Transfer and developmental processes in adult foreign language speech production

JAMES EMIL FLEGE
University of Alabama in Birmingham

RICHARD D. DAVIDIAN
Northwestern University

ADDRESS FOR CORRESPONDENCE
James E. Flege, Department of Biocommunication, University of Alabama in Birmingham,
Birmingham, AL 35296

ABSTRACT
This study tested the hypothesis that factors that shape children's production of their native language (L1) will also influence adults' pronunciation of sounds in a foreign language (L2). The final stops in CVC English words produced by 12 adult native speakers each of English, Polish, Spanish, and Chinese were phonetically transcribed. The frequency with which these stops were devoiced, deleted, or fricativized was tabulated. The Spanish subjects (unlike the Chinese or Poles) showed the effect of a transfer process, producing word-final /b,d,g/ as fricatives. Subjects in all three non-native groups (but not the native English subjects) resembled English-learning children in devoicing word-final /b,d,g/. Subjects whose L1 does not have word-final stops (i.e., the native speakers of Chinese and Spanish) showed another process commonly observed in English L1 acquisition: final stop deletion. A number of language background variables (e.g., age of arrival in the United States) were found not to be significantly correlated with the accuracy of final stop production. These findings suggest that, in addition to transfer processes arising from L1/L2 phonetic and phonological differences, developmental processes similar to those affecting child L1 speech production also influence adult L2 speech production.

INTRODUCTION
For the past decade the syntactic and morphological errors of adult L2 learners have been categorized in terms of L1/L2 differences, universal constraints, and developmental patterns (e.g., Gass, 1982; Snow, 1981; Wode, 1981). The goal of characterizing the sources of L2 pronunciation errors has lagged notably behind.

This may be due in part to the observation that, superficially at least, the learning of L2 pronunciation does not appear to correspond closely to the learn-
ing of morphology and syntax. For example, while rate of learning of syntax and morphology has generally been found to correlate positively with age and/or general cognitive abilities, the same is not true for pronunciation. Older learners may show no lasting superiority in pronouncing foreign languages compared with younger learners, or else be found to pronounce foreign languages less successfully (see, e.g., Snow & Hoefnagel-Höhle, 1978, 1980). To explain this, researchers have pointed to the need to establish new patterns of perceptual processing or modify pre-established patterns of motor control, to the possible existence of a “sensitive period” for speech learning, and to the powerful affective and social effects of small variations in pronunciation (see Flege, 1984b, for a recent review).

In an important recent article, Macken and Ferguson (1981) considered the pronunciation errors of child L1 and adult L2 learners within the same theoretical framework. Both the child and adult learner replace “model” sounds in a target language (L1 or L2) with incorrect “replica” sounds that are to some extent predictable. Like many other researchers (e.g., Flege & Port, 1981) Macken and Ferguson (1981) posit that “transfer” (interference) arising from structural and/or phonetic differences between L1 and L2 influence the speech of adult L2 learners. (See Locke, 1983, for a discussion of this phenomenon in L1 acquisition.) Transfer is presumably the basis for the recognizably different foreign accents of individuals differing in native language background. In addition, Macken and Ferguson (1981) hypothesized that “developmental” processes similar to those observed in L1 speech acquisition also influence L2 speech production as the result of the “reactivation of L1 strategies and processes” (p. 117; cf. Tarone, 1978).

This hypothesis is important for two reasons. First, Mulford and Hecht (1980) hypothesized that L2 mispronunciations arising jointly from transfer and developmental processes will occur more frequently and persist longer than those arising from just one kind of process. Second, the hypothesis represents an important first attempt to differentiate major sources of L2 pronunciation error.

Transfer and developmental processes

Developmental processes describe a patterned relationship between model sounds in the (mature) target language and replica sounds in learner speech. Like Macken and Ferguson (1981), we consider them to be “probablistic” in nature. Many, but not necessarily all, individual learners will show the effect of various developmental processes during speech learning. Developmental processes do not absolutely constrain the production of sounds found in the phonetic surface of the target language, but those that are widespread among normally developing children seem to stem from “universal” constraints on human memory/cognition, or from innate characteristics of the human speech production mechanism or auditory perception.

Despite their “universal” basis, the same developmental processes will not always be seen in the speech of children learning different native languages for the obvious reason that phonetic inventories and rules of sound combination differ cross-linguistically. So, for example, a child learning a language with only
consonant-vowel syllables will never be observed to reduce consonant clusters or delete final obstruents because the opportunity for doing so does not exist. At the same time, such a child will not be provided the opportunity to learn how to produce clusters and word-final obstruents.

Macken and Ferguson (1981) noted that it may be difficult to distinguish developmental from transfer processes. Anecdotal evidence concerning the word-final English stops produced by an Icelandic boy supports the belief that both types of process may jointly affect a single phone. Mulford and Hecht (1980) found that in addition to devoicing word-final /b,d,g/ — something monolingual English-speaking children often do — this child also produced /b,d,g/ with the strong voiceless aspirated release typical of Icelandic but not English stops.

Another problem is that the same phonetic effect might be the result of either a transfer or developmental process, depending on the talker’s previous linguistic experience. For example, final stop devoicing is probably the result of a developmental process when observed in the speech of children learning English as an L1. For adult German learners of English, on the other hand, final devoicing could conceivably be the result of the transfer of the German rule that devoices voiced obstruents in syllable-final position instead of (or in addition to) a developmental process. This leads to the possibility that a developmental process might be misidentified as a transfer process, or vice versa.

Given this, any attempt to distinguish “transfer” and “developmental” processes in adult L2 speech must be preceded by a clear definition of the two. Macken and Barton (1981) did not define the domain of transfer processes (i.e., interference). Their approach suggests that transfer processes pertain to learned aspects of speech behavior, whereas developmental processes pertain to speech patterns that are unlearned and independent of specific linguistic experience. This is consistent with the generally accepted view, mentioned earlier, that “transfer” is what primarily accounts for the recognizably different accents of L2 learners differing in native language background.

Another possibility is that any difference between L1 and L2, including the absence of L2 elements from L1, might result in “transfer.” In this view, pronunciation of English big as [big] by a native speaker of Spanish would be regarded as a transfer error because Spanish has an /i/ but no /I/ phoneme. This definition suggests that the absence of specific learning in L1 acquisition constitutes an important aspect of an individual’s phonological competence that may be transferred into L2 production.

We take the alternative view that what is most likely to be transferred is what has previously been learned. Our view (see Flege & Hillenbrand, 1984) is that the Spanish speaker’s rendition of big as something like [big] results more importantly from the transfer of a previously learned [i]-quality vowel rather than from an inability to produce an [I]-quality vowel. Support for this comes from the general observation that, in the vowel systems of natural languages, the range of allophones permitted as the realization of a vowel category depends on the number of categories within the system. For example, American English /u/ (unlike the /u/ found in French) is sometimes produced as an [y]-quality vowel because there is no /y/ category in English with which it risks being confused.
(Stevens, 1983). To our ears the /i/ of Spanish appears to encompass parts of the vowel space accorded to the /i/ and /I/ of English, suggesting that Spanish speakers are not incapable of producing an [I]-quality vowel.

