neutral» vocal tract setting corresponding to vowel gestures by default.

As regard lexicalization then, it is logical to assume that the language learner will store the distorted input to memory even if acoustically there is no vocalic element. It is clear that the language learner perceives there is, and therefore has no empirically-based motivation for which to question this generalization. Prince & Smolensky (1993), basing their argument in part on Stampe (1972), formalize this concept into their *Lexicon Optimization* strategy which states that if there is no empirical evidence to the contrary, a language learner will always lexicalize the form which coincides most closely with the output form.

With formal study of languages such as English which permit the word-initial /sC/ structure, she may indeed reassess this generalization, or at least learn to suppress the extra vowel gesture during production, but there is no reason for her to ever deduce, based on perceptual experience alone, the non-existence of a segment which for her is quite real. Thus at the time of production, the only grammatically relevant constraint is that which forbids the elimination of input segments, that is faithfulness.

4 A faithfulness account for vowel prosthesis in word-initial /sC/ loan words

The basic premise on which the current proposal is based is that Spanish-speakers form biases for native timing configurations which distort how certain consonant sequences are perceived. It has been shown in past work, that native Spanish-speakers perceive an illusory vowel in all non-native word-initial /sC/ clusters. With no positive output evidence from the native language to contradict the bias that develops over time, it is claimed that the form the language-learner lexicalizes will be that which most closely matches the output, as perceived.

This proposal is quite different from previous proposals treating vowel prosthesis in Spanish in the sense that the appearance of the vowel is not treated as a grammatical operation, but is considered a function of the lexicon. As such then, the only function the grammar plays in producing a prosthesized output is to ensure that no input segments are deleted in the output.

Optimality Theory posits a function of this nature as a dependent of faithfulness, which ensures maximal correspondence between input and output.

BDD-A5302 © 2012 Romania Minor
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 17:58:36 UTC)

With specific regard to vowel prosthesis, a constraint Max-I/O (Maximality-Input/Output) will ensure that all input segments are present in the output:

(12) Max(imality)
Every segment in the input has a correspondent in output (no segment deletion).

Since no phonotactic constraint is violated by the vowel-initial input, there is no sound reason to posit a specific markedness constraint which targets word-initial /sC/ onsets.

As for coordinating the phonological units in time, the following constraints are reintroduced:

(13) *Overlap/Fα,Fβ
Do not overlap the release of a gesture specified for afeature Fα with the plateau of a following associated gesture specified for a feature Fβ.

(14) Coord(inate)-CC (C_1 center, C_2 onset)
Align the temporal center of C_1 with the onset of C_2.

(15) Assoc(iate)-CV
A consonant gesture must have a coordination relationship with the nearest following vowel gesture.

(16) Assoc(iate)-CC
A consonant gesture must have a coordination relationship with adjacent consonant gestures.

(17) Assoc(iate)-C
A (non-nuclear) consonant gesture must not be unassociated.

(18) *Mult(iple)Assoc(iation)
A consonantal gesture must not be associated with multiple vowels. Consonantal gestures associated with different vowels must not be associated with one another.

A complete hierarchy headed by faithfulness will appear as the following:

(19) Max-I/O ≫ *MultAssoc, Assoc-C, *OV/[sib],[+cons], CV-Coord, CC-Coord, DEP ≫ Assoc-CV, Assoc-CC, *OV/[sib],[+cons]

Notice how this hierarchy is capable of coordinating the segments to their corresponding vocalic unit:

© Romania Minor
http://www.romaniaminor.net/ianua/

BDD-A5302 © 2012 Romania Minor
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 17:58:36 UTC)

(20) Input: /estop/

s t p / e o	MAX-I/O	*MULTASSOC	ASSOC-C	*OV/[sib],[+cons]	CV-COORD	CC-COORD	DEP	ASSOC-CV	ASSOC-CC	*OV/[sib],[+cons]
☞ a. [es.top]								*	*	
b. [e.stop]				*(!)						*
c. [st.op]	*(!)			*						*

In this tableau, candidate (20b) is ruled out due to its fatal violation of *OV/[sib],[+cons] which penalizes coordination between /s/ and a following consonant gesture. Candidate (20c) presents an output in which the vocalic gesture is eliminated in the output, a fatal violation to MAX, the superior constraint. Candidate (20a) is deemed optimal due to the fact that (i) all input segments are represented in the output and (ii) that /s/ and /t/ are coordinated to distinct vocalic units.

