This is a contribution from *Perspectives on Phonological Theory and Development. In honor of Daniel A. Dinnsen*.
Edited by Ashley W. Farris-Trimble and Jessica A. Barlow.

Covert contrast in the acquisition of second language phonology

Fred R. Eckman, Gregory K. Iverson and Jae Yung Song
University of Wisconsin–Milwaukee

This paper reports results on the acquisition of the English /s/ – /z/ phonemic contrast by native speakers of Spanish. The central finding is that some of the research participants exhibited a *covert* contrast between these segments in their interlanguage productions. Acoustic analysis revealed that four of the participants produced a statistically reliable distinction between English [s] and [z], however, this difference was not perceived by the transcribers who were phonetically trained, native speakers of English. The existence of a stage of covert contrast in L2 phonology is eminently plausible, given the progressive nature of phonological acquisition, and brings the learning of second-language contrasts into conformity with findings in the areas of L1 acquisition and phonologically disordered speech.

Introduction

Over the last few decades, research on the acquisition of target language (TL) sound patterns by second language (L2) learners has relied almost exclusively (studies such as Flege 1987 notwithstanding) on impressionistic, phonetic transcriptions to document the learners' progress in acquiring the TL phonology (Altenberg & Vago 1983; Broselow, Chen & Wang 1998; Carlisle 1998; Eckman 1981; Eckman & Iverson 1994; Flege 1987; Hammerly 1982; Major 1994; Ritchie, 1968; among many others). These transcriptions have constituted the data for determining the nature of the interlanguage (IL) phonological system being acquired by the learners. In many cases, the transcriptions have shown extensive neutralization of the TL contrasts being learned; in other words, the transcribers did not perceive a distinction in some of the sounds that the research participants were producing.

Until some thirty years ago, there was also this same reliance on phonetic transcriptions in research on child-language (L1) acquisition. However, since seminal work in this area by Macken and Barton (1980), numerous studies on both normal acquisition of their native phonology by children and on children with phonological disorders have shown that the participants often produce contrasts that are not perceived by the adult listeners/transcribers (Forrest, Weismer, Hodge, Dinnsen & Elbert 1990; Gierut & Dinnsen 1986; Macken & Barton 1980; Maxwell & Weismer 1982; Scobbie, Gibbon, Hardcastle & Fletcher 2000). This phenomenon, in which L1 learners produce a statistically reliable distinction between sounds that is not perceived by adults, whether phonetically trained or not, is known as *covert contrast* (as opposed to *overt contrast*, which is perceived by transcribers). To paraphrase Scobbie (1998), the idea behind covert contrasts is that the phonological system of a language may be acquired independently of how that system is implemented phonetically. We will have more to say on this matter below.

The purpose of the present paper is to report preliminary findings of an ongoing investigation into the acquisition of L2 phonemic contrasts. The case at hand centers on the acquisition of the English phonemes /s/ and /z/, as in *sip* versus *zip*, respectively, by native speakers of Spanish. Though [s] and [z] both occur in Spanish, these sounds are allophones of the phoneme /s/, with [z] occurring only before voiced consonants within the same word or phrase, as in *mi*[z]*mo* 'same' and *la*[z] *gatas* 'the (female) cats.' This distribution motivates a rule (or constraint) to the effect expressed in (1) below.

(1) Spanish allophonic pattern of *s*-voicing
/s/ → [voice]/___ [consonantal, voice]

The implementation of this process is variable, however, in that whether it takes place is an apparent function of the rate of speech (Harris 1969; 1983), faster speech favoring voicing; another aspect of the rule's variability is that it is optional (Hualde 2005). And inasmuch as voice onset time (VOT) is a continuum, the degree or extent of the assimilatory voicing itself has been observed to be gradient rather than categorical, ranging from partially voiced [s̬] through weakly voiced [z̥] to thoroughly voiced [z] (Bradley & Delforge 2006; Garcia 2013; Martínez-Gil 2003).[1] As we will suggest below, such variation and gradience appear also to influence the English productions of our Spanish-speaking research participants, some of whom produced a covert, rather than overt, contrast between English /s/ and /z/. We present an acoustic analysis of their utterances showing that four of the

1. In describing the extent of the fricative-voicing phenomenon in Spanish, Martínez-Gil (2003:57) remarks that "I do not know of any compelling evidence suggesting that partial voicing assimilation is a phonological property and not simply a fact of phonetic implementation. In fact, most available descriptions clearly indicate that the process is gradient, and thus typical of phonetic phenomena."

fourteen participants produced a reliable distinction between these fricatives that was not perceived by the native-speaker transcribers.

The paper proceeds as follows. In the next section, we place our study in context by reviewing some of the literature on covert contrasts in the acquisition of English by children who acquire their native language without problems (i.e., normal acquisition), and by children who are phonologically disordered. In accordance with some of our recent work on the acquisition of L2 phonemic contrasts (Eckman & Iverson 2013; Eckman, Iverson, Fox, Jacewicz & Lee 2011), we then present a hypothesis about the role played by s-voicing for Spanish speakers learning English. This is followed by a description of the methodology by which the data were gathered and the reporting of the results as they bear on the hypothesis. We conclude with a discussion of our findings in light of the hypothesis, and with a view toward some pedagogical implications for teaching TL phonemic contrasts.

Background

Covert contrast

Although earlier research had foreshadowed the idea that children acquiring their native phonology were making statistically significant distinctions that were not being perceived by adults (Kornfeld & Goehl 1974; Ohala 1974; Smith 1979), the article by Macken and Barton (1980) is generally cited as being the seminal study to report the stage of a covert contrast in the acquisition of phonology by children. Theirs was a longitudinal study of the acquisition of the English voice contrast by four monolingual children. The authors analyzed the productions of children between the ages of one year, four months (1;4) and two years, four months (2;4). Separate frequency distributions were calculated for the phonemically voiced and voiceless stops at each point of articulation, followed by tests of significance between the mean voice onset time (VOT) values for the voiced and voiceless consonants. In stops, VOT delay or lag is defined as the time, measured in milliseconds (ms), between the release of closure and the beginning of vocal cord vibration in the following vowel (or sonorant consonant). VOT delay is a primary acoustic cue in determining whether stop consonants are perceived by speakers as voiced or voiceless (cf. Iverson & Salmons 1995): in phrase-initial position in English, stops categorized as 'voiced' have a mean VOT value of less than 20 ms. (thus are phonetically voiceless, and largely unaspirated), whereas stops categorized as 'voiceless' show an average VOT lag of 40 ms. or more (thus are prominently aspirated).