The question also arises as to how to define the domain of developmental processes. We define a developmental process as a pattern of model/replica differences commonly observed in the speech of normally developing children that appears to simplify speech output and have a cognitive, physiological, or anatomical basis.

An alternative view is that developmental processes represent a subset of transfer processes. Within the framework of Natural Phonology (Donegan and Stampe, 1979; Stampe, 1969/1979) developmental and transfer processes would not be regarded as distinct. The basic assumption of Natural Phonology is that the phonological system of mature speakers is the residue of innate or "natural" processes that have gone unsuppressed through the period of L1 acquisition because their application is consistent with the elements and contrasts that need to be produced in L1. As noted by Major (1984) processes are defined in Natural Phonology to include not only phenomena operating to produce surface alternations, but also phenomena that generate underlying forms and may apply vacuously, thereby acting as surface phonotactic constraints. For example, adult native speakers of Portuguese – a language with no voiced word-final stops – might be regarded as devoicing English /b,d,g/ because an (unsuppressed) devoicing rule is transferred into English, rather than because a devoicing process is "reactivated" when adult speakers of Portuguese first face the need to produce word-final /b,d,g/ in English.

The primary problem with this view is that it precludes an empirical investigation of the effects of transfer and developmental processes by denying a priori the possibility that the two kinds of processes are distinct. Our primarily functional definition of developmental processes posits that to be considered a developmental process, a pattern of model/replica differences must be observed, and that it is the effect of learning to narrow model/replica differences that is transferred. This definition is consistent with the belief that, for example, the Spanish speaker who trills English /r/ does so for a reason different from that of a French learner who mispronounces the initial fricative in the English word "the."

**Empirical evidence**

It seems reasonable to suppose that the developmental processes that shape L1 acquisition would also affect children’s pronunciation of a foreign language. However, there are two a priori reasons to question whether developmental processes will also affect the foreign language speech production of adults. First, adults presumably have a superior knowledge of their vocal tract and its capabilities than children, as well as a generally greater degree of neuromotor control (Kent, 1976). This may make it easier for adults than for children to produce new sounds in a foreign language. Second, adults have presumably mastered the phonology of their native language to a greater extent than have children. This may make it more likely for adults than for children to transfer aspects of L1 phonology into L2 production. Given the widespread existence
of foreign accent in adult L2 production (see Flege, 1985), it is conceivable that transfer processes may exist to the exclusion of developmental processes in adult L2 speech learning.

Given this, how could it be empirically demonstrated that transfer and developmental processes independently affect adult L2 pronunciation? The first and perhaps most obvious way would be to show that adults learning L2, but not adult native speakers of L2, show some of the same developmental processes as children learning L2 as their native language. Some such evidence has been provided in recent studies showing the existence of final stop devoicing in the speech of adult native speakers of Vietnamese (Gladstein, Membrino, & Flege, 1982; cf. Donegan & Stampe, 1979), Spanish (Eckman, 1981) and Arabic (Flege & Port, 1981). However, these studies and others like them fail to provide convincing evidence of developmental processes in L2 learning for several reasons.

In the case of Vietnamese learners, at least, it could reasonably be argued that what appears to be a developmental process (i.e., the devoicing of /b,d,g/ in word-final position) is actually the result of a transfer process. Vietnamese words do not end in voiced stops but may be terminated by a voiceless unreleased stop. Thus in realizing English /b,d,g/ as phonetically voiceless stops, Vietnamese learners may be substituting the closest possible Vietnamese sound for English /b,d,g/ (i.e., an unreleased /p,t,k/) rather than devoicing /b,d,g/. If so, their mispronunciation could be regarded as the result of transfer.

Another problem of past studies is the lack of adequate English control groups against which to compare the speech performance of non-native speakers. For example, previous research (e.g., Flege, 1982) indicates that children may require many years to establish mature control of the contrast between word-final /b,d,g/ and /p,t,k/, and that even adult native English speakers tend to devoice /b,d,g/ in casual speech (Shockey, 1974). Thus to document a devoicing process in adult L2 production it would not be necessary to show exceptionless behavior, but it would be necessary to show a substantially higher frequency for non-native than for native speakers.

Another shortcoming of some previous studies is the failure to demonstrate the unidirectionality of a process. For example, Flege and Port (1981) found that 12 Saudi learners devoiced word-final /b/ 22% of the time. The same talkers were also perceived to “voice” word-final /p/ 22% of the time. Given the absence of a /p/ phoneme in Arabic and the lower rate of voicing errors for /d/ and /g/, it is possible that the voicing errors observed for /p-b/ were due to confusion of the /p-b/ categories rather than to a developmental process.

Since developmental processes ultimately derive from universal characteristics of human talkers, we assume that they are independent of previous linguistic experience. Based on this assumption, we expect to observe the same developmental processes in the L2 speech of learners of disparate native language backgrounds. Previous studies, however, have not convincingly demonstrated the occurrence of a process across multiple groups of L2 learners. For example, if developmental processes are “reactivated” we would expect Chinese and Spanish learners of English to devoice final stops because neither Chinese nor Spanish has voiced stops in word-final position. Eckman (1981), however, found that Spanish but not Chinese L2 learners devoiced word-final stops.
A second way to demonstrate the existence of developmental processes in adult L2 production would be to show that an L2 element or contrast is deleted or neutralized in accordance with L2 developmental patterns even though L1 possesses a closely similar element or contrast. Evidence for this would be provided, for example, if adult French learners were to neutralize the contrast between word-final English /p,t,k/ and /b,d,g/ despite the existence of a contrast between voiced and voiceless stops in French. Another example of such evidence would be the deletion of word-final /p,t,k/ by native speakers of languages (like Vietnamese) that possess voiceless stops in final position.

Evidence for developmental processes might also be provided in the absence of deletions or the neutralization of an L2 contrast. For example, English-learning children may realize English /r/ as a [w]-like phone. In mature Spanish /r/ is realized as a tap or trill but never, to our knowledge, as a [w]-like phone. The expected transfer process is for Spanish adults to realize English /r/ as a tap or trill. Evidence of a [w]-for-/r/ substitution in Spanish-accented English would strongly support the hypothesis that developmental processes are reactivated in L2 learning.

Evidence of this kind has in fact been reported by Wode (1977). Several German children aged 4–9 years were observed, at least briefly, to realize English /r/ as a [w]-like phone rather than as the uvular [ʁ] of their native language. However, Wode’s study did not provide conclusive evidence both because of its anecdotal nature and because the child L2 learners studied may have been reproducing [w]-for-/r/ substitutions heard in the speech of native-English age-mates. Moreover, what is observed in child L2 learning may not typify adult L2 learning.

The direct observation of developmental processes in adult L2 production, although important, would not in itself convincingly demonstrate that transfer and developmental processes independently affect adult L2 pronunciation. However, we can think of two general ways for demonstrating the hypothesized independence.

