



Perception-based vowel insertion by native Spanish-speaking learners of English

Mark Gibson



Electronic version

URL: <http://journals.openedition.org/tipa/205>
DOI: 10.4000/tipa.205
ISSN: 2264-7082

Publisher

Laboratoire Parole et Langage

Electronic reference

Mark Gibson, « Perception-based vowel insertion by native Spanish-speaking learners of English », *TIPA. Travaux interdisciplinaires sur la parole et le langage* [Online], 28 | 2012, Online since 29 October 2012, connection on 01 May 2019. URL : <http://journals.openedition.org/tipa/205> ; DOI : 10.4000/tipa.205

This text was automatically generated on 1 May 2019.



La revue *TIPA. Travaux interdisciplinaires sur la parole et le langage* est mise à disposition selon les termes de la licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International.

Perception-based vowel insertion by native Spanish-speaking learners of English

Mark Gibson

Si buscas una buena solución y no la encuentras,
consulta al tiempo, puesto que el tiempo es la
máxima sabiduría.
Tales de Mileto

1. Introduction

- 1 It is an intuitively straightforward observation that the emergence of a prosthetic [e] to the left margin of word-initial [sC] sequences in the L2 English speech of native Spanish-speakers is correlated in an intimate way to conventions of L1 syllable structure:

(1)	English word	Spanish output
	stop	[es.tóp]
	sleep	[es.líp]
	smart	[es.márt]
	spot	[es.pót]

--	--	--

- 2 Notice that [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.
- 3 Computationalist¹ approaches rooted in American behaviorist thinking explain vowel prosthesis as the insertion of a post-lexical [e] which results from a top-down, production-based repair process triggered by the interference of some L1 phonological stipulation in the L2 grammar ². In such constructs, it is generally acknowledged that the L1 and L2 grammatical systems function autonomously of one another. The architectures of the individual grammars by design provide no explicit correlation between high-level phonological abstractions and the low-level articulatory procedures involved in prosodic structure.
- 4 The pertinent phonological information which is presumably transferred to the L2 grammar has historically been centered round a feature-based, segmental approach to the phonology, and the means by which this information affects the L2 grammar has been expressed theoretically by way of (1) structure building rules which emerge linearly in the L1, (2) the fulfillment of language-specific L1 syllabic templates, and/or (3) highly ranked universal markedness laws governing L1 structural well-formedness. The vast majority, if not all phonological justifications of vowel insertion hinge on the supposition that [sC] sequences are banned in Spanish complex onsets, yet detrimentally, none offers a coherent description of the phonological conditions on which such an assertion is determined.
- 5 Phonological models of vowel prosthesis have been based to a large extent on the notion that a post-lexical [e] aligns to a non-prosthetic input in order to appease high-level injunctions governing complex onset formation in Spanish. Segment-based approaches to the phonology express the structural well-formedness of complex onset sequences by way of the linear organization of timing-independent spatial units. Most often, the correct arrangement 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 algorithm³. In this conception, there is no linear path which correlates invariable, categorical sonority values with low-level patterns of vocal tract constriction, nor any programmed explanation as to how low-level spatio-temporal relationships involving vocal tract constriction and constriction location affect optimal outputs.
- 6 Apart from the reductionist and quasi-tautological character of the top-down justification of vowel prosthesis, a number of functional problems also arise which challenge the veracity of the purely phonological explanation. The main problem is that models of complex onset formation derived from constraints governing sonority distance are dysfunctional, in the sense that there is no way to justify the emergence of all optimal output sequences while simultaneously explaining the elimination of sub-optimal sequences.⁴ Consider a stipulation 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 [fr], which are wholly admissible complex onset sequences in Spanish, must be explained away as exceptions to the general rule.⁵ This last pitfall is problematic in the sense that it presents grave problems with regard to acquisition and learnability algorithms. Harris (1989) 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 surface optimally, of which only [fr] and [fl] constitute well-formed complex onsets. Neither minimal sonority distance addresses the elimination of [dl] onsets⁶.

- 7 Let us suppose for the sake of argument that a minimal sonority distance of three were a genuine grammatical principle reflecting L1 phonological information related to syllable structure, albeit Spanish-specific. Its high ranking in an immature L2 grammar could, hypothetically, trigger vowel prosthesis in all /sC/ clusters. After all, there is no possible combination of [sibilant/C] clusters which could meet this criterion. However, this explanation cannot account for the fact that a post-lexical vowel does not surface in word-initial [θr] clusters in the L2 English speech of L1 Spanish-speakers: [θrow], and *[eθ .row] *throw*. Notice that according to a minimal sonority distance of three, [θr] would be just as ill-formed as any [sC] sequence.
- 8 The fact of the matter 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 (1) be transferred to the L2 grammar, and later (2) play a role in triggering the emergence of a post-lexical repair process, is theoretically tenuous and functionally unsupported.
- 9 Despite the general ineffectiveness of the previous computationalist approaches, the claim that prosthetic outputs are derived from a non-prosthetic lexical form has prevailed in the recent literature. In this conception, the lexical form of word-initial [sC] clusters matches the non-native output form. Such an assertion ignores strong empirical evidence to show that the perception, and subsequent lexical storage, of continuous non-native speech streams is heavily influenced by the L1 phonology (see Best, 1995; Hallé *et al.*, 1998; Dupoux *et al.*, 1998; Kabak & Idsardi, 2003; Andersen, 2004; Best & Hallé, 2007). Dupoux *et al.* (1998), among many others, show that distortion of the auditory input prompted by the L1 phonology may indeed render an innovative lexical form even if acoustic cues to support the novel representation are nonexistent.
- 10 In the remainder of this article an alternative hypothesis to justify vowel prosthesis is proposed and experimental data corroborating the central premises of this claim are provided. The focal hypotheses underscoring the following study are that (1) vowel prosthesis is a perception-based repair phenomenon, and (2) conditioned by a language-specific bias based on principles of relative timing between gestures in complex onsets. The first claim naturally refutes top-down, production accounts of vowel insertion in word-initial [sC] sequences and is founded on the proposal that cognitive biases for native timing configurations can influence the perception of acoustic inputs.

- 11 The second hypothesis supposes the specific L1 bias to be the byproduct of parameter-dynamic attunement to L1 patterns of spatio-temporal coordination in complex onsets, a proposal which necessarily contradicts segmental accounts of complex onset formation based on sonority hierarchies. It is believed that in the event that an unlicensed, non-native timing pattern is perceived, then a bias for native patterns of intergestural timing will motivate a repair process before lexical storage occurs.

