Overt and covert contrast in L2 phonology

Fred Eckman, Gregory Iverson and JaeYung Song
University of Wisconsin-Milwaukee, USA

This paper reports results on the acquisition of the English /p/–/b/ contrast by native speakers of Arabic. This contrast does not exist in the participants' native language (NL). The central finding of this study is that some of the research participants exhibited a *covert* contrast between these segments in their interlanguage productions. That is, two of the five Arabic-speaking participants who were transcribed as having no contrast between [p] and [b] did, in fact, produce a statistically reliable distinction in voice onset time lags between the two target segments. The existence of such an intermediate stage of covert contrast in the learning of L2 phonology is eminently plausible, in view of the progressive nature of phonological acquisition. Our results help bring the learning of second-language contrasts into conformity with findings of the same phenomenon in the areas of L1 acquisition and phonologically disordered speech.

1. Introduction

It seems reasonably well accepted within the fields of child-language (L1) acquisition, disordered speech and second-language (L2) learning that all language acquirers construct their own, intermediate version of the target (or ambient) language (TL), which they then use to produce and understand utterances of the language being acquired. A key research question in each of these disciplines is what the nature of these learner languages is, and why they are as they are. Over the last thirty years or so, attempts by researchers to understand these systems in L1 acquisition and disordered speech has led to the discovery of a *covert contrast*, a statistically reliable acoustic difference between targeted phonemes that is produced by a language learner, but that is nevertheless not perceived by native speakers of the target language (Macken & Barton, 1980; Gierut & Dinnsen, 1986). This type of distinction has been hypothesized to be an intermediate stage in a learner's progression from making no distinction between contrasting target segments to full phonetic implementation of the contrast that is appropriately perceived by transcribers and native speakers of the TL (Scobbie, Gibbon, Hardcastle & Fletcher, 2000). The

Journal of Second Language Pronunciation 1:2 (2015), **254–278**. DOI 10.1075/jslp.1.2.06eck
ISSN 2215–1931 / E-ISSN 2215–194X

findings on covert contrasts in L1 acquisition and disordered speech have had both theoretical and practical implications for those fields of inquiry. During this same period, research on L2 acquisition has taken a different path. Aside from our own recent study documenting the production of a covert contrast between English /s/ and /z/ by native speakers of Spanish (Eckman, Iverson & Song, 2014), we are aware of no research on covert contrasts in L2 acquisition.[1]

The purpose of this paper is to report findings of an ongoing investigation into the acquisition of L2 phonemic contrasts, and thereby to document the stage of a covert contrast by L2 learners of English. The case at hand involves the production of the /p/–/b/ contrast by adult learners of English whose native language (NL) is Arabic, a language that does not have this contrast because the sounds in question are in complementary distribution. In Arabic, [p] is an allophone of /b/, occurring only before voiceless obstruents, whereas [b] occurs elsewhere.

We wish to emphasize here that our goal in this paper is simply to attest the production of a covert contrast by L2 learners, thereby bringing findings in second-language phonology in line with research in L1 acquisition and disordered speech, fields which have over the past three decades shown the importance of covert contrasts in phonological acquisition. It remains to be shown by future research in this area how widespread covert contrasts are and whether they constitute a necessary, intermediate stage on the path to full implementation of the TL contrast.

The remainder of the paper is structured as follows. In the next section, we place our study in context, first, by reviewing two of the seminal studies on covert contrasts in the acquisition of English by normal children and by children who are phonologically disordered, and second, by supplying the necessary details of our own work in this area. After providing a description of [p] and [b] in Arabic, we state our hypothesis, give an outline of the methodology by which the data were gathered, and then report the results as they bear on the hypothesis. We conclude with a discussion of our findings in light of our claims, and with a view toward some of the implications of the results for the acquisition of TL phonemic contrasts.

2. Background

Until some thirty years ago, there was widespread reliance on phonetic transcriptions in research on L1 acquisition and disordered speech. However, since groundbreaking work in this area by Macken & Barton (1980), numerous studies on both acquisition of their native phonology by typically developing children, and on children with phonological disorders have shown that the participants often produced

1. We are aware of only one other publication, an abstract by Lim & Oh (2008).

contrasts that were not perceived by the adult listeners or transcribers (Macken & Barton, 1980; Maxwell & Weismer, 1982; Gierut & Dinnsen, 1986; Forrest et al., 1990; Scobbie et al., 2000). This phenomenon, in which L1 learners produce a statistically reliable distinction that is not perceived by adults, whether phonetically trained or not, was termed a *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.

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 pioneering study to report a covert contrast in the acquisition of phonology by children. Theirs was a longitudinal study of the acquisition of the English voicing contrast by four monolingual children. The authors analyzed the productions of participants 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 is defined as the time, usually measured in milliseconds (ms), between the release of closure and the beginning of vocal cord vibration in the following vowel (or sonorant consonant). In some languages, such as English, the onset of vocal cord vibration occurs after the release of the consonant, in which case VOT is stated as a VOT lag. In other languages, such as Arabic, in which the onset of vocal cord vibration occurs before the release of the stop, the VOT is stated as a VOT lead. VOT 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 (Lisker & Abrahamson, 1964), and thus are phonetically voiceless, and largely unaspirated, whereas stops categorized as 'voiceless' show a mean VOT of 60 ms or more, and 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 all of these values 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.

In the ensuing years, there have been a number of studies on the acquisition of covert contrasts among phonologically disordered children. Limitations of space will allow us to describe only two. One of the earliest was by Maxwell and Weismer (1982), who studied a misarticulating boy at the age of three years, eleven months (3;11). Initial testing of the child's articulation revealed that his phonemic inventory maintained a contrast between oral and nasal consonants, but that [b] and [d] were the only obstruents he pronounced. Follow-up articulations from this child were analyzed using wide-band spectrograms for each utterance in order to measure stop-closure duration, voicing during closure, and VOT. Results of this analysis showed that the test words could be partitioned into three groups: words beginning with [d] in adult speech ([d]-words), words beginning with other voiced obstruents in adult speech ([+voice]-words), and words beginning with voiceless obstruents in adult speech ([-voice]-words). One-tailed *t*-tests showed a statistically significant difference between the mean VOT of the [d]-words and the [+voice]-words, and between the [d]-words and the [-voice]-words. In short, the child was implementing a three-way consonantal contrast that was perceived by native-speaking transcribers as /d/.

In another early study on misarticulating children, Gierut and Dinnsen (1986) analyzed two children, aged 4;6 and 4;3. Phonetic transcription of the children's utterances indicated that both 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 three decades, it has 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. Scobbie (1998) presents a thorough listing and review of a large number of such studies, and more recently,

Berti (2010) and Munson et al. (2010) have discussed covert contrast in disordered speech.