The first approach involves an examination of the speech of a single group of L2 learners. One could demonstrate that L2 phones undergo the joint influence of a developmental and transfer process. To provide this kind of evidence, it would be necessary to show that an L2 element or contrast is deleted or neutralized according to an L2 developmental pattern (even though L1 does not possess an analogous element or contrast) and according to an L1 phonological pattern (even though this pattern is not present in mature native L2 speech).

The second approach involves the comparison of multiple groups of L2 learners. One could show that the frequency of a deletion or neutralization in L2 speech (e.g., production of /b,d,g/ as [p,t,k]) is greater for learners in whose native language the same phenomenon results from the application of a phonological rule than it is for learners whose L1 has no such rule but also lacks [b,d,g] in word-final position. The second approach is based on the assumption that native speakers of both kinds of language would begin learning English without previous experience producing [b,d,g] in word-final position. If developmental processes are reactivated and affect adult L2 speech production, the developmental process of devoicing should affect the speech of learners of both
L1 backgrounds equally. If a learned L1 devoicing rule is also transferred into L2 production, learners whose L2 possesses a devoicing rule should devoice more frequently than learners whose L1 does not.

Developmental processes

To establish that a model/replica pattern in adult L2 production is the result of a developmental process, it is first important to establish that the pattern in question truly represents a developmental process in the speech of children learning L2 as a native language. With this in mind, we will describe two developmental processes affecting the production of word-final stop by English-learning children, as well as a model/replica difference pattern that does not occur with sufficient frequency to be regarded as a developmental process.

Final stop deletion. Children learning English are frequently observed to produce consonant-vowel-consonant (CVC) words in such a way that the final consonant is not perceptually evident. Like other developmental processes, this process may be said to reduce the phonetic “complexity” of words (Ingram, 1976; Macken & Ferguson, 1981). The CV syllable is often considered to be the easiest of syllables to produce because of its early appearance in babbling and its widespread occurrence in human languages (Tarone, 1981). Many children who are learning English as their native language delete final stops, at least during the earliest stages of speech learning. Normally developing children generally succeed in producing word-final stops by the age of 2–3 years (Smith & Stoel-Gammon, 1983), although the deletion of final stops may persist longer in pathological or speech-delayed populations (Hodson & Paden, 1981; Smith & Stoel-Gammon, 1983; Weismer, Dinnsen, & Elbert, 1981).

Final stop devoicing. English-learning children are frequently observed to produce word-final /b,d,g/ so that they are perceived as /p,t,k/. Final stop devoicing appears to be greatly reduced in frequency by the age of 3–4 years (Smith & Stoel-Gammon, 1983), but may persist somewhat longer in the speech of children whose speech development is noticeably delayed (Hodson & Paden, 1981; Ingram, 1981). Like final consonant deletion, final consonant devoicing may be viewed as resulting in a less “complex” (or “marked”) sound. The aerodynamic difficulty inherent in maintaining the transglottal pressure difference needed for glottal pulsing in obstruents is often cited as a cause of final devoicing. Consistent with this is the observation that many human languages have voiceless but not voiced stops in word-final position, final devoicing appears in the speech of certain types of aphasic individuals, and more final voiceless than voiced final stops are heard in the babbled syllables of prelinguistic infants (MacNeilage, Hutchinson, & Lasater, 1981).

Fricativization of stops. One model/replica pattern that is much less commonly observed in L1 speech development is the realization of stops as fricatives. Anterior fricatives are seldom observed in babbling and are often acquired relatively late – usually after homorganic stops – because they may pose special articulatory and perceptual difficulties for young learners. For this reason Jakob-
son (1968) posited that the stopping of fricatives is a universal process affecting L1 acquisition. The realization of fricatives as stops has frequently been observed in the speech of children learning English, although this process appears to be more apt to affect word-initial than word-final fricatives (Edwards, 1979; Ferguson, 1975; Ingram et al., 1980). Smith (1973) reports that an English-learning child realized voiceless final fricatives as voiced stops (e.g., “bud” for “bus”), but did not seem to show the reverse pattern, i.e., the realization of stops as fricatives. Although the fricativization of stops has been observed in some children (e.g., Ferguson, 1975), it appears to be much less common than the stopping of fricatives and is not regarded as a developmental process affecting English speech production by young children (see, e.g., Ingram, 1976).

Cross-language differences

We chose to study the English produced by native speakers of Spanish, Polish, and Chinese because phonological differences between these languages and English permitted us to test the hypothesized role of transfer and developmental processes as outlined earlier. We focused on production of word-final stop consonants because it has been hypothesized (Mulford & Hecht, 1980; cf. Macken & Ferguson, 1981, p. 118) that stops may be more likely than other phone classes to exhibit the dual influence of developmental and transfer processes.

From the perspective of the hypothesis being tested, the most important difference among Polish, Spanish, and Chinese concerns the variety of phones permitted to occur in word-final position (Brady, 1975; Cheng, 1973; Harris, 1969). A wide variety of consonants are permitted in the final position of Polish words, including nasals, alveolar and palatal fricatives, affricates, and stops. Most Spanish words, on the other hand, end in a vowel. The consonant phones permitted to occur in the final position of Mexican Spanish words are more restricted. These include a lateral (/l/), taps and trills, nasals, and dental or slightly retroflexed fricatives. The stop consonants /p,t,k/ and /b,g/ do not occur in native words, but a few may be found in the final position of loanwords and proper names of Catalonian origin. The stop /d/ is ordinarily realized as a voiced or voiceless fricative. Chinese permits even fewer word-final consonants than Spanish. The few CVC syllables that do exist end in /r/ or a nasal consonant.

Polish and Spanish possess phonological rules that might affect native speakers’ production of English (see Port & Mitleb, 1980). Polish is analyzed as having a phonological rule that devoices obstruents in word-final position. This rule applies without exception to voiced stop phonemes occurring in absolute utterance-final position. Its application is more restricted for word-final stops occurring in utterance-medial positions.

Harris (1969, p. 44) proposed a rule for Spanish that devoices obstruents occurring in utterance-final position. However, given the limited occurrence of voiced obstruents in word-final position, this rule has more restricted application than the Polish devoicing rule, and may not apply to all dialects of Mexican Spanish. A much more productive rule of Spanish is one that results in the realization of the stop consonants /b,d,g/ as homorganic fricatives. As formulated by Harris (1969, p. 40), this rule applies to stops found in intervocalic
position and to word-initial stops following a word that terminates in /l/ or a nasal. It also appears that this rule may apply to the few words ending in voiced stops (e.g., "sed," "club," or "zigzag"). Recent evidence (Macken, 1980) also suggests, contrary to Harris' formulation, that the rule may apply to word-initial tokens of /b,d,g/ occurring in utterance-initial position.

In summary, native speakers of Polish, Spanish, and Chinese bring to English widely divergent experience in producing stops in word-final position. Children learning Chinese need not learn to terminate words with stop consonants, and Polish children need learn only to produce voiceless stops in word-final position. The situation is less clear for children learning Mexican Spanish, where stops appear to have marginal status. Native words do not end in /p,t,k,b,g/, but do end in /d/. All of these stops, however, appear to be subject to rules of fricativization and possibly devoicing, so it is probably the case that monolingual native speakers of Mexican Spanish do not learn to produce stops in word-final position.