2. Perceptual epenthesis

- 12 The claim that the perception of a continuous speech stream can be influenced by the L1 phonology of a listener/perceiver was first introduced in Dupoux *et al.* (1998) and elaborated on in Kabak & Idsardi (2003), among many others. Specifically, the data show that an illusory vowel can be perceived regardless of whether the acoustic properties of the speech stream involve a vocalic unit. Dupoux *et al.* (2001) and Dehaene-Lambertz *et al.* (2000) neurolinguistic study corroborate the general findings outlined in Dupoux *et al.* (1998) concerning the emergence of an epenthetic unit in Japanese, while Kabak & Idsardi (2003) study of the perception of vocalic intrusion by Korean informants suggests that the illusory vowel is correlated in a programmed way to conventions of L1 syllable structure.
- 13 A more recent study by Hallé *et al.* (2008) corroborates the claim made by Kabak & Idsardi (2003) concerning the role of syllable structure in the emergence of the vocalic unit. In a series of perception tests, a group of native Spanish-speaking informants were presented stimuli which contained an illicit word-initial [sC] sequence. The researchers found that in a majority of the cases, an illusory vowel was perceived by the L1 Spanish informants. The authors concluded that the emergence of [e] was in fact correlated to the ill-formedness of the [sC] onset, although no account was offered addressing the conditions on which the illicitness of the [sC] cluster was determined.
- 14 Inherent to the specific interpretation of the concept of perceptual epenthesis outlined in this article is the viewpoint that high-level linguistic abstraction specific to individual languages and universal anatomical constraints imposed on the speech and auditory organisms interact in a unified system. The non-native auditory input is perceived by the listener and later distorted by a system of high-level biases construed from the systematic patterns of L1 syllable structure before lexical storage occurs.
- 15 Below, an argument is laid which proposes that the phonological conditions which trigger the perception of a prosthetic vowel to the left of word-initial [sC] clusters by L1 Spanish-speakers are correlated to the systematic articulatory patterns of intergestural timing in Spanish complex onsets. Seen in this way then, the misperception of the word initial [sC] cluster is an upshot of the incomplete acquisition of L2 phasing relationships. It is believed that parameter-dynamic attunement patterns during L1 learning translate productive stipulations related to native relative timing patterns into phonological biases which are accessed by listeners in order to assess acoustic inputs.

2.1 Patterns of relative timing in English and Spanish complex onsets and the (mis)perception of timing relations

- 16 It has been argued in past work that Spanish permits two basic phasing patterns by which to organize two-consonant sequences (Bradley & Schmeiser, 2003; Gibson, in progress).

Two-consonant sequences whose phonetic properties permit the gestures to be arranged globally in relation to a following vocalic target may form complex onsets, while [CC] sequences in which no coordination relationship with the following vocalic target is possible must be bonded to different vocalic units:

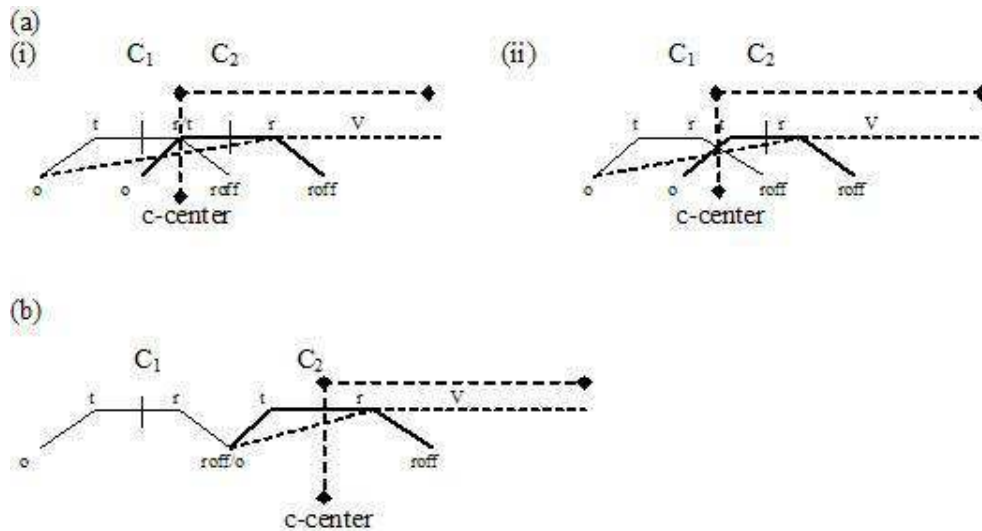


Fig. 1. The phasing relations in (a) represent the two permissible patterns of intergestural timing in Spanish complex onsets. The timing relation in (b) represents a first gesture which is arranged linearly in relationship to the preceding vowel and a second gesture which is arranged globally in relation to the following vocalic target. In (b), the first consonant constitutes a coda gesture and the second consonant constitutes an onset gesture.

- 17 Independent acoustic data suggest that [sC] clusters in Spanish do not show signs of gestural coordination in complex onsets. A number of studies, however, reveal that the gestures in English [sC] clusters are consistently arranged such that the c-center of the cluster is phased globally in relation to the following vocalic target (c-center effect) (Browman & Goldstein, 1988; Honorof & Browman, 1995; Byrd 1995). X-ray microbeam data illustrate that the relative timing of the latencies for the c-center to the vocalic target in tokens such as *said*, *payed* and *spayed* are all relatively stable. Even though the tongue tip gesture for [s] in *spayed* occurs before the same gesture in *sayed*, and the lip gesture in [p] in *spayed* is realized later than the same gesture in *paid*, the c-center remains remarkably stable, meaning that in English complex onsets, as opposed to Spanish, the initial sibilant can be arranged globally (in-phase) in relation to the following vocalic target (Browman & Goldstein, 1988).
- 18 The previous studies illustrate that Spanish and English differ with respect to the relative timing patterns in complex onsets. The phonetic characteristics of the consonant gestures in [sC] clusters do not permit in-phase coupling in Spanish, while English permits onset phasing in many [sC] sequences. It is expected that when a non-native input containing impermissible timing configurations is perceived, a bias derived from native patterns of intergestural timing will prevent a correct mapping from the continuous acoustic input to discrete phonological units before lexical storage takes place.

2.2 The development of L1 biases derived from systematic patterns of spatio-temporal coordination

- 19 In this section it is assumed that stable patterns of relative timing are encoded cognitively as coordination relations between gestures, constituting a formal part of the productive grammar, and lead to the development of biases which listeners access in order to map acoustic inputs to representational units before lexical storage occurs. If humans decode acoustic cues as vocal tract gestures, then the natural upshot is that the relative timing relations between those articulatory movements is also (1) discernible, and (2) cognitively pertinent.
- 20 Concerning (1) and (2), Baker *et al.* (2007) show that speakers of SSBE are able to discern between phonemically-identical acoustic inputs on the basis of phonetic-level duration alone, suggesting a patent relationship between timing, cognition and lexical access. In another case, Nazzi *et al.* (1998) exposed human infants to recordings of low-pass filtered sentences which only preserved the prosody. They found that the infants' sucking rate increased initially and diminished when the infants became accustomed to the speech rate, whereupon the rate class was changed and sucking rate increased, showing that sensitivity to timing and other temporal aspects of speech constitute a natural process of phonological development.
- 21 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.
- 22 With regard to syllable structure, Bertoncini & Mehler (1981) discovered that two month olds were more inclined 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.
- 23 It has been shown that L1 biases can affect the perception and acquisition of non-native features. It is believed that the distortion of L2 auditory inputs stems in part from the fact that during L1 phonological development, normal neural maturation leads to the loss of neuroplasticity, which in turn hampers the attunement of L2 phonological structures (Long, 1990; Scovel, 2000). Flege (1995), Iverson *et al.* (2003), Kuhl (2000), McClelland (2001) Pallier *et al.* (1997, 2003), and Sebastián-Gallés & Soto-Faraco (1999), in fact, go a step further and suggest that L1 phonological biases may actually impede L2 phonological acquisition. Iverson *et al.* (2003) argue that the distortion of L2 auditory characteristics

takes place at an early phonetic or late auditory level of processing and becomes increasingly problematic to reverse later in life.

