The development of skill in producing word-final English stops: Kinematic parameters

James Emil Flege
Department of Biocommunication, University of Alabama at Birmingham, Birmingham, Alabama 35294

(Received 22 January 1988; accepted for publication 14 July 1988)

It was hypothesized that native English adults would be more skillful in producing word-final English /p/ and /b/ than native English children who, in turn, would be more skillful in doing so than adult native speakers of a language (Mandarin Chinese) that does not possess word-final stops. A video tracking system was used to monitor lip and jaw movements. The subjects in all three groups made vowels significantly longer before /b/ than /p/, but the effect seen for the English subjects was three times as large as the Chinese subjects' effect and depended less on differences in lip closing velocity for /b/ and /p/. The English subjects also showed a difference in duration between /a/ and /i/ that was twice as large as the difference seen for the Chinese subjects. Of the three groups, only the English adults showed significantly greater displacement and peak movement velocity for the final stop consonant of /hap/ than /bab/. This suggested that their central phonetic representations specified a more forceful constriction of the lips for /p/ than /b/. The English adults seemed to compensate more effectively for a bite block in producing the final stops in /hip/ and /bib/. The results obtained for the English children were intermediate to those obtained for the English and Chinese adults, which is consistent with the hypothesized experience-based differences in level of skill.

PACS numbers: 43.70.Aj, 43.70.Ep

INTRODUCTION

Recent studies of speech motor control have generally adopted one of two broad perspectives. Some researchers (e.g., Kelso et al., 1986) take the position that movement of the articulators used to form speech sounds is regulated to a large extent by the same forces and constraints that govern goal-directed movements of the limbs. In this view, speech motor control is a specific instance of a more general problem. Other investigators (e.g., Smith and McLean-Muse, 1986, 1987) have explicitly or implicitly regarded the production of speech sounds as being shaped to some extent by the information specified in central phonetic representations, which may change slowly as the result of language-specific experience. In this view, speech is in some sense "special."

One danger in making abstract, central representations the object of inquiry in speech research is the temptation to attribute any observable, systematic differences to differences in central representations (Baars, 1986). It is not necessarily the case that all measurable differences need to be controlled directly by the nervous system. The value of an action theoretical perspective is that it provides constraints on what control parameters are attributed to central representations (Smythe and Wing, 1984). However, at least some of the differences observed are likely to be based on differences in central representations. This study sought to determine what language-specific information might be encoded in the central representations for English /p/ and /b/ by comparing the speech of talkers in three groups differing in English-language experience.

A. Differences between /p/ and /b/

Both /b/ and /p/ are implemented by constricting the lips, but adult native speakers of English differentiate them in several ways. The supraglottal cavity is enlarged actively during the constriction of /b/ but not /p/, apparently to sustain closure voicing (Flege et al., 1987). The voiceless stop /p/ is produced with a laryngeal devoicing gesture that suppresses closure voicing. It is said to be produced with relatively greater "muscular effort" than /b/, which may be manifested in terms of a relatively longer period of bilabial constriction and greater force of constriction (Belasco, 1953; Kozhevnikov and Chistovich, 1965; Ohman, 1967; Chomsky and Halle, 1968; MacNeilage, 1972; Flege and Brown, 1982). Lubker and Parris (1970) noted greater pressure of bilabial contact for word-final /p/ than /b/. Electromyographic studies have shown a higher level of muscular activity for /p/ than /b/, although the differences noted were often small, and intersubject variability was typically large (Harris et al., 1965; Fromkin, 1966; Lubker and Parris, 1970; Sussman et al., 1973).

One might expect greater force of constriction to be accompanied by greater movement velocity. It appears that native speakers of English do move the upper lip, lower lip, and jaw (or some combination thereof) more rapidly in forming post-vocalic /p/ than /b/, although the differences noted in previous studies have often been small, sometimes nonsignificant, and prone to great intersubject variability (Chen, 1970; Tuller and Kelso, 1984; Sussman et al., 1973; Kim, 1972, reported in MacNeilage, 1972; see, also, Kozhevnikov and Chistovich, 1965). More recently, Smith and
McLean-Muse (1988) reported significantly greater peak closing velocities for /p/ than /b/ in the speech of adults and children between the ages of 4 and 11 years (see also Van Summers, 1987). Interestingly, the magnitude of the
/p/ — /b/ velocity difference tended to increase with age, suggesting that specification of the underlying parameter(s) responsible for the velocity difference may change with English-language experience.

It is well known that the differences between English /p/ and /b/ are accompanied by differences in preceding vowel duration, at least in slow, careful speech. Chen (1970) noted that the size of this "consonant voicing effect" on preceding vowel duration is larger in English than other languages with voiced and voiceless obstruents in word-final (or post-stressed) position (see also Belasco, 1953; Lisker, 1974; Raphael et al., 1980; Mack, 1982). He hypothesized that the consonant voicing effect might derive at least in part from the fact that Chinese adults sustain voicing for shorter durations in /b/ than /p/ for native English adults. The jaw was also held in a quasi steady-state position significantly longer before the final consonant closing gestures for /b/ than /p/. Smith and McLean-Muse (1988) estimated the duration of labial closing gestures for the word-final /p/ and /b/ tokens spoken by English adults and children. The duration differences could account for only about one-third of the consonant voicing effect on vowel duration that was measured acoustically.

These last two findings suggest that the relatively large consonant voicing effect seen in English is not due entirely to differences in the durations of consonant closing for /p/ and /b/. The effect in English may represent the exaggeration of a smaller and seemingly universal (but see Flege and Port, 1981) effect that is conditioned by differences in closing velocities. Klatt (1976) proposed that the large consonant voicing effect in English derives from the application of a phonological rule (see also Klatt, 1973). A phonological rule might operate by making lip closing begin later in the vowel for words ending in /b/ than /p/ (Fromkin, 1966; Kelso and Tuller, 1987; Van Summers, 1987).

