



NIH Public Access

Author Manuscript

J Speech Lang Hear Res. Author manuscript; available in PMC 2009 November 25.

Published in final edited form as:

J Speech Lang Hear Res. 2008 April ; 51(2): 321–332. doi:10.1044/1092-4388(2008/024).

Physiological Indices of Bilingualism: Oral–Motor Coordination and Speech Rate in Bengali–English Speakers

Rahul Chakraborty, Lisa Goffman, and Anne Smith

Purdue University

Abstract

Purpose—To examine how age of immersion and proficiency in a 2nd language influence speech movement variability and speaking rate in both a 1st language and a 2nd language.

Method—A group of 21 Bengali–English bilingual speakers participated. Lip and jaw movements were recorded. For all 21 speakers, lip movement variability was assessed based on productions of Bengali (L1; 1st language) and English (L2; 2nd language) sentences. For analyses related to the influence of L2 proficiency on speech production processes, participants were sorted into low- ($n = 7$) and high-proficiency ($n = 7$) groups. Lip movement variability and speech rate were evaluated for both of these groups across L1 and L2 sentences.

Results—Surprisingly, adult bilingual speakers produced equally consistent speech movement patterns in their production of L1 and L2. When groups were sorted according to proficiency, highly proficient speakers were marginally more variable in their L1. In addition, there were some phoneme-specific effects, most markedly that segments not shared by both languages were treated differently in production. Consistent with previous studies, movement durations were longer for less proficient speakers in both L1 and L2.

Interpretation—In contrast to those of child learners, the speech motor systems of adult L2 speakers show a high degree of consistency. Such lack of variability presumably contributes to protracted difficulties with acquiring nativelike pronunciation in L2. The proficiency results suggest bidirectional interactions across L1 and L2, which is consistent with hypotheses regarding interference and the sharing of phonological space. A slower speech rate in less proficient speakers implies that there are increased task demands on speech production processes.

Keywords

bilingual; speech motor control; variability; Bengali

Changes in movement parameters, such as variability and duration, have been used to evaluate the maturation of the speech production system (e.g., Sharkey & Folkins, 1985; A. Smith & Goffman, 1998; A. Smith & Zelaznik, 2004; B. L. Smith, 1978). A robust finding in all of these studies has been that both variability and duration of movement decrease with development. More recently, it has also been observed that language variables interact with speech production processes in complex ways. For example, speech motor variability for a short phrase embedded in a sentence increases as a function of changes in length and syntactic complexity for children, but not for adult speakers (Kleinow & Smith, 2000; Maner, Smith, & Grayson, 2000). Such results suggest that there are bidirectional interactions between motor parameters, such as variability and duration, and cognitive–linguistic processing (A. Smith & Goffman, 2004).

Adults who are second language (L2) learners present an interesting test case of such language–motor interactions, because they are at an advanced level of motor and language learning in the first language (L1) but lack equivalent experience in L2. To our knowledge, very few investigators (Flege, 1988; Zsig, 2003) have explored the nature of speech motor control in bilingual speakers. In this article, we evaluate how speech motor variability and rate relate to language proficiency in a group of adult second language learners.

In the L2 learning literature, there are competing hypotheses about how L1 and L2 interact; for example, the language subsystems may be shared across L1 and L2 (Paradis, 1993) or may be independent (De Bot, 1992). Regardless of how individual language components interact, at the level of speech implementation, later learners show significant differences from early learners as evidenced by increased interference of native phonetic and, perhaps, other linguistic components. For example, age of L2 exposure influences production accuracy (Asher & Garcia, 1969; Flege & Fletcher, 1992; Flege, Munro, & MacKay, 1995; Flege, Yeni-Komshian, & Liu, 1999; Long, 1990; Moyer, 1999; Patkowski, 1990) and acoustic duration (Munro & Derwing, 1995; Riggenbach, 1991).

Individuals who acquire a second language early rather than late in life exhibit relatively nativelike pronunciation (Asher & Garcia, 1969; Long, 1990; Flege & Fletcher, 1992; Flege, Munro, & MacKay, 1995; Flege et al., 1999; Moyer, 1999; Patkowski, 1990). Thus, the presence of a detectable nonnative accent is considered evidence of L1 influence on L2 (Flege et al., 1995, 1999; Piske, MacKay, & Flege, 2001). It may be hypothesized that the more established L1 is at the onset of L2 acquisition, the greater the influence it will exert on L2. However, it is not known where in the speech production process these language-specific differences in implementation originate, even though it is clear that speech production characteristics, such as foreign accent, arise in part through the implementation process.

Variability as an Index of Language Experience

Speech movement variability is defined, for the present purposes, as the spatial and temporal consistency of a set of movement trajectories from the repeated performance of a task (Ackermann, Hertrich, & Scharf, 1995; Adams, Weismar, & Kent, 1993; Bernstein, 1967; Ostry, Cooke, & Munhall, 1987; A. Smith & Goffman, 1998; Smith, Johnson, McGillem, & Goffman, 2000). Although movement variability is a complex construct, it is often viewed as an indicator of adaptive flexibility, with decreased consistency observed during the acquisition of a new behavior (A. Smith & Goffman, 1998; A. Smith & Zelaznik, 2004; Thelen & Smith, 1994). Such flexibility is also seen in developmental studies, which reveal that young children are more variable than adults across a range of speech production tasks (Goffman, 1999; A. Smith & Zelaznik, 2004; Walsh & Smith, 2002). Systematic changes in variability continue through adolescence (A. Smith & Zelaznik, 2004; Walsh & Smith, 2002), with increased consistency attributed to both language and motor experience. According to Thelen and Smith, variability could be considered an adaptive mechanism that allows individuals the flexibility to modify strategies to suit environmental demands. Similar observations about the relation of variability to language experience may be extended to L2 learning in adults.

Within individual speakers, specific aspects of language have also been found to have differential effects on movement variability. For example, in the phonological domain, monolingual English-speaking adults and children show greater spatiotemporal variability in the production of trochaic (i.e., strong–weak) than iambic (i.e., weak–strong) words (Goffman, 1999; Goffman & Malin, 1999). It has also been reported that increases in sentence length and syntactic complexity lead to commensurate increases in the spatiotemporal variability in single articulator trajectories (Maner et al., 2000; Sadagopan & Smith, in press). These results

demonstrate that spatiotemporal movement variability is sensitive not only to language experience but also to the specific dimensions of language that an individual is producing.

In earlier experiments examining variability in sets of speech movement trajectories, movements of a single articulator were analyzed (A. Smith & Goffman, 1998; A. Smith, Goffman, Zelaznik, Ying, & McGillem, 1995; Walsh & Smith, 2002). However, essentially the same analysis can be applied to difference trajectories, such as the lip aperture signal (upper lip–lower lip movement; A. Smith & Zelaznik, 2004). In this case, the variability index reflects the trial-to-trial consistency in interarticulator coordination. Interestingly, the index of coordinative stability among three articulators (upper lip, lower lip, and jaw) shows greater consistency than that of a single component (upper lip) or that of the lower lip/jaw complex (A. Smith & Zelaznik, 2004; Walsh & Smith, 2002). *Lip aperture* is defined as the dynamically varying distance between the upper and lower lips. The lip aperture variability measure clearly decreases with maturation and provides a sensitive index of oral–motor coordination for speech. Given that dynamic variation in lip aperture has important effects on the speech acoustic output (A. Smith & Zelaznik, 2004), we elected to use this measure in the present study.