Based on analysis of the VOT values in their participants' productions of initial stops, Macken and Barton identified three stages of acquisition. The first was one in which the children did not produce a voice contrast in any stop consonants, as the VOT values for both voiced and voiceless consonants fell within the short lag range of adult speech. In the second stage, the children produced a statistically significant VOT contrast between the voiced and voiceless stops, but these values all fell within the adult perceptual categories of English voiced stop phonemes. In other words, the VOT distinction that the children were making was not sufficiently great to be perceived by adults, but was nevertheless statistically reliable. The presence of such a covert contrast stage suggested that the children were aware of the voicing contrast, although their implementation of VOT was not yet adult-like. The final stage was one in which the children's production of a VOT contrast resembled that produced by adults. Additionally, Macken and Barton found that the children acquired the VOT contrast at some points of articulation before others.

In the ensuing years, there have been a number of studies on the acquisition of covert contrasts among phonologically disordered children. One of the earliest was by Gierut and Dinnsen (1986), who analyzed two children, aged 4;6 and 4;3. Phonetic transcription of their utterances indicated that both children were producing the same kinds of errors. Specifically, according to the transcriptions, both children failed to make a voice contrast in word-initial stops. Based on an acoustic analysis of VOT and closure duration, the authors found that one of the children was producing statistically significant differences between initial voiced and voiceless stops in both VOT and closure duration measurements, although these distinctions could not be perceived by adult listeners.

In the following two and a half decades, it seems to have become widely accepted in research on both the acquisition of L1 phonology and on phonological disorders that there is a need to move from listener-oriented to speaker-oriented data (Hewlett 1988). Numerous studies investigating stages of covert contrast have been carried out in both research domains, on a myriad of phonological contrasts involving a large number of acoustic cues, including, but not limited to, amplitude, differential vowel duration, formant analysis, pitch and VOT. The reader is referred to Scobbie (1998) for a thorough listing and review of a large number of such studies.

In sum, the ample research on the acquisition of covert contrasts in phonological acquisition is sufficient to conclude that making such a contrast is a well-documented, intermediate stage in acquiring phonemic distinctions. We suppose, therefore, that covert contrasts should be attested in the acquisition of contrasts in second-language phonology, too.

Allophonic splits

As we have detailed in other work, there is evidence that learning to contrast TL sounds which are allophones of the same phoneme in the native language (NL) involves suppressing the NL allophonic rule and results in implicationally related patterns of acquisition. We have shown this claim to be supported for the acquisition of the English /s/ – /š/ contrast by native speakers of Korean (Eckman & Iverson 2013). In the current study, we present evidence that the same pattern of acquisition occurs in the case of native Spanish speakers acquiring English /s/ and /z/, and, further, that these patterns are also attested by learners who make the contrast covertly.

In the case of native speakers of Spanish learning to distinguish [s] and [z] phonemically, we assume that, in the early stages of English acquisition, the rule (or implementation strategy) relating [s] and [z] as allophones of /s/ transfers into the IL grammar and applies to the pronunciation of TL words, causing errors. This IL rule is subject to the general phonological constraints that pertain to primary-language grammars, including the *derived environment effect* (Kiparsky 1982), according to which rules effecting the substitution of one phoneme for another apply only in so-called derived environments.[2] In the context of SLA, this means that an NL rule such as (1) above, when transferred into the IL, begins by applying across-the-board. The IL then moves through a stage in which the rule's application persists only in the inter-morphemic environment (i.e., applies only when the following voiced consonant is separated from /s/ by an intervening morpheme boundary, as in *seriousness*), but is suppressed morpheme internally. Ultimately, the application of the rule is suppressed by the L2 learner, and does not apply at all in the IL. Conversely, rules relating allophones of the same phoneme may apply in both basic and derived environments, without regard for morphological structure.[3] Thus, derived environments consist in portions of words that contain a representation to which a rule would apply inter-morphemically, i.e., the segments in question are separated by a morpheme boundary, whereas basic environments are found in mono-morphemic words which contain the appropriate segments for application of the rule.

2. *Derived environment* refers to a context for the application of a phonological rule where the crucial representation needed for the rule to be applicable includes a morpheme boundary. An example of a derived environment in the case of (1) above would be a situation where a morpheme boundary exists between /s/, the segment to which the rule applies, and the following voiced consonant.

3. Cho (1999; 2001) has shown that, in Korean, the effect of allophonic palatalization on /n/ before /i/ is gradient, with a greater palatalizing effect when the following /i/ begins a new morpheme than when it is in the same morpheme. It may be that the variable voicing of Spanish /s/ is implemented similarly, with generally more penetration of voicing into /s/ before a hetero-morphemic voiced consonant than before one in the same morpheme. Though we do not have data on that in Spanish, this parallels the pattern some of the L2 learners show in English, which then might suggest transfer from Spanish.

The derived environment effect holds implications for IL phonologies in which an L2 learner must acquire a TL contrast between two segments, e.g., English /s/ and /z/, that are allophones of the same phoneme in the NL. Transfer of the rule relating these NL allophones, before the TL contrast between them has been established in the IL, leads L2 learners to err across-the-board, applying the NL allophonic rule in all environments where it can be applied, irrespective of morphological structure. For Spanish-speaking learners of English, specifically, fricative voicing would take place (incorrectly) both in mono-morphemic *Christmas* and poly-morphemic *seriousness*. But as learners acquire the TL contrast in some words (*Christmas* now with [s] rather than [z], *business* still with [z]), thereby introducing these sounds into the IL lexicon as phonemes, application of the NL allophonic rule becomes restricted to derived environments (*seriousness* still with [z]).[4] Ultimately, the learner may be able to suppress the application of the NL allophonic rule altogether and thus acquire the contrast in all environments (*seriousness* with [s] vs. *noiseless* with [z]). In view of the derived environment effect, however, a fourth, logically possible pattern is excluded, viz., that in which the learner suppresses the application of the NL allophonic rule only in derived environments (yielding *seriousness* with [s], but *Christmas* with [z]).

Given this background, the specific hypothesis we test here is the following:

(2) Hypothesis
Acquisition of the English /s/ – /z/ contrast by Spanish-speaking learners will be sensitive to morphological structure in a manner consistent with the derived environment effect.

Two observations about this hypothesis bear mention. The first is that we will determine whether our participants have acquired the /s/ – /z/ contrast either overtly or covertly. A participant will be credited with an overt contrast in a given phonological environment (e.g., initially, inter-morphemically etc.) if the participant produces, on the basis of the phonetic transcriptions, at least 80% target-like productions for both [s] and [z] in the specified environment.[5] If, according to the transcriptions, the participant fails to reach the 80% threshold on either or both [s] and [z] in an environment, then we conclude that the participant lacks the relevant overt contrast in that environment.

4. From a historical perspective, words such as *Christmas* and *business* consisted of two morphemes, though there is no evidence that native speakers of English today treat them as consisting of two morphemes. Moreover, the protocol we followed required the participants to add the suffixes in question, *-ness* and *-less,* as they pronounced the word, thus forcing the participants to make an utterance morphologically composite and therefore, derived.

5. The 80% threshold for acquisition of a structure in L2 acquisition was first used, to the best of our knowledge, in Cancino et al. 1978, and has been employed extensively since that time.