5 Conclusion

In this article I have claimed that /e/ insertion in non-native word-initial /sC/ sequences does not result from a productive operation of the Spanish grammar. Instead, I addressed results of a series of perception tests which suggest that Spanish-speakers actually perceive an illusory vowel even if there is no indication of such a segment in the acoustic environment. Past studies were reviewed which confirm that the perception of an illusory vowel may be triggered by high-level phonological biases referencing syllable structure. These data formed the basis of the gestural-based analysis offered here in which it was shown how a phonological bias can be derived from a hierarchy of coordination constraints which regulate the spatio-temporal coordination of Spanish syllables.

Another important claim was made concerning the specific vowel which inserts in these cases. I suggested an intimate link between the perception of

BDD-A5302 © 2012 Romania Minor
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 17:58:36 UTC)

an illusory vowel and the articulatory setting for Spanish. Although still in its infancy, such an explanation could provide interesting new insight into the systemic interaction of articulatory action and the perception of action in speech. This topic will be readdressed in future studies.

References

Andersen, Raquel (2004): «Phonological acquisition in preschoolers learning a second language via immersion: A longitudinal study.» *Clinical Linguistics and Phonetics* 18(3): 183–210.

Barry, W. (1992): «Comments on chapter 2 (Browman and Goldstein).» In: Gerard J. Docherty; D. Robert Ladd [ed.]: *Papers in laboratory phonology II: gesture, segment, prosody*. Cambridge: Cambridge University Press, 65–67.

Bertinetto, Pier Marco (2004): «On the undecidable syllabification of /sC/ clusters in Italian: converging experimental evidence.» *Italian Journal of Linguistics* 16: 349–372.

Bertoncini, Josiane; Mehler, Jacques (1981): «Syllables as units in infant speech perception.» *Infant Behavior and Development* 4: 247–260. DOI: 10.1016/S0163-6383(81)80027-6.

Best, Catherine T. (1995): «A direct realist view of cross-language speech perception.» In: Winifred Strange [ed.]: *Speech perception and linguistic experience: Theoretical and methodological issues*. Baltimore (MD): York Press, 171–204.

Bradley, Travis; Schmeiser, Benjamin. (2003): «On the phonetic reality of Spanish /ɾ/ in complex onsets.» In: Paula Kempchinsky; Carlos-Eduardo Piñero [ed.]: *Theory, practice and acquisition: Papers from the 6th Hispanic Linguistics Symposium and the 5th Conference on the Acquisition of Spanish and Portuguese*. Sommerville (MA): Cascadilla Press, 1–20.

Browman, Catherine P.; Goldstein, Louis M. (1986): «Towards an Articulatory Phonology.» *Phonology Yearbook* 3: 219–252.

Browman, Catherine P.; Goldstein, Louis M. (1988): «Some notes on syllable structure in Articulatory Phonology.» *Phonetica* 45(2–4): 140–155. DOI: 10.1159/000261823.

Browman, Catherine P.; Goldstein, Louis M. (1995): «Dynamics and articulatory phonology.» In: Robert F. Port; Timothy van Gelder [ed.]: *Mind as motion: Explorations in the dynamics of cognition*. Cambridge (MA): MIT Press, 175–193.

Browman, Catherine P.; Goldstein, Louis M. (2000): «Competing constraints on intergestural coordination and self-organization of phonological structures.» *Les Cahiers de l'ICP, Bulletin de la Communication Parlée* 5: 25–34.