- 24 Results from recent research addressing the discrimination of the Catalan vowels [e] and [ɛ] by early Spanish-speakers corroborate the claims originally made in Flege (1995) regarding the role of L1 phonological biases in L2 phonological acquisition. Data from a series of discrimination tasks reveal that native Spanish-speakers exhibit a marked difficulty in discerning between [e] and [ɛ] in native Catalan words, suggesting a strong interference of the L1 phonological system in L2 perception (Bosch *et al.*, 2000; Pallier *et al.*, 1997, 2001; Sebastián-Gallés & Soto-Faraco, 1999). In another study treating the effects of L1 syllable structure in the L2, Matthews & Brown (2004) suggest that the emergence of epenthetic segments is conditioned directly by L1 biases for native arrangements. The authors argue that the acoustic signal may be distorted upon encountering segmental sequences unlicensed by the L1 phonology. Interestingly, the authors found that distortion of the auditory input decreases when the duration of interstimulus intervals was varied and that auditory perturbations may be evaded when listeners have access to more detailed acoustic representations.
- 25 The claim that biases for specific timing configurations may influence the perception of an acoustic input has ample precedence outside of language. Heuer & Schmidt (1988) show that invariance in relative timing of certain motor patterns is a strategic phenomenon provoked by biases for particular temporal configurations which are dependent on individual spatial organizations. In Repp's study (1992), a group of musically knowledgeable informants was presented with a musical excerpt in which the relative timing relations between the notes had been modified. The results of a series of detection tasks designed to reveal the informants' expectations of specific timing patterns showed that the listeners routinely accessed prior knowledge derived from the cognitive representation of micro-timing structure in music in order to make their determinations.
- 26 The study of the effects of timing relations in the perception of language is still in its infancy. A bias for native coupling configurations is derivationally tenable if certain statistical heuristics and Kolmogorovian variables are accounted for. The first and most observable extrinsic factor is that all complex onsets in Spanish adhere to a strict paradigm of temporal coupling. This seems to be a general principle found in most languages studied. Differences in permissible sequences between languages can be accounted for by the fact that the spatial characteristics of the phonetic units across languages may vary and that languages may be more or less strict in their adherence to phasing principles.
- 27 From the point of view of L1 acquisition, this means that all (100%) of the output data available to a language learner exhibit identical patterns of in-phase coupling. The conglomeration of sequences whose spatial characteristics permit in-phase coupling can be represented cognitively as (χ) , in which (τ) represents the temporal relationship between the spatial units (g) . In this abstraction (τ) is not only a sub-component of (χ) , but rather one of its defining characteristics. The fact that (χ) always contains (τ) , and that (τ) does not exist autonomously of (χ) means that the probability of codifying a rule: $(\chi) = \text{the coordination of } (g) \text{ in a relation } (\tau)$, as a productive grammatical function are better than casual ($p \text{ value} < 0.05$). From the point of view of perception then, a rule can be deduced which states that all gestures in [CCV] sequences must coordinate in line with

(τ). The fact that (τ) constitutes a consistent reality of onset formation available to all perceivers, means the viability of such a rule is theoretically credible (Linnell, 1982).

- 28 But a bias (β) for (τ) cannot be justified by extrinsic factors alone. In order to be evolutionarily optimal, it must be shown that (β) serves a functional purpose as well. After all not all systematic low-level patterns must be functionally represented. With the exception of extremely rare cases, for example, most articulatory patterns are realized during the expulsion of air from the respiratory system. It would be folly to suggest that such systematicity were likewise projected onto the grammar. It is merely an anatomic, and perhaps anthropophonic concurrence. In order to qualify as a bias which is capable of distorting the auditory signal, (β) must (1) make some qualitative improvement on the economics of the production/perception loop, and (2) aid in the translation of low-level motor movement to high-level phonological abstraction, ultimately facilitating lexicon construction and access. In the case of a bias for native coupling patterns, (1) is fulfilled if (τ) can provide a more rapid identification of complex gestural sequences, while (2) is accomplished if it can be shown that (τ) correlates low-level motor patterns to phonological representation and later meaning. Below, the results of three distinct tasks designed to test these hypotheses are reviewed.

3. Perception of illusory syllables in [sC]/[sCr/l] clusters by L1 Spanish-speakers

- 29 Here the two central claims of this article are tested employing a series of off-line perception tests. First, the proposal that vowel prosthesis in the L2 English speech of native Spanish-speakers is a perception-based phenomenon is addressed. A syllable-counting test was designed in order to uncover the quantity of syllables a native Spanish-speaker perceives in tokens containing word-initial [sC] clusters. It is expected that if vowel prosthesis is indeed a perceptual phenomenon, then native Spanish-speakers should perceive an extra syllable in the target stimuli. Additionally, it is proposed that if phonological conditions based on patterns of relative timing are responsible for the perception of an illusory vowel, say as opposed to a paradigm based on sonority distances, then there should be no significant difference between the perception of an extra syllable in stimuli containing word-initial [s+stop], [s+nasal] and [s+liquid] sequences. Finally a set of identification and discrimination tasks were created in order to assess the viability of the claim for a bias based on systematic patterns of relative timing relations in complex onsets.

3.1 Syllable-counting task (unmodified stimuli)⁷

3.1.1 Research questions and hypotheses

- 30 The current study seeks to shed light on the topic of vowel prosthesis in word-initial [sC] sequences by L1 Spanish learners of English. The principal assumption is that the emergence of an illusory vowel is the upshot of an L1 bias favoring native phasing relations between gestures in complex onsets. When a non-native input is perceived, it is hypothesized that the structure will be processed in accordance with L1 phasing patterns, meaning that [s] will be perceived as being arranged linearly (anti-phase $\phi 180^\circ$) in

relation to a preceding non-existent vocalic unit, and the following consonant will be perceived as a singular onset gesture.

- 31 In order to corroborate the notion that perception of the illusory vowel is derived from characteristics of spatio-temporal coordination, as opposed to segmental properties, the current study also seeks to uncover whether there is any scientifically significant difference between the perception of [e] in [s+stop], [s+nasal] and [s+liquid] clusters, which there should be if stipulations based on sonority sequences were a viable cognitive reality. It is to be expected that, if sonority sequencing is indeed responsible for the perception of a prosthetic [e], there should be a marked difference between the perception of [e] in word-initial [s+stop] sequences and [s+liquid] clusters. Since the sonority difference between [s] and [l] is greater than that between [s] and unvoiced stops, then the perception of a prosthetic [e] in the former cases should be substantially lower than in the latter cases. If, however, complex onset formation is a matter of intergestural timing, then no significant difference should emerge between the aforementioned sequences, since no combination of gestures in Spanish [sC] sequences permit onset coupling. That is, [sl] and [s+stop] combinations are equally illicit if intergestural timing is the chief determiner of complex onset formation.