The second important aspect of (2) is that evidence of a derived environment effect is the acquisition of the relevant contrast, either overtly or covertly, in only a basic environment, or in both a basic and derived environment, but not in only a derived environment. Thus, all of our participants should evince one of the acquisition patterns outlined above, and none should exhibit the excluded pattern (contrast only in a derived environment).

Methodology

Stimuli

In order to test the hypothesis in (2), we elicited productions on the /s/ – /z/ contrast from fourteen L2 learners of English, all of whom were native speakers of Spanish associated with the University of Wisconsin–Milwaukee. In an attempt to enlist participants who had a relatively wide range of English proficiency, we recruited volunteers for the research both from the English as a Second Language Program and from native speakers of Spanish in the campus community. Thus, the participants varied widely in age (18 to 49), in how long they had studied English, in their length of residence in the United States, in their nation of origin and in their overall command of English. All participants were paid a small fee for their participation in the project.

Within this context it is important to point out that our hypothesis is independent of the level of English proficiency for our research participants. In other words, the hypothesis simply asserts that, with respect to the /s/ – /z/ contrast, the IL of any Spanish-speaking learner of English will fall into one of the three predicted stages of acquisition. Therefore, the hypothesis is testable regardless of the English background of the participants.

The stimuli used to elicit the productions consisted of a set of 90 words, 60 of which were targets (listed in the Appendix) and 30 of which were fillers. All are existing lexical items in English, and each target word contained /s/ or /z/ in one of three different positions in a morphologically basic word, and additionally in a morphologically derived environment. In the morphologically basic words, the positions of occurrence for the /s/ or /z/ are initially before a vowel (e.g., *sip/zip*), medially following a vowel and before a voiced or voiceless consonant (e.g., *Christmas/business*), and word-finally following a vowel (e.g., *pass/buzz*). The morphologically derived environment is the position following a vowel and preceding a voiced consonant at the juncture of another morpheme, either the suffix *-ness* or *-less* (*faceless/noiseless/seriousness*).

Several custom programs were written in MATLAB for the purposes of the present study. A program that controlled the recordings displayed on a computer screen a set of pictures, clues, and commands, such as "Wait" or "Speak", that were designed to guide the participant and the experimenter through the elicitation of each word. Words were elicited, not by giving their spelling, but by displaying an image depicting the object or idea in question for both basic (picture, say, of a face to elicit the word *face*) and derived forms (same picture, but with the cue, "without," appearing on the screen one-half second after appearance of the image to elicit the word *faceless*). If participants did not immediately recognize the word or concept being depicted, they were given on-screen clues, or definitions, and, if need be, a recorded model of the word's pronunciation. The participants were also given a practice exercise with a different set of words using the same cues in order to ensure that they could correctly produce the intended derived words with the appropriate suffix, in this case, *-less*. The stimuli were presented in a pseudo-randomized order in that all basic forms were elicited before their related derived forms. The elicitations were recorded directly onto a hard disc drive at the sampling rate of 44.1 kHz. Participants spoke into a head-mounted microphone at a distance of one inch from the lips and produced the set of 90 words twice, both during the same session.

Transcriptions

The data were collected at the University of Wisconsin–Milwaukee and then transferred to another major, mid-western university via file transfer protocol where they were transcribed by an assistant who was blind to the hypothesis and unaware of the intended target segments. The transcriber listened to the utterances in question and focused either on a consonant in word-initial position or a word-medial consonant occurring before the suffixes *-less* or *-ness*. The transcriber's task was then to choose from a menu of several choices: (1) [z], (2) [d] voiced alveolar stop, (3) [s], (4) [t] voiceless alveolar stop, or (5) other. If option "(5)" was selected, the transcriber was also required to enter the segment or to make a comment on what was heard. The completed transcriptions were then returned to UW-Milwaukee where they were scored.

For the purposes of testing the hypothesis, we consider the basic environment to be exemplified by two phonological contexts: in words containing [s] and [z] in word-initial position before a vowel, and in words with those sounds in word-medial position following a vowel and before a consonant. Derived environments are found in words in which either [s] or [z] occurs before the suffix *-less* or *-ness*, as in *seriousness, noiseless, baseless,* etc. [6]

6. In frequently occurring morphemes, fossilization on the *status quo ante* pronunciation may result in what appears to be the application of the rule in a basic environment (e.g., *Chris̲tmas* with [z] rather than [s]), which at this stage we take to be lexicalized pronunciations. Support for this interpretation would come from additional data involving novel words with morpheme-internal pre-consonantal [s] (e.g., *Islip*, *parsnip*) or nonce words of the same type.

We elicited a total of ten words with word-initial [s] and ten words with word-initial [z], as well as an additional ten words with word-medial [s] and ten with word-medial [z]. For the occurrence of [s] and [z] in the derived environment, participants produced a total of twenty words in which target [s] occurred before the suffix *-less* or *-ness*, and the same number of words in which target [z] occurred before the same suffixes.[7] A participant thus produced a total of 60 target words. In order for a participant's IL to be credited with having a contrast between [s] and [z] in a given environment, the performance on the productions had to reach the 80% criterial threshold for both [s] and [z] in that environment, as already noted above. If a participant's target-like pronunciations reached the criterial threshold on only one of the segments in a given environment, or did not reach criterion on either segment, the participant's IL grammar was scored as lacking the contrast in that environment. For example, a participant had to produce [s] in at least eight of the ten words in which [s] occurred in initial position before a vowel, and likewise for [z], in order for the IL to be accorded having the /s/ – /z/ contrast in the basic environment.

Acoustic analysis

As we are also investigating whether any of our participants made a covert contrast between [s] and [z], we performed an acoustic analysis on the relevant sound files. We begin by considering the acoustic measures for distinguishing between [s] and [z] in general, and then proceed to the particular parameters that we employed.

Previous literature has demonstrated the robust effect of voicing on the duration of fricative noise, with a longer period of frication occurring for voiceless than for voiced fricatives. For example, in a corpus study, Crystal and House (1988) found that the duration of voiceless fricatives was overall 47 ms longer than that of voiced fricatives (97 ms for voiceless fricatives versus 50 ms for voiced). Similarly, Stevens, Blumstein, Glicksman, Burton and Kurowski (1992) reported that the duration of the voiceless fricative [s] was about 30 ms longer than that of the voiced fricative [z] in intervocalic position (108 ms for the former, 78 ms for the latter). The voicing of fricatives is also known to affect the duration of the preceding vowel. In English, vowels are typically about 100 ms longer before voiced than before voiceless obstruents (House 1961; but also see Crystal & House 1988 who report this difference only in utterance-final position).

7. We also elicited tokens of [s] and [z] in word-final position; however, we are not reporting these data because none of the subjects even approached the criterial 80% threshold on these productions.