Byrd, Dani (1995): «C-Centers revisited.» *Phonetica* 52(4): 263–282. DOI: 10.1159/000262183.

© Romania Minor
http://www.romaniaminor.net/ianua/

BDD-A5302 © 2012 Romania Minor
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 17:58:36 UTC)

Byrd, Dani (1996): «Influences on articulatory timing in consonant sequences.» *Journal of Phonetics* 24(2): 209–244. DOI: 10.1006/jpho.1996.0012.

Clements, G.N. (1990): «The role of the sonority cycle in core syllabification.» In: John Kingston; Mary E. Beckman [ed.]: *Papers in Laboratory Phonology, Volume 1: Between the grammar and physics of speech*. Cambridge: Cambridge University Press, 283–333.

Colina, Sonia (1995): «A constraint-based analysis of syllabification in Spanish, Catalan, and Galician.» Ph.D. dissertation. University of Illinois at Urbana-Champaign. URL: <http://hdl.handle.net/2142/21187>.

Davidson, Lisa (2003): «The atoms of phonological representation: Gestures, coordination and perceptual features in consonant cluster phonotactics.» Ph.D. dissertation. Johns Hopkins University, Baltimore (MD). URL: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.111.2227>.

Dehaene-Lambertz, G.; Dupoux, E.; Gout, A. (2000): «Electrophysiological correlates of phonological processing: A cross-linguistic study.» *Journal of Cognitive Neuroscience* 12(4): 635–647. DOI: 10.1162/089892900562390.

Dupoux, Emmanuel; Kakehi, Kazuhiko; Hirose, Yuki; Pallier, Christophe; Mehler, Jacques (1999): «Epenthetic vowels in Japanese: A perceptual illusion?» *Journal of Experimental Psychology: Human Perception and Performance* 25(6): 1568–1578. DOI: 10.1037/0096-1523.25.6.1568.

Gafos, Adamantios I. (2002): «A grammar of gestural coordination.» *Natural Language & Linguistic Theory* 20(2): 269–337. DOI: 10.1023/A:1014942312445.

Gibson, Mark (2010): «La fonología de español: un enfoque desde la Teoría de Optimidad.» Ph.D. dissertation. Universidad Complutense de Madrid. URL: <http://eprints.ucm.es/11031/>.

Gibson, Mark (2012): «Perception-based vowel insertion by native Spanish-speaking learners of English.» *TIPA. Travaux interdisciplinaires sur la parole et le langage* 28. URL: <http://tipa.revues.org/205>.

Gick, Bryan (2002): «An X-ray investigation of pharyngeal constriction in American English schwa.» *Phonetica* 59(1): 38–48. DOI: 10.1159/000056204.

Gick, Bryan; Wilson, Ian; Koch, Karsten; Cook, Clare (2004): «Language-specific articulatory settings: Evidence from inter-utterance rest position.» *Phonetica* 61(4): 220–233. DOI: 10.1159/000084159.

Goldstein, Louis; Byrd, Dani; Saltzman, Elliot (2006): «The role of vocal tract gestural action units in understanding the evolution of phonology.» In: Michael A. Arbib [ed.]: *Action to language via the mirror neuron system*. Cambridge: Cambridge University Press, 215–249.

Goldstein, Louis; Chitoran, Ioana; Selkirk, Elisabeth (2007): «Syllable structure as coupled oscillator modes: Evidence from Georgian vs. Tashlhiyt Berber.» In: J. Trouvain; W.J. Barry [ed.]: *Proceedings of the XVI International Congress of Phonetic Sciences*. Saarbrücken: Universität des Saarlandes, 241–244.

BDD-A5302 © 2012 Romania Minor
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 17:58:36 UTC)

Goodman, Mara; Jusczyk, Peter; Bauman, Angela (2000): «Developmental changes in infants' sensitivity to internal syllable structure.» In: Michael B. Broe; Janet B. Pierrehumbert [ed.]: *Papers in laboratory phonology V: Acquisition and the lexicon*. Cambridge: Cambridge University Press, 228–239.