3.1.2 Design

- 32 The current study is a component of a larger, two-year long longitudinal endeavor which seeks to uncover the intuitive knowledge that L1 Spanish-speaking children have of their native language and how this information affects L2 acquisition. The pertinent portion of this research program presented here involves a series of off-line perception tests in which a set of phonotactically illicit (restricted to word-initial [sC] sequences) nonce sequences are dispersed throughout a list of admissible word-initial onset clusters.

3.1.3 Informants

- 33 The informants were L1 Spanish-speaking children aged ten to eleven. Preliminary interviews were conducted to test for auditory impairment. No informant showed any indication of hearing impairment. In total there were fifty informants.
- 34 In order to test the informants' level of English, the informants were prompted to talk freely about a favorite hobby or activity. Later, each informant was asked to read a passage in English in which examples of the target sequences appeared. Although the levels of vowel insertion varied, all informants inserted a prosthetic vowel before the target clusters in at least half of the cases.
- 35 The informants were asked to complete a questionnaire regarding their specific contact with English and time spent in an English speaking country. The majority have never had any formal experience outside of the core curricular English classes given at their primary school, which translates to roughly 960 total hours of instruction⁸. Of the fifty informants, nineteen had a private English tutor outside of the school, although no data were collected to confirm that the teacher was an L1 English speaker. Three of the informants had spent significant time, defined here as extending longer than two months, in an English-speaking country. Of those three, one spoke English with his mother and caretaker, both of whom were native English speakers from the Philippines.

- 36 As regard to formal experience dealing with knowledge of syllables, all informants were aware of what a syllable was and could detect the quantity of syllables with accuracy in native Spanish words.

3.1.4 Stimuli and subjects' task

- 37 A total of fifty stimuli containing both target nonce stimuli and phonotactically licit structures were recorded by a native English speaker in a sound proof studio with minimal noise. Nonce tokens were used because possible familiarity with the orthography of a genuine word may affect how the subjects perceive the stimuli (Ziegler & Ferrand, 1998). The only phonotactically illicit sequences involved onsets. All codas were licit gestures in both Spanish and English. Vowel gestures were alternated in each token and coincided with prototypical rules of English syllable structure. That is, closed syllables had a short vowel while open syllables contained a long vowel. Four seconds were allowed between each token in order to allow the informants time to mark their responses on the answer sheet provided. The subjects listened to the target stimuli on loudspeakers connected to an interactive blackboard in a sound-proof room. The volume of the recording was set to a comfortable sound level. Although the informants listened as a group (2 groups of 25 informants), a monitor instructed beforehand that no talking was permitted. The informants were dispersed sufficiently throughout the room as to prevent them from copying their answers from another subject. The groups were instructed on the task they were to perform, that each stimulus would be heard only one time and that they had four seconds to mark their response.
- 38 The subjects were given an answer sheet on which to mark their responses. They were told to mark the number of syllables each stimulus contained. The answer sheet listed numbers from 1-5 which corresponded with the number of syllables. Responses were marked by circling the number which corresponded with the number of syllables the subjects heard. As the informants were children, they were told that a prize would be given to the "winner" in order to maximize attention.

3.1.5 Results

- 39 The results of the tests were by and large consistent with the forecasted hypothesis regarding the perception of an extra syllable in all tokens with a word-initial [sC]/[sCr/l] sequence. Only responses which included the perception of one extra syllable were admitted. In each group, there were a minor number of divergent responses which were singularly discarded before the statistical analyses were performed. Divergent responses included those in which the quantity of syllables perceived was more than one, or those in which the quantity perceived was smaller than the actual correct response:

Tab. 1. Table of results of syllable-counting task

[sp]		[st]	[sk]	[spr]	[spl]	[str]	[skr]	[sm]	[sn]	[sl]
2	#tokens	2	2	1	1	2	2	2	2	2
1	#syllables	2	3	1	1	2	3	1	1	1
1	per token	2	2			1	3	1	2	2
14	Correct	15	19	11	6	21	20	25	18	14
6	responses	16	15			9	22	16	14	11
34	Responses	32	28	37	40	26	26	22	31	34
42	extra σ	32	32			37	23	30	30	38
29.16 %	% correct	31.91 %	40.42 %	22.91 %	13.04 %	44.68 %	43.47 %	53.19 %	36.73 %	29.16 %
12.24 %		33.33 %	31.91 %			19.56 %	48.88 %	34.78 %	31.11 %	22.44 %
72.34 %	% with	68.08 %	59.57 %	77.08 %	86.95 %	55.31 %	56.52 %	46.80 %	63.26 %	70.83 %
85.71 %	extra σ	66.66 %	68.08 %			80.43 %	51.11 %	65.21 %	68.18 %	77.55 %

- 40 As table 1 illustrates, out of a total of 846 (18 tokens x 50 possible responses, - discarded examples) possible responses for all stimuli, 574 (67.84%) were perceived as containing an extra syllable. The [s+stop] sequences entailed 285 total responses (6 tokens x 50 informants, - discarded examples). Of these 285 possible responses, 200 were perceived to have an extra syllable (70.17%). There were 186 possible responses for [s+nasal] stimuli, of which the perception of an extra syllable appeared in 113 (60.75%) of the cases. Out of 97 total responses for tokens containing a word-initial [sl] sequence, 72 (74.22%) were believed to contain an additional syllable. The average mean of incorrect responses between the three groups is calculated at 68.37% and the standard deviation for each group was as follows: [s+stop] = 1.8; [s+nasal] = 7.62; [s+liquid] = 5.85. A one-way ANOVA comparing the percentages for the perception of an added syllable in [s+stop], [s+nasal] and [s+liquid] reveals no significant difference between the three group; p value > 0.05.
- 41 Regarding the tokens in which the target stimuli did not appear, and therefore in theory should be marked with a certain accuracy, 158 (9.86%) responses out of a possible 1600 total responses were misperceived by the subjects. It is interesting to note that, contrary to the misperceived responses for the target stimuli, the incorrect responses for the phonotactically legal tokens were highly sporadic, and even inconsistent from one subject to another. When comparing the percentages for misperceived responses for the two groups (target stimuli and phonotactically licensed tokens), a two-sample z-test of proportions with a .05 critical alpha level renders the difference significant: z-value 29.733.
- 42 Predictably, the learners who had spent significant time in an English-speaking country fared better statistically with regard to the perception of an additional syllable. Two of the informants perceived an illusory syllable in only 8% of the total stimuli, while the other subject who had spent significant time in an English-speaking country perceived an extra syllable in 38% of the target stimuli. All were significantly lower than the global average of 67.84%. Nevertheless, one subject who had never spent any time in an Anglophone country scored the highest with a score of 100% for all target and non-target stimuli.