Specific hypotheses

If both transfer and developmental processes affect adult L2 production, we would expect to observe the following differences among the four subject groups:

H1: Native speakers of Spanish will realize word-final stops as fricatives more frequently than native speakers of English, Polish, or Chinese. If this hypothesis is supported it would demonstrate an effect of transfer, especially if fricativization affected /b,d,g/ but not /p,t,k/ (as specified by the rule of Spanish; Harris, 1969). In testing this hypothesis, the English, Polish, and Chinese groups will all serve as control groups. We do not expect to see fricativization in the speech of native speakers of these languages because none of them, like Spanish, possess a fricativization rule.

H2: Native speakers of Polish, Spanish, and Chinese will devoice /b,d,g/ more frequently than native speakers of English. If supported, this would demonstrate that the developmental process of final stop devoicing affects adult L2 speech production.

H3: Native speakers of Chinese, and perhaps Spanish, will delete final stops more frequently than native speakers of Polish or English. This would demonstrate that the developmental process of final stop deletion affects L2 production. In testing this hypothesis, both the Polish and English groups serve as control groups since both languages possess word-final stops.

H4: Native speakers of Polish will devoice /b,d,g/ significantly more often than speakers of Chinese, and perhaps Spanish. If supported, this would support the hypothesis that transfer of a devoicing rule and the developmental process of devoicing jointly affect adult L2 production.

H5: Native speakers of Spanish will both fricativize and devoice /b,d,g/. If supported, this would also demonstrate the joint effect of transfer and developmental processes.
METHODS

Subjects

Subjects for this study were 12 normal adult native speakers each of English, Spanish, Chinese, and Polish. The native English speakers were all undergraduate students at Northwestern University in Chicago.

Talkers in the three non-native groups all satisfied the selection criterion of speaking fluent but obviously accented English. The Spanish and Polish native speakers were recruited from intermediate-level ESL classes at a local community college. The Chinese subjects were located through personal contact by the first author. This group included native speakers of Mandarin (2), Taiwanese (9), and Haka (1). The Taiwanese and Haka native speakers all reported they had learned Mandarin at school in Taiwan.

In a study such as this it would be desirable to ensure that all non-native subjects have similar motivation and opportunity to learn English, as well as similar aptitude for learning to pronounce foreign languages. Unfortunately, we knew of no instrument that could be used effectively to evaluate subjects for these factors, nor could we think of a way accurately to assess the amount of subjects’ “intake” (meaningful use) of English.

We did, however, use a language background questionnaire to provide several gross measures of linguistic experience. This information is summarized in Table 1. The Spanish group consisted of 11 Mexicans and 1 Salvadoran. As shown in the table, they arrived in the United States at an average age of 20 years and had lived there for an average of 7 years. The Poles arrived in the United States somewhat later (at an average age of 34), and had spent only about 3 years there. The Chinese subjects entered the United States at an average age of about 26 years, and had lived there about 5 years at the time of the study. According to self-report, the Poles used English somewhat more frequently on a daily basis than either the Spanish or Chinese subjects, and the Chinese subjects had more formal English instruction than did subjects in the other two non-native speaker groups. No exact information concerning the substance or quality of the non-native subjects’ English instruction is available.

Speech material

Each talker produced the following English words:

- cub, pod, bag, bug, dog
- cup, pot, back, buck, doc

These minimal pair test words were chosen primarily for their picturability. The list included three /k-g/ pairs but only one /p-b/ and /t-d/ pair in order to maximize the likelihood that devoicing would be observed if it does affect L2 production. Previous research (e.g., Smith, 1978) suggests that velar stops may be devoiced more frequently than labial or alveolar stops for aerodynamic and anatomical reasons.

Data were elicited by five randomized presentation of line drawings representing the 10 test words. The subjects were instructed to produce the test words both
Table 1. *Factors relating to non-native subjects' experience with English*<sup>a</sup>

<table>
<thead>
<tr>
<th>Subject</th>
<th>Sex</th>
<th>Age</th>
<th>Age of arrival</th>
<th>Years in U.S.</th>
<th>Months of English</th>
<th>Use of English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chinese</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>M</td>
<td>32</td>
<td>24</td>
<td>7.0</td>
<td>63</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>M</td>
<td>28</td>
<td>26</td>
<td>2.2</td>
<td>72</td>
<td>7</td>
</tr>
<tr>
<td>3</td>
<td>M</td>
<td>34</td>
<td>19</td>
<td>15.0</td>
<td>90</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>M</td>
<td>29</td>
<td>28</td>
<td>1.1</td>
<td>72</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>F</td>
<td>26</td>
<td>26</td>
<td>0.5</td>
<td>90</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>F</td>
<td>20</td>
<td>17</td>
<td>3.0</td>
<td>27</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>M</td>
<td>29</td>
<td>23</td>
<td>6.0</td>
<td>63</td>
<td>5</td>
</tr>
<tr>
<td>8</td>
<td>M</td>
<td>36</td>
<td>35</td>
<td>0.6</td>
<td>90</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>F</td>
<td>31</td>
<td>30</td>
<td>1.2</td>
<td>63</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>F</td>
<td>34</td>
<td>26</td>
<td>6.0</td>
<td>54</td>
<td>3</td>
</tr>
<tr>
<td>11</td>
<td>M</td>
<td>31</td>
<td>29</td>
<td>1.6</td>
<td>72</td>
<td>7</td>
</tr>
<tr>
<td>12</td>
<td>M</td>
<td>24</td>
<td>24</td>
<td>8.2</td>
<td>63</td>
<td>5</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>30</td>
<td>26</td>
<td>3.7</td>
<td>68</td>
<td>5.8</td>
</tr>
<tr>
<td>Polish</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>M</td>
<td>36</td>
<td>35</td>
<td>1.0</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>M</td>
<td>28</td>
<td>27</td>
<td>0.8</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>F</td>
<td>38</td>
<td>35</td>
<td>3.0</td>
<td>24</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>M</td>
<td>46</td>
<td>41</td>
<td>5.0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>M</td>
<td>52</td>
<td>49</td>
<td>2.5</td>
<td>24</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>M</td>
<td>44</td>
<td>43</td>
<td>0.8</td>
<td>27</td>
<td>7</td>
</tr>
<tr>
<td>7</td>
<td>F</td>
<td>38</td>
<td>32</td>
<td>6.0</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>8</td>
<td>M</td>
<td>42</td>
<td>29</td>
<td>13.0</td>
<td>54</td>
<td>7</td>
</tr>
<tr>
<td>9</td>
<td>F</td>
<td>41</td>
<td>38</td>
<td>2.5</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>F</td>
<td>32</td>
<td>30</td>
<td>2.0</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>11</td>
<td>F</td>
<td>26</td>
<td>23</td>
<td>3.0</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>12</td>
<td>M</td>
<td>23</td>
<td>22</td>
<td>0.9</td>
<td>148</td>
<td>6</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>37</td>
<td>34</td>
<td>3.4</td>
<td>27</td>
<td>5.4</td>
</tr>
<tr>
<td>Spanish</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>M</td>
<td>30</td>
<td>25</td>
<td>4.5</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>M</td>
<td>24</td>
<td>15</td>
<td>9.0</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>3</td>
<td>M</td>
<td>35</td>
<td>25</td>
<td>10.0</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>M</td>
<td>18</td>
<td>17</td>
<td>1.0</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>M</td>
<td>25</td>
<td>16</td>
<td>9.0</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>6</td>
<td>F</td>
<td>29</td>
<td>25</td>
<td>4.0</td>
<td>38</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>M</td>
<td>48</td>
<td>27</td>
<td>21.0</td>
<td>28</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>F</td>
<td>27</td>
<td>10</td>
<td>17.0</td>
<td>48</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>M</td>
<td>19</td>
<td>15</td>
<td>4.0</td>
<td>54</td>
<td>6</td>
</tr>
<tr>
<td>10</td>
<td>M</td>
<td>25</td>
<td>19</td>
<td>6.0</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>11</td>
<td>M</td>
<td>26</td>
<td>24</td>
<td>2.0</td>
<td>18</td>
<td>4</td>
</tr>
<tr>
<td>12</td>
<td>F</td>
<td>28</td>
<td>27</td>
<td>1.0</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>28</td>
<td>20</td>
<td>7.4</td>
<td>19</td>
<td>4.8</td>
</tr>
</tbody>
</table>