Over this same period of time, research on the acquisition of target-language sound patterns by L2 learners has proceeded along two distinct paths, neither of which has led to the finding of the production of a covert contrast. The first path is shown by a vast literature reporting careful studies that document L2 phonetic development, status and implementation, both in the areas of perception and production, of the sound inventories of numerous TLs (Bohn & Flege, 1992; Flege, 1987, 1990, 1993, 1995; Leather & James 1991; Munro & Derwing, 1997). This work has led to the formulation of at least two important frameworks in the acquisition of L2 speech, the Speech Learning Model (Flege, 1995) and the Perceptual Assimilation Model (Best, 1995).

A number of phonetic studies, as reported in Flege (1991), have documented that nonnative speakers of English may produce target-language stops with VOT values that are intermediate between those of the native language and those of the TL. For example, Flege (1991) found that adult, native speakers of Spanish who were L2 learners of English produced English stop consonants with VOT values that were intermediate (i.e., they were a "compromise" to use Flege's term) between those produced by monolinguals of Spanish and those produced by monolinguals of English. However, we do not know how the consonants produced by the L2 learners in Flege's study were perceived by native speakers of the TL.

The second path in research on the acquisition of L2 phonologies has relied almost exclusively on phonetic transcriptions to document the learners' progress in acquiring the TL phonology (Altenberg & Vago, 1981; Broselow, Chen & Wang, 1998; Carlisle, 1998; Eckman, 1981; Eckman & Iverson 1994; Flege, 1987; Hammerly, 1982; Major, 1994; Ritchie, 1968). The phonetic transcriptions used by these studies 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, meaning that the transcribers did not perceive a distinction in some of the sounds that the research participants were uttering.

Based on the large amount of research on the acquisition of covert contrasts both in L1 phonological acquisition and in disordered speech, it seems plausible that a stage of covert contrast could be documented in the acquisition of L2 phonology, and in fact, we did attest a stage of covert contrast in our investigation of the acquisition of the English /s/–/z/ contrast by native speakers of Spanish, as reported in Eckman et al. (2014).[2]

2. One of the anonymous reviewers questioned whether it was possible to use the term "covert contrast" in L2 exactly as we do in describing L1 acquisition, because L2 acquisition differs from

Spanish contains both [s] and [z], but these sounds are allophones of /s/ in Spanish, with [z] occurring only before voiced consonants and [s] occurring elsewhere. Therefore, native speakers of Spanish acquiring English would need to learn how to distinguish [s] and [z] phonemically. In order to test the hypothesis that at least some of our Spanish-speaking participants would exhibit a covert contrast in acquiring the English /s/–/z/ distinction, we elicited from 14 native speakers of Spanish a total of 60 English words containing the /s/–/z/ contrast in three different environments, including medially following a vowel and preceding a voiced consonant (e.g., *sip*, *zip*, *business*, *Christmas*, *pass*, *buzz*). After the data were collected and transcribed by research assistants who were blind to the hypothesis, we performed an acoustic analysis on the relevant sound files by measuring the percent of voicing that overlapped with the fricative noise for each token of the [s] and [z] targets.

According to the transcribers, seven of the 14 participants produced an overt contrast between [s] and [z] in at least one of the targeted positions, and seven of the subjects produced no contrast between [s] and [z] in any of the word-positions. However, according to the acoustic analysis, four of the seven participants who were transcribed as producing no contrast between [s] and [z] did, in fact, produce a statistically significant difference in the percentage of overlap during frication between target [s] and [z]. In other words, four of the subjects produced a covert contrast between English /s/ and /z/, supporting our hypothesis.

Before proposing the hypothesis for the current study, we wish to address two important questions raised by one of the anonymous reviewers. The first is whether there is any theory of second-language acquisition that would predict the occurrence of a covert contrast. And the second is why one should expect covert contrasts in second-language learners, where unlike L1 acquisition, there often exists, as is true for our participants, a context in which the L2 learners are acquiring the TL in a tutored setting with ample support from their instructors. We will respond to each question in turn.

We are not aware of any theoretical framework, in either first or second language acquisition, which predicts the existence of a covert contrast. However, two important facts are pertinent here. The first is that since the beginning of the generative era, an assumption held by virtually all researchers into language acquisition

L1 acquisition in that the former does not always result in complete mastery of the TL phonetic output, whereas L1 acquisition does. However, we believe that "covert contrast" can be applied in a meaningful way to describing second-language acquisition because there is nothing in the use of this term that implies whether or not the learner will gain mastery over the TL phonetic output. Rather, a covert contrast in the present context simply refers to a stage in which the learner is implementing the phonemic distinction in a way that is not correctly perceived by native speakers of the TL.

is that learners must create a grammar. The second relevant fact is that a major goal of all acquisition studies within this context has been to document and explain the intermediate stages of learner-systems that are constructed on the path to full acquisition. Thus, for example, Berko (1958), Brown (1970) and Tyler and Nagy (1989), to cite just a few, have addressed the developmental stages of morphology in child-language acquisition; Dulay and Burt (1973), among others, documented the sequence of acquisition for grammatical morphemes in child second-language acquisition. Morgan (1987), Bardovi-Harlig (1987) and Bialystok (1987) have all reported on the stages of the acquisition of various words and phrases by L1, L2, and bilingual learners, respectively. Bellugi (1971), Klima and Bellugi (1966) for L1 acquisition, and Wode (1981) for L2 learning, have described and attempted to explain the various steps involved in the acquisition of English questions. And finally, the studies and discussions in Slobin (1985) focus on the various pathways followed in the cross-linguistic development of L1 acquisition. Covert contrasts in L2 phonology, it seems to us, fit into this context insofar as they are hypothesized to constitute an intermediate stage in the acquisition of TL phonemic distinctions.

The second question raised by the reviewer can be addressed along this same line of reasoning. Numerous studies in second-language acquisition, including those cited in the previous paragraph, have investigated learning by participants who were studying the TL in a tutored context, reporting that such learners, just as do those learning the TL naturalistically, evince intermediate stages of acquisition. In other words, there seems to be no compelling reason not to expect that L2 learners enrolled in a language program would exhibit stages of learning.

Moreover, articulatory movements and resulting acoustic parameters are necessarily continuous. However, listeners tend to perceive sounds that lie along an articulatory or acoustic continuum as belonging to distinct categories, a well-known phenomenon termed categorical perception (Liberman, Harris, Hoffman, & Griffith, 1957). According to categorical perception, sounds that fall within the range for the same category will be perceived as belonging to that category even when they are systematically different from each other.

Before language learners are able to acquire the correct acoustic space for contrasting target sounds, and to implement them appropriately, they may produce sounds that fall within the same category. For children, articulatory skills and controls that are still developing could be a limiting factor that affects their performance. For adult L2 learners, the problem might be attributed more to learning the correct acoustic ranges for the target segments, especially if the ranges are different from their native language. However, independent of the domains and limiting factors, we can make a prediction based on categorical perception that some sounds will be perceived as undifferentiated by adult listeners in the case

of L1 acquisition, or by native speakers of the target language in the case of L2 acquisition, even though the sounds are in fact physically different.

Thus, in our view, it is an open, empirical question as to whether L2 learners exhibit a covert contrast, and while it is not the case that covert contrasts are necessarily predicted by any theoretical framework, it is certainly reasonable to hypothesize that they occur.