3.1.6 Discussion

- 43 The data indicate that (1) L1 Spanish speakers perceive an illusory syllable in a vast majority of the target stimuli, and (2) the perception of the extra syllable seems to be independent of simple segmental organization.
- 44 The hypothesis that vowel prosthesis in word-initial [sC] sequences is a perception-based process conditioned by L1 biases pertaining to syllable structure is feasible. The results indicate that the perception of an illusory syllable (1) is prevalent in all target sequences, and (2) is unsupported in non-target stimuli. Although more studies will be necessary in order to pinpoint the exact conditions on which the L1 bias is based, the explanation provided here founded on spatio-temporal coordination is corroborated by the current study.
- 45 The idea of an L1 bias derived from language-specific patterns of intergestural timing relations is circuitously supported by the results of the study. 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 qualitative 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.
- 46 Evidence not presented above but derived from a different portion of the tests further supports the argument in favor of a bias derived from native patterns of intergestural timing in complex onsets. Results of a reading and writing section which accompanied the syllable-counting task provide proof that L1 Spanish-speakers are intuitively aware that the [e] they perceive has no orthographic, or perhaps phonological, corollary in English. In a task in which a visual prime was provided, the informants were presented with a list of response options consisting of [esC], [seC], [sC] sequences and asked to indicate the token they heard. The results of these tests were remarkable in that the ratio of correct responses was significantly higher when visual stimuli were introduced. This suggests that a substantial number of the informants are conscience of the fact that the [e] they perceive is not mapped to any grapheme/phonological unit. It is reasonable to assume that this knowledge was not accessed, however, during the auditory task. It seems that even though the informants are aware of the linear/spatial organization of the individual units, they still have not fully acquired how these spatial components coordinate in time.

3.2 Discrimination and identification tasks

- 47 The results of the previous syllable-counting task show that native Spanish-speakers perceive an illusory syllable in all word-initial [sC] sequences. The results from a follow-up identification task reveal that the illusory syllable involves the emergence of a vocalic unit [e] to the left margin of the [sC] cluster. The fact that no significant difference between [s+stop], [s+nasal] and [sl] sequences was detected intuitively supports the secondary claim of this paper which asserts that the perception of an illusory vowel is

motivated by a phonological bias based on the relative timing of gestures in complex onsets.

- 48 During the attunement process, speakers/learners refine output timing relations until the output configurations they produce, Ψ^{OUT} , correspond to those found in the phonetic environment. In such a scenario, the temporal coordination of low-level motor tasks could be projected onto the grammar as high-level stipulations which make reference to certain articulatory actions over time. Once established, it is logical that these grammatical stipulations lead to the development of biases for specific native timing configurations which (1) aid in discerning perceptually salient cues used for gesture recovery in speech perception, and (2) later distort the perception of non-native arrangements. In the following section, results from a series of identification and discrimination tasks which seem to support this last claim are reviewed.

3.2.1 Discrimination task

- 49 A 4I-oddity discrimination test was performed in order to determine whether native Spanish perceivers are sensitive to relative timing patterns in complex onsets. The objective of the task was to determine if relative timing relations between complex onset gestures are perceptually salient. If the perception of native timing configurations can be proven to provide a more economic way to translate acoustic cues to phonological form, then the functionality of a bias for native phasing patterns is substantiated. It is instinctively plausible that such a paradigm does indeed provide a more economic translation since the perceiver is only burdened by identifying a singular timing relation, as opposed to identifying two separate segments. Afterward, the claim that temporal coordination between units is cognitively significant will be addressed.

3.2.1.1 Design

- 50 A total of twenty nonce stimuli containing complex onsets were recorded in a sound-proof studio by a 34 year-old native Spanish-speaker from Madrid, Spain. All standard phonotactic conventions of Spanish were observed when designing the nonce words. The timing of the individual units was thereupon modified using a manipulation command on Praat, affecting the global temporal organization of the sequence. Only the temporal characteristics of the units were affected. All spatial characteristics of the gestures were respected. That is stop closure and overlap patterns were left unmodified.
- 51 Four stimuli were presented to the informant. Three of the four stimuli were presented to the listener exactly as they were recorded by the speaker. One of the four tokens had been modified. The position of the modified stimulus was randomized in a list of the unmodified tokens. Half a second was provided between each token and the subject was given a short break after the tenth set of stimuli.

3.2.1.2 Informants and tasks

- 52 A total of thirty informants were chosen randomly from the original fifty children selected for the syllable-counting task. The children were instructed that their task was to listen to four stimuli provided on the recording and indicate which token was not like the rest by clicking on the number of the token which they perceived as different. The subjects performed the task individually in a sound-proof room on a portable computer

with a built-in auditory device. They were instructed that once the test began, no questions were permitted and the test could not be stopped. A follow-up test was then administered in which both modified and unmodified stimuli were presented again and the subject was asked to make a goodness judgment for the two inputs.

3.2.1.3 Results

- 53 The results of the discrimination task reveal that relative timing relations in complex onsets are perceptually salient. Of the 600 possible responses (20 stimuli x 30 subjects), 540 (90%) were judged correctly by the informants, while 60 (10%) of the total responses were assessed incorrectly. A one sample t-test renders this difference significant: $p < 0.05$. In a follow-up test, the informants were asked to provide a goodness judgment for each of the stimuli. In 100% of the cases, the native Spanish-speaking informants judged the unmodified stimuli more favorably than the modified tokens.

3.2.2 Identification tasks

- 54 A series of forced-choice identification tasks were performed in order to determine if L1 Spanish-speakers access information derived from systematic patterns of intergestural timing in order to map acoustic streams to categorical units before lexical storage occurs. For a bias based on stable phasing patterns to be evolutionarily plausible, it must be shown that the timing relations between gestures are not only functionally correlated with low-level motor coordination, but also linked to high-level phonological abstraction, ultimately facilitating lexical construction and access. The global hypothesis is that perceivers will misinterpret acoustic inputs in which the temporal specifications of the individual units in complex onset gestures have been altered.

3.2.2.1 Design

- 55 The identification task was designed in order to test whether L1 Spanish-speaking children were able to decipher an acoustic input when the relative timing relations of the gestures were modified and if temporal relations between onset gestures in complex onsets form part of an L1 Spanish-speaker's intuitive phonological knowledge pertaining to syllable structure. A basic hypothesis is tested which predicts that L1 Spanish-speakers will misinterpret inputs containing complex onsets in which the temporal measurements of the units have been modified, corroborating the claim that the temporal coordination between units is cognitively and grammatically relevant.
- 56 The test contained a total of thirty stimuli containing complex onsets. All the complex sequences used in the target stimuli appear in native Spanish words. Of the thirty tokens, the relative timing of the consonants in the complex onsets were modified in fourteen tokens. The remaining sixteen stimuli were presented as recorded. Modification of the target stimuli was carried out using a manipulation command on Praat. As the objective of this study was not to discover at which point communication of the acoustic input is ruptured, the modifications of the stimuli were not performed in equal intervals. The target stimuli and unaltered tokens were randomized in such a way as to avoid the appearance of two identical stimuli types in sequence.
- 57 As opposed to the stimuli in the syllable-counting task, the target tokens chosen for the identification task were authentic Spanish words containing a word-initial complex

onset. A special effort was made to choose stimuli for which minimal triplets existed on the basis of the onset gestures. For example in one instance, the modified target stimulus *plana* (target stimuli) ‘flat’ was presented and the informants were asked to choose the word they heard from a list which contained the words *pana*, *lana* and *plana*, ‘corduroy’, ‘wool’ and ‘flat’ respectively, all authentic words.