<sup>a</sup> "Age of arrival" refers to the age at which subjects arrived in an English-speaking environment; "months of English" refers to formal English language instruction; "use of English" refers to subjects' estimation on a 10-point rating scale of their daily use of English.
in isolation and at the end of a short carrier phrase on each trial. The subjects were required to produce an identifying number before each utterance to avoid later confusion concerning the intended test word. For example, in response to a picture of “back” they were to respond: “Number Seven [pause] Back [pause] Now I say back.” The subjects were encouraged to repeat test words that did not meet with their satisfaction.

The experimenter attempted not to say the test words before and during data collection in order to avoid providing subjects with an overt model. If a test word was unfamiliar, the subject was simply told the spelling of that word. The subjects needed to demonstrate the ability to name all 10 test words without hesitation before data collection began. Awareness of the phonological identity of word-final stops was periodically tested during data elicitation by asking subjects to spell the test word they had just spoken. They always succeeded in correctly spelling the test words, suggesting they were aware of the place and voicing contrasts present in word-final position.

The speech material was recorded on high quality equipment (Sony Model TCD5M) for later analysis. A constant mouth-to-microphone distance was ensured for the English, Spanish, and Polish subjects by using a head-set microphone (Shure Model SM12A), and for the Chinese subjects by holding a microphone (Nakamichi Model CM-300) at a constant distance of about 15 cm from the talker’s mouth.

**Analysis**

Two trained listeners (JEF and RDD) independently transcribed stops in the final 40 utterances produced by the English, Polish, and Spanish subjects; only one judge (JEF) transcribed the speech of the Chinese subjects. The speech material was presented to these listeners at a comfortable level via a loudspeaker. After an initial transcription, the listener(s) replayed the tape several times until satisfied that each stop had been correctly transcribed.

This provided a total of 80 word-final stops (10 words × 4 repetitions × 2 positions) for each of the 48 talkers tested. Based on these transcriptions, the final consonant in each word was tabulated as (1) present or omitted, (2) voiced or voiceless, and (3) continuant or noncontinuant.

Subjectively, we found that it was sometimes difficult to distinguish unreleased stops from stops that had been deleted. Given the nature of the hypothesis being tested, we adopted the most conservative possible criterion for judging a stop to have been deleted. If the termination of a word gave any indication of place of articulation (including glottal place), a consonant was considered to be present in final position. This criterion may have led us to underestimate somewhat the frequency of final stop deletion, for untrained listeners would probably have been less discerning.

It was sometimes also difficult to judge whether final stops were phonetically voiced or voiceless. The non-native subjects, especially the Poles, sometimes produced final stops that seemed to be phonetically intermediate to the voiced and voiceless stop categories produced by the native English subjects. At times, they appeared to produce /b,d,g/ with the syllable termination and voicing
characteristics of English \([b,d,g]\) but with the voiceless aspirated release associated with English \([p,t,k]\). This observation is consistent with the view of some phonologists that Polish has a devoicing rule that "neutralizes" the distinction between \(/b,d,g/\) and \(/p,t,k/\).

Despite these difficulties, the transcriptions appear to have been adequately reliable. There was nearly 100% agreement on all three features between the two listeners for the native English speakers. These subjects realized \(/b,d,g/\) and \(/p,t,k/\) with the correct voicing and continuancy feature in nearly every instance. There was 91% agreement for the Polish subjects after the first listening, and 97% after further listening. For stops produced by the Spanish subjects, the two listeners agreed 65% of the time after an initial listening, and 89% of the time after further listening. Most disagreements concerning stops produced by the native Spanish speakers pertained to the continuancy feature.

In those few instances (about 135, or 5%) when the two judges continued to disagree, the judgment of a third trained listener who was unaware of the hypotheses being tested was accepted.

To estimate the reliability of transcriptions made for the Chinese talkers, the single listener (JEF) transcribed 234 randomly chosen stops two weeks after the first transcription. These stops were given the same phonetic label 91% of the time.

Previous research (Eckman, 1981; Tarone, 1980) suggests that L2 learners whose LI does not possess word-final stops may insert a vowel after a word-final stop, thereby creating a CVCV form. In only one instance out of the approximately 3,840 test words examined was a vowel perceived to have been inserted after the word-final stop. There will therefore be no further discussion of para-gogic vowel insertion.

The dependent measure of this study was the percentage of times each talker devoiced, deleted, or fricativized word-final stops. Preliminary analyses indicated that there were no significant differences in the frequency of these phonetic effects either as a function of place of articulation or position (i.e., isolated words versus words occurring at the end of the carrier phrase). The data for each subject have therefore been pooled. Stops could be deleted a maximum of 80 times and fricativized a maximum of 80 times (provided none were deleted). Voiced stops could be devoiced a maximum of 40 times (5 test words \(\times\) 4 productions \(\times\) 2 positions), provided none were deleted.

The significance of between-group differences in the frequency of the three phonetic effects were tested by a nonparametric ANOVA based on ranks (Kruskal-Wallis, alpha = .001). The Mann-Whitney U test (alpha = .01) was used to make pairwise comparisons between each of the four subject groups when the effect of Subject Group reached significance.