B. Design and hypotheses

The speech production of three groups of subjects differing in English-language experience was compared. An important assumption underlying this study was that speaking skill increases as talkers gain experience speaking and hearing word-final stops. It was assumed that, even though the 7-year-old native English children who were examined could produce a perceptually effective contrast between /p/ and /b/, they had considerably less experience doing so and might, therefore, be less skillful than the native English adults (their mothers).

In most studies of speech development, the performance of normal young adults, who presumably represent the highest possible level of skill, is compared to that of children in several age groups. One potential problem with this approach is that orofacial morphological differences might lead to between-group differences in the absence of differences in skill (see the results for upper lip displacements presented below). The performance of English adults were therefore also compared to a group of adult non-native speakers.

The non-native speakers were native speakers of Chinese who learned English as adults in the United States, where they had lived for 6 years on average. They had considerably less experience producing word-final stops than either the English adults or children because their native language, Mandarin, does not have obstruent consonants in word-final position (Maddieson, 1984). Based on many previous studies of second-language speech production (Flege, 1988), it was hypothesized that the Chinese adults would be found to be less skillful than the native English adults.

It should be noted that this design introduces a potential confounding factor that differs from the one just mentioned, viz., that phonetic implementation rules might be transferred from Chinese into English (Flege and Eefting, 1988). However, one consideration suggests that cross-language phonetic interference did not undermine the utility of comparisons between the English and Chinese subjects. Previous research has shown that Chinese speakers of English have difficulty producing stops in the final but not the initial position of English words (Taroné, 1980; Eckman, 1981; Flege and Davidian, 1984; Heyer, 1986; Flege et al., 1987). This suggests that Chinese adults who learn English do not necessarily use the phonetic implementation rules established for producing /p/ and /b/ in the initial position of Chinese words in the final position of English words.

Previous research has shown that native speakers of Chinese may not produce a perceptually effective contrast between word-final /p/ and /b/ in English (Flege and Davidian, 1984). However, based on the observation of significant oral air pressure differences between /p/ and /b/, Flege et al. (1987) inferred that adult native Chinese speakers of English produced /p/ and /b/ with different laryngeal articulations. The Chinese adults sustained voicing for shorter durations in /b/ closure intervals than native English children who, in turn, produced less closure voicing than English adults. This finding was consistent with the differences in skill level that are hypothesized to differentiate the three subject groups examined in the present study.

No evidence is available concerning kinematic aspects of Chinese adults' production of English word-final stops. If native speakers of English learn to constrict the lips more forcefully for /p/ than /b/, then one might expect the English adults to show relatively greater labial displacement and peak closing velocity for /p/ than /b/. The magnitude of the /p/ — /b/ differences might also increase according to the hypothesized skill levels, that is, be greater for English adult than English children than Chinese adults.

Since adult Chinese speakers of English differentiate word-final English /p/ and /b/ phonetically (Flege et al., 1987), they should make vowels significantly longer before /b/ than /p/. However, if the consonant voicing effect they produce is much smaller than the English native speakers' because it derives from differences in the duration of consonant closing gestures (see Chen, 1970), then stronger correlations should exist between the duration of labial closing
gestures and preceding vowel duration for the native Chinese than English subjects. If the Chinese but not the native speakers show significant correlations between labial closing and vowel duration, it would suggest that the consonant voicing effect seen for native speakers of English is represented centrally.

The subjects in this study produced words first at a normal speaking rate and then at a self-selected fast rate. The purpose of this experimental manipulation was to stress the speech production mechanism in order to better delineate potential between-group differences. Previous kinematic studies, most of which examined prevocalic stops, have shown inconsistent effects of a speaking rate increase on the duration, displacement, and velocity of the bilabial closing gestures (Abbs, 1973; Hughes and Abbs, 1976; Kelso et al., 1985; Caligiuri and Abbs, 1985; Smith and McLean-Muse, 1987). It is not known whether an increase in speaking rate will result in an increase in bilabial closing velocity for word-final stops and, if so, whether a rate increase will affect differentially the closing velocities for /p/ and /b/.

The subjects also produced the test words in a fixed-jaw condition in which mandibular movement was severely restricted. Previous research, most examining word-initial (prevocalic) stops, has demonstrated that native speakers of English manage to achieve bilabial constriction in speech produced with a bite block, although the inferior–superior location of the constriction changes according to the width of the bite block (Folkins and Canty, 1986; Smith, 1987). Smith and McLean-Muse (1987) found that children tended to increase velocity in labial stops, whereas adults tended to decrease velocity in a way that corresponded to the increases and decreases in articulatory displacements. It is, therefore, possible that smaller (or larger) /p/ —/b/ differences might be observed in the fixed- than free-jaw condition, and that between-group differences found in one condition might not be observed in the other condition.

I. METHODS
A. Subjects and speech material

The ten subjects in each of three groups who participated as paid subjects differed in age and/or native language background. The English adult group (designated EA) consisted of females from Alabama with a mean age of 34.2 years (s.d. = 4.0) who spoke only American English. The four boys and six girls in the English child group (EC) had a mean age of 7.4 years (s.d. = 0.8). They were offspring of the subjects in the English adult group. The native Chinese subjects (CA) were adult female native speakers of Mandarin with a mean age of 32.4 years (s.d. = 6.4) who had lived for 0.5–18 years in the United States (M = 6.4 years, s.d. = 5.8). Four were from Taiwan, and six were from mainland China. A language background questionnaire indicated that, although the Chinese subjects had studied English in China, their first massive exposure to native-produced English occurred when they arrived in the U.S. They could all communicate in English, but each spoke with an obvious foreign accent in the author’s opinion.

Realization of the phonetic contrast between /p/ and /b/ was examined in /bip/—/bib/ and /bap/—/bab/. The vowels /i/ and /a/ were chosen because they differ according to jaw height and the size of the mouth aperture (Fromkin, 1964) and because Chinese has an /i/ and /a/ (Maddieson, 1984).