Speech Rate as an Index of Language Experience

The second aspect of movement implementation hypothesized to be influenced by language experience is speaking rate, inferred in the present study from the duration of the entire movement sequence for the sentence. Slower speech rates have been argued to relate to increased processing load, which may be related to reduced language experience (Munro & Derwing, 1995; Rigganbach, 1991; B. L. Smith, 1978; Tingley & Allen, 1975) as well as a number of other cognitive linguistic or motor demands (Dromey & Benson, 2003). As was the case with variability, speech rate increases with maturation in children (Kent & Forner, 1980; A. Smith & Goffman, 1998; Smith & Zelaznik, 2004; B. L. Smith, 1978; Tingley & Allen, 1975; Walsh & Smith, 2002).

Speech rate has also been examined in L2 learners. Munro and Derwing (1995) reported that when Mandarin speakers who had been in Canada for 4 years produced English, their utterances were significantly longer than those of native speakers of English. Rigganbach (1991) also reported such rate differences and suggested that L2 speaking rate was faster for more advanced L2 learners. It is important that these reported results are on global measures of overall utterance duration rather than on measures of internal, shorter segments (e.g., consonant or vowel durations). Timing is a complex phenomenon, and many articulatory events lead to acoustic changes in duration (Fukaya & Byrd, 2005).

In summary, the purpose of this experiment is to examine how, at the movement implementation, or “articulator” level (De Bot, 1992), age of exposure and proficiency in L2 influence movement variability and speech rate. Our predictions are as follows:

1. For all speakers, oral movement consistency is expected to be greater in L1 (Bengali) than in L2 (English) utterances, reflecting more proficient production capacities in L1.
2. Bilingual speakers who were exposed to L2 early and are more proficient in L2 will have greater consistency of coordination in L2 production and will show higher speech rates compared with their less experienced adult peers.

Method

Participants

Twenty-one Bengali–English speakers (referred to as *bilinguals*) between the ages of 23 and 42 years participated in this study. Standard Bengali was the first language of all of the Bengali–English speakers, and they had lived in the United States for a range of 8 months to 10 years. All of the bilinguals were born and brought up in the same geographical region of Calcutta and used the same urban Bengali dialect. Thus, Bengali was the primary language of communication, and the participants lived in India for 22–32 years. For all bilingual participants, both parents were native Bengali speakers. Participant selection was based on their medium of instruction at the primary and secondary school level. Of the 21 Bengali–English speakers, 11 received their primary and secondary education in schools where the primary medium of instruction was English, and 10 received their primary and secondary education in schools where the primary medium of instruction was Bengali. Therefore, on the basis of the age of academic immersion in English, 11 participants were exposed to English early (from the primary school level), and the remaining 10 participants were exposed to English late (only at or after the college level). All 21 participants received their graduate education in English. At the time of this study, all speakers were part of an active Bengali community in West Lafayette, Indiana, and spoke Bengali on a regular basis. All participants reported no history of speech, language, reading, or neurological problems and also passed a pure-tone hearing screening at 20 dB at 0.5 kHz, 1.0 kHz, 2.0 kHz, 4.0 kHz, and 6.0 kHz.

To measure language proficiency (Guion, 2005), all participants were given the Speaking Grammar and Listening Grammar subtests of the Test of Adolescent and Adult Language—Third Edition (TOAL-3; Hammill, Brown, Larsen, & Wiederholt, 1994). These subtests both require knowledge of complex English syntax. The Speaking Grammar subtest is a sentence imitation task (e.g., “The boy who won the race is a friend of mine and helped us coach the younger runners”). For the Listening Grammar subtest, participants had to choose which two of three sentences mean the same thing (e.g., “Should Vickie like it, we will buy it,” “We will buy it only if Vickie likes it,” or “If we buy it, Vickie should like it”). Because this test was not standardized on bilingual speakers, only raw scores are reported. For the 11 participants exposed to English early, the Listening Grammar scores ranged from 21 to 33 ($M = 28.8$, $SD = 4.5$). Speaking Grammar scores for this group ranged from 16 to 24 ($M = 19.7$, $SD = 2.8$). The 10 speakers in the late exposed group showed Listening Grammar scores in the 11–31 range ($M = 20.1$, $SD = 7.8$). Speaking Grammar scores for the late group ranged from 4 to 17 ($M = 8.9$, $SD = 5.0$).

Because we wanted to obtain subgroups that clearly differed in L2 oral proficiency, we further classified the 21 participants on the basis of their oral English performance. As Figure 1 shows, the 21 participants were sorted into three groups (7 participants/group) on the basis of their performance on the Speaking Grammar section of the TOAL-3. The three groups were classified as having high English skills (scores between 20 and 24; $M = 21.5$, $SD = 1.4$), middle English skills (scores between 14 and 18; $M = 16.1$, $SD = 1.2$), and low English skills (scores between 4 and 9; $M = 6$, $SD = 1.9$). For statistical tests of more general differences between L1 and L2, data for all 21 participants were included. For tests of potential effects of time of L2 immersion and L2 proficiency on speech motor variability and speech rate, only data from the 7 early/high and 7 late/low participants were included. This strategy seemed appropriate to increase the probability of observing the predicted effects of these two variables.

It should be noted that in previous studies, proficiency groups have often been sorted on the basis of age of immersion in a second language, often indexed by age of arrival (e.g., Flege, 1995, 1998; Weber-Fox & Neville, 1996). However, in many environments, including India, the nature of bilingualism is more complex. Individuals are exposed to multiple languages,

often from early childhood. However, their bilingualism is not balanced—that is, they are not similarly dominant in L1 and L2. In the present study, all speakers ranked Bengali as their most proficient language. Age of initial exposure to L2 was similar across individuals, but the degree of that exposure was not. Because of this acquisition scenario, a different approach needs to be taken in analyses of proficiency. It is insufficient to merely judge using the conventional criteria associated with age of arrival. Therefore, we took the conservative approach of using converging evidence from a laboratory measure of language proficiency and reported age of academic immersion in our assignment to proficiency groups.

Sentence Stimuli

Stimuli were three English sentences and three Bengali sentences. The three English sentences were as follows: “Buy Bobby a puppy,” “Mommy bakes pot pies,” and “Vicky feeds the fat pony.” The first two sentences were used in the A. Smith and Zelaznik (2004) study of 180 children and adults. These were selected for the present investigation because a large normative database on monolingual English speakers is available. The analogous Bengali sentences were as follows: “Baba pepe khabe” ([baba pepe k^habe]; Daddy will eat papaya), “Papur ma marbe” ([papur ma marbel]; Papu’s mother will punish him), and “P^honi p^hul b^halobashe” ([p^honi p^hul b^halobashe]; Phoni loves flowers). Bengali sentences were constructed such that each had one English comparator. For example, “Buy Bobby a puppy” and “Baba pepe khabe” both have six syllables and five bilabial stops (three voiced and two voiceless). “Mommy bakes pot pies” and “Papur ma marbe” both have five syllables and five bilabial stops (two nasals, two voiceless orals, and one voiced oral). The third pair of sentences was designed to include sounds that were native to one language but not the other. Bengali does not have labiodental fricatives (i.e., /f/ and /v/). For English, bilabial aspirated voiced and voiceless stops (i.e., /b^h/ and /p^h/) are not contrastive. The language-specific English sentence was “Vicky feeds the fat pony.” In Bengali, the sentence “P^honi p^hul b^halobashe” did not share all segmental contrasts with English. Bengali and English stimulus sentences were recorded by native speakers of the respective languages for auditory presentation.