RESULTS

Transcriptions

Figure 1 displays the average frequency with which subjects in the four groups deleted, devoiced, or fricativized word-final stops. Data for individual subjects are presented in Table 2.
Table 2. The frequency with which individual native speakers of English, Polish, Chinese and Spanish deleted, devoiced, and fricativized the final stops in CVC English words

<table>
<thead>
<tr>
<th></th>
<th>Deletion</th>
<th></th>
<th></th>
<th>Devoicing</th>
<th></th>
<th></th>
<th>Fricativization</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>82</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>97</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>50</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
<td>82</td>
<td>70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td></td>
<td>40</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td></td>
<td>84</td>
<td>41</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>4</td>
<td></td>
<td>100</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>x</td>
<td>0%</td>
<td>2.3%</td>
<td>3.4%</td>
<td>0%</td>
<td>48.3%</td>
<td>29.5%</td>
<td>43%</td>
<td>0.3%</td>
</tr>
</tbody>
</table>
The frequency of stop deletion is displayed in the left panel of Figure 1. None of the native English or Polish subjects consistently deleted final stops. According to our conservative criterion (see above), subjects in the Spanish group deleted a total of 31 (3%) word-final stops, ranging from 0–1% (for six subjects) to 10–14% (for two subjects). Subjects in the Chinese group deleted a total of 20 (2%) final stops, ranging from 0–1% (for six subjects) to 5% (for four subjects).

As predicted, the subjects whose native language does not possess word-final stops (i.e., the Chinese and Spanish subjects) deleted word-final stops more frequently than did subjects whose native language does possess final stops (i.e., the English and Polish subjects). The effect of the factor Subject Group on the frequency of final stop deletion was significant ($H = 18.2, df = 3, p < .001$). The Chinese and Spanish subjects deleted stops more often than the American and Polish subjects, but there was no difference between the Chinese and Spanish subjects ($p < .01$).

The frequency of stop devoicing is displayed in the middle panel of Figure 1. None of the stops produced by the English subjects were perceived to have been devoiced. However, subjects in all three non-native groups were perceived to devoice /b,d,g/. The Polish subjects did so 48% of the time, the Spanish subjects 43% of the time, and the Chinese subjects 30% of the time.

Statistical tests confirmed the prediction that subjects in the three non-native groups would devoice /b,d,g/ more frequently than the native English speakers. The effect of the factor Subject Group on the frequency of devoicing was significant ($H = 26.2, df = 3, p < .001$). Pairwise between-group comparisons indicated that all three groups of non-native subjects devoiced /b,d,g/ more frequently than the native English speakers ($p < .01$), and that there was no difference between any of the non-native subject groups. This last finding failed to confirm the hypothesis that the Polish (and possibly Spanish) subjects would devoice /b,d,g/ more frequently than the Chinese subjects as the result of the joint effect of a transfer and developmental process.
A consideration of individual data (Table 2) indicates that final stop devoicing was by no means an inevitable consequence of language background. There was a wide range of variability among individual subjects in each speaker group. Two subjects in the Chinese group, six in the Polish group, and five in the Spanish group devoiced /b,d,g/ more than 50% of the time. There were four subjects in the Polish group and four in the Chinese group who devoiced /b,d,g/ less than 10% of the time.

The mean frequency with which subjects in the four groups realized /b,d,g/ and /p,t,k/ as fricatives is displayed in the right panel of Figure 1. The Spanish subjects fricativized word-final stops about 19% of the time, ranging from 0–5% (for four subjects) to 44% (one subject). Subjects in the English, Polish, and Chinese groups, on the other hand, fricativized word-final stops less than 2% of the time.

The effect of the factor Subject Group on the frequency of fricativization was significant ($H = 18.5, df = 3, p < .001$). The Spanish group fricativized stops more frequently than the English, Polish, and Chinese groups ($p < .01$). This confirms the hypothesis that the Spanish subjects would fricativize word-final stops more frequently than subjects in each of the other three groups as the result of a transfer process.

One other aspect of the data supports the interpretation that the fricativization of stops by the Spanish subjects was due to transfer from Spanish rather than to some other source of error. According to Harris (1969, p. 40) the rule that fricativizes stops in Spanish applies to voiced but not voiceless phonemes. The Spanish talkers in this study fricativized English /b,d,g/ 158 times, as against only 13 times for /p,t,k/. The frequency of fricativization was significantly greater for voiced than voiceless stops ($p < .001$ by chi-square analysis).

**Untrained listeners**

One reasonable question to ask is whether the transcriptions of the two trained listeners correspond to the perception of listeners without training in phonetics. To determine this, we examined perception by untrained listeners of the voicing feature of final stops.

Three of the sentence-final stops produced by 8 randomly selected English, Polish, and Spanish subjects were presented to 10 untrained listeners, all monolingual English-speaking college students who were paid for their services. These untrained listeners were to judge the voicing feature of word-final stops. They indicated their choice by circling one of two possible responses on a specially prepared answer sheet (e.g., “bag” or “back”).

The results of this listening test are presented in Table 3. These results again demonstrate the non-native subjects’ difficulty in producing perceptually effective voiced stops in the final position of English words. The voicing feature of the final stops produced by the native English subjects was correctly identified 98% of the time. The voicing feature of stops produced by the Poles was correctly identified only 67% of the time, and those of the Spanish subjects, only 70% of the time. Most misidentifications pertained to /b,d,g/ rather than /p,t,k/.

A correlational analysis revealed that judgments of the untrained listeners closely agreed with judgments made by the two trained listeners. The percentage
Table 3. The frequency with which 10 untrained English-speaking listeners correctly identified the voicing characteristic of final stops (/b,d,g,p,t,k/) in English CVC words produced by 8 native speakers each of English, Polish, and Spanish.\(^a\)

<table>
<thead>
<tr>
<th>Speaker</th>
<th>cup</th>
<th>cub</th>
<th>pot</th>
<th>pod</th>
<th>doc</th>
<th>dog</th>
<th>buck</th>
<th>bug</th>
<th>back</th>
<th>bag</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>30</td>
<td>29</td>
<td>30</td>
<td>30</td>
<td>29</td>
<td>30</td>
<td>30</td>
<td>28</td>
<td>30</td>
<td>28</td>
<td>30</td>
</tr>
<tr>
<td>3</td>
<td>30</td>
<td>30</td>
<td>27</td>
<td>30</td>
<td>28</td>
<td>30</td>
<td>28</td>
<td>27</td>
<td>25</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>5</td>
<td>28</td>
<td>29</td>
<td>28</td>
<td>28</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>20</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>6</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>28</td>
<td>30</td>
<td>20</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>7</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>28</td>
<td>30</td>
<td>30</td>
<td>29</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>8</td>
<td>29</td>
<td>30</td>
<td>25</td>
<td>30</td>
<td>24</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>11</td>
<td>27</td>
<td>30</td>
<td>29</td>
<td>30</td>
<td>29</td>
<td>30</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
</tr>
<tr>
<td>12</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>28</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>234</td>
</tr>
<tr>
<td>Polish</td>
<td>218</td>
</tr>
<tr>
<td>Spanish</td>
<td>191</td>
</tr>
</tbody>
</table>

\(^a\)The maximum number of correct responses was 30 (3 repetitions \(\times\) 10 listeners).