B. Instrumentation

A video tracking system described previously (McCulcheon et al., 1977) was used to monitor lip and jaw movements. It consisted of a modified Cohu video camera with a Vivitar macro lens that was able to detect changes in the X-Y location of light sources at a 100-Hz rate with 16-bit resolution. The light sources were small (4.0-mm diameter) glass beads that reflected light emitted from a ring of lamps surrounding the camera. The system was able to track illuminated beads attached to the lips and jaw at velocities much greater than those typically seen in the production of /p/ and /b/. The system was calibrated using beads separated by known distances. In the configuration used in the study, measurement resolution was limited to one video scan line or 0.27 mm.

The four beads tracked in this study were positioned in the midsaggital plane, as shown in Fig. 1. The beads designated UL and LL were attached at the vermillion border of the upper and lower lips. Bead J was fixed to a rigid cantilever attached to a dental splint. The approximately 3-mm-thick splint that was fabricated individually for each subject conformed to a cuspid and bicuspid tooth on the right side of the mouth. The cantilever was bent so that it exited the oral cavity near the corner of the mouth without deforming noticeably the upper or lower lip. Bead H was fixed to a vertical cantilever attached to a pair of special glasses to provide a fixed reference point for determining vertical movements of UL, LL, and J. It was located on a plane roughly parallel to the occlusal plane about 25 mm anterior to UL and LL.

![FIG. 1. Illustration of the placement of light-reflecting beads in the midsaggital plane. UL was fixed to the upper lip, LL to the lower lip, H to a rigid cantilever attached to the head via glasses, and J to a cantilever attached to mandibular teeth.](image-url)
C. Procedures

The data were acquired in a sound booth during a single session lasting about 45 min. The subjects were seated comfortably in a dental chair equipped with a custom-built cephalostat. They were familiarized with the speech material to be produced, and the four beads were attached using a drop of medical adhesive. The subjects produced eight sets of 25 utterances (13 words ending in /b/ and 10 ending in /p/). Four sets were produced at a normal rate followed by four sets at the fast rate. Following the practice of Folkins and Canty (1986), the free-jaw condition preceded the fixed-jaw condition at each speaking rate, and the order of vowel (/i/ vs /a/) was counterbalanced within the four rate x condition combinations.

Within each set, the subjects said only one of two possible words preceded by the word “say.” Each utterance was prompted by the presentation of a 3- × 5-in. card with one of the words written on it (for example, “beep” or “beep”). This procedure was adopted because it avoided the need for a live-voice model and obviated confounding effects that might have derived from differences in reading ability.

The jaw was immobilized in a slightly higher-than-normal position in the fixed-jaw condition for the words with /i/, and in a substantially higher-than-normal position for the words with /a/ (see below). The talkers held a custom-made acrylic bite block between bicuspid and premolar teeth on the left side of the mouth while producing /bip/ and /bib/. A thinner bite block was used with the children than adults (6 vs 8 mm) to compensate for the children’s smaller oral cavities. The talkers in all three groups were asked to bite down on the splint holding the jaw cantilever while producing /bap/ and /bab/. This was equivalent to clenching an approximately 3-mm bite block.

D. Segmentation

The vertical locations (in 0.27-mm video scan lines) of the four beads and the derived separation (in millimeters) between the beads on the upper and lower lips were displayed on a graphics terminal. As shown in Fig. 2, The vowel interval (“V”) was defined as extending from the beginning of a rapid increase in lip separation associated with release of the word-initial /b/ to the end of the decrease in lip separation distances associated with the bilabial constriction for the word-final /b/. The bilabial closing gestures associated with the word-final stop (“C”) were defined as extending from the beginning of the marked decrease in lip separation to the minimum lip separation associated with bilabial constriction. Note that little change is evident here for bead H because the head was held by the cephalostat. Little movement is evident for bead J because the utterance illustrated here was spoken with a bite block.

The duration of intervals V and C was calculated by multiplying the number of samples within each interval by 10 ms. The vertical displacements of the upper lip, lower lip, and jaw were calculated by multiplying the change in scan lines (with reference to bead H) from the beginning to the end of the bilabial closing interval by 0.27. Displacement of the lower lip was calculated by subtracting jaw displacement from net lower lip displacement. Net bilabial displacement was calculated by adding the displacements seen for the upper and lower lips. The peak velocity (in mm/s) of the bilabial closing movements was estimated by dividing the difference in displacement between each pair of adjacent data points by 10 ms, then picking the largest value in interval C.

E. Analyses

Mean duration, displacement, and velocity values were calculated from the first eight /p/ and /b/ tokens meeting the segmentation criteria. The mean values obtained for the normal- and fast-rate samples were submitted to separate mixed-design ANOVAs in which consonant voicing (/p/,/b/), condition (free-jaw, fixed-jaw) and vowel (/i/, /a/) were repeated measures. Significant interactions involving the between-subjects factor group were explored by...
tests of simple main effects, and by post hoc comparisons using the Newman–Keuls test. An alpha level of 0.05 was used for testing significance.

II. RESULTS

A formal perceptual appraisal of the word-final stops spoken by the talkers in the three groups was not carried out, but it appeared to the author that the talkers in all three groups produced audible word-final bilabial stops in both the free- and fixed-jaw conditions. This subjective observation for stops spoken in the free-jaw condition by the Chinese adults agrees with previous studies examining the word-final English stops produced by adult native Chinese speakers of English (Flege and Davidian, 1984; Flege et al., 1987). The fixed-jaw results for the English adults agree with studies examining the bite-block production of bilabial stops by native speakers (e.g., Folkins and Canty, 1987).

A. Duration

The first two questions addressed were whether the magnitude of the duration difference between the vowels preceding /b/ and /p/ would increase with English-language experience and, if the vowel duration differences might be accounted for by differences in how long it took the lips to close for /b/ and /p/.

1. Vowel duration

The mean vowel durations shown in the left panels of Fig. 3 have been averaged across the free- and fixed-jaw conditions because the condition factor was nonsignificant at both the normal and fast speaking rates \( F(1,27) = 0.12, 0.00 \). The mean durations of the bilabial closing gestures for /p/ and /b/ shown in the right panels of Fig. 3 have also been averaged across the two conditions because the condition factor was nonsignificant at both rates \( F(1,27) = 0.10, 0.06 \).