Given this context, we believe that covert contrasts should be more prevalent in second-language phonology than what the L2 literature has heretofore reported. Consequently, we propose the hypothesis in (1).

(1) Hypothesis
In their acquisition of a target-language phonemic distinction, some second-language learners may implement the TL contrast acoustically in a way that is not perceived by native speakers of the TL.[3]

We intend to test the hypothesis in the current study by attesting the production of a covert contrast in the L2 acquisition of English /p/ and /b/ by native speakers of Arabic.

We begin with some phonetic background on the stops in these two languages. Both English and Arabic have the three points of articulation for stops, labial, coronal and velar, which can be categorized differently with respect to their VOT ranges. According to Lisker & Abramson (1964), the mean VOT values for English word-initial voiceless stops, depending on point of articulation, vary from 58 ms for [p] to 80 ms for [k]. The mean VOT lags for English voiced stops are from one ms for [b] to 21 ms for [g]. On the other hand, Arabic voiced stops are articulated with vocal vibration during the closure, and therefore have negative VOT values, averaging from −40 to −90 ms, depending on place of articulation (Yeni-Komshian et al., 1977). The VOT value for Arabic [b] varies from −40 ms to −80 ms, depending on the following vowel. Arabic voiceless stops are produced with short-lag VOT values ranging from 15 to 35 ms, also depending on point of articulation of the stop, and on the following vowel. Arabic has no /p/ phoneme,

3. One question that may be raised by the statement of the hypothesis in (1) is why we are asserting that L2 learners "may" produce a covert contrast between the targeted sounds, rather than claiming that all participants "will" exhibit such a contrast. Our view is that it seems reasonable to expect that all participants will pass through an intermediate stage of covert contrast as they progress to the final state of producing the contrast overtly. However, whether we are able to attest a stage of covert contrast for all participants seems to us to be simply a question of timing. And given that a) we have no evidence that covert contrast is a necessary stage in the progression to overt contrast, and b) that we are not equipped in the present study to test such a claim, we must leave the question open for future work, and be content with testing the hypothesis in (1).

as it is an allophone of /b/, with [p] occurring only before voiceless obstruents, and [b] occurring elsewhere. Thus, in Arabic, only the voiced, and not the voiceless, bilabial stop, [b], occurs word-initially, the environment on which we focused our attention for the participants' production of the English /p/–/b/ contrast.

3. Methodology

To address the claim in (1) above, we elicited productions of English words containing the /p/–/b/ contrast in word-initial position (e.g., *pie*, *buy*) from nine native speakers of Arabic who were learning English as a second language.[4] In an attempt to enlist participants who had a relatively wide range of English proficiency, we recruited subjects both from the English as a Second Language Program at a Midwest University, and from the campus community at large. Six of the participants were enrolled in the university's ESL Program, and three were recruited from credit courses taught on campus. Seven of the participants were from Saudi Arabia and two were from Jordan. Both varieties of the regional vernacular of Arabic in these countries lack a /p/-/b/ contrast, where [p] is an allophone of /b/. The participants varied in age from 19 to 24 years (mean 22 years, four months), and all had studied English in their own country for at least four years. Their length of residence in the United States varied from six to fifteen months, and their overall command of spoken English was estimated by the administrators of the ESL program to be mid- to high-intermediate. All subjects were paid a fee for their participation in the project.

Within this context, it is important to point out that our hypothesis is not necessarily dependent on the level of English proficiency of our research participants, though, other things being equal, one would expect that more proficient learners of English (however measured) would be more likely to have acquired the contrast in question. However, our hypothesis simply asserts that, with respect to the contrast that we are investigating, the IL of any Arabic-speaking learner of English can be analyzed on the basis of the productions of the respective learner, and that our claim is testable regardless of the level of English proficiency of the participants.

All of the stimuli used to elicit the productions from the participants are existing lexical items in English exhibiting the relevant distinction. There were 64 words total, 21 of which targeted the contrast in word-initial position, 11 beginning with

4. One of the anonymous reviewers commented that the sample size was too small to draw any conclusion about covert contrast. In this respect, it is important to note that we are not attempting to draw any inferences across our participants that would depend on the size of the group. Rather, our goal is simply to attest the production of a covert contrast in L2 acquisition.

/p/ and 10 with /b/. The remaining words fell into three categories, the first being minimal pairs for /p/ and /b/ in word-final position, the second consisting of minimal pairs for /k/ and /g/ in onset and coda positions, and the third having minimal pairs for /t/ and /d/ in onset and coda position. The first two categories were fillers in the sense that we wanted to distract our participants from the focus of the elicitation, which was word-initial /p/ and /b/. On the other hand, the /t/–/d/ words acted as control items, because Arabic, the participants' NL, has a contrast between /t/ and /d/. Thus, we expected the participants to have no difficulty implementing a voice contrast in these TL consonants. Our goal was to test whether the participants made a voice contrast on TL /p/ and /b/. The words used in the elicitation are shown in the Appendix.

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 subject and the experimenter through the elicitation of each word. We elicited the words, not by giving their spellings, but by displaying an image depicting the object or idea in question on the screen. If subjects 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.[5] The subjects were also given a practice exercise with a different set of words in order to ensure that they understood the directions. The stimuli were presented in a randomized order, and were recorded directly onto a hard disc drive at the sampling rate of 44.1 kHz. Subjects spoke into a head-mounted microphone at a distance of one inch from the lips.

All of the data were collected and were then transcribed by a total of five research assistants, all of whom were phonetically trained native speakers of English and were unaware of the hypothesis or the goal of the study. The transcribers were given the instructions to listen to the recordings of the target words and control items and to provide a broad phonetic transcription of the initial consonants using the International Phonetic Alphabet. To provide a more reliable transcription, we had the five transcribers independently make close phonetic transcriptions of all the word-initial consonants for all three of the voice contrasts viz., for /p/–/b/, /k/–/g/ and /t/–/d/. This procedure produced a total of five transcriptions for each token. The final transcription of a token was entered to reflect the majority of the transcriptions (i.e., at least three out of the five transcriptions), and the reliability figure was computed on the basis of these transcriptions. For example, if the word *bull* was transcribed with an initial /b/ by three of the transcribers and with an

5. The option of a recorded model was never used, as all participants were able to determine the targeted word on the basis of the picture and/or the definitions.

Table 1. The reliability scores, in percentages, for the grand mean, means for each consonant, and for each of the participants on the six consonants. The participant identification numbers are listed in the first column, with the relevant consonants listed across the top of the table. The mean reliability calculations for each segment across the nine participants are shown at the bottom of the table, with the overall mean presented in the lower corner of the table.

Participant	initial /p/	initial /b/	initial /k/	initial /g/	initial /t/	initial /d/	
6001	98	100	100	100	100	100	
6002	88	93	100	100	100	100	
6003	96	98	100	100	96	100	
6004	96	84	98	100	100	100	
6005	100	98	100	100	100	100	
6006	90	91	100	100	100	100	
6007	96	95	98	100	100	100	
6008	100	98	100	100	100	100	
6009	72	95	100	100	100	100	
Mean	92.9	94.5	99.5	100	99.6	100	
Grand mean							97.7

initial /p/ by the other two transcribers, the final transcription for *bull* was entered as /b/ with a reliability figure of 60%, reflecting the three out of five transcriptions that were in agreement.