Apparatus

For the kinematic data collection, each participant sat on a chair, 6 ft from a Northern Digital (Waterloo, Ontario, Canada) Optotrak 3020 three-camera system that records movement in three dimensions. This system tracks movement by sensing the location of small (approximately 7-mm) infrared light-emitting diodes (IREDs) and then recording their positional changes (A. Smith et al., 2000). Eight IREDs were used. Four IREDs were attached to the forehead, upper lip, lower lip, and jaw at the midline by using medical adhesive. The forehead, upper lip, and lower lip IREDs were attached to the skin surface, and the jaw IRED was attached at midline to a small, lightweight, L-shaped splint projected inferiorly under the chin (A. Smith et al., 2000). The remaining four IREDS were mounted to modified sports goggles. These four IREDS were used to form a reference frame for the head motion. Data from these four reference IREDS were used to compute the three-dimensional axes for the head. Motions of the upper lip, lower lip, and jaw were calculated using this head coordinate system as a reference. This process corrected for head motion artifact (A. Smith et al., 2000). Only the superior–inferior dimension of movement was analyzed, as this is the primary dimension for the speech sounds included in these stimuli. The sampling rate for the kinematic data was 250 Hz.

The speech acoustic signal was digitized at 8 kHz by the Optotrak Data Acquisition Unit. This signal provided an acoustic record that was synchronized to the movement data. The acoustic signal was used offline to ensure that the kinematic records corresponded with the appropriate speech sample. A separate video recording was obtained to judge participants’ productions, and only fluent sentences were included in the analysis. Because many of these speakers were

late L2 learners, stable substitution errors or allophonic variants were included in the analysis. When an utterance included a disfluency or an unstable articulation error, the next consecutive utterance was selected. In the English sentences, 52 (8%) sentences were excluded due to disfluency, whereas in Bengali, 11 (2%) were excluded. A smaller number of articulatory errors occurred, with a total of 3 (0.5%) in English and 16 (2%) in Bengali sentences.

Data Recording and Procedure

Each experimental session lasted approximately 1 hr, which also included hearing screening and administration of the two subtests of the TOAL-3. The session was divided into two blocks. In the first block the target language was Bengali, and a native Bengali speaker who was exposed to English relatively late interacted with the participants. The target language was English in the second block, and a monolingual English speaker conducted the session. The experimental stimuli were also recorded by two native speakers: one a Bengali speaking late learner of English, and the second a monolingual English speaker. To minimize the potential influence of L2 on L1, the order of the sessions was fixed with the Bengali session always occurring first.

After attaching the IREDS, the first block started. To establish the context of the targeted language for each block, reading and speaking tasks in the target language preceded the data collection. Participants initially read aloud a short paragraph in the target language (e.g., in Bengali, a 1,000-word passage taken from a popular Bengali newspaper; in English, the “Grandfather” passage), followed by answering a general question (e.g., “What are your future plans?” in Bengali or “What are the similarities and differences between the Indian and American educational systems?” in English). For the experimental stimuli, each participant was instructed to listen to the target sentence carefully and then to repeat it verbatim, using his or her normal rate, loudness, and accent pattern. The three sentences were presented in a randomized order through a loudspeaker at a comfortable loudness level. After collecting 10 fluent tokens for each of the three target sentences for that language, each block ended.

Data Analysis

Kinematic signals from upper lip, lower lip, and jaw were imported with the accompanying speech acoustic signal into a signal processing program written in MATLAB (The Mathworks, 1994). All kinematic signals were low-pass filtered (10 Hz cutoff) in forward and backward directions using a Butterworth filter. The three-point difference method was used to compute velocity (Winter, 1979). To segment the data for each sentence, continuous displacement signals from the lower lip marker and the associated velocity record were displayed.

Waveforms corresponding to the target stimuli were extracted from each displacement file. Onsets and offsets were extracted using peak velocity for initial and final labial consonants (A. Smith et al., 1995). For example, as shown in Figure 2, the movement onset in “Baba pepe khabe” was peak velocity of lower lip opening for “Ba” in “Baba,” and the offset was peak velocity opening for “be” in “khabe.” The onset and offset points were determined from the lower lip marker only, and these indices were used to segment time-locked data from the upper lip and jaw signals. To ensure that the kinematic data segmented for analysis corresponded to the target speech sample, the onset and offset points were used to play back the speech acoustic signal from that time interval. Sentence duration was determined as the total duration of the lower lip movement signal for each of the 10 utterances in the Bengali and English conditions. For each utterance type, there were 10 displacement-time signals extracted for the upper lip, lower lip, and jaw. To derive the lip aperture signal, the lower lip signal was subtracted from the upper lip signal. This difference signal reflects the coordination of upper lip, lower lip, and jaw to produce a target oral opening over time. It is assumed to reflect the action of a higher order functional synergy to achieve a “dynamically targeted lip aperture” (A. Smith & Zelaznik, 2004).

Using the lip aperture signal, a variability index was computed for each of the participants and for each sentence. This index is referred to as *lip aperture variability* and serves as a measure of consistency across repeated task performance (A. Smith & Zelaznik, 2004). To compute the lip aperture variability index, 10 lip aperture signals for each condition were amplitude and time normalized (see Panel B of Figure 3 for an illustration of normalized data). For each lip aperture record, amplitude normalization was accomplished by subtracting the mean of the record and then dividing by the standard deviation (A. Smith et al., 1995). For linear time normalization, a cubic spline procedure was used to interpolate each displacement record onto a constant length axis of 1,000 points (A. Smith et al., 2000). The standard deviation was computed at 2% intervals across the set of 10 normalized waveforms for each condition. As shown in Panel C of Figure 3, the resultant 50 standard deviations were summed to determine the variability of lip aperture separately for each participant in each condition. The lip aperture variability index reflects the underlying movement patterns in the absence of linear changes in amplitude and duration, whereas any nonlinear variations remain and contribute to the variability index. Overall, if the set of 10 normalized difference signals produced for the target sentence is highly convergent, the variability index will be low for lip aperture, suggesting that the speaker has a highly consistent underlying coordination system. Time and amplitude normalization remove the effects of absolute differences in loudness and speech rate and allow the underlying spatiotemporal patterning to emerge.

Results

Analysis Design

The first analysis included all 21 participants and asked explicitly about speech movement variability of Bengali versus English. In this case, a within-participant analysis of variance (ANOVA) was performed, with language (Bengali vs. English) and sentence (three for each language) as the within-participant variables. The second set of analyses concerned the effects of language proficiency on variability and speech rate. Because of the complex language learning environment in India, a conservative approach was adopted. Only speakers who showed both high proficiency on the TOAL-3 and early exposure to English were included in the early/high-proficiency group. In contrast, only those who performed poorly on the TOAL-3 and who were exposed to English late were categorized as late/low proficiency. In this case, the between-groups variable was proficiency level (early/ high vs. late/low), and the within-group variables were language (Bengali vs. English) and sentence. The three different Bengali and English sentences were nested within language.