Another important correlate of the voiced-voiceless contrast for fricatives is the duration of voicing during the fricative noise. In other words, the amount of overlap between the fricative noise and the period of voicing has been shown to be systematically greater for voiced fricatives than for voiceless fricatives. This measure has been successfully used to examine, for example, the partial devoicing of [z], which seems to be virtually universal in utterance-final or pre-pausal position (e.g., Smith 1997). However, it should be noted that fully voiced fricatives with 100% overlap (i.e. fricatives in which the time of vocal cord vibration overlaps completely with the period of frication) are relatively difficult to produce, and are uncommon cross-linguistically. As the airflow from the lungs is interrupted by the constant closing and opening of the glottis for voiced fricatives, it is physiologically difficult to maintain, at the same time, both voicing and the high velocity of airflow necessary for the turbulent noise characteristic of a fricative.

The acoustic measure that we used for distinguishing the production of [s] and [z] by our participants was the percent of voicing that overlapped with the fricative noise for each token of these segments. For the target fricative consonant in each word, we calculated the percent of the fricative noise duration during which the vocal folds were vibrating. On this measure, zero percent indicates that the frication noise did not overlap at all with the observable vocal fold vibration (as indicated by periodicity in the waveform and vertical striations in the spectrogram). Alternatively, 100 percent indicates that the overlap between frication and voicing was complete.

It is worth making several points about the acoustic measure that we employed. First, since some of the [s] and [z] segments that we analyzed occurred in word-initial position, it was not possible to use the difference in the duration of the vowel before a voiceless versus voiced fricative as a distinguishing measure. Second, the measure that we employed is appropriate for our data in that the target words were not controlled for the number of segments they contained; rather, the words were chosen on the basis of how easily and recognizably they could be pictured on a computer screen. Because the duration of segments varies depending on the number of segments in the word (cf., e.g., Lehiste 1972), we used proportional measurements rather than raw numbers. Thus, for any target segment, we measured the percent of the fricative noise that overlapped with voicing.

Figure 1 below shows the examples of three different renditions of the same word *zee* by three of our Spanish-speaking participants. Two graphs are represented in each of Figures (1a – c), with the top half showing the waveform and the bottom half showing the spectrogram. The vertical dotted lines in the graphs mark the beginning of the fricative noise (A), the onset of voicing (B), and the end of the fricative noise (C). In Figure (1a), the fricative noise of the [z] fully overlaps with

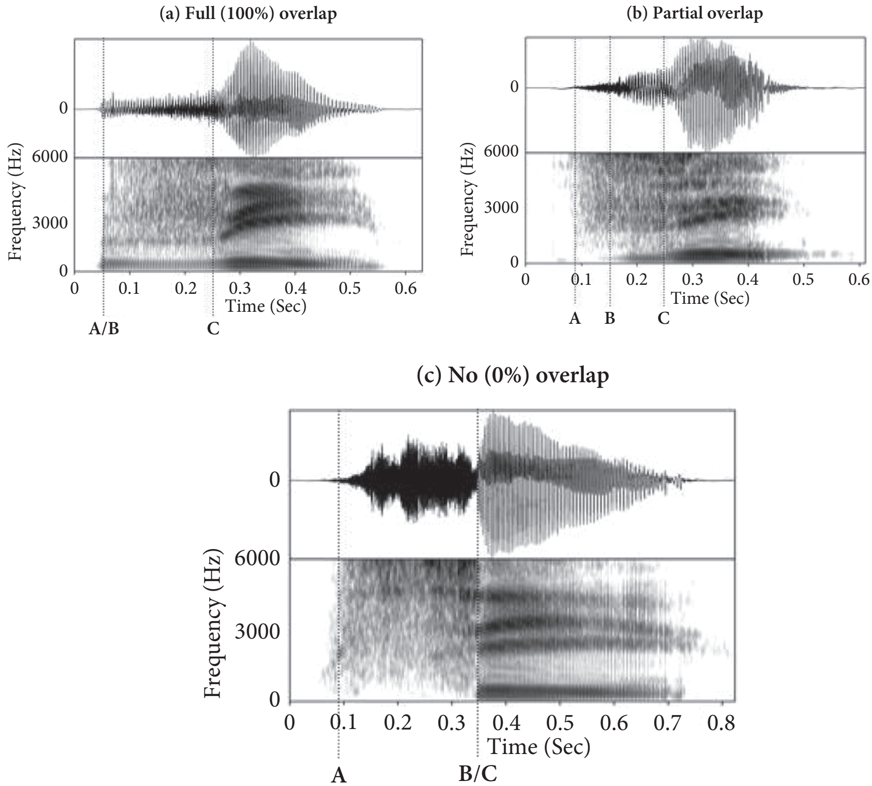

Figure 1. Representative waveform and spectrogram for the word *zee* [zi] produced by Spanish speakers. "A" indicates the beginning of fricative noise; "B" shows the beginning of voicing; and "C" marks the end of fricative noise. The interval between "B" and "C" is the fricative noise overlapping with voicing

voicing, as indicated by the simultaneous onset of fricative noise and the voicing, with voicing continuing through to the onset of the vowel at approximately 0.25 ms. The second example, Figure (1b), shows partial overlap of the fricative noise and associated voicing, as frication begins at about 0.1 ms and lasts until about 0.25 ms, and voicing begins at 0.15 ms. Figure (1c) exemplifies no overlap between the fricative noise and voicing, as voicing does not begin until the onset of the vowel. The acoustic coding was carried out by two trained assistants using Praat (Boersma & Weenink 2005). Visual information from the spectrogram and waveform, as well as auditory information, were used to determine the beginning and end of fricative noise and voicing.

Results

We report findings first with respect to whether the participants exhibited an overt contrast between [s] and [z], then whether the participants evidenced a covert contrast between these segments.

Table 1 shows the participants' productions, according to the transcribers, expressed as a percentage of target-like performance, in words containing [s] and [z] in the three relevant phonological environments: initially, medially and inter-morphemically.

Table 2 translates the results from Table 1 into categorical representations as to whether a given participant's IL showed a voice contrast between [s] and [z] in his/her English productions, again in the three specified phonological environments.