The maximum number of correct responses was 30 (3 repetitions \(\times\) 10 listeners). of times the final stops produced by the subset of 24 English, Polish, and Spanish subjects were judged by the untrained listeners to be “voiced” was calculated. These scores were correlated with the percentage of times the same stops had been transcribed as “voiced” by the two trained listeners. The correlation was highly significant (\(R = .921, p < .001\)), indicating that results of the transcriptional analysis were generalizable to other native English-speaking listeners.

**Effect of language background**

We found that the Polish subjects did not devoice /b,d,g/ more frequently than the Chinese subjects. One possible reason this hypothesis was not confirmed is
that the native speakers of Polish may have made greater progress in learning to pronounce English than the Chinese subjects, thereby removing a between-group difference that might have been observed earlier. This could have been the case if, for example, the Poles had greater motivation, opportunity, or aptitude for learning L2 pronunciation.

The non-native subjects in this study were all normal adults who had studied English in elementary and/or high school, immigrated to the United States in late adolescence or adulthood, and spoke fluent (albeit accented) English. It should be pointed out, however, that these subjects were selected largely on the basis of availability. As a result, it should be apparent from Table 1 that subjects in the three non-native groups were not perfectly matched. For example, the Chinese talkers arrived in the United States at an earlier age than the Poles, and had received more formal instruction in English.

A multiple correlation analysis was performed to determine whether these quantifiable differences between subject groups affected the results presented earlier. The frequency with which each of the 36 non-native speakers were perceived (by the trained listeners) to devoice /b,d,g/ was correlated with the language background facts summarized in Table I. All other things being equal, we might expect a relatively low rate of devoicing from subjects who were relatively young, had arrived in an English-speaking environment at a relatively early age, had a great deal of formal instruction in English, and/or spoke English relatively often according to self-report.

The multiple correlation did not confirm these hypotheses. The coefficient of multiple correlation was .37 (df = 5,31), which accounted for only a nonsignificant 14% of the variance. This suggests that the between-group disparities evident in Table 1 were not responsible for the between-group differences noted earlier. This result also suggests that such gross estimates of L2 experience may not be a good predictor of L2 pronunciation success.

DISCUSSION

This study confirmed the hypothesis (Macken & Barton, 1981) that developmental processes are “reactivated” when adults attempt to produce L2 sounds not found in their native language. Production of word-final English stops by adult L2 learners showed the influence of two developmental processes – final stop devoicing and final stop deletion – which affect the speech of children learning English as their native language.

It was hypothesized that native speakers of Polish, Spanish, and Chinese would devoice English /b,d,g/ because, in learning their native language, they had not previously learned to do so. We found that subjects in the Chinese, Polish, and Spanish groups devoiced word-final /b,d,g/ significantly more often than subjects in the native English-speaking control group. It was hypothesized that subjects in the Spanish and Chinese groups would delete word-final English stops more often than subjects in the Polish and English control groups because Spanish possesses very few stops in final position and Chinese has none. This hypothesis was also confirmed.

Many children learning English as a native language have been observed to
delete and devoice word-final stops. These patterns of model/replica differences are so common they are accepted as ‘‘developmental’’ processes. It seems reasonable to think that the production of word-final stops, especially voiced ones, requires the establishment of certain motor skills once cognitive and/or sensory targets have been established. Although the non-native subjects’ level of skill acquisition cannot be directly discerned from the present results, it appears that many of them had not yet established the same degree of mastery in producing final stops as adult native speakers of English.

The present study also confirmed the hypothesis that the Spanish subjects would realize stops as fricatives more frequently than subjects in the other three groups. We concluded that this substitution pattern was due to a transfer rather than developmental process for three reasons. First, the fricativization of stops is clearly not a developmental process in the speech of children learning English as their native language. Second, of the four native language backgrounds represented in this study, only Spanish has a productive phonological rule that changes stops to fricatives. Third, the phonological rule of Spanish is formulated to affect /b,d,g/ but not /p,t,k/ (Harris, 1969, p. 40). We found that the Spanish talkers fricativized /b,d,g/ 10 times more often than /p,t,k/.

One might argue that the devoicing of /b,d,g/ by the Spanish subjects was also the result of a transfer process. However, we can be reasonably sure it was the result of a developmental process. We have no certain information regarding the frequency of word-final /p,t,k/ in the speech of monolingual native speakers of Mexican Spanish. However, Mexican Spanish has no native words ending in /p,t,k/, making it unlikely that the Spanish subjects used a Spanish phone as a substitute for English [b,d,g]. This inference is further confirmed by the observation that the Chinese subjects, whose native language has no stops in word-final position, also devoiced /b,d,g/.

An inference can be made from these data concerning the relative ‘‘strength’’ of the final stop deletion and devoicing processes. We observed a much higher frequency of devoicing than deletion. For example, the Chinese subjects deleted only about 2% of the stops examined, whereas they devoiced 30% of (undeleted) voiced stops. A similar difference in frequency was also apparent for the Spanish subjects.

How can this difference in frequency be explained? One possibility (see Methods section) is that, in transcribing word-final stops, the two listeners were more sensitive to devoicing than deletion. Another possibility is that adult L2 learners delete final stops only in the earliest stages of learning, just as children ordinarily delete stops only in the beginning stages of L1 acquisition. If so, we might have observed a higher rate of final stop deletion had we examined adult subjects who had just begun to learn English. The most likely possibility, in our opinion, is that the tendency to delete final stops is more easily overcome than the tendency to devoice /b,d,g/ in both L1 acquisition and L2 learning. Smith and Stoel-Gammon (1983) noted that final stop deletion had largely disappeared from the spontaneous speech of normal English-learning children by the age of 3 years, whereas these children devoiced word-final /b,d,g/ about 75% of the time.

We proposed two specific tests of the hypothesis that transfer and developmental processes independently influence adult L2 speech production. One con-
cerned additivity. It was hypothesized that if Polish subjects transferred the devoicing rule of their native language into English, they would devoice /b,d,g/ more frequently than subjects in the Chinese and Spanish groups. This hypothesis was based on the assumption that the developmental process of devoicing would equally affect subjects in all three non-native groups since all lacked experience producing voiced word-final stops prior to their exposure to English. This hypothesis was not confirmed. Although the Poles devoiced /b,d,g/ more frequently than the Spanish and Chinese subjects, the difference between subject groups was not significant.

The second test of the hypothesized independence of transfer and developmental processes was that word-final stops produced by a single group of subjects would simultaneously manifest the effect of a developmental and transfer process. This test was met in data provided by the native speakers of Spanish. The Spanish subjects, but not the English, Polish, or Chinese subjects, were observed to fricativize English stops in word-final position. As discussed earlier, this is clearly the result of a transfer process. Also as mentioned earlier, the Spanish talkers showed the influence of the developmental processes of final stop deletion and devoicing.

It should be noted that the transcriptional data presented here provide no direct indication of which syllable-final acoustic characteristics of words produced by non-native subjects led to the perception of deletion, devoicing, or fricativization. Although the transcriptions were shown to be adequately reliable, it is important to remember that the transcriptional method of assessing speech behavior is subjective in nature. Moreover, it may tend to obscure certain continuously varying aspects of articulation that are relevant to the process of speech learning by adults (Flege, 1980) or children (Walsh, 1974) because it is based on a limited number of discrete categories.