Comparison of Variability Across English and Bengali

All 21 participants were included in this analysis. Figure 4 shows lip aperture variability values for the three English sentences and the three Bengali sentences. As illustrated, there was no main effect of language, $F(1, 20) = 0.50, p = .49$, and of sentence, $F(2, 40) = 2.07, p = .14$, on lip aperture variability. Native Bengali speakers were not more variable in L2 sentences. The Language \times Sentence interaction was significant, $F(2, 40) = 5.04, p = .01$. Post hoc tests using Tukey's honestly significant difference (HSD) procedure revealed that speakers were more variable for "P^honi p^hul b^halobashe" (bff) than for "Vicky feeds the fat pony" (evf) and also more variable for "P^honi p^hul b^halobashe" (bff) than for "Baba pepe khabe" (bbp).

Effects of Time of L2 Exposure and L2 Proficiency

Movement variability as a function of L2 proficiency—The lip aperture variability index for the early/high- and late/low-proficiency groups is presented in Figure 5. Early/high and late/low groups did not differ, $F(1, 12) = 1.11, p = .31$. In addition, the effect of language was not significant: English and Bengali sentences were produced with similar degrees of lip aperture consistency, $F(1, 12) = 0.39, p = .54$. There was a trend toward a Proficiency Group

× Language interaction, $F(1, 12) = 4.28, p = .06$, with the early/high group showing marginally less consistency in Bengali compared with the late/ low group. There was no main effect of sentence, $F(2, 24) = 2.58, p = .10$, but there was a Sentence × Proficiency Group interaction, $F(2, 24) = 8.68, p = .001$. Post hoc testing (the Tukey's HSD) showed that for the late/low group, the sentences with contrasts that are not native to both English and Bengali (i.e., "Vicky feeds the fat pony" and "P^honi p^hul b^halobashe") were more variable than those that included segments native to both Bengali and English. A three-way Language × Proficiency Group × Sentence interaction was also observed, $F(2, 24) = 3.58, p = .04$. Post hoc testing using the Tukey's HSD procedure revealed that the early/high group produced the sentence "P^honi p^hul b^halobashe" (LAVAR = 16.53) more variably than "Vicky feeds the fat pony" (LA VAR = 10.23). Although visual inspection suggests that there is a difference between proficiency groups in the production of the English sentence containing the nonnative segments (i.e., "Vicky feeds the fat pony"), the high degree of variability likely led to no significant effect in the post hoc.

Speech rate as a function of L2 proficiency—Mean durations for the three English and three Bengali stimuli for early/high and late/low groups are shown in Figure 6. As with the lip aperture variability analysis, a mixed ANOVA was performed. For speech rate, there was a significant effect of proficiency group, $F(1, 12) = 17.73, p = .001$, with early/high speakers showing a faster rate than late/low. Bengali and English sentences were produced at equivalent speech rates, $F(1, 12) = 2.76, p = .12$, with no Language × Proficiency Group interaction, $F(1, 12) = 0.26, p = .62$. Not surprisingly, because of inherent differences in length, there was a significant effect of sentence, $F(2, 24) = 125.74, p < .0001$. There was no Sentence × Proficiency Group interaction, $F(2, 24) = 2.61, p = .09$.

To determine whether speech rate variability was influenced by language proficiency in L1 or L2, the standard deviations of the movement durations were also assessed. There were no significant main effects or interactions: proficiency group, $F(1, 12) = .26, p = .62$; language, $F(1, 12) = 2.64, p = .13$; Language × Proficiency Group, $F(1, 12) = 0.42, p = .53$; sentence, $F(2, 24) = 0.12, p = .89$; and Sentence × Proficiency Group, $F(2, 24) = 1.08, p = .36$. Overall, movement duration variability was not affected by proficiency or L1 or L2 parameters.

Discussion

Variability in L1 Versus L2

There are no earlier direct kinematic studies of speech motor performance in bilingual speakers. Therefore, we had little empirical basis for predicting the outcomes of the present experiment. The first analysis was completed on the entire group of 21 participants to assess whether L2 learners as a whole showed differences in English versus Bengali utterances. We predicted that in L1, bilingual speakers would show more consistent coordinative patterns as compared with L2. Even though Bengali was the L1 for all participants, contrary to our expectations, in general, variability for Bengali sentences was not lower than for English sentences. Surprisingly, in their productions of the sentences "Buy Bobby a puppy" (BBAP) and "Mommy bakes pot pies" (MBPP), these bilingual speakers showed similar values of lip aperture variability to a large group of monolingual English speaking adults (A. Smith & Zelaznik, 2004). Specifically, lip aperture variability for BBAP was 12.44 ($SD = 2.91$) in the A. Smith and Zelaznik dataset and for MBPP was 11.08 ($SD = 2.28$). In the present study, the lip aperture variability values for the late/low-proficiency group was 11.17 ($SD = 1.91$) for BBAP and 12.96 ($SD = 2.70$) for MBPP. For the early/high group, lip aperture variability was 14.76 ($SD = 2.51$) for BBAP and 14.13 ($SD = 4.28$) for MBPP. Thus, the present L2 learners, even those who were less proficient English speakers, relied on highly stable motor patterns for their repeated productions of these sentences and were comparable in their consistency to monolingual English speakers. The

underlying speech motor systems of adult L2 speakers show decreased flexibility, which presumably influences their difficulties with reorganizing their speech movements to approximate nativelike pronunciation in L2.

For production of the Bengali sentences, these speakers demonstrated similar lip aperture variability values to their English correlates. It is important to note that although we were careful to select sentences that matched in segmental, syllabic, and lexical structure, there may be intrinsic differences (such as in prosodic structure or syntactic or lexical complexity) between Bengali sentences and their English analogs. Despite cross-language similarities and differences, it seemed that our participants had already developed a stable pattern of movement control that they used in both L1 and L2.

Clearly, many of our initial hypotheses were not supported. Perhaps this should not be surprising because although young children can easily acquire a second language with nativelike proficiency, adults—even when immersed for many years—often maintain nonnative characteristics, such as a pronounced accent. Electrophysiological evidence reveals that many aspects of neural organization for language differ in individuals immersed in a second language even in later childhood. For example, compared with earlier immersed L2 learners, in Chinese–English bilinguals exposed after age 7, there are differences in the latency and distribution of event-related potentials elicited in language processing tasks (Weber-Fox & Neville, 1996). The present findings are analogous in that adult second language learners do not flexibly alter their speech production in their L2. Unlike young children, adults use highly stable movement patterns for the production of accented speech. This is compatible with Thelen and Smith's (1996) observation that, in development, a loss of variability also leads to a loss of flexibility.

Variability as an Index of Language Proficiency

The analysis of the effects of proficiency on speech motor coordination included only those speakers categorized as early/high ($n = 7$) and late/low ($n = 7$) based on (a) age of academic immersion in English and (b) the syntactic production score on the TOAL-3. Based on the lack of significant main effects for proficiency level or language, it appears that both early/high and late/low adult L2 learners use stable, repeatable motor patterns in their production of L2. In their production of English sentences, these speakers are restricted to their already established movement routines and execute those routines with consistency.