Table 1. Subjects' performance, expressed as a percentage of target-like productions, on [s] and [z] in word-initial position before a vowel, word-medial position following a vowel and preceding a voiced consonant, and in inter-morphemic position before the suffix *-ness* or *-less*

Subjects	Word-initial		Word-medial		Inter-morphemic	
	[s]	[z]	[s]	[z]	[s]	[z]
No contrast						
3017	100	0	10	100	20	85
3018	100	0	80	40	50	20
3020	100	50	30	90	25	65
3021	100	0	0	100	5	85
3022	100	40	30	100	5	90
3025	100	0	80	30	80	15
3029	100	60	30	60	60	85
Contrast in basic environment						
3019	100	100	0	100	30	95
3026	100	90	40	80	40	80
3027	100	80	20	50	40	95
3028	100	90	80	90	70	40
3030	100	90	30	70	45	100
3031	80	90	20	60	20	100
Contrast medially & in derived environment						
3024	100	30	90	80	85	85

Table 2. Subjects' performance, according to whether or not they reached the criterial threshold of 80% (Yes), or whether they fell below the threshold (No), on [s] and [z] in initial position before a vowel, medial position following a vowel and preceding a voiced consonant, and in inter-morphemic position before the suffix *-ness* or *-less*

Subjects	Contrast initial	Contrast medially	Contrast inter-morphemically
No contrast			
3017	No	No	No
3018	No	No	No
3020	No	No	No
3021	No	No	No
3022	No	No	No
3025	No	No	No
3029	No	No	No
Contrast in basic environment			
3019	Yes	No	No
3026	Yes	No	No
3027	Yes	No	No
3028	Yes	Yes	No
3030	Yes	No	No
3031	Yes	No	No
Contrast medially & in derived environment			
3024	No	Yes	Yes

The findings presented in Table 1 support the hypothesis, which claims that the presence of the contrast in derived environments implies the contrast in basic environments, but not vice versa. All of the participants fall into one of the permitted patterns and none of the participants evinces the excluded pattern. The first set of participants in the table represents the "no contrast" group in that none of them exhibits the /s/ – /z/ contrast in any of the three environments. This is because there are no scores at or above the 80% threshold on both [s] and [z] in any of these environments. The second group of participants in the table evidences the contrast in the basic environment, specifically, in word-initial position, or in word-initial and word-medial positions, but not in the derived (inter-morphemic) environment. The final participant, 3024, is classified as having the contrast in both basic and derived environments, with 3024 evincing the contrast in both a basic (word-medial) and a derived (inter-morphemic) environment. Thus, the first three groupings of the participants in Table 1 are consistent with the hypothesis in (2): seven participants do not have the contrast in any environment; six participants show the contrast in the basic environment only; and participant 3024 has the contrast in both the basic and derived contexts. Moreover, none of the participants exhibits the excluded stage of acquisition, i.e., showing the contrast only in the derived, intermorphemic environment while lacking the contrast in the basic environment.

The productions of participant 3024, while consistent with the hypothesis in maintaining the contrast in both the derived and basic environments, are a bit of an anomaly, because this participant maintains the contrast in word-medial position, but lacks the contrast word-initially. The general expectation among phonologists is that a word-medial contrast would also occur word-initially. We will have more to say about this participant's performance below.

We now focus on results showing that some participants make a covert contrast between English [s] and [z]. Findings are reported first for the group of participants, as is customary, then for individuals, as is necessary in the context of our hypothesis. Our hypothesis makes claims about the status of a learner's IL grammar, stating that acquisition of the /s/ – /z/ contrast will exhibit a derived environment effect. IL grammars are mental systems whose placement in time and space is in the mind of individual learners. Therefore, we must test such claims about the state of an interlanguage grammar using individualized data, simply because there is no IL grammar of a group of people, at least not one that can be situated in time and space, just as there is no mind of a group.

Group results

In order to examine whether our native speakers of Spanish were making an acoustic distinction between [s] and [z], we compared the percent of the fricative noise overlapping with voicing for [s] and [z] using paired t-tests. Before running these t-tests, we examined the distribution of the data. Because the distribution turned out to be skewed to the right, the data were log-transformed in order to better approximate a normal distribution. When transforming the data to logarithmic values, following convention, we first added a constant, 1, to all raw percentages (some of which were 0, indicating no overlap) so that all percentages were greater than 0. This procedure was to avoid dealing with the log of 0, which cannot be defined.

The group results for all of the participants for the paired t-tests using the log-transformed data are shown in Table 3.

The findings suggest that the percent of fricative noise that overlapped with voicing for [s] significantly differs from that for [z] word-initially and intermorphemically, suggesting that [z] is significantly more voiced than [s] in these positions. However, the difference was not significant in the basic-environment of word-medial position following a vowel and preceding a voiced consonant, i.e., where the fricatives were produced.

Table 3. Group results. (Note: The mean and SD are based on raw percentages. The statistical results are based on log-transformed data.)

	[s]		[z]		
Position	**Mean**	**SD**	**Mean**	**SD**	**Statistics**
Word-initial	2.27	5.66	43.90	38.76	t(13) = −5.87, $p < 0.00$
Word-medial	15.67	15.61	19.03	15.25	t(13) = −1.62, $p = 0.13$
Inter-morphemic	16.51	18.33	22.51	23.85	t(13) = −2.27, $p < 0.05$

Individual results

Next we examined which of the individual speakers showed a difference in the degree of fricative voicing between [s] and [z]. To this end, we ran unpaired t-tests for each speaker, using values from individual target words for each participant. As with the group data, the raw percentages were log-transformed before running the unpaired t-tests. Table 4 compares the results from the acoustic analysis and phonetic transcriptions in word-initial position.

Table 4. Subjects' performance on [s] and [z] in initial position before a vowel. (Note: The mean and SD are based on raw percentages. The statistical results are based on log-transformed data.)

	Acoustical analysis results					*Transcription results*	*Contrast*
	[s]		[z]				
	Mean	**SD**	**Mean**	**SD**	**Statistics**		
3017	.28	.42	.15	.23	$t(8) = .48, p = .64$	No	No
3018	.43	.58	5.88	12.94	$t(8) = -.66, p = .53$	No	No
3020	.00	.00	16.13	22.22	$t(8) = -2.28, p = .05$	No	Covert
3021	.18	.34	.66	.66	$t(8) = -1.42, p = .19$	No	No
3022	2.17	1.12	17.47	19.47	$t(8) = -1.87, p = .10$	No	No
3025	.33	.53	5.57	11.88	$t(8) = -.87, p = .41$	No	No
3029	1.37	.58	47.96	27.23	$t(8) = -5.65, p < .001$	No	Covert
3019	.00	.00	90.00	22.36	$t(8) = -32.77, p < .001$	Yes	Overt
3026	.00	.00	90.00	22.36	$t(8) = -32.77, p < .001$	Yes	Overt
3027	.56	.29	64.01	39.62	$t(8) = -5.94, p < .001$	Yes	Overt
3028	1.10	1.86	75.29	36.22	$t(8) = -8.20, p < .001$	Yes	Overt
3030	.00	.00	100.00	.00	$t(8) = -2296.81, p < .001$	Yes	Overt
3031	21.59	43.76	85.84	21.75	$t(8) = -3.58, p < 0.01$	Yes	Overt
3024	3.76	3.41	15.62	31.27	$t(8) = -.44, p = .67$	No	No

The results show that all 6 participants who reached the criterial threshold of 80% (as indicated by "Yes" under Transcription results in Table 2) also showed a significant difference in the percent of the fricative noise overlapping with voicing between [s] and [z]. That is, these participants exhibited differences between [s] and [z] in word-initial position both in the acoustic analysis and in the phonetic transcriptions, thereby making the contrast overtly. In addition, there were two participants who showed significant differences in the percent of the fricative noise overlapping with voicing, but fell below the threshold of 80%: 3020, 3029. That is, these participants produced a statistically reliable word-initial distinction between [s] and [z] in terms of fricative-voicing overlap that was not perceived by transcribers using the 80% criterion, indicating that they made the contrast covertly. All other Spanish speakers showed no difference between word-initial [s] and [z] on the basis of both acoustic analysis and phonetic transcriptions, suggesting that they exhibited no contrast between the two sounds.