Hallé, Pierre A.; Best, Catherine T. (2007): «Dental-to-velar perceptual assimilation: A cross-linguistic study of the perception of dental stop+/l/ clusters.» *Journal of the Acoustical Society of America* 121(5): 2899–2914. DOI: 10.1121/1.2534656.

Hallé, Pierre A.; Segui, Juan; Frauenfelder, Uli; Meunier, Christine (1998): «Processing of illegal consonant clusters: A case of perceptual assimilation?» *Journal of Experimental Psychology: Human Perception and Performance* 24(2): 592–608. DOI: 10.1037/0096-1523.24.2.592.

Harris, James (1983): *Syllable structure and stress in Spanish: A nonlinear analysis*. Bloomington (IA): Indiana University Linguistics Club.

Hermes, Anne; Grice, Martine; Mücke, Doris; Niemann, Henrik (2008): «Articulatory indicators of syllable affiliation in word initial consonant clusters in Italian.» In: *Papers from the 8th International Seminar on Speech Production*. Strasbourg, 433–436. URL: <http://phonetik.phil-fak.uni-koeln.de/uploads/tx_sibibtex/issp2008-102_02.pdf>.

Hillenbrand, James (1983): «Perceptual organization of speech sounds by infants.» *Journal of Speech and Hearing Research* 26(2): 268–282.

Honikman, Beatrice (1964): «Articulatory settings.» In: David Abercrombie; D.B. Fry; P.A.D. MacCarthy; N.C. Scott; J.L.M. Trim [ed.]: *In honour of Daniel Jones: Papers contributed on the occasion of his eightieth birthday, 12 September 1961*. London: Longman, 73–84.

Honorof, Douglas; Browman, Catherine (1995): «The center or edge: How are consonant clusters organized with respect to the vowel?.» In: K. Elenius; P. Branderud [ed.]: *Proceedings of the XIIIth International Congress of Phonetic Sciences, Stockholm, Sweden*. Stockholm: KTH; Stockholm University, 552–555.

Hume, Elizabeth (2002): «Labial unmarkedness in Sri Lankan Portuguese Creole.» *Phonology* 19(3): 441–458. DOI: 10.1017/S0952675703004421.

Itô, Junko (1989): «A prosodic theory of epenthesis.» *Natural Language & Linguistic Theory* 7(2): 217–259. DOI: 10.1007/BF00138077.

Jusczyk, P.W.; Friederici, A.D.; Wessels, J.M.I.; Svenkerud, V.Y.; Jusczyk, A.M. (1993): «Infants' sensitivity to the sound patterns of native language words.» *Journal of Memory and Language* 32(3): 402–420. DOI: 10.1006/jmla.1993.1022.

Kabak, Baris; Idsardi, William (2003): «Syllabically conditioned perceptual epenthesis.» *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 29: 233–244. URL: <http://elanguage.net/journals/bls/article/view/3411>.

© Romania Minor
http://www.romaniaminor.net/ianua/

BDD-A5302 © 2012 Romania Minor
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 17:58:36 UTC)

Kelso, J.A. Scott (1995): *Dynamic patterns: The self-organization of brain and behavior*. Cambridge (MA): MIT Press.

Krakow, Rena Arens (1989): «The articulatory organization of syllables: A kinematic analysis of labial and velar gestures.» Ph.D. dissertation. Yale University.

Krakow, Rena Arens (1999): «Physiological organization of syllables: a review.» *Journal of Phonetics* 27(1): 23–54. DOI: 10.1006/jpho.1999.0089.

Leben, William Ronald (1973): «Suprasegmental phonology.» Ph.D. dissertation. Massachusetts Institute of Technology. URL: <http://hdl.handle.net/1721.1/16364>.

Lombardi, Linda (2002): «Markedness and the typology of epenthetic vowels.» Ms. ROA 578. URL: <http://roa.rutgers.edu/article/view/588>.