Turning to L1—the Bengali productions—no differences were expected across proficiency groups. All speakers were immersed in and used Bengali while residing in India at least through young adulthood, and all continued to use Bengali in their daily lives in the United States. Surprisingly, we found a trend for the high-proficiency L2 speakers to be more variable in L1. Although this result was unexpected, it may be that a form of interference (e.g., Flege, 1987, 1988, 1995; Flege, Frieda, & Nowaza, 1997; Oyama, 1976) explains why L2 proficiency influences L1. The term *interference* suggests influence of one language on another (Weinreich, 1953). However, the nature, strength, and direction of such interference probably depends on how established and elaborated the two languages are within a speaker's linguistic repertoire, which in turn depends on factors such as age of L2 learning (e.g., Bailystok & Miller, 1999; Yeni-Komshian, Flege, & Liu, 2000) and overall usage of L1 and L2. Evidence of such linguistic interference has been reported in prosodic (Archibald, 1995; Weinreich, 1953), temporal (MacKay & Flege, 2004), and segmental (e.g., Flege, 1987, 1988, 1995) domains. In the present study, it appears that interference was observed at the movement implementation level, in that variability in L1 increased in the more proficient speakers of L2. The less proficient speakers did not show such an increase in coordinative variability in L1, suggesting differential effects as a function of language experience. It appears that factors related to L2 influence L1. This explanation seems more likely than one related to language attrition (Major, 1992) because

most of these speakers reported Bengali as their most proficient language (6 out of 7 in both early/high-proficiency and late/low-proficiency groups), and all continued to communicate regularly in Bengali. However, to further pursue this interesting relation, more specific contributions to observed changes in speech motor implementation, such as segmental or prosodic variables, need to be explicitly manipulated.

Other aspects of production were also influenced by degree of proficiency. More detailed analyses of interactions in the results suggested that sentences containing phonetic segments not shared across languages were treated differently by both early/high and late/low groups. A post hoc analysis of the Proficiency Group \times Sentence interaction revealed that, for the late/low group, the sentences containing segments not shared across Bengali and English (i.e., “Vicky feeds the fat pony” and “Phoni p^hul b^halobashe”) were produced with greater articulatory variability compared with the sentences that included only segments common to both languages. Thus, as is consistent with Flege’s (1999) notion of a “common phonological space,” those segments shared across languages were produced with relatively less variability. Flege further argued that the greater dissimilarity observed across L1 and L2, the more likelihood of the formation of a new category. These late/low-proficiency speakers produced the segments that are contrastive only in English (i.e., [f, v]) with relatively high articulatory variability. What is a bit more surprising in the present results is that L1 was also implicated in the possible formation of a new category—that is, sentences containing segments exclusive to Bengali were also produced more variably. Consistent with the interference hypothesis described previously, this suggests that there may be bidirectional reorganization as a learner acquires a new phonology; L1 and L2 may both be affected.

There was also a three-way Language \times Proficiency Group \times Sentence interaction. As in the prior case, this interaction also implicated the sentences containing nonshared segments, but this time for the early/high group. These more proficient speakers produced the Bengali sentence “Phoni p^hul b^halobashe” with more variability than the English sentence “Vicky feeds the fat pony.” The late/low group did not show this difference. A possible explanation for this result is that these speakers were especially sensitive to nonnative articulatory targets. Other studies have reported such decreased movement variability for targets that are relatively difficult to acquire and produce (Goffman, Heisler, & Chakraborty, 2006; Heisler, 2004). For example, English-speaking children and adults show more variability in their production of early developing and highly frequent trochees than of later developing and less frequent iambs (Goffman & Malin, 1999). Further study is needed to evaluate if increased linguistic constraints (e.g., native and nonnative changes in syntactic, prosodic, or segmental parameters) could lead to differences in coordinative variability as a function of language experience.

Speech Rate as an Index of Language Experience

Speech rates were computed as the overall duration of the movement sequence for each sentence. The early/ high-proficiency group members produced significantly faster speech rates than the late/low-proficiency group members for both L1 and L2. This relation between L2 speaking proficiency and speaking rate is consistent with prior research on bilingual speakers. For example, Munro and Derwing (1995) reported that their Mandarin participants (who lived in Canada for 4 years) spoke more slowly than native English participants. Guion, Flege, Liu, and Yeni-Komshian (2000) examined 240 native speakers of both Italian and Korean (selection was based on their age of arrival in Canada) and suggested that, as the age of learning L2 increases, the L2 speaking rate becomes slower.

From these studies, it has been inferred that processing load influences utterance duration. The present results suggest that a processing load explanation is overly simplistic. The early/high group members produced significantly shorter utterance durations not only in L2 (English) but also in L1 (Bengali). As previously discussed, it may be that these group differences relate to

complex and bidirectional interactions between L1 and L2 that differentially influence high- and low-proficiency speakers (Flege, Yeni-Komshian, & Liu, 1999; Fox, 1996; Weber-Fox & Neville, 1996). Bailystok and Miller (1999) suggested that L1's influence on L2 might be asymmetric, and that L1 might have a relatively stronger impact on late versus early L2 learners. Again, more data are needed to fully assess the nature of bidirectional interference (Flege, 1995, 1998, 1999; Yeni-Komshian et al., 2000) in bilinguals.

Most prior investigations have not considered L1 duration. However, in a recent study, MacKay and Flege (2004) demonstrated that early bilinguals (i.e., who were immersed in a native English-speaking community in childhood) produced shorter durations for L2 than for L1 utterances and that late bilinguals exhibited the opposite pattern. Our durational findings differ from such results probably because, first, all of our participants' age of arrival in a native English-speaking country was between 22 and 32 years. Second, unlike the usual age of arrival in a predominantly L2-speaking community, for our study, age of academic immersion in English was a criterion used to classify the two proficiency groups.

In summary, the present results provide new insights into how L1 and L2 interact in L2 learners. It is important that the nature of bilingualism in our participants was different from that generally studied, in that the course of the L2 exposure was domain specific but protracted. Clearly, bilinguals are not homogenous. The present results relate to a unique language learning environment. In India, individuals are often exposed to a particular dialect of English to varying degrees as part of the educational process. Thus, the present results need to be viewed in reference to specific attributes of Bengali–English language learning.

Acknowledgments

This research was supported by the National Institute on Deafness and Other Communication Disorders Grant DC04826. We are grateful to Janna Berlin, Laurence Leonard, Christine Weber-Fox, and Howard Zelaznik for invaluable assistance with many phases of this work.

References

- Ackermann H, Hertrich I, Scharf G. Kinematic analysis of lower lip movements in ataxic dysarthria. *Journal of Speech and Hearing Research* 1995;38:1252–1259. [PubMed: 8747818]
- Adams SG, Weismer G, Kent RD. Speaking rate and speech movement velocity profiles. *Journal of Speech and Hearing Research* 1993;36:41–54. [PubMed: 8450664]
- Archibald. Phonological acquisition and phonological theory. Hillsdale, NJ: Erlbaum; 1995.
- Asher JJ, Garcia R. The optimal age to learn a second language. *The Modern Language Journal* 1969;53:334–341.
- Bailystok E, Miller B. The problem of age in second language acquisition: Influence from language, structure, and task. *Bilingualism: Language and Cognition* 1999;2:127–145.
- Bernstein, NA. The coordination and regulation of movements. Oxford, England: Pergamon Press; 1967.
- De Bot K. A bilingual processing model: Levelt's “speaking” model adapted. *Applied Linguistics* 1992;13:1–24.
- Dromey C, Benson A. Effects of concurrent motor, linguistic, or cognitive tasks on speech motor performance. *Journal of Speech, Language, and Hearing Research* 2003;46:1234–1246.
- Flege JE. The production of “new” and “similar” phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics* 1987;15:47–65.
- Flege JE. The development of skill in producing word-final /p/ and /b/: Kinematic parameters. *Journal of the Acoustical Society of America* 1988;84:1639–1652. [PubMed: 3209769]
- Flege, JE. Second-language speech learning: Findings and problems. In: Strange, W., editor. *Speech perception and linguistic experience: Theoretical and methodological issues*. Timonium, MD: York Press; 1995. p. 233–273.