We interpreted the transcription results as showing that participants exhibited an overt contrast between the phoneme pairs in question if they were transcribed as correctly producing each member of the contrast at least 80% of the time in word-initial position. This criterion has been widely used in second-language acquisition research, and dates back at least to the seminal work by Cancino, Rosansky and Schumann (1978). It forms the basis for concluding that the participants produced the targeted segments systematically enough for the contrast in question to be considered part of their IL.[6]

6. One might propose, as did one of the anonymous reviewers, that it would be more reliable to compute a d-prime value for the transcriptions instead of using the 80% criterial threshold. The d-prime statistic, however, is not appropriate in this case, because d-prime is used to determine whether there is any bias on the part of the perceiver in the perception of some signal (sound), where the target sound being perceived is known. In the current case, the utterances produced by the participants are not known, as that's what we are asking the transcribers to determine. Thus d-prime is not appropriate.

On the other hand, we did run a one-sample t-test comparing a participant's percentage score on the transcriptions against the chance value of 50%, which would be the score if the

General reliability measures, shown in Table 1, were calculated for each of the six consonants in question by averaging the reliability figures from the five transcribers across all of the items for each of the nine participants. As can be seen in Table 1, the reliability figures for both /p/ and /b/ are above 92%, and the overall reliability calculation across all six segments was 97.7%. All of these figures are deemed acceptable.

We now present the results of the study.

4. Results

In this section we present our findings as they bear on the hypothesis in (1), showing first the results from the transcriptions in Table 2, and then the results of the acoustic analysis in Figure 1. We will conclude that a participant produced a covert contrast if the acoustic facts indicate a statistically reliable difference in the VOT measurements for targeted /p/ compared to /b/, but there is no difference between the segments according to the transcriptions.

Table 2 shows, according to the final transcriptions, the percentage of target-like productions for each of the voiced-voiceless pairs of English stops in word-initial position. Columns 4, 7 and 10 indicate, according to the 80% criterial threshold, whether the participant produced the contrast in question.

The productions of all of the subjects evidenced an overt phonemic distinction between /k/ and /g/ and between /t/ and /d/ according to the criterion. However, only four of the participants exhibited the /p/–/b/ contrast overtly; five participants did not. Of the five who did not evidence the contrast, two failed to reach the 80% criterion on both /p/ and /b/ (6002 & 6007); three participants did not reach the threshold on only one of the sounds in question. Specifically, 6004 did not produce the segment /b/ to the 80% level, and two participants, 6006 and 6009, did not reach this threshold with their productions of /p/.

We also performed an acoustic analysis on the sound files of the targeted /p/–/b/ contrast, as well as on /t/ and /d/. The latter contrast was used as a control to ensure that any participant's failure to systematically pronounce a contrast between /p/ and /b/ was not simply an artifact of an inability to produce a voice

transcribers were perceiving the participants' utterances randomly. Our results did not change on the basis of the t-test, in that all of the participants that were perceived as either producing or not producing the contrast under the 80% criterion scored the same under the t-test. In fact, for the number of tokens produced by the participants, the level above 50% that yielded statistical significance for the t-test was close to 80% for both /p/ and /b/. Finally, it is worth noting that our results do not change if we use a 70% or a 90% criterial threshold for the transcriptions.

Table 2. Results in percentages of target-like transcriptions for English voiceless and voiced obstruents in word-initial position. Contrasts indicated using 80% as the criterial threshold.

Subject	/p/	/b/	/p/–/b/ contrast	/t/	/d/	/t/–/d/ contrast	/k/	/g/	/k/–/g/ contrast
6001	100	91	Yes	100	100	Yes	100	100	Yes
6002	60	64	No	100	100	Yes	100	100	Yes
6003	90	100	Yes	100	100	Yes	100	100	Yes
6004	100	64	No	100	100	Yes	100	100	Yes
6005	100	100	Yes	100	100	Yes	100	100	Yes
6006	20	100	No	100	100	Yes	100	100	Yes
6007	70	64	No	100	100	Yes	100	100	Yes
6008	100	100	Yes	100	100	Yes	100	100	Yes
6009	60	91	No	100	100	Yes	100	100	Yes

contrast between stops in general. For the acoustic coding, we used a set of coding conventions that employed both auditory and visual information from the spectrogram and waveform. The analysis for both pairs of segments was straightforward: for each of the word-initial stops, we measured the VOT from the release of the stop closure to the vibration of the vocal folds (i.e., the onset of voicing of the following vowel). The coding was carried out by two trained assistants using Praat (Boersma & Weenink, 2005). All of the words, with the exception of *table*, were monosyllabic, consisted of roughly the same number of segments, and had similar phonetic structure.

The results of the acoustic analysis for /p/ and /b/ are shown in Figure 1 and those for /t/ and /d/ are given in Figure 2.

4.1 Group results

In order to determine whether our native speakers of Arabic were making a distinction between [p] and [b] we compared the speakers' VOT values of these consonants using a paired t-test. Each speaker had average VOT values for [p] and [b], and these were paired in the t-test. We did the same comparisons between [t] and [d] as a control. The results showed that the mean VOT for voiceless stops was significantly longer than for voiced stops. Specifically, the mean VOT for /p/ was 45 ms ($SD = 23$), which was significantly longer than that for /b/, 9 ms ($SD = 13$), ($t(8) = 4.02$, $p < 0.01$). Likewise, the mean VOT for /t/ was 67 ms ($SD = 20$) versus the mean for /d/, four ms ($SD = 24$), ($t(8) = 5.89$, $p < 0.001$). Thus, analysis of the

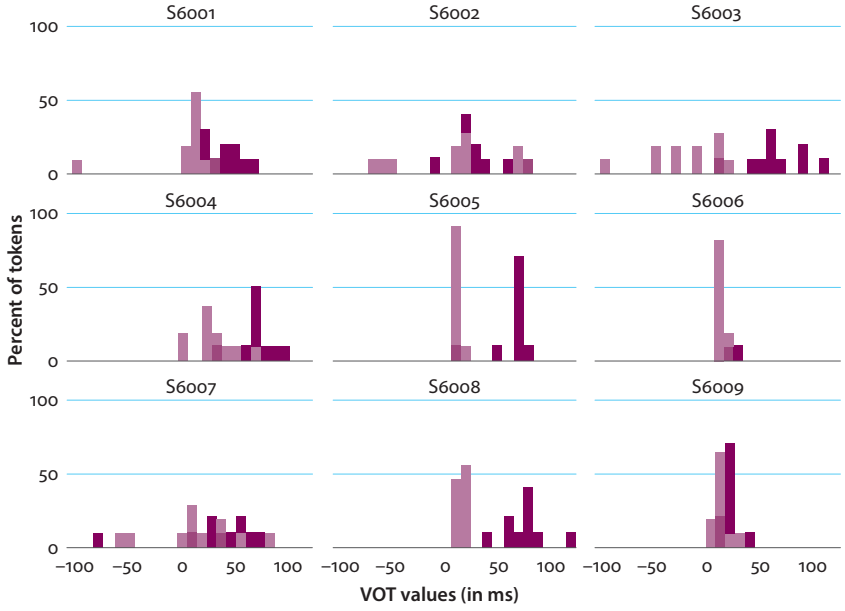

Figure 1. The distribution of VOT measurements for /b/ (light bars) and /p/ (dark bars) as produced by each participant. The VOT values are binned every 10 ms. The identification numbers of participants are shown for each histogram.