We turn now to the participants' performance on [s] and [z] in the basic context of medial position following a vowel and preceding a voiced consonant, shown in Table 5. Twelve participants who did not exhibit the /s/ – /z/ contrast in this environment on the basis of the phonetic transcriptions also did not show a difference in the percent of the fricative noise overlapping with voicing between [s] and [z]. However, there were two participants (3024, 3028) who, according to the phonetic transcriptions, scored at or above the 80% threshold on both [s] and [z] but did not show differences acoustically. Though we did not perform additional acoustic analyses on these participants, we infer that 3024 and 3028 must have implemented the contrast between [s] and [z] in some other way that was perceived by the transcribers.

Table 6 compares the results from the acoustic analysis and phonetic transcriptions in inter-morphemic position. Although none of the participants (except for 3024) reached the 80% threshold on the basis of the phonetic transcriptions, three of the participants (3019, 3029, 3030) made the contrast covertly in this environment by showing a significant difference in the percent of fricative noise overlapping with voicing between [s] and [z].

To summarize this section, the results of our acoustic analysis revealed that four participants (3020, 3029, 3019, and 3030) maintained in their IL a voice contrast between [s] and [z] that was not perceived by the research assistants, and therefore was not transcribed. In other words, these four participants evidenced a covert contrast in that the percentages of voicing that overlapped with the fricative noise for [s] and [z] were statistically different in tokens containing [s] and [z] word-initially and inter-morphemically following a vowel preceding either the

Table 5. Subjects' performance on [s] and [z] in medial position following a vowel, and before a voiced consonant. (Note: The mean and SD are based on raw percentages. The statistical results are based on log-transformed data.)

	Acoustical analysis results					*Transcription results*	*Contrast*
	[s]		[z]				
	Mean	**SD**	**Mean**	**SD**	**Statistics**		
3017	58.4	34.52	69.69	33.07	$t(8) = .86, p = .41$	No	No
3018	9.43	7.18	10.58	6.99	$t(8) = .94, p = .37$	No	No
3020	6.76	8.13	7.90	5.95	$t(8) = .03, p = .98$	No	No
3021	9.49	4.69	19.34	21.50	$t(8) = .26, p = .80$	No	No
3022	11.29	9.77	11.82	13.10	$t(8) = -.23, p = .83$	No	No
3025	2.24	2.10	3.39	3.09	$t(8) = -.24, p = .82$	No	No
3029	7.20	3.63	4.53	2.02	$t(8) = .72, p = .49$	No	No
3019	25.84	30.84	62.46	33.05	$t(8) = -.18, p = .87$	No	No
3026	6.54	5.71	8.33	6.00	$t(8) = -1.47, p = .18$	No	No
3027	21.26	4.01	19.67	5.04	$t(8) = .51, p = .63$	No	No
3028	7.54	1.37	6.83	2.27	$t(8) = -.71, p = .50$	Yes	No
3030	5.26	2.59	21.43	24.51	$t(8) = -.40, p = .70$	No	No
3031	55.66	23.38	63.36	13.75	$t(8) = -.40, p = .70$	No	No
3024	4.26	2.53	5.75	5.06	$t(8) = -.38, p = .72$	Yes	No

suffix *-ness* or *-less*. To our knowledge, evidence of such a covert contrast in the acquisition of second-language pronunciation has not been reported[8].

Discussion

Though the findings of our study are subject to certain limitations (segmentally disparate pairings, acoustic measures used are limited), they nevertheless point to several interesting and important implications for both second-language acquisition theory and for L2 pedagogy.

As stated at the outset, the segments [s] and [z] occur in both English and Spanish: in Spanish as allophones of the same phoneme (/s/), but as contrasting

8. The only exception to this claim that we have encountered is Lim and Oh (2007), a published abstract from a conference, which clearly encompasses the idea of covert contrast in SLA, but has not been widely circulated.

Table 6. Subjects' performance on [s] and [z] in inter-morphemic position before the suffix *-ness* or *-less*. (Note: The mean and SD are based on raw percentages. The statistical results are based on log-transformed data.)

	Acoustical analysis results					*Transcription results*	*Contrast*
	[s]		[z]				
	Mean	SD	Mean	SD	Statistics		
3017	59.24	34.16	39.55	42.02	$t(18) = -.75, p = .46$	No	No
3018	5.83	3.06	4.16	2.68	$t(18) = -.63, p = .53$	No	No
3020	11.88	9.92	10.94	7.99	$t(18) = -1.28, p = .22$	No	No
3021	14.22	14.96	12.06	11.17	$t(18) = -1.10, p = .29$	No	No
3022	11.20	6.26	13.45	9.12	$t(18) = .47, p = .65$	No	No
3025	1.44	1.36	3.60	5.97	$t(18) = -.53, p = .60$	No	No
3029	9.29	3.69	7.90	4.04	$t(18) = 2.08, p = 0.05$	No	Covert
3019	26.03	23.15	40.02	42.48	$t(18) = -2.7, p < .05$	No	Covert
3026	5.98	4.23	37.31	32.67	$t(18) = -1.07, p = .30$	No	No
3027	17.91	6.30	15.77	4.30	$t(18) = .86, p = .40$	No	No
3028	6.87	1.06	10.5	9.48	$t(18) = 1.02, p = .32$	No	No
3030	7.86	3.08	16.73	24.78	$t(18) = -3.19, p < 0.01$	No	Covert
3031	36.53	35.62	48.14	39.04	$t(18) = -1.15, p = .27$	No	No
3024	5.13	6.26	6.29	4.69	$t(18) = .05, p = .96$	Yes	No

phonemes (/s/ versus /z/) in English. Native speakers of Spanish acquiring English must therefore implement two major changes in their IL grammar relative to their NL: first, they must develop the voiced sibilant fricative [z] as a phoneme in contrast to /s/, and second, they must suppress imposition of the NL allophonic pattern in the IL. Transferring the NL distribution of [s] and [z] into a learner's IL would cause that learner to err on TL words containing [z] in all environments except before voiced consonants. Specifically, such learners would err on all TL words containing [z], except for those where that segment occurs before a voiced consonant, as in *business* or *prizeless*. Thus, learning to contrast /s/ – /z/ in the IL of native speakers of Spanish would ostensibly entail their making an allophonic split, that is, the learners must separate the NL allophones [s] and [z] into two phonemes in the IL (Eckman & Iverson 2013). The case at hand, however, is complicated by the fact that the Spanish allophonic pattern produced by (1), which voices /s/ to [z] before voiced consonants, is optional, or gradiently variable, and subject to rate of speech (Bradley & Delforge 2006; Harris 1983; Hualde 2005).