The talkers in all three groups made vowels longer before /b/ than /p/, but the magnitude of the difference was about three times greater for the native than the non-native speakers. The native English adults and the native English children showed a larger vowel duration difference than the Chinese adults at the normal rate (83 and 62 ms vs 23 ms) and fast rate (47 and 42 ms vs 14 ms). The simple main effect of voice proved to be highly significant for all six group \( \times \) speaking rate combinations \( (p < 0.001) \). It is likely, therefore, that the significant group \( \times \) voice interactions obtained at both speaking rates were due to differences between groups in the magnitude of the consonant voicing effects \( F(2,27) = 9.87, 4.39 \).

The talkers in all three groups made /a/ longer than /i/. A significant group \( \times \) vowel interaction was obtained for normal-rate vowels because the native English adults and native English children showed a larger /a/ -- /i/ difference than the Chinese adults (33 and 42 ms vs 12 ms). \( F(2,27) = 4.42 \). A significant group \( \times \) vowel interaction was not obtained at the fast rate, however, because the magnitudes seen there for the English adults and children more closely resembled those seen for the Chinese adults (35 and 30 ms vs 20 ms). \( F(2,27) = 1.22 \).

2. Labial closing duration

The lips took significantly longer to close for /b/ than /p/ at the fast (100 ms vs 96 ms) but not at the normal rate (110 ms vs 108 ms). \( F(1,27) = 15.9, 2.33 \). A significant /p/ -- /b/ difference was not obtained at the normal rate.
because the English adults and the Chinese adults showed significantly longer closing gestures for /b/ than /p/ (5 and 6 ms, respectively), whereas the English children showed a significant 5-ms effect in the opposite direction. The difference between the adults and children resulted in a significant group × voice interaction \( F(2,27) = 7.45 \).

### 3. Relation between vowel and labial closing duration

To help determine why the native English adults and children showed a much larger consonant voicing effect than the Chinese adults, the 20 mean vowel and consonant closing duration values obtained for the three talker groups in each of the eight vowel × speaking rate × jaw condition combinations were submitted to correlation analyses. If the small effect of the native Chinese, but not the larger effect of the native English, talkers, derived from differences in how rapidly the lips were approximated for /b/ and /p/, the correlations between vowel and consonant closing durations should be greater for the Chinese than English subjects.

The results presented in Table I are largely consistent with this hypothesis. The native Chinese adults showed moderately strong average correlations between vowel and consonant durations for words spoken in the free- and fixed-jaw conditions (\( r = 0.676, 0.524 \)). The English adults and English children showed smaller average correlations in the two conditions (\( r = 0.306, 0.293; r = -0.010, 0.230 \)). For words spoken in the free-jaw condition, the Chinese adults showed significant correlations for words with /i/ and /a/ spoken at the fast rate, and the English adults showed a significant correlation for fast-rate words with /i/. When the values for the normal-rate words with /i/ and /a/ were analyzed together, the Chinese adults showed a significant correlation \( r(38) = 0.539, p = 0.0002 \), but not the English adults or the English children \( r(38) = 0.261, -0.105, p > 0.05 \).

For words spoken in the fixed-jaw condition, the Chinese adults showed significant correlations for the words with /i/ spoken at the normal rate and for the words with /i/ and /a/ spoken at the fast rate. The English adults and children, on the other hand, showed significant correlations only for fast-rate words with /i/.

### B. Displacement

The next question to be addressed was whether there would be greater displacement for /p/ than /b/ as the result of a relatively more forceful constriction for the voiceless stop.

<table>
<thead>
<tr>
<th></th>
<th>Free-jaw condition</th>
<th>Fixed-jaw condition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EA</td>
<td>EC</td>
</tr>
<tr>
<td>/i/-NOR ( r )</td>
<td>0.328</td>
<td>-0.282</td>
</tr>
<tr>
<td>( p )</td>
<td>0.157</td>
<td>0.229</td>
</tr>
<tr>
<td>/a/-NOR ( r )</td>
<td>0.180</td>
<td>-0.171</td>
</tr>
<tr>
<td>( p )</td>
<td>0.449</td>
<td>0.470</td>
</tr>
<tr>
<td>/i/-FAS ( r )</td>
<td>0.611</td>
<td>0.399</td>
</tr>
<tr>
<td>( p )</td>
<td>0.004</td>
<td>0.081</td>
</tr>
<tr>
<td>/a/-FAS ( r )</td>
<td>0.105</td>
<td>0.034</td>
</tr>
<tr>
<td>( p )</td>
<td>0.660</td>
<td>0.868</td>
</tr>
</tbody>
</table>

![FIG. 4](image-url) The mean displacement of the upper lip in word-final /p/ and /b/ tokens spoken in the context of /i/ and /a/ at two rates in the free- and fixed-jaw conditions by native English children (EC), native English adults (EA), and native Chinese adults (CA), in ms.
1. Individual articulators

The mean displacements of the upper lip, lower lip, and jaw are shown as a function of vowel context and speaking rate in Figs. 4–6. The values obtained in the free- and fixed-jaw conditions are presented separately in these figures and those to follow because the condition factor often interacted significantly with other factors in the ANOVAs.

Figure 4 shows the mean displacement of the upper lip for /p/ and /b/. The lack of main effects of voice or voice-group interactions indicated that talkers in the three groups did not move the upper lip significantly more for /p/ than /b/. A significant group main effect was obtained for fast-rate words, however, because the native English children moved the upper lip significantly more than the English adults and the Chinese adults (4.0 mm vs 2.3 and 2.6 mm),
and a nonsignificant trend in the same direction was noted for normal-rate words (4.0 mm vs 2.7 and 3.4 mm) \[F(2,27) = 5.83, 2.61\]. Significant voice \(\times\) vowel interactions were obtained at both speaking rates because the upper lip moved more for \(/p/\) than \(/b/\) in the context of \(/a/\), whereas the reverse was true in the \(/i/\) context \[F(1,27) = 14.1, 12.8\].