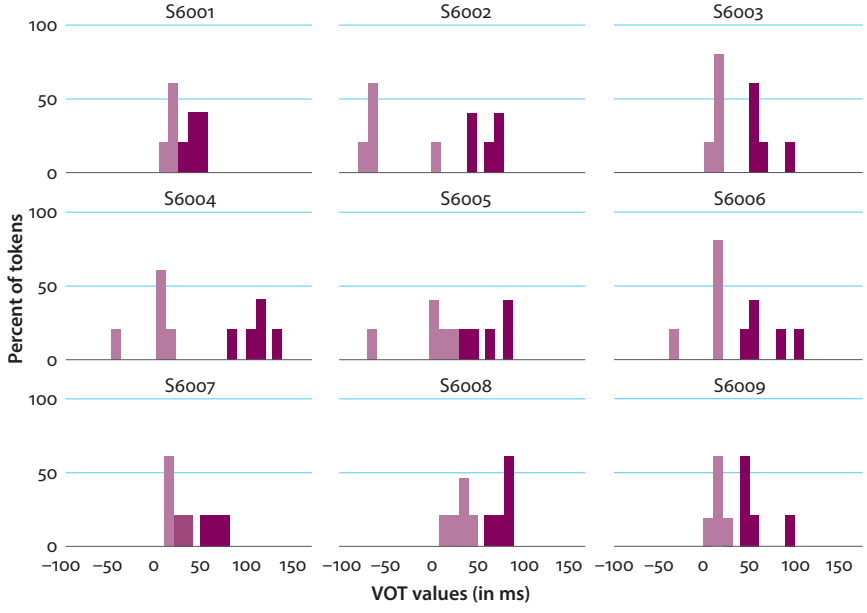

Figure 2. The distribution of VOT measurements for /d/ (light bars) and /t/ (dark bars) as produced by each participant. The VOT values are binned every 10 ms. The identification numbers of participants are shown for each histogram.

aggregated productions of our subjects evidenced a significant distinction in VOT between /p/ and /b/, and also between /t/ and /d/. One point worth noting regarding the VOT lags is that the mean VOT for /p/ for the group, 45 ms, is slightly shorter than the mean VOT for English /p/, which is 58 ms (Lisker & Abramson 1964). The VOT values represented by these means could have contributed to the fact that five of our participants were not perceived by the transcribers as successfully producing a contrast between /p/ and /b/.

4.2 Individual results

Next we examined the results for the individual participants to determine which, if any, showed a difference in VOT between /p/ and /b/, and as a control, between /t/ and /d/. We ran permutation tests for each speaker, using values from individual target words for each participant. Because our sample size is somewhat small, with an unequal number of target words for /p/ vs. /b/, we chose to use the nonparametric permutation test. Nonparametric tests do not require the assumption that the population distribution is normal. In a permutation test, the distribution of a statistic is derived from the observed data themselves by carrying out a large number of permutations. Then the observed mean difference between the two categories (e.g., /p/ vs. /b/) can be compared to this distribution to calculate a p-value. For further information on the permutation test, see Good (2000). The statistical analysis was carried out using R Version 3.1.1 (R foundation for Statistical Computing, 2009).

The results for /p/ and /b/, as well as those for /t/ and /d/, are shown in Tables 3 and 4, respectively. Table 4 shows that all of the speakers produced a significant VOT difference for /t/ and /d/, and that this distinction was also perceived by the transcribers. However, not all of the participants evidenced a distinction between /p/ and /b/, as seen in Table 3. More specifically, four of the subjects, 6001, 6003, 6005 and 6008, produced a reliable VOT difference between /p/ and /b/ that was perceived by the transcribers. We labeled this kind of performance as an *overt contrast*. Three subjects, 6002, 6006 and 6007 did not produce a VOT distinction for /p/ and /b/, and the transcribers failed to record a systematic difference in these segments for these three participants, who were labeled as having *no contrast*. Finally, two subjects, 6004 and 6009, produced a reliable difference in VOT for /p/ and /b/, but this difference was not perceived by the transcribers. We designated the productions of these participants as showing a *covert contrast*.

Several comments need to be made with respect to these findings. We consider the results for the individuals to be more important than those for the group. A covert contrast, just as is the case for an overt contrast, is a property of a learner's grammar, in this context, a learner's IL grammar. Because grammars are mental

Table 3. Results for each subject according to VOT calculations, permutation test, and interpretation of the acoustic analysis and transcriptions of /p/ and /b/ data. Column 1 lists the participants. Columns 2 through 5 give the mean VOT values and standard deviations (SD). The values for the significance level from the permutation tests are shown in the next column. The last three columns present the interpretations of the results according to whether a) the participant in question produced an acoustic distinction, b) the transcribers perceived a distinction, and c) the subject systematically produced no contrast, an overt contrast or a covert contrast.

Subject	VOT for /p/		VOT for /b/		Permutation test results p value	Differences between /p/ & /b/		
	Mean	SD	Mean	SD		Acoustic distinction	Transcription	Contrast?
6001	39	18	1	34	$p < 0.001$	Yes	Yes	Overt
6002	30	26	9	53	$p = 0.269$	No	No	No
6003	63	27	−22	35	$p < 0.001$	Yes	Yes	Overt
6004	70	20	27	20	$p < 0.001$	Yes	No	*Covert*
6005	64	21	12	4	$p < 0.001$	Yes	Yes	Overt
6006	14	6	13	4	$p = 0.794$	No	No	No
6007	34	45	15	44	$p = 0.358$	No	No	No
6008	76	21	16	7	$p < 0.001$	Yes	Yes	Overt
6009	19	7	11	7	$p = 0.014$	Yes	No	*Covert*

systems, being placed in time and space in the mind of the speaker/hearer, they are necessarily individual. There is no grammar of a group, at least not one that can be placed in time and space, because there is no mind of a group. Group results show us whether these Arabic speakers make a significant difference between /p/ and /b/ in VOT. However, group results cannot indicate anything about the status of individual IL grammars.