- Flege, JE. The role of subject and phonetic variables in L2 speech acquisition. In: Gruber, KD.; Higgins, D.; Olsen, K.; Wysocki, T., editors. *Papers for the 34th annual meeting of the Chicago Linguistic Society II: The panels*. Chicago: Chicago Linguistic Society; 1998. p. 101-132.
- Flege, JE. Age of learning and second language speech. In: Birdsong, D., editor. *Second language acquisition and the critical period hypothesis*. Mahwah, NJ: Erlbaum; 1999. p. 101-132.
- Flege JE, Fletcher KL. Talker and listener effects on degree of perceived foreign accent. *Journal of the Acoustical Society of America* 1992;91:370-389. [PubMed: 1737886]
- Flege J, Frieda E, Nowaza T. Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics* 1997;25:169-186.
- Flege J, Munro M, MacKay I. Factors affecting degree of perceived foreign accent in a second language. *Journal of the Acoustical Society of America* 1995;97:3125-3134. [PubMed: 7759652]
- Flege JE, Yeni-Komshian G, Liu H. Age constraints on second language acquisition. *Journal of Memory and Language* 1999;41:78-104.
- Fox E. Cross-language priming from ignored words: Evidence for a common representational system in bilinguals. *Journal of Memory and Language* 1996;35:353-370.
- Fukaya T, Byrd D. An articulatory examination of word-final flapping as phrase edges and interiors. *Journal of the International Phonetic Association* 2005;35:45-58.
- Goffman L. Prosodic influences on speech production in children with specific language impairment and speech deficits: Kinematic, acoustic, and transcription evidence. *Journal of Speech, Language, and Hearing Research* 1999;42:1499-1517.
- Goffman L, Heisler L, Chakraborty R. Mapping of prosodic structure onto words and phrases in children's and adults' speech production. *Language and Cognitive Processes* 2006;21:25-47.
- Goffman L, Malin C. Metrical effects on speech movements in children and adults. *Journal of Speech, Language, and Hearing Research* 1999;42:1003-1015.
- Guion SG. Knowledge of English word stress patterns in early and late Korean-English bilinguals. *Studies in Second Language Acquisition* 2005;27:503-533.
- Guion SG, Flege JE, Liu SH, Yeni-Komshian GH. Age of learning effects on the duration of sentences produced in a second language. *Applied Psycholinguistics* 2000;21(2):205-228.
- Hammill, DD.; Brown, VL.; Larsen, SC.; Wiederholt, JL. *Test of Adolescent and Adult Language*. Vol. 3. Austin, TX: Pro-Ed; 1994.
- Heisler, L. The influence of neighborhood density and phonotactic probability on representational and production factors in word learning. *Purdue University*; 2004. Unpublished doctoral dissertation
- Kent, RD. The biology of phonological development. In: Ferguson, CA.; Menn, L.; Stoel-Gammon, C., editors. *Phonological development: Models, research, implications*. Timonium, MD: York Press; 1992. p. 65-90.
- Kleinow J, Smith A. Kinematic correlates of speaking rate changes in stuttering and normally fluent adults. *Journal of Speech, Language, and Hearing Research* 2000;43:521-536.
- Long M. Maturational constraints on language development. *Studies in Second Language Acquisition* 1990;12:251-286.
- MacKay I, Flege J. Effects of the age of second language learning on the duration of the first and second language sentences: The role of suppression. *Applied Psycholinguistics* 2004;25:373-396.
- Major RC. Losing English as a first language. *Modern Language Journal* 1992;76:190-208.
- Maner K, Smith A, Grayson L. Influences of utterance length and complexity on speech motor performance in children and adults. *Journal of Speech, Language, and Hearing Research* 2000;43:560-573.
- Moyer A. Ultimate attainment in L2 phonology. *Studies in Second Language Acquisition* 1999;21:81-108.
- Munro M, Derwing T. Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech* 1995;38:389-406.
- Ostry DJ, Cooke JD, Munhall KG. Velocity curves of human arm and speech movements. *Experimental Brain Research* 1987;68:37-46.
- Oyama S. A sensitive period of the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research* 1976;5:261-283.

- Paradis M. Linguistic, psycholinguistic, and neurolinguistic aspects of interference in bilingual speakers: The activation threshold hypothesis. *International Journal of Psycholinguistics* 1993;9:133–145.
- Patkowski M. Age and accent in a second language: A reply to James Emil Flege. *Applied Linguistics* 1990;11:73–89.
- Piske T, MacKay I, Flege J. Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics* 2001;29:191–215.
- Riggenbach H. Toward an understanding of fluency: A microanalysis of nonnative speaker conversations. *Discourse Processes* 1991;14:423–441.
- Sadagopan N, Smith A. Developmental changes in the effects of utterance length and complexity of speech motor variability. *Journal of Speech, Language, and Hearing Research*. (in press)
- Sharkey SG, Folkins JW. Variability of lip and jaw movements in children and adults: Implications for the development of speech motor control. *Journal of Speech and Hearing Research* 1985;28:8–15. [PubMed: 3982001]
- Smith A, Goffman L. Stability and patterning of speech movement sequences in children and adults. *Journal of Speech, Language, and Hearing Research* 1998;41:18–30.
- Smith, A.; Goffman, L. Interaction of motor and language factors in the development of speech production. In: Massen, B.; Kent, R.; Peters, H.; Van Leieshout, P.; Hulstijn, W., editors. *Speech motor control in normal and disordered speech*. New York: Oxford University Press; 2004. p. 227–252.
- Smith A, Goffman L, Zelaznik H, Ying G, McGillem C. Spatiotemporal stability and patterning of speech movement sequences. *Experimental Brain Research* 1995;104:493–501.
- Smith A, Johnson M, McGillem C, Goffman L. On the assessment of stability and patterning of speech movements. *Journal of Speech, Language, and Hearing Research* 2000;43:277–286.
- Smith A, Zelaznik H. The development of functional synergies for speech motor coordination in childhood and adolescence. *Developmental Psychobiology* 2004;45:22–33. [PubMed: 15229873]
- Smith BL. Temporal aspects of English speech production: A developmental perspective. *Journal of Phonetics* 1978;6:37–67.
- The MathWorks. MATLAB (Version 6.1) [Computer software]. Natick, MA: Author; 1994.
- Thelen, E.; Smith, L. A dynamic systems approach to the development of cognition and action. Cambridge, MA: MIT Press; 1994.
- Tingley BM, Allen GD. Development of speech timing control in children. *Child Development* 1975;46:186–194.
- Walsh B, Smith A. Articulatory movements in adolescents: Evidence for protracted development of speech motor control processes. *Journal of Speech, Language, and Hearing Research* 2002;45:1119–1133.
- Weber-Fox C, Neville H. Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience* 1996;8:231–256.
- Weinreich, U. Languages in contact. New York: Linguistic Circle of New York; 1953.
- Winter, DA. The biomechanics of human movement. New York: Wiley; 1979.
- Yeni-Komshian GH, Flege JE, Liu S. Pronunciation proficiency in the first and second languages of Korean–English bilinguals. *Bilingualism* 2000;3(2):131–149.
- Zsiga EC. Articulatory timing in a second language. *Studies in Second Language Acquisition* 2003;25:399–432.