Figure 5 shows the mean displacement of the lower lip. Neither the main effect of group nor the group \(\times\) voice interaction reached significance for normal-rate words. A significant four-way interaction was obtained for fast-rate words because the native English adults moved the upper lip significantly more for \(/p/\) than \(/b/\) in the fixed-jaw condition following \(/a/\) (7.4 vs 6.9 mm) \[F(2,27) = 6.49\]. Tests of the simple simple main effect of voice revealed that none of the other \(/p/\) \(\rightarrow\) \(/b/\) differences was significant. Significant voice \(\times\) vowel interactions were obtained at the normal and fast speaking rates because the lower lip moved slightly more for \(/p/\) than \(/b/\) in the context of \(/a/\), whereas a tendency in the opposite direction was evident in the \(/i/\) context \[F(1,27) = 12.3, 6.07\].

Figure 6 shows that the jaw moved little in the fixed-jaw condition, and that it moved substantially more following \(/a/\) than \(/i/\) in the free-jaw condition. A four-way interaction was obtained for normal-rate words because the English adults moved the jaw significantly more for \(/p/\) than \(/b/\) in the free-jaw condition following \(/a/\) (3.4 vs 2.7 mm) \[F(2,26) = 4.01\]. Tests of simple simple main effects revealed that none of the other \(/p/\) \(\rightarrow\) \(/b/\) differences at the normal rate was significant. Neither the main effect of group nor the group \(\times\) voice interaction reached significance for fast-rate words.

2. Net lower lip and bilabial displacements

The results for the individual articulators showed no systematic differences in displacement for \(/p/\) and \(/b/\), although the English adults showed significantly greater movements of the lower lip and jaw for \(/p/\) than \(/b/\) in one instance each following \(/a/\). This suggests the possibility that a more systematic difference between subject groups might exist when the combined movements of the lower lip + jaw, and of the upper lip + lower lip complex, are considered.

Figure 7 shows the combined displacement for the lower lip + jaw. A significant group \(\times\) voice \(\times\) vowel interaction was obtained at the normal rate because the three groups showed divergent \(/p/\) \(\rightarrow\) \(/b/\) differences in the context of \(/a/\) \[F(2,27) = 6.58\]. None of the three groups showed a significant difference in labial displacement between \(/p/\) and \(/b/\) in the context of \(/i/\). In the context of \(/a/\), however, the English adults and the English children moved the lower lip significantly more for \(/p/\) than \(/b/\) (8.3 vs 7.3 mm, 8.1 vs 7.3 mm), whereas the Chinese adults moved it significantly less (8.3 vs 8.7 mm). Neither the group factor nor the group \(\times\) voice interaction was significant at the fast rate.

The results for net bilabial displacement closely resembled those obtained in the analyses of lower lip + jaw movements and will therefore not be displayed. A significant group \(\times\) vowel \(\times\) voice interaction was obtained in the examination of the normal-rate stops \[F(2,27) = 4.61\]. No significant between-group differences existed in the \(/i/\) context while in the \(/a/\) context, on the other hand, the English adults and English children moved the lips significantly more in forming \(/p/\) than \(/b/\) (11.1 vs 9.9 mm, 11.9 vs 11.0

![Figure 7](image)
jaw displacements were only 0.2 mm less on average in the jaw by the bite block resulted in very little compensation by groups, the lower lip moved slightly less (0.1 mm) rather than more in the fixed- than free-jaw condition. Somewhat surprisingly, net bilabial displacement increased by 0.6 mm due to an 0.9-mm increase in the movement of the upper lip.

These results suggested that the small perturbation of the jaw by the bite block resulted in very little compensation by the lower lip but in overcompensation by the upper lip.

When producing the words with /a/, the talkers clenched an approximately 3-mm-thick dental splint that positioned the jaw in a substantially higher-than-normal position. As a result, the jaw moved 3.1 mm less in the fixed-than in the free-jaw condition. The upper lip showed no compensation for this reduction in jaw movement. In fact, it moved 0.4 mm less on average in the fixed- than free-jaw condition. Net bilabial displacement was only 1.4 mm less in the fixed-than in the free-jaw condition, however, because lower lip displacement increased by 2.2 mm in the fixed-jaw condition. It thus appeared that reducing jaw movement to a relatively great extent resulted in compensation by the lower lip alone.

The next question to be addressed was whether the patterns of compensation just described applied equally to the talkers in all three groups. To determine this, the percentage of bilabial displacement due to movements of the upper lip and the lower lip was calculated for /p/ and /b/ tokens spoken at the normal rate. (A preliminary analysis showed that compensatory strategies at the fast rate were comparable to those seen at the normal rate.) A total of 240 percentage scores was obtained for stops spoken in each vowel context at the two speaking rates by dividing the displacement seen for the upper lip and lower lip (with jaw component removed) by net bilabial displacement. These scores were then submitted to separate group X lip X voice X condition ANOVAs for each vowel.

Figure 8 shows the mean percentage of total bilabial displacement due to upper lip and lower lip movements in stops spoken in the free- and fixed-jaw conditions. The upper lip contributed less to total bilabial displacement than the lower lip in both the /i/ and /a/ contexts (34% vs 62%, 30% vs 60%) [F(1,27) = 49.4, 54.2]. Significant lip X condition interactions were obtained in the /i/ and /a/ contexts because of the differences in the compensatory responses of the two lips [F(1,27) = 14.8, 43.7]. In the /i/ context, the upper lip percentage increased in the fixed-jaw condition by 7%, whereas the lower lip percentage decreased by 5%. In the /a/ context, on the other hand, the upper lip percentage remained the same, whereas the lower lip contribution increased by 25%.