3.2.2.2 Informants and tasks

- 58 A total of thirty informants were chosen randomly from the original fifty selected for the syllable-counting task. Each child was instructed that their task was to listen to the stimuli provided and indicate their answer by clicking on the option provided to them on the computer screen. They were instructed that once the test began, no questions were permitted and the test could not be stopped.

3.2.2.3 Results

- 59 The results of the identity tasks were generally consistent with the expectations expressed in the hypothesis, although the rate of misperception was lower than anticipated. By and large, lexical access seemed to be inhibited by modifying the timing relations between the acoustic units, supporting the view that timing relations in complex onsets are cognitively and grammatically significant in Spanish:

Tab. 2. Results from forced-choice identity task

	Target stimuli (14 stimuli x 30 subjects= 420)		Unmodified stimuli (16 stimuli x 30 subjects= 480)	
	Correct responses	Incorrect responses	Correct responses	Incorrect responses
Total number	324	96	445	35
% of total	77.14 %	22.85 %	92.71 %	7.3 %

- 60 As table 2 indicates, 96 (22.85%) of the target stimuli were misinterpreted, presumably as a result of the timing modifications made. Misinterpretation of the lexical signal was highest in the stimuli which underwent greater temporal alterations, although the data pertaining to the incremental differences and misperception of the stimuli will be formalized in a subsequent study. Importantly, misperception of the signal was greatly reduced in the non-target stimuli, 7.3% (35 incorrect responses out of a total of 480 possible replies). Of the total 480 responses for the non-target tokens, 445 were perceived correctly. Upon comparing the percentages for incorrect responses in the two stimuli groups (target stimuli and non-modified tokens), a two-sample z-test of proportions with a .05 critical alpha level renders a z-value of 6.443, or significant.

3.2.3 Discussion

- 61 The results of the identification and discrimination tasks show that the modification of temporal characteristics of the units in complex onsets impedes a correct mapping from acoustic input to representational units. Based on these data, the claims that L1 Spanish listeners exploit temporal information in deciphering categorical representation from phonetic inputs and encode phonological information derived from the temporal arrangement of phonetic gestures seem feasible. Sensitivity to temporal aspects of spatial gestures in complex onsets does indeed seem to form part of the intuitive cognitive knowledge L1 Spanish-speakers have regarding native syllable structure. The development of a bias codifying information derived from native timing patterns would thus be an economical and effective mechanism by which language learners may begin to build an initial lexicon.

4. Conclusions

- 62 At the outset of this article a claim was made that vowel prosthesis in word-initial [sC] sequences in the non-native English speech of L1 Spanish-speakers is a perception-based repair process conditioned by L1 patterns of intergestural timing in complex onsets. Essentially, it was claimed that non-native timing configurations between spatial gestures can lead to the erroneous perception of an illusory vowel. The second assertion implies that perceivers not only access phonological information regarding the linear organization of spatial gestures, but also exhibit sensitivity to how these spatial characteristics coordinate in time.
- 63 Subsequently, the results of a series of perception tests were reviewed which seem to corroborate the central claims of this article. The initial syllable-counting task revealed that in a high number of tokens, native Spanish speakers do indeed perceive an illusory vowel in most word-initial [sC] sequences, corroborating the general findings outlined in Hallé & colleagues (2008). The fact that no scientifically significant difference was uncovered between the perception of an illusory vowel in [s+stop], [s+nasal] and [sl] sequences superficially substantiates the argument that perceivers utilize patterns of intergestural timing in order to map acoustic units to categorical representations before lexical storage occurs.
- 64 Later, results from a set of discrimination and identification tasks showed that the temporal organization of the acoustic units is perceptually salient and cognitively pertinent. Although, at this point it would be premature to categorically pronounce that a bias based on intergestural timing relations were solely responsible for the perception of an illusory vowel in non-native [sC] sequences, this study has laid the cornerstone for future research in this arena.
- 65 As regard to the acquisition of relative timing patterns, this study provides evidence to suggest that the acquisition of native timing patterns forms a crucial part of phonological development. It has also been shown that information derived from these patterns (1) seems to form part of native speaker's intuitive phonological knowledge, and (2) is accessed in order to assess innovative inputs.
- 66 Finally, the results of this current study offer insight into the acquisition of non-native timing relations and the strategies language learners must utilize in order to attain

phonological accuracy in the L2. If the assumptions regarding the effects of native timing relations on L2 phonological acquisition are correct, then attaining native-like accuracy cannot be understood merely as a convention of suppressing certain units. Rather, it seems that the acquisition of phonetic accuracy in the L2 depends on the active processes of learning and storing L2 coordination patterns to cognition, and later accessing this information in order to assess and produce non-native timing configurations.

BIBLIOGRAPHY

- Abrahamsson, N. (1999) Vowel Epenthesis of /sC(C)/ Onsets in Spanish/Swedish Interphonology: A Longitudinal Case Study, *Language Learning*, 49, 3, p. 473–508.
- Altenberg, E. (2005) The judgment, perception, and production of consonant clusters in a second language, *International Review of Applied Linguistics*, 43, p. 53–80.
- Andersen, R. (2004) Phonological acquisition in preschoolers learning a second language via immersion: a longitudinal study, *Clinical Linguistics and Phonetics*, 18, 3, p. 183–210.
- Baker, R., Smith, S., Hawkins, S. (2007) Phonetic differences between mis- and dis- in English prefixed and pseudo-prefixed words, *Proceedings of the 16th International Congress of Phonetic Sciences*, Saarbrücken, Germany, Aug 6–10.
- Bertoncini, J. & J. Mehler (1981) Syllables as units in infant speech perception, *Infant Behavior and Development*, 4, p. 247–260.
- Bertschinger, N. & T. Natschläger (2004) Real-Time Computation at the Edge of Chaos in Recurrent Neural Networks, *Neural Computation*, 16, 7, p. 1413–1436.
- Best, C. (1995) A direct realist view of cross-language speech perception, in Strange, W. (ed) *Speech Perception and Linguistic Experience: Theoretical and methodological issues*, Baltimore, MD: York Press, p. 171–204.
- Best, C. & P. Hallé (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, p. 2899–2914.
- Biswal, B. & C. Dasgupta (2002) Neural Network Model for Apparent Deterministic Chaos in Spontaneously Bursting Hippocampal Slice, *Physical Review Letters*, 88, p. 88–102.
- Bosch, L., Costa, A., Sebastián-Gallés, N. (2000) First and second language vowel perception in early bilinguals, *European Journal of Cognitive Psychology*, 12, p. 189–221.
- Boudaoud, M. & W. Cardoso (2009) Vocalic [e] Epenthesis and Variation in Farsi-English Interlanguage Speech, *Concordia Working Papers in Applied Linguistics*, 2, p. 1–34.
- Bradley, T. & B. Schmeiser (2003) On the Phonetic Reality of Spanish /ɾ/ in Complex Onsets, in Kempchinsky, P. & C. Piñero, (eds) *Theory, Practice and Acquisition*, Somerville, MA: Cascadia Press, p. 1–20.
- Browman, C. & L. Goldstein (1988) Some Notes on Syllable Structure in Articulatory Phonology, *Phonetica*, 45, p. 140–155.