The individual results are important in that they attest a covert contrast, robustly in participant 6004, and perhaps marginally in 6009.[7] There are two reasons why this is true. First, as we can determine from Table 3, the mean VOT distinction between /p/ and /b/ for participant 6004 is 43 ms (70 ms for /p/ and 27 ms for /b/), whereas for 6009 the mean difference is eight ms (19 ms for /p/ and 11 ms

7. One of the anonymous reviewers pointed out that participants 6004 and 6009 were two of three participants who mastered one but not both of the targeted phonemes according to the transcribers (cf. Table 2). The reviewer then asked whether mastery of only one of the two phonemes was necessary in order to produce a covert contrast. Whereas this is an intriguing question, we have no data bearing on this matter and so must leave it open for future research. Moreover, there do not appear to be any obvious demographic similarities between participant 6004 and any of the other participants who exhibited either no contrast or an overt contrast.

Table 4. Results for each subject according to VOT calculations, permutation test, and interpretation of the acoustic analysis and transcriptions of /t/ and /d/ data. Column 1 lists the participants. Columns 2 through 5 give the mean VOT values and standard deviations (SD). The values for the significance level from the permutation tests are shown in the next column. The last three columns present the interpretations of the results according to whether a) the participant in question produced an acoustic distinction, b) the transcribers perceived a distinction, and c) the subject systematically produced no contrast, an overt contrast or a covert contrast.

Subject	VOT for /t/		VOT for /d/		Permutation test results p value	Differences between /t/ & /d/		
	Mean	SD	Mean	SD		Acoustic distinction	Transcription	Contrast?
6001	46	9	15	7	$p < 0.01$	Yes	Yes	Overt
6002	61	12	−52	34	$p < 0.01$	Yes	Yes	Overt
6003	68	18	12	2	$p < 0.01$	Yes	Yes	Overt
6004	112	18	−3	22	$p < 0.01$	Yes	Yes	Overt
6005	62	20	−1.6	37	$p < 0.01$	Yes	Yes	Overt
6006	68	24	3	22	$p < 0.01$	Yes	Yes	Overt
6007	50	23	19	9	$p = 0.04$	Yes	Yes	Overt
6008	79	9	31	10	$p < 0.01$	Yes	Yes	Overt
6009	54	21	14	8	$p < 0.01$	Yes	Yes	Overt

for /b/). Although this distinction for 6009 is statistically reliable, it is nevertheless small. Moreover, the permutation test result for 6009 is no longer significant after we apply the Bonferroni correction, which requires us to divide the alpha level of .05 by the number of null hypotheses we are testing, in this case, nine. This brings the significance level for each hypothesis test to .005, which means that the results for 6004 are statistically significant after we apply the correction, but those for 6009 are not. As a consequence, we are able to attest an L2 covert contrast between English /p/ and /b/ on the part of participant 6004.

To sum up this section briefly, the assistants transcribed the productions of five of the nine participants as not evidencing a distinction between /p/ and /b/. However, VOT measurements indicated that at least one of the five clearly produced a reliable distinction between /p/ and /b/. We now turn to the discussion of these findings.

5. Discussion

There are several implications of our findings that we would like to discuss. The first is that attesting a covert contrast among second-language learners brings research on the acquisition of L2 phonemic contrasts in line with findings from work on disordered speech and child-language acquisition. This, in turn, suggests that transcription data for L2 phonological analyses must be augmented, whenever feasible, by acoustic analyses. The same conclusion was reached some time ago for studies on L1 acquisition (Macken & Barton 1980) and disordered speech (Maxwell & Weismer 1982; Hewlett 1988). A second implication of our results is that the implementation of a covert contrast by L2 learners is predictable from the Interlanguage Hypothesis (ILH). Therefore, despite not being commonly reported in the L2 phonology literature, covert contrasts are to be expected, and have consequences for how the L2 acquisition of phonemic contrasts should be characterized. And third, we will conclude this section by taking up two questions for future research on covert contrasts, viz., what we would anticipate finding from longitudinal studies, and whether covert contrasts have implications for learners' ability to perceive the phonemic distinction in question.

The seminal research on covert contrasts in L1 acquisition was conducted by Macken & Barton (1980), who, as outlined above, carried out a longitudinal study on the acquisition of voice contrasts in word-initial stops by English-learning children. In measuring the VOT lags in the productions of their participants, the authors identified three stages in the acquisition of a voice contrast in stops, one of which involved the children making a systematic acoustic distinction that was not perceivable to native listeners. The significance of this finding is that it contributes to our understanding and characterization of language learning as a process that does not necessarily move in one step from a stage where no contrast is evidenced to one in which the distinction is fully implemented. Instead, the claim is that learners proceed through an intermediate step of a covert contrast, where the learner's developing system shows evidence of acquiring the relevant *function*, that is, the phonemic distinction, before acquiring the pertinent *form*, namely, the implementation of the contrast in a target-like way. The idea that language learners acquire functions before they learn forms is long-standing, and harks back in L1 acquisition at least to the chapters in Slobin (1985), and in L2 acquisition, to the work of Bardovi-Harlig (1992), among others.

Moreover, this result is to be expected, as it is consistent in terms of an acoustic dimension with the documentation of numerous intermediate stages of acquisition in other areas of the grammar. Thus, for example, it is well known that children acquire some phonemic contrasts in onsets before codas, that learners regularize aspects of morphology, such as irregular plurals, before they learn the

irregular forms, and that child learners systematically produce non-target-like forms of some sentence types, such as questions, before acquiring the adult forms.

Covert contrasts have also played an important role in characterizing learner systems in research on disordered speech. The study by Gierut and Dinnsen (1986) documented that the utterances of two children who produced identical errors on word-initial stops could be distinguished in terms of the learners' underlying grammar. Acoustic measures showed that one of the participants was in fact producing a covert contrast between the target phonemes in terms of VOT and closure duration, whereas the other participant was not. In addition, the authors argued that the child-grammar that evidenced the covert contrast could be shown to be much closer to the adult grammar, and could be placed on a trajectory of normal acquisition. These findings, in turn, had consequences for remediation strategies.

In the same vein, attesting a covert contrast in second-language acquisition stands to give us greater insight into the learner systems of L2 acquirers. An interesting question is whether learners who produce a covert contrast between TL phonemes also evidence concomitant differences in other areas of the target phonology. For example, one could investigate whether a second-language learner who instantiates a covert contrast between TL phonemes is better at perceiving the distinction in question compared to a learner who shows no evidence of the contrast. This is clearly a question for future research, and as we have no data to report on this matter at this time, we must leave the question open. The second implication of our findings that we wish to discuss is the fact that, if the ILH is defensible, we should expect to find learners with systems that have covert contrasts. The ILH makes the claim that all L2 learners internalize a system that enables them to speak and understand utterances in the TL, and that this system can be independent of both the learner's NL and the TL (Gass, Behney & Plonsky, 2013; Tarone, 2006). An IL system with a covert contrast is independent of both the NL and TL in that it is not NL-like, because the NL lacks the contrast, and it is not TL-like, because the contrast is implemented in a way that is not perceived by native speakers of the TL.