A difference in compensatory strategy between the English adults and the two less experienced groups of subjects was evident for stops spoken in the /i/ context. The English adults showed a nonsignificant increase of 2% in the fixed-jaw condition for both the upper lip and lower lip. The English children and the Chinese adults, on the other hand, showed significant increases in the upper lip contribution in the fixed-jaw condition (8%, 12%). The English children showed a nonsignificant decrease of 4% in the lower lip percentage, and the Chinese adults showed a significant decrease of 12%. These differences yielded a significant group X lip X condition interaction [F(2,27) = 4.24]. The three-way interaction did not reach significance for stops in the /a/ context, indicating that talkers in the three groups compensated in a similar manner in the fixed-jaw condition [F(2,27) = 0.69].

D. Peak velocity

The final question to be addressed was whether the lips and jaw moved more rapidly for /p/ than /b/.
1. Individual articulators

The peak velocity of upper lip movements for /p/ and /b/ did not differ significantly at either the normal or fast rate (85 vs 83 mm/s, 84 vs 85 mm/s) \( F(1,27) = 0.43, 0.76 \). Neither the main effect of group nor any interaction involving this factor reached significance at either rate.\(^5\)

The lower lip closed significantly more rapidly for /p/ than /b/ at the fast rate (112 vs 105 mm/s) but not at the normal rate (103 vs 98 mm/s) \( F(1,26) = 19.6, 4.05 \). Neither the main effect of group nor any interaction involving this factor reached significance at the normal or fast rate.\(^6\)

A significant voice \( \times \) vowel interaction was obtained at the normal rate because the /p/ \( - /b/ \) difference was greater in the /a/ than /i/ context (105 vs 97 mm/s, 100 vs 99 mm/s) \( F(1,26) = 6.64 \).\(^7\)

The analysis of peak jaw velocity was confined to stops spoken in the free-jaw condition following /a/ since the jaw moved so little following /i/ (see above) and was largely immobilized in the fixed-jaw condition. Jaw movement velocity was significantly greater for /p/ than /b/ at both the normal and fast speaking rates (60 vs 53 mm/s, 55 mm/s vs 50 mm/s) \( F(1,27) = 9.63, 4.94 \). Neither the group factor nor the group \( \times \) voicing interaction reached significance at either rate.\(^8\)

2. Net lower lip \( + / b / \) jaw and bilabial closing velocity

There were no significant differences between the three groups in movement velocity for /p/ and /b/ when the lower lip and jaw were considered separately, but inspection of Fig. 9 reveals a between-group difference in the net velocity of the lower lip + jaw complex. For words spoken at both the normal and fast rates, the lower lip moved significantly more rapidly for /p/ than /b/ (134 vs 128 mm/s, 140 vs 131 mm/s) \( F(1,26) = 5.47, 19.5 \). A group \( \times \) voice \( \times \) vowel interaction was obtained because no group showed a significant /p/ \( - /b/ \) difference in the /i/ context, whereas just one group—the English adults—showed a significant /p/ \( - /b/ \) difference in the /a/ context (153 vs 134 mm/s) \( F(2,26) = 3.32 \). Neither the main effect of group nor any interaction involving the group factor reached significance at the fast rate.\(^7\)

An analysis of the velocity at which the lips moved together in forming bilabial closure revealed that peak bilabial velocity was not significantly greater for /p/ than /b/ at the normal rate (203 vs 197 mm/s) although it was significantly greater for /p/ than /b/ at the fast rate (207 vs 197 mm/s) \( F(1,26) = 3.03, 11.5 \). Neither the effect of group nor any interaction involving this factor reached significance at either rate.\(^9\) One other measure of bilabial velocity was computed to further test the hypothesis that the magnitude of the /p/ \( - /b/ \) velocity difference increases with English-language experience. Since the lips and especially the jaw move anteriorly, "diagonal" displacements based on movements in both the anterior–posterior and inferior–superior dimensions were calculated. The velocity values derived from movements in two dimensions were about 3 mm/s greater on average than velocities based on vertical movements alone. The results obtained in analyzing these values did not differ from those already reported, and so they will not be presented here.

III. DISCUSSION

The present study examined kinematic parameters associated with the production of /p/ and /b/ in the final position of English words by talkers differing in English-language experience. Consonant–vowel–consonant words were

---

FIG. 9. The mean peak velocity of lower lip + jaw movements in word-final /p/ and /b/ tokens spoken in the context of /i/ and /a/ at two rates in the free- and fixed-jaw conditions by native English children (EC), native English adults (EA), and native Chinese adults (CA), in mm/s.
spoken at a normal then a self-selected fast rate, and with the jaw free to move normally followed by a fixed-jaw condition. The study yielded a number of between-group differences that seem to have derived from underlying differences in the central representations for stop consonants, which, themselves, are likely to have been shaped by experience hearing and saying the sounds of English.

A. Duration

Like adult native Chinese speakers of English examined in previous research (Flege and Davidian, 1984), the Chinese adults in the present study tended to devoice /b/. Despite this, they made vowels significantly longer before /b/ than /p/. This strongly suggests that their tendency to devoice was due to how they realized /b/, not to a lack of awareness that /p/ and /b/ contrast phonetically in the final position of English words. Flege et al. (1987) reached the same conclusion concerning speakers of Chinese. The Chinese adults in that study (two of whom participated in the present study) did not sustain glottal pulsing significantly longer in the closure interval of word-final /b/ than /p/. They did, however, resemble native speakers of English in producing /p/ with significantly greater peak oral pressure than /b/, which suggested awareness of the difference in laryngeal articulation distinguishing /p/ and /b/ in the final position of English words. The talkers in all three groups in the present study made vowels significantly longer before /b/ than /p/, but the Chinese adults' 19-ms consonant voicing effect was only one-third the size of the native English adults' and children's effects. This finding agrees with the results obtained by Mack (1982) for native and French speakers of English. The difference between the native and French speakers observed by Mack (1982) could be attributed to the transfer of French timing patterns into English since the small consonant voicing effect seen in the French-accented English was similar in magnitude to the one seen in French words spoken by French monolinguals (see Flege, 1980; Flege and Port, 1981; Flege and Eefting, 1988). However, the small consonant voicing effect observed for the Chinese adults in the present study can not be attributed to cross-language phonetic interference since Mandarin has neither voiced nor voiceless obstruents in word-final position.