- Byrd, D. (1995) C-Centers revisited, *Phonetica*, 52, p. 263-282.
- Byrd, D. (1996) Influences on articulatory timing in consonant sequences, *Journal of Phonetics*, 24, p. 209-244.
- Cairns, C. & M. Feinstein (1982) Markedness and the Theory of Syllable Structure, *Linguistic Inquiry*, 13, 2, p. 193-226.
- Clements, G. N. (1990) The role of the sonority cycle in core syllabification, in Kingston, J. & Beckman, M. (eds) *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*, Cambridge: Cambridge University Press.
- Colina, S. (1995) *A constraint-based analysis of syllabification in Spanish*, Catalan and Galician, Doctoral dissertation, University of Illinois at Urbana-Champaign.
- Colina, S. (2009) *Spanish Phonology: A Syllabic Perspective*, Washington D.C.: Georgetown University Press.
- Dehaene-Lambertz, G., Dupoux, E., Gout, A. (2000) Electrophysiological correlates of phonological processing : a cross-linguistic study, *Journal of Cognitive Neuroscience*, 12, 4, p. 635-647.
- Dupoux, E., Hirose, Y., Kakehi, K., Pallier, C., & Mehler, J. (1998) Epenthetic Vowels in Japanese: a Perceptual Illusion?, *Journal of Experimental Psychology: Human Perception and Performance*, 25, 6, p. 1568-1578.
- Dupoux, E., Pallier, C., Kakehi, K., Mehler, J. (2001) New evidence for prelexical phonological processing in word recognition, *Language and Cognitive Processes*, 5, 16, p. 491-505.
- Flege, J. E. (1995) Second language speech learning: theory, findings, and problems, in Strange, W (ed) *Speech Perception and Linguistic Experience: Issues in Cross-language Research*, York: York Press, p. 233-277.
- Gibson, M., (in progress) *Intersegmental timing as a determiner of complex onset formation in Spanish*.
- Goodman, M. B. & Jusczyk, P. W., (2000) The onset of sensitivity to internal syllable structure, in Pierrehumbert, J. & M. Broe (eds), *Papers in Laboratory Phonology V*, Cambridge: Cambridge University Press, 228-239.
- Gouskova, M. (2001) *Falling sonority onsets, loanwords and syllable contact*, [ROA-491, Rutgers Optimality Archive, <http://roa.rutgers.edu>].
- Hallé, P., Segui, J., Frauenfelder, U. Meunier, C. (1998) Processing of illegal consonant clusters: A case of perceptual assimilation?, *Journal of Experimental Psychology: Human Perception and Performance*, 24, 2, p. 592-608.
- Hallé, P., Dominguez, A., Cueto, F. & Segui, J. (2008) Phonological Mediation in Visual Masked Priming: Evidence from Phonotactic Repair, *Journal of Experimental Psychology: Human Perception and Performance*, 34, 1, p. 177-192.
- Harris, J. (1969) *Spanish Phonology*, Cambridge, MA: MIT Press.
- Harris, J. (1983) *Syllable structure and stress in Spanish: a nonlinear analysis*, Bloomington, IA: Indiana University Linguistics Club.
- Harris, J. (1989) Our present understanding of Spanish syllable structure, in Bjarkman, P & Hammond, R. (eds) *American Spanish pronunciation: theoretical and applied perspectives*, Washington D.C.: Georgetown University Press, 151-169.

- Hermes, A., Grice, M., Mücke, D. & Niemann, H. (2008) Articulatory indicators of syllable affiliation in word initial consonant clusters in Italian, *Papers from the 8th International Seminar on Speech Production*, Strasbourg, France, p. 433-436.
- Heuer, H. & R. A. Schmidt (1988) Transfer of learning among motor patterns with different relative timing, *Journal of Experimental Psychology: Human Perception and Performance*, 14, 2, p. 241-252.
- Hillenbrand, J. (1983) Perceptual organization of speech sounds by infants, *Journal of Speech and Hearing Research*, 26, p. 268-282.
- Honorof, D., N. & C. Browman (1995) The center or edge: How are consonant clusters organized with respect to the vowel?, in Elenius, K. & P. Branderud (eds) *Progress in Theoretical Biology*, New York: Academic Press, p. 303-338.
- Itô, J. (1986) *Syllable theory in prosodic phonology*, Doctoral dissertation, University of Massachusetts at Amherst.
- Itô, J. (1989) A Prosodic Theory of Epenthesis, *Natural Language and Linguistics Theory*, 7, p. 217-260.
- Iverson, P., Kuhl, P. K., Yamada, R. A., Diesch, E., Tohkura, Y., Kettermann, Siebert, C. (2003) A perceptual interference account of acquisition difficulties for non-native phonemes, *Cognition* 87, B47-B57.
- Jusczyk, P. W., Friederici, A. D., Wessels, J., Svenkerud, V. Y. & Jusczyk, A. M. (1993) Infants' sensitivity to the sound patterns of native language words, *Journal of Memory and Language*, 32, p. 402-420.
- Kabak, B. & Idsardi, W. (2003) Syllabically conditioned perceptual epenthesis, in Nowak, P. et al (eds) *Proceedings of the Berkeley Linguistics Society* 29, p. 233-245.
- Kiparsky, P. (1979) Metrical Structural Assignment is Cyclic, *Linguistic Inquiry* 10, 3, p. 421-441.
- Krakow, R. A. (1989) *The articulatory organization of syllables: A kinematic analysis of labial and velar gestures*, Doctoral dissertation, Yale University.
- Krakow, R. A. (1999) Physiological organization of syllables: a review, *Journal of Phonetics*, 27, p. 23-54.
- Kuhl, P. K. (2000) A new view of language acquisition, *Proceedings from the National Academy of Science U.S.A.* 97, p. 11850-11857.
- Leben, W. (1973) *Suprasegmental Phonology*, Doctoral dissertation, MIT. New York: Garland (1979).
- Linnell, P. (1982) The concept of phonological form and the activities of speech production and perception, *Journal of Phonetics*, 10, p. 37-72.
- Long, M. H. (1990) Maturational constraints on language development, *Studies in Second Language Acquisition*, 12, p. 251-285.
- Martínez-Gil, F. (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, p. 201-246.
- Martínez-Gil, F. (1997) Obstruent vocalization in Chilean Spanish: A serial versus a constraint-based approach, *Probus*, 9, p. 165-200.
- Mateus, M. & E. d'Andrade (2000) *The Phonology of Portuguese*, Oxford: Oxford University Press.
- Matthews, J. & C. Brown (2004) When intake exceeds input: Language specific perceptual illusions induced by L1 prosodic constraints, *International Journal of Bilingualism*, 8, p. 5-27.