As a consequence, native speakers of Spanish learning the English /s/ – /z/ contrast would seem to have a greater possibility of splitting their NL allophones [s] – [z] into separate phonemes by suppressing the pattern resulting from (1) than would L2 learners who have to suppress an allophonic rule that is obligatory (Eckman & Iverson 2013). We would therefore expect that those research participants who have acquired a contrast between /s/ and /z/, in at least one position, not to impose the NL pattern arising from (1). In other words, we would expect that those learners who know that the TL contrasts /s/ and /z/ in, say, word-initial position, would not transfer the NL pattern in (1) to the IL, because imposing this pattern would neutralize the /s/ – /z/ contrast. As we can see from Table 2, however, this is not the case, as participants 3019, 3026, 3027, 3028, 3030, and 3031 produce the contrast word-initially, yet – except for 3019 and 3030, who show the contrast covertly in inter-morphemic position – they continue to follow the NL pattern by voicing /s/ to [z] before a hetero-morphemic voiced consonant. The behavior of these six participants would fall into place if it turned out that NL allophonic patterns are not optional when implemented in an IL, or if their elicitations were produced at a speech rate that caused the rule or constraint in (1) to be invoked consistently. For now, we must leave the question open.

The second point we would like to make related to second-language acquisition theory concerns the derived-environment effect. As outlined above, the derived-environment effect is a constraint according to which any learner acquiring an allophonic split will make the relevant contrast in a derived environment only if that learner also makes the contrast in a basic environment. This constraint obtains for our participants, and interestingly, holds also for those participants making a covert contrast. Thus, participants 3029, 3019 and 3030 evince a covert contrast between [s] and [z] in inter-morphemic position. These same participants also show that contrast in basic environments, either overtly, as in the case of participants 3019 and 3030, or only covertly, as with 3029.

Within this context, we should point out an apparent anomaly in the productions of 3024. Although this participant maintains a contrast in both a basic and a derived environment, the contrast in the basic environment occurs in word-medial position but not, as we would expect, also in word-initial position. Because the transcriptions for this participant ran counter to this widespread expectation that a voice contrast in word-medial position would entail such a contrast word-initially, the investigators listened to the sound files of the tokens containing word-initial [s] and [z] for this participant. Whereas the words containing initial [s] were pronounced by this participant with what was clearly a voiceless sibilant fricative, the tokens containing initial [z], conversely, in most cases sounded more voiced than voiceless, and in other instances were clearly voiced.[9] However, because our stated protocol was to accept the transcriptions as they came to us via file transfer protocol from the assistants at a Midwestern university, we nevertheless based our findings on these renderings of the utterances. Further analysis may show that participant 3024 nonetheless does evince a word-initial contrast between [s] and [z], along with the noted word-medial contrast.

Pedagogical implications

We now turn to three pedagogical implications of our findings. The first is that the relative difficulty that L2 learners have in splitting NL allophones into TL phonemes was first pointed out, to our knowledge, by Lado (1957: 15), when he claimed that this learning situation constituted maximum difficulty. Our findings can shed some light on the explanation of this difficulty in that the derived environment effect constrains learning such that acquisition of a contrast in derived environments implies that the learner will have the contrast in basic environments. This allows, of course, for the L2 learners to show the contrast in basic contexts, yet lack the contrast (that is to say, to continue to err systematically) in derived contexts. We can only speculate at this point, but observation of this learning pattern may have caused linguists and language teachers, including Lado, to take note of this difficulty.

The second pedagogical implication is that acquisition of a phonemic contrast, especially one involving an allophonic split, is a function of phonological environment. The fact that an L2 learner has acquired a TL phonemic contrast in a given environment does not mean that the learner has also acquired that contrast in some other environment. Our findings indicate the contrary: a learner may systematically evidence a contrast in one environment, and just as systematically lack that contrast elsewhere. Moreover, our results also show that there is a relationship between the existence of certain kinds of contrast and the environments in which those contrasts occur.

The final pedagogical consequence of these findings is that learners may make a phonemic contrast covertly. If our results, which still must be considered preliminary, are viewed in the same light as work on covert contrasts in L1 acquisition and phonologically disordered speech, then the presence of covert contrasts among L2 learners points to an intermediate stage of acquisition. On this interpretation,

9. The protocol for the elicitation of the data from the subjects directed us, in cases where subjects produced more than a single token of the target word, to accept only the last uttered token. Subject 3024 produced two or more tokens of four of the ten words, and pronounced the majority of them with a voiced sibilant fricative. If all of these utterances were considered, 3024 would have reached the 80% criterion in the pronunciation of word-initial [z].

it would be expected that some L2 learners progress from a state of no contrast and then pass through a stage of covert contrast before ultimately arriving at a (final) stage of overt contrast. Indeed, an intermediate stage of covert contrast appears to be necessary, as it takes time to learn, progressively, the phonetic implementation of target phonemes.

Conclusion

This paper has reported findings indicating that the acquisition of an L2 phonemic contrast may involve, for at least some learners, an intermediate stage of covert contrast. Some L2 learners may implement a TL phonemic distinction in a way that can be reliably measured acoustically, but which is not perceived by native speakers of the TL, even those who are phonetically trained. The existence of an intermediate stage of covert contrast in the learning of L2 phonology is eminently plausible, in view of the progressive nature of this task, and brings the acquisition of second-language contrasts into conformity with findings of the same phenomenon in the areas of L1 acquisition and phonologically disordered speech.

Acknowledgement

This work from was supported in part for the first two authors by a grant from the National Institutes of Health 1 R01 HD046908-05. The positions expressed in this paper are those of the authors, and do not necessarily reflect those of NIH. We thank Carolyn Barry and Alison Garcia for their role in conducting the acoustic analysis. Any remaining errors are the fault of the authors.

References

Altenberg, E. & Vago, R. 1983. Theoretical implications of an error analysis of second language phonology production. *Language Learning* 33: 427–447.

Boersma, P. & Weenink, D. 2005. PRAAT: Doing phonetics by computer, Version 4.4.07. <http://www.praat.org/>

Bradley, T.G. & Delforge, A.M. 2006. Systemic contrast and the diachrony of Spanish sibilant voicing. In *Historical Romance Linguistics: Retrospectives and Perspectives* [Current Issues in Linguistic Theory 274], D. Arteaga & R. Gess (eds), 19–52. Amsterdam: John Benjamins.

Broselow, E., Chen, S. & Wang, C. 1998. The emergence of the unmarked. *Studies in Second Language Acquisition* 20: 261–280.

Cancino, H., Rosansky, E. & Schumann, J. 1978. The acquisition of English negative and interrogatives by native speakers of Spanish, in *Second Language Acquisition,* E. Hatch (ed). Rowley MA: Newbury House.

Carlisle, R.S. 1998. The acquisition of onsets in a markedness relationship: A longitudinal study. *Studies in Second Language Acquisition* 20: 245–260.