Participants: 21 Bilinguals, Bengali (L1) and English (L2)

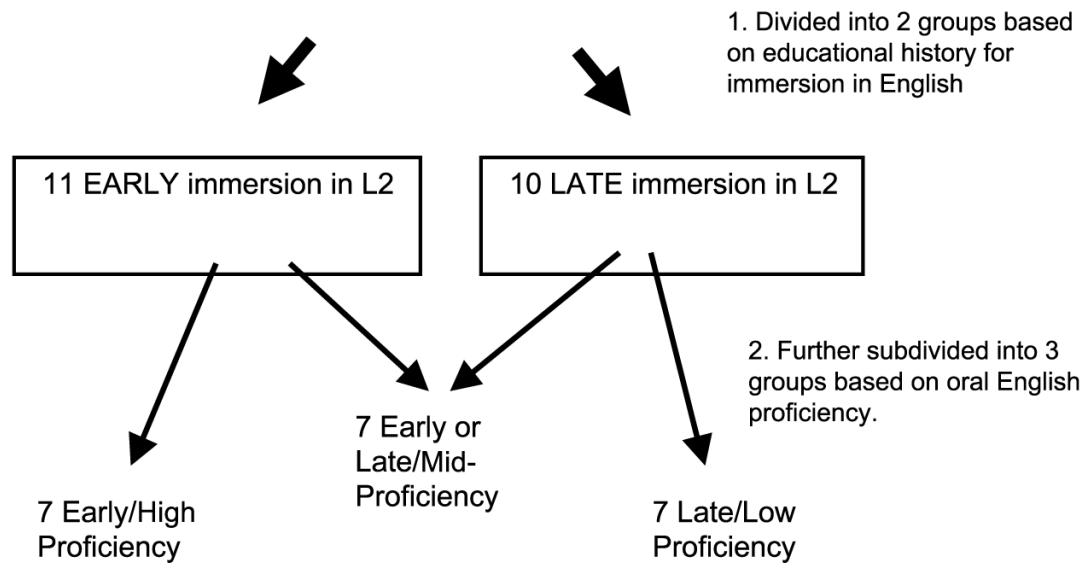


Figure 1.
Illustration of participant selection and analysis strategy.

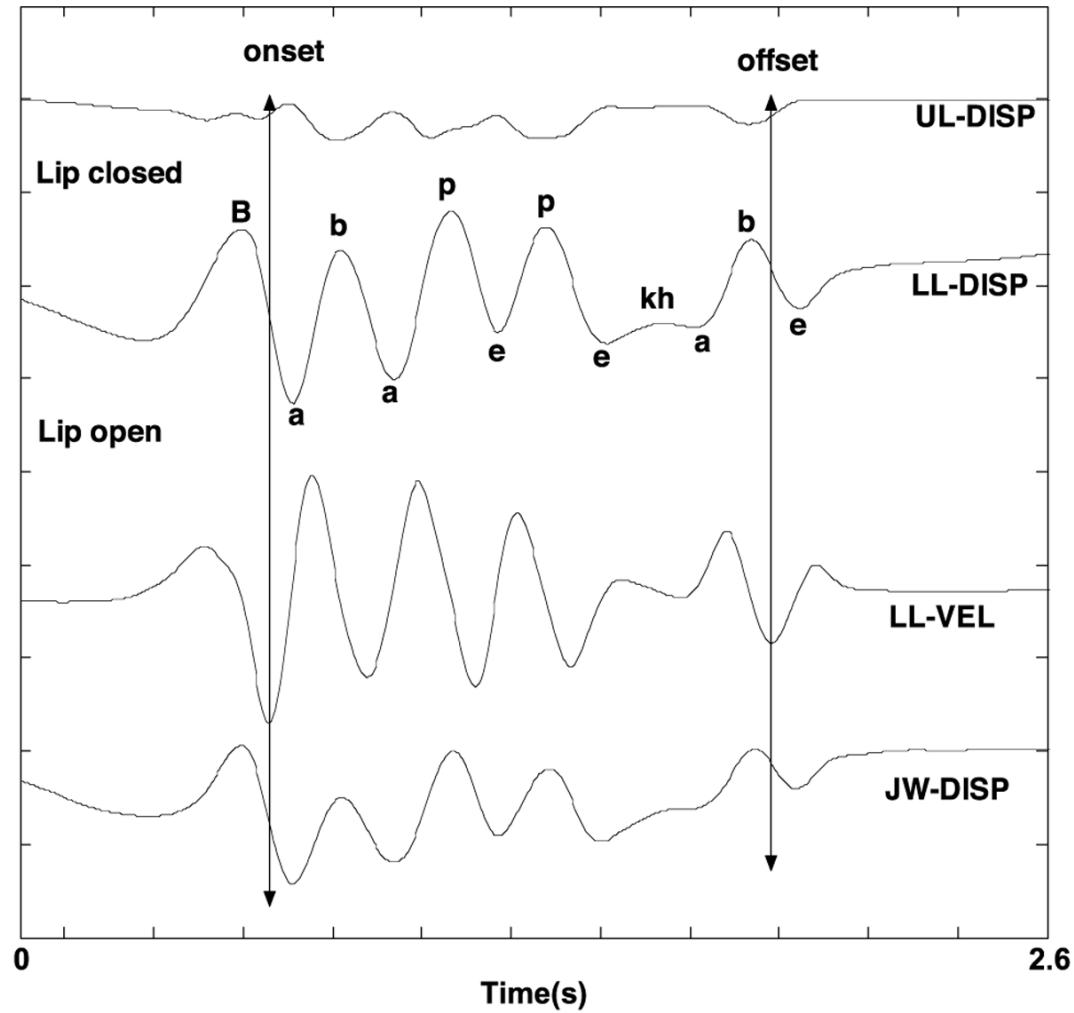
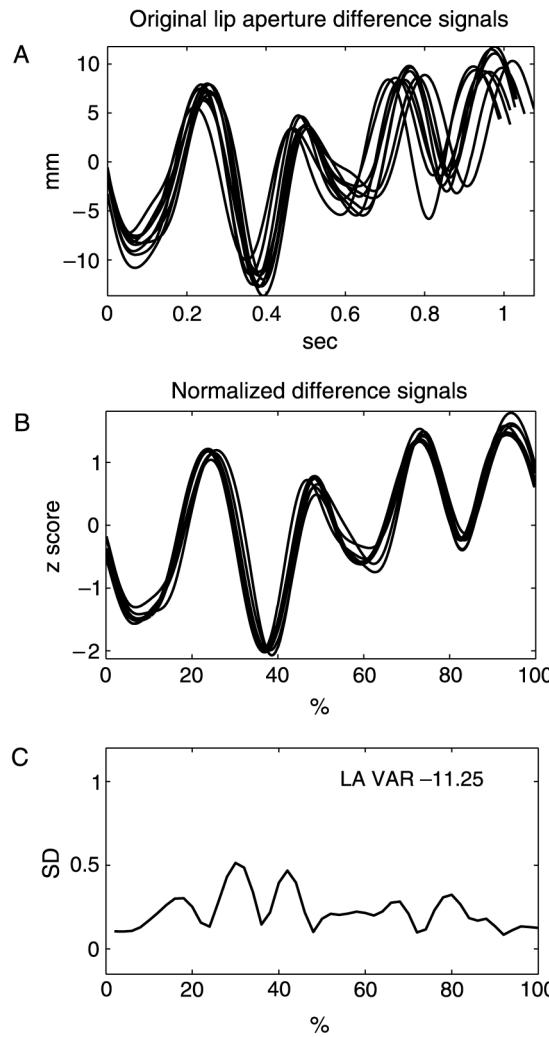


Figure 2.