- McClelland, J. L. (2001) Failures to learn and their remediation: A Hebbian Account, in McClelland, & Siegler, J.L. (eds) *Mechanisms of Cognitive Development: Behavioral and Neural Approaches*, Erlbaum, p. 97– 211.
- Musacchia G., Sams, M., Skoe E., Kraus N. (2007) Musicians have enhanced subcortical auditory and audiovisual processing of speech and music, *Proceeding of the National Academy of Science of the U.S.A.* 104, 40, p.15894–15898.
- Nazzi, T., Bertoncini, J., Mehler, J. (1998) Language Discrimination by Newborns: Towards an understanding of the role of rhythm, *Journal of Experimental Psychology: Human Perception and Performance*, 24, p. 1-11.
- Pallier, C., Bosch, L., Sebastián-Gallés, N. (1997) A limit on behavioral plasticity in speech perception, *Cognition*, 64, p. B9–B17.
- Pallier, C., Dahan, S., Poline, J. B., LeBihan, D., Argenti, A. M., Dupoux, E., Mehler, J. (2003) Brain imaging of language plasticity in adopted adults: Can a second language replace the first?, *Cerebral Cortex*, 13, p. 155–161.
- Parker, S. (2002) *Quantifying the Sonority Hierarchy*, Doctoral dissertation. University of Massachusetts.
- Prince, A. & P. Smolensky (1993) *Optimality Theory: Constraint Interaction in Generative Grammar*, Malden, MA: Wiley-Blackwell.
- Repp, B. H. (1992) Probing the cognitive representation of musical time: structural constraints on the perception of timing perturbations, *Cognition*, 44, 3, p. 241-281.
- Scovel, T. (2000) A critical review of the critical period research, *Annual Review of Applied Linguistics*, 20, p. 213–223.
- Sebastián-Gallés, N. & S. Soto-Faraco (1999) Online processing of native and non-native phonemic contrasts in early bilinguals, *Cognition*, 72, p. 111– 123.
- Selkirk, E. (1981) Epenthesis and degenerate syllables in Cairene Arabic, *MIT Working Papers in Linguistics*, 3, p. 209-232.
- Shaw, J. A., Gafos, Adamantios, I., Hoole, P. & Zeroual, C. (2009) Syllabification in Moroccan Arabic: evidence from patterns of temporal stability in *Articulation, Phonology*, 26, p. 187-215.
- Shepherd, M. (2003) *Constraint Interactions in Spanish Phonotactics: An Optimality Theory Analysis of Syllable Level Phenomena in the Spanish Language*, M.A. Thesis, California State University.
- Sproat, R. & O. Fujimura (1993) Allophonic variation in English /l/ and its implications for phonetic implementation, *Journal of Phonetics*, 21, p. 291-311.
- Werker, J. F. & R. C. Tees (1984) Cross-language speech perception: Evidence for perceptual reorganization during the first year of life, *Infant Behavior and Development*, 7, p. 49–63.
- Wheeler, M. (1979) *The Phonology of Catalan*, Oxford: Blackwell.
- Wright, R. (2004) A review of perceptual cues and cue robustness, in Hayes, B., Kirchner, R., Steriade, D. (eds) *Phonetically Based Phonology*, Cambridge: Cambridge University Press, p. 34-57.
- Ziegler, J. & L. Ferrand (1998) Orthography shapes the perception of speech: The consistency effect in auditory word recognition, *Psychonomic Bulletin and Review*, 5, 4, p. 683-689.

NOTES

1. Here 'computationalist' refers to approaches which understand cognition and the mind as information processing systems as in the computational theory of mind.
2. For a review of computationalist accounts of L2 vowel insertion, see Abrahamsson's (1999) study of L1 Spanish-speakers learning Swedish, Boudaoud and Cardoso's (2009) account of L1 Farsi-speakers learning English, and Altenberg's (2005) study of vowel insertion in the L2 English speech of L1 Spanish-speakers. For a good review of vowel insertion from various theoretical perspectives see Harris (1969, 1983), Wheeler (1979, 2005), Mateus and d'Andrade (2000) Kiparsky (1979), Cairns and Feinstein (1982), Selkirk (198), Itô (1986, 1989) and Shepherd (2003).
3. 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. This topic has been vigorously debated in the Autosegmental literature and is discussed specifically for Spanish in Harris (1989), Clements (1990), Colina (1995), Martínez-Gil (1996, 1997), Parker (2002) and Wright (2004) among others.
4. 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.
5. 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 [θr] appears with no surface-level modification, *thriller* [θrí.ler] as well as the phonetic adaptation of *Heathrow* realized as [hi.θrow], 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.
6. In most analyses, /dl/ is prevented by a ban preventing contiguous segments to share the certain [place] specifications (see *Obligatory Contour Principle*. Leben 1973).
7. Another series of on-line syllable-counting tests were administered in which listeners were asked to identify the quantity of syllables in stimuli in which the timing relations between acoustic units were modified. Due to the increased complexity of the methodology and results, the findings of this portion of the study will be published in a separate article.
8. The informant's primary school imparts 210 hours of English classes per year. At the time of the tests, the children were entering their twentieth week of their fifth year at the school, which yields, with perfect attendance, a total of 960 course hours in English.

ABSTRACTS

A hypothesis is proposed which claims that vowel prosthesis before word-initial [sC] clusters by L1 Spanish learners of English is a perception-based repair phenomenon conditioned by L1 biases for native patterns of intergestural timing in complex onsets. Data from a series of off-line perception tests involving fifty L1 Spanish-speaking children confirm the erroneous perception of an extra syllable in a majority of nonce tokens containing the target stimuli. Results of a set of follow-up discrimination and identification tasks prove that Spanish-speakers access information derived from native timing patterns in order to map continuous speech streams to categorical units before lexical storage occurs, supporting the viability of a bias derived from native timing configurations. The findings are of interest since they bear directly on contemporary research in the acquisition of L1 phasing relations and their effects on L2 phonological acquisition.

Nous présentons une hypothèse qui propose que l'épenthèse de la voyelle [e] devant la suite de consonnes [sC] en début de mot, observée chez les hispanophones lorsqu'ils parlent anglais, est un phénomène de perception qui provient d'une préférence pour certaines relations de coordination spatio-temporelle dans les attaques complexes. Les résultats d'une série d'épreuves de perception indiquent que les hispanophones perçoivent un [e] dans tous les *stimuli* qui contiennent la suite inadmissible de consonnes [sC] en début de mot. Les résultats d'autres épreuves de discrimination et d'identification corroborent l'idée que l'interprétation erronée des suites de consonnes est liée à la codification cognitive des relations spatio-temporelles systématiques qui proviennent de l'espagnol. Les résultats de cette étude sont intéressants parce qu'ils nous renseignent sur l'acquisition des relations spatio-temporelles dans la première langue, et sur la manière dont ces relations altèrent l'acquisition de la phonologie d'une deuxième langue.

INDEX

Keywords: L1 biases, relative timing patterns, perceptual epenthesis, Spanish

Mots-clés: épenthèse, relations de coordination, espagnol

AUTHOR

MARK GIBSON

Universidad de Navarra
mgibson@unav.es