Cho, T. 1999. Specification of Intergestural Timing and Gestural Overlap: EMA and EPG Studies, MA thesis, University of California, Los Angeles.

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

Crystal, T.H. & House, A.S. 1988. Segmental durations in connected-speech signals: Current results. *Journal of the Acoustical Society of America* 83: 1553–1573.

Eckman, F. 1981. On the naturalness of interlanguage phonological rules. *Language Learning* 31: 195–216.

Eckman, F. & Iverson, G. 1994. Pronunciation difficulties in ESL: Coda consonants in English interlanguage. In *First and Second Language Phonology,* M. Yavas (ed.), 251–266. San Diego CA: Singular.

Eckman, F., Iverson, G., Fox, R.A., Jacewicz, E. & Lee, S. 2011. Explicit training and implicit learning of L2 phonemic contrasts. In *Implicit and Explicit Conditions, Processes and Knowledge in SLA Bilingualism*, C. Sanz & R. P. Leow (eds), 159–174. Washington DC: Georgetown University Press.

Eckman, F. & Iverson, G. 2013. The role of the native language in the acquisition of L2 phonemic contrasts. *Studies in Second Language Acquisition* 35: 1–26.

Flege, J.E. 1987. The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics* 15: 47–65.

Forrest, K., Weismer, G., Hodge, M., Dinnsen, D. A. & Elbert, M. 1990. Statistical analysis of word-initial /k/ and /t/ produced by normal and phonologically disordered children. *Clinical Linguistics and Phonetics* 4: 327–340.

Garcia, A. 2013. Allophonic Variation in the Spanish Sibilant Fricative. PhD dissertation, University of Wisconsin at Milwaukee.

Gierut, J.A. & Dinnsen, D. 1986. On word-initial voicing: Converging sources of evidence in phonologically disordered speech. *Language and Speech* 29: 97–114.

Hammerly, H. 1982. Contrastive phonology and error analysis. *IRAL* 20: 17–32.

Harris, J. 1969. *Spanish Phonology*. Cambridge MA: The MIT Press.

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

Hewlett, N. 1988. Acoustic properties of /k/ and /t/ in normal and phonologically disordered speech. *Clinical Linguistics and Phonetics* 2: 29–45.

House, A.S. 1961. On vowel duration in English. *Journal of the Acoustical Society of America* 33: 1174–1182.

Hualde, José Ignacio. 2005. *The Sounds of Spanish*. Cambridge: CUP.

Iverson, G. & Salmons, J. 1995. Aspiration and Laryngeal Representation in Germanic. *Phonology* 12: 369–396.

Kiparsky, P. 1982. Lexical phonology and morphology. In *Linguistics in the Morning Calm,* I.S. Yang (ed.), 3–91. Seoul: Hanshin.

Kornfeld, J.R. & Goehl, H. 1974. A new twist to an old observation: Kids know more than they say. In *Papers from the Tenth Annual Meeting of the Chicago Linguistic Society, Parasession on Natural Phonology,* 210–219. Chicago IL: CLS.

Lado, R. 1957. *Linguistics Across Cultures*. Ann Arbor MI: University of Michigan Press.

Lehiste, I. 1972. The timing of utterances and linguistic boundaries. *Journal of the Acoustical Society of America* 51: 2018–2024.

Lim, J.-A. & Oh, M. 2007. Covert contrast in second language acquisition. In *Proceedings of the 16th International Conference on Korean Linguistics*, International Circle of Korean Linguistics (ed.), 202–203. New York NY: SUNY at Binghamton & Cornell University.

Macken, M.A. & Barton, D. 1980. A longitudinal study of the acquisition of the voicing contrast in American-English word-initial stops, as measured by voice onset time. *Journal of Child Language* 7: 41–74.

Major, R. 1994. Chronological and stylistic aspects of second language acquisition of consonant clusters. *Language Learning* 44: 655–680.

Martínez-Gil, F. 2003. Resolving rule-ordering paradoxes of serial derivations: An optimality theoretical account of the interaction of spirantization and voicing assimilation in Peninsular Spanish. In *Theory, Practice, and Acquisition*, P. Kempchinsky & C-E. Piñeros (eds), 40 – 67. Somerville MA: Cascadilla Press.

Maxwell, E.M., & Weismer, G. 1982. The contribution of phonological, acoustic, and perceptual techniques to the characterization of a misarticulating child's voice contrast for stops. *Applied Psycholinguistics 3*: 29–43.

Ohala, J.J. 1974. Experimental historical phonology. In *Historical Linguistics,* II: *Theory and Description in Phonology. Proceedings of the 1st International Conference on Historical Linguistics. Edinburgh, 2–7 Sept. 1973*, J.M. Anderson & C. Jones (eds), 353–389. Amsterdam: North Holland.

Ritchie, W. 1968. On the explanation of phonic interference. *Language Learning* 18:183–197.

Scobbie, J.M. 1998. Interactions between the acquisition of phonetics and phonology. In *Papers from the 34th annual regional meeting of the Chicago Linguistic Society,* Vol. II, M.C. Gruber, D. Higgins, K. Olson & T. Wysocki (eds), 343–358. Chicago IL: CLS.

Scobbie, J.E., Gibbon, F., Hardcastle, W.J. & Fletcher, P. 2000. Covert contrast as a stage in the acquisition of phonetics and phonology. In *Papers in Laboratory Phonology,* V: *Language Acquisition and the Lexicon*, M. Broe & J. Pierrehumbert (eds), 194–203. Cambridge: CUP.

Smith, B.L. 1979. A phonetic analysis of consonant devoicing in children's speech. *Journal of Child Language* 6: 19–28.

Smith, C.L. 1997. The devoicing of /z/ in American English: Effects of local and prosodic context. *Journal of Phonetics* 25: 471–500.

Stevens, K.N., Blumstein, S.E., Glicksman, L., Burton, M. & Kurowski, K. 1992. Acoustic and perceptual characteristics of voicing in fricatives and fricative clusters. *Journal of the Acoustical Society of America* 91: 2979–3000.

Appendix

Target words

1. sink
2. see
3. sick
4. seat
5. scissors
6. cessna
7. muesli
8. Christmas
9. Presley
10. isthmus
11. anxious
12. serious
13. close
14. famous
15. base
16. loose
17. face
18. price
19. horse
20. house
21. zebra
22. zero
23. zee
24. zoo
25. zip
26. grizzly
27. asthma
28. plasma
29. business
30. paisley
31. bruise
32. rose
33. wise
34. news
35. nose
36. noise
37. haze
38. breeze
39. prize
40. glaze
41. seriousness
42. houseless
43. anxiousness
44. baseless
45. looseness
46. faceless
47. horseless
48. closeness
49. famousness
50. priceless
51. newsless
52. breezeless
53. hazeless
54. wiseness
55. roseless
56. bruiseless
57. prizeless
58. wiseness
59. noseless
60. glazeless