These plots show the upper lip (UL-DISP), lower lip (LL-DISP), and jaw (JW-DISP) displacements along with lower lip velocity (LL-VEL) profiles, during a single production of the sentence “Baba pepe khabe.” The LL-VEL profile was used to segment the onset and offset points, as illustrated by the vertical line.

**Figure 3.**

Top panel shows original lip aperture trajectories for 1 speaker's 10 productions of the sentence "Buy Bobby a puppy." The middle panel shows those same 10 lip aperture signals after amplitude and time normalization. The bottom panel shows the cumulative sum of the standard deviation values (i.e., the lip aperture variability index [LA VAR]) that were obtained at 2% intervals from the 10 normalized trajectories.

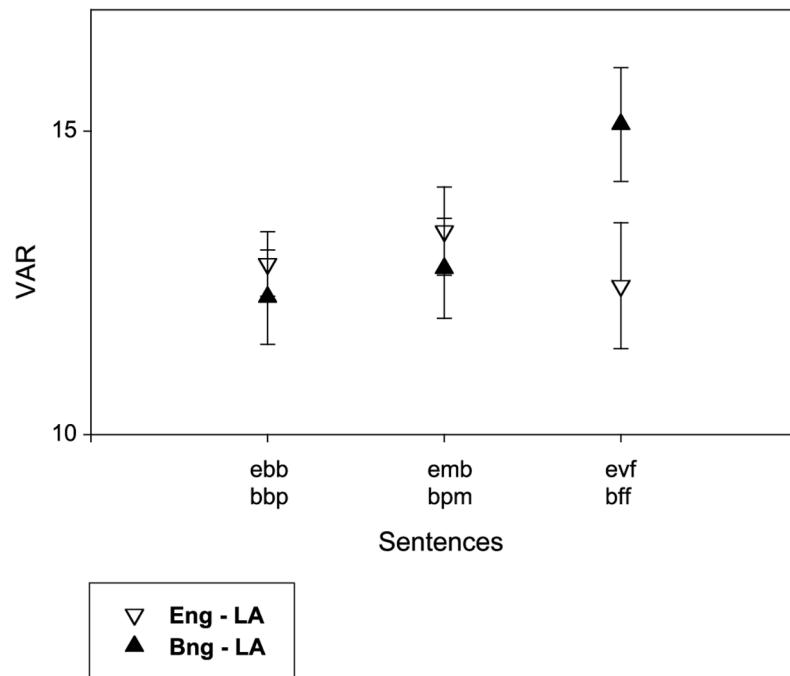
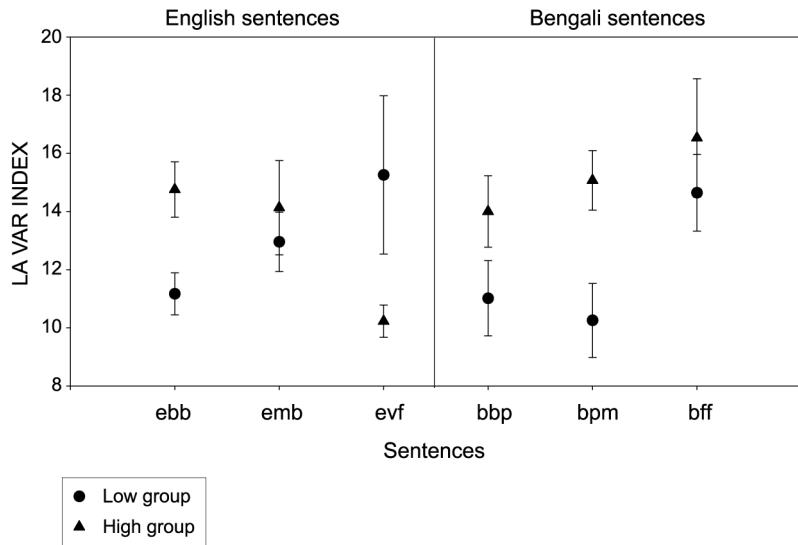
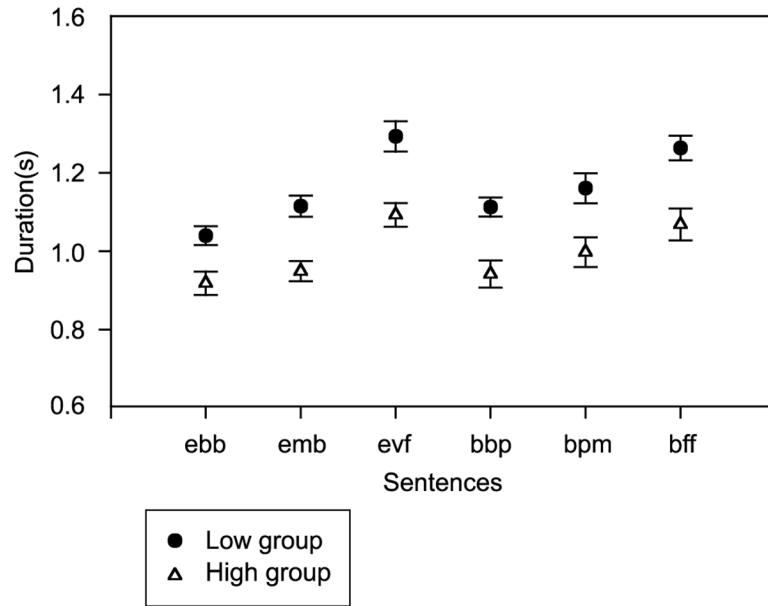


Figure 4.

Lip aperture (LA) variability (VAR) means and standard errors (represented by bars) obtained from all 21 participants. The English sentences are “Buy Bobby a puppy” (ebb), “Mommy bakes pot pies” (emb), and “Vicky feeds the fat pony” (evf). The Bengali sentences are “Baba pepe khabe” (bbp), “Papur ma marbe” (bpm), and “*P^honi p^hul b^halobashe*” (bff). Open symbols refer to the English sentences, and filled symbols to the Bengali sentences.

**Figure 5.**

Lip aperture variability means and standard errors (represented by bars) obtained from the late/low (low; circles) and early/high (high; triangles) groups for all six sentences. The English sentences are “Buy Bobby a puppy” (ebb), “Mommy bakes pot pies” (emb), and “Vicky feeds the fat pony” (evf). The Bengali sentences are “Baba pepe khabe” (bfp), “Papur ma marbe” (bpm), and “Phoni phul bhalobashe” (bff).

**Figure 6.**

Sentence duration means (in seconds) and standard error bars for all six sentences from the late/low (low; circles) and early/ high (high; triangles) groups. The English sentences are “Buy Bobby a puppy” (ebb), “Mommy bakes pot pies” (emb), and “Vicky feeds the fat pony” (evf). The Bengali sentences are “Baba pepe khabe” (bbp), “Papur ma marbe” (bpm), and “Phoni p^hul b^halobashe” (bff).