The consonant voicing effects seen here for the native and Chinese subjects may have had different bases. The small effect seen for the Chinese adults seems to have derived, at least in part, from differences in bilabial closing velocity for /p/ and /b/, as hypothesized by (Chen, 1970). Only for the Chinese adults was the duration of bilabial closing significantly longer for /b/ than /p/ at the normal rate. For words spoken in the free-jaw condition at the normal rate, the Chinese but not native English subjects showed a significant correlation between vowel and consonant closing durations. The much larger consonant voicing effects seen for the native English speakers may have been due to the application of a phonological rule that prolonged vowels preceding /b/ (Klatt, 1976). However, it is important to note that the /b/ - /p/ labial closing duration difference seen for the Chinese adults was smaller than their consonant voicing effect at the normal rate (6 vs 23 ms), so part of the effect seen in their speech may have been due to an emerging phonological rule.

Lehiste (1970, p. 18; see, also, Lindblom, 1967) hypothesized that low vowels are "universally" longer than high vowels because greater displacements (especially that of the relatively massive mandible) are needed for low than high vowels. Despite this, the native Chinese subjects produced a difference in duration between /a/ and /i/ that was less than one-half the size of the difference seen for the native speakers of English. They made /a/ just 12 ms (5%) longer than /i/, which resembled a nonsignificant 22-ms (8%) effect reported by Ren (1985) for Mandarin Chinese (Ren, 1985) much more closely than the effects seen here and in previous studies for native speakers of English (Lehiste, 1970; Port and Rotunno, 1979). One might speculate that, in this instance, the Chinese subjects transferred a pattern of Chinese vowel timing into English. If so, then the relatively large vowel height effect of the native English subjects may derive in part from temporal differences encoded in central phonetic representations (Smith and Gartenberg, 1984; Keating, 1985). A firm conclusion cannot be reached based on the data presented here, however, because the between-group differences might also have derived from unexamined differences in vowel quality or mandibular height.

B. Displacement and peak velocity

It is likely that the fundamental phonetic goal for /p/ and /b/, namely, constriction of the lips, is the same but these stops may be differentiated in other ways (see the Introduction). The results provided strong, albeit indirect, support for the hypothesis that the talkers with the greatest English-language experience would close the lips more forcefully for /p/ than /b/ (Belasco, 1953). None of the three groups showed a significant difference in the displacement of the upper lip for /p/ and /b/. Of the three groups, only the native English adults showed significantly greater movement of the lower lip and jaw for /p/ than /b/ (but only in one instance each in the eight possible condition x vowel x rate combinations). Both the English adults and the English children, but not the Chinese adults, moved the lower lip + jaw complex significantly more for /p/ than /b/ at the normal rate following /a/. (Similar results were obtained for net bilabial displacement.) Displacement was measured in this study by tracking the time-varying position of small beads attached to the upper and lower lips. It appears that the relatively greater displacement for /p/ than /b/ occurred because the lips continued moving together longer after the constriction for /p/ than /b/, suggesting more forceful compression of the lips /p/.

An examination of peak movement velocity suggested that an alternative explanation for the /p/ - /b/ displacement difference, namely, that the native English subjects opened the mouth wider during vowels preceding /p/ than /b/ (see Sec. I), is not correct. Of the three subject groups, only the native English adults showed greater peak velocity of movement of the lower lip + jaw complex for /p/ than /b/. The velocity difference, which occurred in the same context as the observed difference in displacement, is just
what one would expect if the lower lip + jaw complex moved more forcefully for /p/ than /b/.

The results suggest that the central representations of subjects in the three groups for /p/ (and possibly /b/) differed. An inference that can be drawn from the study is that the native English adults' central representation specified a more forceful constriction of the lips for /p/ than their representation for /b/. A more forceful constriction does not appear to have been specified in the Chinese subjects' representation for word-final /p/, presumably because they had considerably less experience with it than the native English adults. The pattern of results obtained for the native English children suggested that their central representations were in some sense intermediate to those of the native English and Chinese adults. This is consistent with the fact that they had much less experience producing word-final stops than the English adults, but more experience than the Chinese adults.

One might speculate that fewer English children than English adults had central representations that specified more forceful bilabial closing for /p/ than /b/, or that the parameters leading to the observed kinematic results differentiated /p/ from /b/ less in the children's than adults' representations. Both of hypotheses are consistent with the adult-child differences observed by Flege et al. (1987). They inferred the existence of a gesture(s) that enlarged actively the supraglottal cavity for /b/ but not /p/ in word-final position from an observation of the oral air pressure waveform associated with those stop consonants. The frequency with which English children's /b/ tokens were classified as having been produced with a cavity-enlarging gesture(s) was intermediate to the frequencies for English and Chinese adults.

The conclusion that talkers learn to produce English word-final /p/ more forcefully than /b/ must be qualified in two ways. First, it is unclear why the velocity and displacement differences between /p/ and /b/ that were observed in the context of /a/ were not also seen in the /i/ context. Perhaps the fact that net bilabial displacements were significantly greater for stops following /a/ than /i/ played a role. Second, the level of muscle activity in the stops examined here was not observed directly using electromyography. The level of muscle activity in word-final /b/ and /p/ has not yet been investigated systematically (Fromkin, 1966), so it would be useful to further test the hypothesis that talkers learn to differentiate /p/ from /b/ according to force of constriction by placing surface electrodes on the lips. The peak velocity results presented here and by Smith and McLean-Muse (1988) suggest that the /p/ - /b/ force difference would increase gradually through childhood. The general pattern of results obtained in studies of second-language speech production (see Flege, 1988) suggest that even highly proficient adult non-native speakers may not show higher muscle activity levels for English /p/ than /b/.