The groundwork that would eventually alter the focus of much research on L2 acquisition was laid in the early 1970s with the proposal, made independently by three scholars (Corder, 1971; Nemser, 1971; Selinker, 1972), that all L2 learners internalize a "learner language", that they employ to produce and understand utterances in the TL. The empirical claim embodied in the ILH is that there exist patterns produced by L2 learners that cannot be attributed to the NL because these patterns are not part of the NL grammar. They also cannot be due to the TL because they are not evidenced in the TL. However, because the L2 utterances in

question are systematic, they must derive from some underlying system, which is postulated to be the learner's IL.

Empirical evidence in support of the ILH has long been reported in the SLA literature, and can be shown in several ways. On the one hand, such supporting evidence can involve documenting an L2 regularity, such as a pattern of consonant clusters, that is not present in either the NL or TL. Support for the ILH can also derive from showing that some set of L2 facts can be explained only by postulating some construct, such as a rule or constraint, that is not motivated for either the NL or TL. Such findings have been reported in various domains over the years, including word-final devoicing (Altenberg & Vago 1983), consonant clusters in onsets and codas (Carlisle 1998), resumptive pronouns in relative clauses (Hyltenstam 1984), verb-second patterns in main clauses (Schwartz & Sprouse 2000), and wh-scope marking (Schulz 2011), among others. In each case the researchers showed that the learner's IL grammar differed systematically from both the NL and TL. Similarly, our results show that the IL phonologies of participants 6004 and 6009, specifically the system of phonemic contrasts, differ systematically from their NL, Arabic, and from the TL, English. This finding suggests that the TL system of contrast, specifically, the TL /p/-/b/ contrast, which is not part of the NL, is being implemented in a way that is different from the TL. This result is consistent with the idea that IL grammars can be independent of both the NL and TL.

Finally, we wish to conclude this section by making the assertion, albeit speculatively, that a stage of covert contrast is not simply a possible stage in the acquisition of L2 phonemic distinctions, but that it is a necessary step in the acquisition process. This certainly seems reasonable to us, especially in view of the intermediate stages that have been documented in the acquisition of other grammatical constructions and concepts, as we mentioned above. Thus, we suggest that it is unlikely that an L2 learner proceeds from an IL grammar manifesting the absence of a contrast to an IL fully implementing the distinction in a single step. The problem with this position, of course, is that our current state of knowledge makes it impossible to test this claim empirically, as there could well be L2 learners who, though they appeared not to exhibit a stage of covert contrast, they in fact did evince such an intermediate stage, but only for a very short time.

6. Conclusion

In this paper we have reported the results from our study of the L2 productions of the English /p/–/b/ contrast by native speakers of Arabic. Perhaps the most important implication of the finding that some of our participants produced a covert contrast between these segments is that the acquisition of second-language

phonemic contrasts seems to pattern in the same way as does the learning of contrasts in child-language acquisition and in disordered speech. This result, in turn, supports the more general conclusion that covert contrasts are likely part and parcel of all phonological learning.

Acknowledgement

This work from was supported in part for the first two authors by a grant from the National Institutes of Health R01 HD046908-05. The positions expressed in this paper are those of the authors, and do not necessarily reflect those of NIH. We wish to thank Hanyong Park for consulting with us on the statistics and to acknowledge the following for their role in assisting with the study: Carolyn Barry, Alana Dust, Bridget Foley, Alison Garcia, Heather Kohl, Kacy Kreger, Cady Pyne, and Valerie Stoffer. We are also grateful to four anonymous reviewers whose comments and suggestions have helped shape the final version of this paper. Finally, any errors, omissions or inconsistencies are the fault of the authors.

Bibliography

Altenberg, E., & Vago, R. (1983). Theoretical implications of an error analysis of second language phonology production. *Language Learning*, 33, 427–447. DOI: 10.1111/j.1467-1770.1983.tb00943.x

Bardovi-Harlig, K. (1987). Markedness and salience in second-language acquisition. *Language Learning*, 37, 385–407. DOI: 10.1111/j.1467-1770.1987.tb00577.x

Bardovi-Harlig, K. (1992). The relationship of form and meaning: A cross-sectional study of tense and aspect in the interlanguage of learners of English as a second language. *Applied Psycholinguistics*, 13, 253–278. DOI: 10.1017/S0142716400005634

Bellugi, U. (1971). Simplification in children's language. In R. Huxley & D. Ingram (Eds.), *Methods and models in language acquisition* (pp. 95–119). New York: Academic Press.

Berko, J. (1958). The child's learning of English morphology. *Word*, 15, 150–177.

Berti, L.C. (2010). Contrast and covert contrast in the speech production of children. *Pro-fono Revista de Atualizacao Cientifica*, 22, 531–536. DOI: 10.1590/S0104-56872010000400029

Best, C. 1995. A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171–204). Baltimore: York Press.

Bialystok, E. (1987). Words as things: Development of word concept by bilingual children. *Studies in Second Language Acquisition*, 9, 133–140. DOI: 10.1017/S0272263100000437

Boersma, P., & Weenink, D. (2005). PRAAT: Doing phonetics by computer (Version 4.4.07) [Computer program]. Available at http://www.praat.org/.

Bohn, O –S., & Flege, J.E. (1992). The production of new and similar vowels by adult German learners of English. *Studies in Second Language Acquisition*, 14, 131–158. DOI: 10.1017/S0272263100010792

Broselow, E., Chen, S., & Wang, C. (1998). The emergence of the unmarked. *Studies in Second Language Acquisition*, 20, 261–280. DOI: 10.1017/S0272263198002071

Brown, R. (1970). *A first language*. Cambridge, MA: Harvard University Press.

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

Carlisle, R.S. (1998). The acquisition of onsets in a markedness relationship: A longitudinal study. *Studies in Second Language Acquisition*, 20, 245–260. DOI: 10.1017/S027226319800206X

Corder, S. P. 1971. Idiosyncratic dialects and error analysis. *IRAL* 9, 149–159.

Dulay, H., & Burt, M. (1973). Should we teach children syntax. *Language Learning*, 23, 245–258. DOI: 10.1111/j.1467-1770.1973.tb00659.x

Eckman, F. (1981). On the naturalness of interlanguage phonological rules. *Language Learning*, 31, 195–216. DOI: 10.1111/j.1467-1770.1981.tb01379.x

Eckman, F., & Iverson, G. (1994). Pronunciation difficulties in ESL: Coda consonants in English interlanguage. In M. Yavas (Ed.), *First and second language phonology* (pp. 251–266). San Diego: Singular Publishing Group, Inc.

Eckman, F.R., Iverson, G.K., & Song, J.Y. (2014). Covert contrast in the acquisition of second language phonology. In A. Farris-Trimble & J. Barlow (Eds.), *Perspectives on phonological theory and acquisition: Papers in honor of Daniel A. Dinnsen* (pp. 25–48). Amsterdam: John Benjamins. DOI: 10.1075/lald.56.05eck

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.

Flege, J.E. (1990). English vowel production by Dutch talkers: More evidence for the "similar" versus "new" distinction. In J. Leather & A. James (Eds.), *New Sounds 90: Proceedings of the Amsterdam Symposium on the Acquisition of Second-language Speech* (pp. 255–293). Amsterdam: University of Amsterdam.