Martínez-Gil, Fernando (1996): «El principio de la distancia mínima de sonoridad y el problema de la vocalización consonántica en el español dialectal de Chile.» *Hispanic Linguistics* 8(3): 201–246.

Martínez-Gil, Fernando (1997): «Obstruent vocalization in Chilean Spanish: A serial versus a constraint-based approach.» *Probus* 9(2): 167–202. DOI: 10.1515/prbs.1997.9.2.167.

Nam, Hosung; Saltzman, Elliot (2003): «A competitive, coupled oscillator model of syllable structure.» In: Maria-Josep Solé; Daniel Recasens; Joaquín Romero [ed.]: *Proceedings of the 15th International Congress of Phonetic Sciences. Barcelona, Spain. August 3–9, 2003*. Bellaterra: Universitat Autònoma de Barcelona, 2253–2256. URL: <http://www.haskins.yale.edu/Reprints/HL1307.pdf>.

Öhman, S. (1967): «Peripheral motor commands in labial articulation.» *STL-QPSR* 8(4): 30–63. URL: <http://www.speech.kth.se/prod/publications/files/qpsr/1967/1967_8_4_030-063.pdf>.

Parker, Stephen G. (2002): «Quantifying the sonority hierarchy.» Ph.D. dissertation. University of Massachusetts. URL: <http://www.ai.mit.edu/projects/dm/theses/more/parker02.pdf>.

Perkell, Joseph S. (1969): *Physiology of speech production: Results and implications of a quantitative cineradiographic study*. Cambridge (MA): MIT Press.

Prince, Alan; Smolensky, Paul (1993): «Optimality Theory: Constraint Interaction in Generative Grammar.» Ms. Rutgers University. ROA 537. URL: <http://roa.rutgers.edu/article/view/547>.

Saltzman, Elliot L.; Munhall, Kevin G. (1989): «A dynamical approach to gestural patterning in speech production.» *Ecological Psychology* 1(4): 333–382. DOI: 10.1207/s15326969eco0104_2.

Shaw, Jason A.; Gafos, Adamantios; Hoole, Philip; Zeroual, Chakir (2009): «Syllabification in Moroccan Arabic: Evidence from patterns of temporal stability in articulation.» *Phonology* 26(1): 187–215. DOI: 10.1017/S0952675709001754.

BDD-A5302 © 2012 Romania Minor
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 17:58:36 UTC)

Shepherd, Michael Andrew (2003): «Constraint interactions in Spanish phonotactics: An optimality theory analysis of syllable level phenomena in the Spanish language.» MA thesis. California State University. ROA 639. URL: <http://roa.rutgers.edu/article/view/649>.

Sproat, Richard; Fujimura, Osamu (1993): «Allophonic variation in English /l/ and its implications for phonetic implementation.» *Journal of Phonetics* 21: 291–311.

Stampe, David (1972): «How I spent my summer vacation (A dissertation on natural phonology).» Ph.D. dissertation. University of Chicago.

Werker, Janet F.; Tees, Richard C. (1984) : «Cross-language speech perception: Evidence for perceptual reorganization during the first year of life.» *Infant Behavior and Development* 7:49–63. [Republished in *Infant Behavior and Development* 25(1): 121–133 (2002). DOI: 10.1016/S0163-6383(02)00093-0]

Wright, Richard (2004): «A review of perceptual cues and cue robustness.» In: Bruce Hayes; Robert Kirchner; Donca Steriade [ed.]: *Phonetically based phonology*. Cambridge: Cambridge University Press, 34–57.

Mark Gibson
Universidad de Navarra
Facultad de Filosofía y Letras
1330 Edificio de Bibliotecas
E-31080 Pamplona – Navarra
Spain

© Romania Minor
http://www.romaniaminor.net/ianua/

BDD-A5302 © 2012 Romania Minor
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 17:58:36 UTC)