C. Stops produced with a bite block

A comparison of stops produced in free-jaw and fixed-jaw conditions revealed another difference between talkers differing in level of English-language experience for stops spoken in the context of /i/. The percentage of net bilabial displacement due to movement of the upper lip and lower lip was calculated for the stops spoken in both conditions. There was a difference between the English adults, on the one hand, and the English children and native Chinese adults, on the other hand. The English adults showed nonsignificant 2% increases in the lower and upper lip percentages when producing stops with a bite block. The native English children and Chinese adults showed reductions in the percentage contribution of the lower lip (4% and 12%, respectively) and significant increases in the upper lip percentages (8% and 12%, respectively).

Lower lip movements for stops following /i/ were reduced by 1.2 mm in the fixed-jaw condition for the Chinese adults. They were not reduced for the English children, and were increased by 0.9 mm for the English adults. It appears that the increases in the percentage of net bilabial displacement due to upper lip movements seen for the English children and Chinese adults was due to a reduced contribution of the lower lip, but it is not clear why they, but not the English adults, showed a lower lip reduction. The between-group differences could not be explained on the basis of differences in the extent to which jaw movements were reduced by the bite block. The percentage of net bilabial displacement due to the jaw was reduced less for the native Chinese than English adults (0.0% vs 3.6%), but it was also less for the English adults than English children (3.6% vs 4.0%).

The results suggest that talkers with the greatest English-language experience (the English adults) were able to compensate more effectively with the lower lip for a small perturbation of the jaw than the talkers with less English-language experience (the English children and especially the Chinese adults). The between-group differences probably derived from how much previous experience talkers in the three groups had with word-final English stops rather than from overall differences in speech motor control. Recall that similar differences were observed between the English adults and children as between the English adults and Chinese adults. It is uncertain why between-group differences were not observed for stops spoken in the /a/ context. It would, therefore, be useful to examine compensation for a bite block that minimally alters the jaw position to determine if the lack of a between-group difference in the /a/ context was due to the extent of the jaw perturbation.

D. Conclusions

The study provided evidence for development of the speech motor skills needed to realize the phonetic contrast between /p/ and /b/. Only the most experienced of three subject groups, the adult native speakers of English, showed greater lower lip displacement and peak movement velocity for /p/ than /b/. It can be inferred that they, but not the Chinese adults and perhaps not the English children, closed the lips more forcefully for /p/ than /b/. This suggests that a more forceful constriction for /b/ than /p/ was specified in the English adults' central phonetic representations, and that this specification was derived from their experience.
with stop consonants in word-final position. The results also suggested that the English adults were able to compensate more effectively for small perturbations of the jaw by a bite block than were the English children and Chinese adults. Taken together, the results suggest that the central representations needed for word-final stop consonants are refined slowly as the result of experience hearing and speaking them.

ACKNOWLEDGMENTS

This research was supported by a grant from the NINCDS (NS-20-572). The author thanks S. Smith for the analysis software, A. Hornsby and J. Byoun for data reduction and analysis, and S. Fletcher for comments on an earlier draft of this article.

1The analyses of upper lip movements yielded significant vowel × condition interactions at the normal and fast rates because the upper lip moved more in the context of /a/ than /i/ in the free-jaw condition but less in the /a/ than /i/ context in the fixed-jaw condition [F(1,27) = 76.2, 49.2].

2The analyses of lower lip movements yielded significant voice × condition interactions at both rates because the lower lip moved slightly more for /p/ than /b/ in the free-jaw but not the fixed-jaw condition [F(1,27) = 7.94].

3The analyses of jaw movements yielded a significant voice × condition interaction at the fast rate because the jaw moved more for /p/ than /b/ in the free-jaw condition (1.6 mm vs 1.4 mm) but, as expected, showed no difference in the fixed-jaw condition [F(1,27) = 7.02]. The difference in jaw displacements that existed as a function of vowel context in the free- but not fixed-jaw condition yielded significant condition × vowel interactions at both speaking rates [F(1,26) = 36.9, 33.6].

4The analyses of peak upper lip velocity yielded significant vowel × condition interactions at both the normal and fast rates because the upper lip moved more rapidly in the fixed- than free-jaw condition in the /i/ context but less rapidly in the /a/ context [F(1,27) = 20.8, 13.7].

5The analyses of peak lower lip velocity yielded significant vowel × condition interactions at the normal rate because the English children and Chinese adults moved the lower lip more in the free- than fixed-jaw condition (7.4 vs 6.6 mm, 8.2 vs 7.0 mm), whereas the English adults showed no difference between the two conditions [F(2,27) = 5.68].

6The analyses of peak lower lip + jaw velocity showed that lower lip velocity was significantly greater in the /a/ than /i/ context in the fixed-jaw condition, whereas the reverse was true in the fixed-jaw condition [F(1,26) = 25.8, 30.2].

7The analyses of peak lower lip + jaw velocity showed that lower lip velocity was significantly greater in the /a/ than /i/ context at both the normal and fast rates (140 vs 122 mm/s, 142 vs 129 mm/s) [F(1,26) = 26.5, 13.9], and significantly more rapid in the fixed- than free-jaw condition (134 vs 129 mm/s, 138 vs 133 mm/s) [F(1,26) = 3.97, 19.5]. A group × condition interaction was obtained at the normal rate because the English adults moved the lower lip more rapidly in the fixed- than free-jaw condition, the English children showed a nonsignificant trend in the opposite direction, and the Chinese adults showed a significant difference in the opposite direction (143 vs 135 mm/s, 121 vs 129 mm/s, 121 vs 137 mm/s) [F(2,26) = 6.55].

8The analyses of peak bilabial closure velocity yielded a significant condition × group interaction at the normal rate because the English adults but not the English children or Chinese adults showed greater velocity in the fixed- than free-jaw condition (210 vs 199 mm/s, 201 vs 196 mm/s, 189 vs 200 mm/s) [F(2,26) = 3.45]. Significant condition × vowel interactions were obtained at the normal and fast rates because velocity was greater in the fixed- than free-jaw condition in the /i/ context, whereas the reverse was true in the /a/ context [F(1,26) = 5.92, 4.15].