Flege, J.E. (1991). Age of learning affects the authenticity of voice onset time (VOT) in stop consonants produced in a second language. *Journal of the Acoustical Society of America*, 89, 395–411. DOI: 10.1121/1.400473

Flege, J.E. (1993). Production and perception of a novel, second-language phonetic contrast. *Journal of the Acoustical Society of America*, 93, 1589–1608. DOI: 10.1121/1.406818

Flege, J.E. (1995). Second language speech learning: Theory, findings and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). Baltimore: York Press.

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. DOI: 10.3109/02699209008985495

Gass, S., Behney, J., & Plonsky, L. (2013). *Second language acquisition: An introductory course*. New York: Routledge.

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

Good, P. (2000). *Permutation tests: A Practical guide to resampling methods for testing hypotheses*. Heidelberg: Springer.

Hammerly, H. (1982). Contrastive phonology and error analysis. *IRAL*, 20, 17–32. DOI: 10.1515/iral.1982.20.1-4.17

Hewlett, N. (1988). Acoustic properties of /k/ and /t/ in normal and phonologically disordered speech. *Clinical Linguistics and Phonetics*, 2, 29–45. DOI: 10.3109/02699208808985242

Hyltenstam, K. (1984).The use of typological markedness conditions as predictors in second language acquisition: The case of pronominal copies in relative clauses. In R. Andersen (Ed.), *Second languages* (pp. 39–58). Rowley, Mass: Newbury House Publishers.

Iverson, G., & Salmons, J. (1995). Aspiration and laryngeal representation in Germanic. *Phonology*, 12, 369–396. DOI: 10.1017/S0952675700002566

Klima, E., & Bellugi, U. (1966). Syntactic regularities in the speech of children. In J. Lyons & R. Wells (Eds.), *Psycholinguistic papers* (pp. 183–208). Edinburgh, UK: University of Edinburgh Press.

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* (pp. 210–219). Chicago Linguistic Society.

Leather, J., & James, A. (1991). The acquisition of second-language speech. *Studies in Second Language Acquisition*, 13, 305–341. DOI: 10.1017/S0272263100010019

Liberman, A.M., Harris, K.S., Hoffman, H.S., & Griffith, B.C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, 54(5), 358–368. DOI: 10.1037/h0044417

Lim, J., & Oh, M. (2008). Covert contrast in second language acquisition. *Proceedings of the 16th International Conference on Korean Linguistics* (pp. 202–203). International Circle of Korean Linguistics. New York: SUNY at Binghamton & Cornell University.

Lisker, L., & Abrahamson, A. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20, 527–565.

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. DOI: 10.1111/j.1467-1770.1994.tb00634.x

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. DOI: 10.1017/S0142716400004148

Morgan, J. (1987). Structural packaging in the input to language learning: Contributions of prosodic and morphological marking of phrases to the acquisition of language. *Cognitive Psychology*, 19, 498–550. DOI: 10.1016/0010-0285(87)90017-X

Munro, M.J., & Derwing, T. (1997). Accent, intelligibility and comprehensibility: evidence from four L1s. *Studies in Second Language Acquisition*, 19, 1–16.

Munson, B., Edwards, J., Schellinger, S.K., Beckman, M.E., & Meyer, M.K. (2010). Deconstructing phonetic transcription: Covert contrast, perceptual bias, and an extraterrestrial view of Vox Humana. *Clinical Linguistics and Phonetics*, 24, 245–260. DOI: 10.3109/02699200903532524

Nemser, W. 1971. Approximative systems of foreign language learners. *IRAL* 9,115–123.

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

R Foundation for Statistical Computing. (2009). R (Version 3.1.1) [Computer software]. Retrieved from http://www.r-project.org/.

Ritchie, W. (1968). On the explanation of phonic interference. *Language Learning*, 18, 183–197. DOI: 10.1111/j.1467-1770.1968.tb00206.x

Schulz, B. (2011). Syntactic creativity in second language English: wh-scope marking in Japanese-English interlanguage. *Second Language Research*, 27, 313–341. DOI: 10.1177/0267658310390503

Schwartz, B., & Sprouse, R. (2000). When syntactic theories evolve: Consequences for L2 acquisition research. In J. Archibald (Ed.), *Second language acquisition and linguistic theory* (pp. 156–186). Malden, MA: Blackwell Publishers.

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

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

Selinker, L. 1972. Interlanguage. *IRAL* 10, 209–231.

Slobin, D., (Ed.). (1985). *The cross-linguistic study of language acquisition, volumes 1 & 2: Theoretical issues*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Smith, B.L. (1979). A phonetic analysis of consonant devoicing in children's speech, *Journal of Child Language*, 6, 19–28. DOI: 10.1017/S0305000900007595

Tarone, E. (2006). Interlanguage. In K. Brown (Ed.), *Encyclopedia of language and linguistics*, (2nd ed., vol. 5) (pp. 747–752). Oxford: Elsevier.

Tyler, A., & Nagy, W. (1989). The acquisition of English derivational morphology. *Journal of Memory And Language*, 28(6), 649–667. DOI: 10.1016/0749-596X(89)90002-8

Wode, H. (1981). *Second language learning 1: An integrated view of language acquisition*. Tübingen: Gunter Narr.

Yeni-Komshian, G., Caramazza, A., & Preston, M. (1977). A study of voicing in Lebanese Arabic. *Journal of Phonetics*, 5, 35–48.

Appendix. Words used to elicit the /p/–/b/, /k/–/g/ & /t/–/d/ contrasts

1. peek	17. cab	33.table	49. bay
2. pole	18. cub	34. time	50. bowl
3. pay	19. cob	35. tear	51. bee
4. pier	20. cake	36. tea	52. beak
5. pig	21. cop	37. two	53. big
6. pear	22. cough	38. feet	54. bear
7. pea	23. cup	39. wheat	55. bye
8. pull	24. kick	40. seat	56. beer
9. pie	25. cap	41. meat	57. bull
10. pin	26. truck	42. gate	58. bin
11. mop	27. leak	43. hat	59. bed
12. cup	28. cheek	44. golf	60. fog
13. cap	29. gab	45. gate	61 dog
14. gap	30. give	46. leg	62. hug
15. dime	31. deer	47. door	63. desk
16. seed	32. feed	48. head	64. cloud

Authors' addresses

Prof. Dr. Fred R. Eckman
Department of Linguistics
University of Wisconsin
UWM Johnston Hall 117
Milwaukee, 53021
Wisconsin
USA

eckman@uwm.edu

Prof. Gregory K. Iverson
University of Wisconsin-Milwaukee
Dept. of Foreign Langs. & Lings.
P.O. Box 413
Milwaukee, WI 53201-0413
USA

iverson@uwm.edu

Jae Yung Song
University of Wisconsin–Milwaukee
Department of Communication
Sciences & Disorders
2400 E. Hartford Ave. Enderis Hall
Room B30
Milwaukee, WI 53211
USA

songjy@uwm.edu