Stop devoicing was the most common model/replica difference observed here. Despite the existence of a great deal of previous research concerning perception of the contrast between word-final English /b,d,g/ and /p,t,k/, we can offer no certain insight concerning why the non-native subjects were perceived to devoice /b,d,g/.

One of the present authors (Flege, 1980; Flege & Port, 1981) examined the word-final stops produced by 12 adult Saudi Arabian learners of English. A listening test revealed that the voicing feature of about one-fourth of the labial stops examined was misidentified by native English listeners. Acoustic analyses revealed that the native English control group reliably produced /b/ but not /p/ with glottal pulsing, produced /p/ with a longer closure interval than /b/, and made vowels longer before final /b/ than /p/. The Saudi subjects generally failed to produce the same acoustic distinctions between /p/ and /b/. However, unpublished correlational analyses failed to identify a significant relationship between any of the acoustic dimensions examined and the perceptual responses of the native English-speaking listeners.

This failure to relate specific acoustic dimensions in naturally produced syllables to phonetic category judgments testifies to the complexity of the perceptual processes underlying the perception of word-final stops. Previous research (see Hillenbrand, Ingrisano, Smith, & Flege, 1984, for a recent review) suggests that
multiple dimensions affect voicing judgments for stops in word-final position. These include the duration and quality of the preceding vowel, the presence/absence of glottal pulsing during the stop closure interval, the duration of formant transitions leading into stop closure, the rate of decay of energy in the syllable, and the frequency of the first formant at its offset.

It is well known that vowel duration is a sufficient but not necessary perceptual cue to the voicing characteristic of a word-final stop consonant, and many researchers (e.g., Raphael, 1972) have regarded it as the primary perceptual cue to the English stop voicing contrast. We did not measure the duration of vowels produced by subjects in this study because the observed deletion, fricativization, and devoicing of final stops would have rendered cross-category comparisons invalid. Our subjective impression, however, was that the non-native subjects tended not to lengthen vowels before /b,d,g/ (relative to /p,t,k/). This is consistent with the results of previous studies of adult L2 production (e.g., Elsendoorn, 1983; Flege & Port, 1981; Mack, 1982) showing that adult L2 learners do not make vowels much longer before English /b,d,g/ than /p,t,k/, as English speakers do.

However, we must caution the reader against reaching the conclusion that the devoicing reported here was due to an incorrect specification of vowel duration by the non-native subjects. In a recent study, Elsendoorn (1984) observed that Dutch speakers did not produce the differential vowel duration noted in native English speech. To compensate for this, he lengthened vowels in English words ending in /b/ and /d/ produced by native speakers of Dutch so that they matched temporal values observed in the speech of native English speakers. This corrective procedure did not significantly decrease the frequency with which the Dutch-produced /b/ and /d/ were perceived to be devoiced.

We observed that the Spanish subjects devoiced final stops about twice as often as they fricativized final stops, raising the issue of whether developmental processes persist longer in L2 speech production than transfer processes. This is tantamount to asking whether it takes an adult learner longer to establish new forms of speech motor behavior for the production of L2 sounds than to stop using speech patterns established for the production of sounds in the native language.

This could be tested by using the multiple baseline training paradigm. Native speakers of Spanish could be trained to produce word-final /b,d,g/ as stops rather than fricatives while, at the same time, native speakers of Polish (or Chinese) are given comparable training directed at the voicing feature of /b,d,g/. A significantly greater decrease in fricativization (a transfer process) than devoicing (a developmental process) over baseline levels would support the hypothesis that transfer processes are more easily ameliorated than developmental processes.

Another question raised by this study concerns the perceptual effect of transfer and developmental processes. Is one kind of process more readily detectable (Flege, 1984a) than the other, and does one kind contribute more importantly to degree of accent? English-speaking adults are accustomed to hearing word-final stops devoiced in the speech of children and (to a lesser extent) other adults. They are much less accustomed to hearing final stops realized as fricatives by
other native speakers of English, including children. This leads to the prediction that, all else being equal, fricativization will be detected more easily than devoicing, and perhaps contribute to a higher degree of perceived accent. If this were empirically confirmed, it would help set a rational agenda for speech training in foreign language classrooms.

We noted a great deal of individual variation among the non-native speakers in terms of frequency of devoicing, deletion, and fricativization of final stops. A correlational analysis revealed no significant relationship between a number of language background variables (such as age of arrival in the United States and amount of formal training in English) and frequency of devoicing. Although this absence of a relationship does not rule out the possibility that some difficult-to-quantify factor such as "intake" (meaningful input) would have been significantly correlated with frequency of devoicing, it suggests that amount of L2 experience does not account for how well adults learn to pronounce sounds in a foreign language. Further investigation of the effect of experience on L2 pronunciation is clearly needed.

In summary, this study indicates that there are at least two major sources of pronunciation errors by adults speaking a foreign language: the transfer of learned patterns of articulatory behavior from L1 into L2 speech production, and the effect of unlearned developmental process that ultimately derive from universal constraints on human cognition/memory or physiological and anatomical constraints on the speech production mechanism. Adults' presumably superior knowledge of the human vocal tract and its capabilities does not seem to guarantee immediate success in pronouncing unfamiliar L2 sounds. In order to produce certain L2 sounds and syllables authentically, adults - like children learning their native language - must acquire certain articulatory skills in order effectively to produce /b,d,g/ and /p,t,k/ in word-final position. This finding is consistent with the growing consensus (e.g., Flege, 1980) that speech learning by children and adults is more similar than has been previously recognized.

ACKNOWLEDGMENTS
This research was supported in part by NIH grants NS07100 and NS20572 to the first author. An abbreviated report of these results was presented at the 1983 meeting of the American Speech-Language-Hearing Association in Cincinnati, Ohio. The authors would like to thank Dr. Julie Weissman of Triton Community College for her help in recruiting adult non-native speakers of English, and Drs. E. Broselow, S. Fletcher, and B. Smith for their helpful comments on an earlier draft of this article.

NOTES
1. Native English speakers make vowels longer in open than in closed syllables, and longer before word-final fricatives than before stops. It would therefore not be reasonable directly to compare vowels occurring before fricativized tokens of /b,d,g/ to vowels occurring before fully stopped tokens of /p,t,k/. Even if such a comparison were to be undertaken, there would be major difficulty in segmentation. The end of the "vowel" interval in CVC syllables terminating by a stop can be determined on
the basis of rapid changes in intensity and/or spectrum of the acoustic signal. However, the fact that /b,d,g/ and /p,t,k/ were not always realized as stops probably means that different criteria would need to be applied to subjects in different groups (e.g., those who deleted stops versus those who did not). This would clearly be unacceptable.

REFERENCES


The detection of foreign accent by American listeners. *Journal of the Acoustical Society of America*, 1984, in press. (a)


Smith, B. L. Effects of place of articulation and vowel environment on “voiced” stop consonant production. *Glossa*, 1978, 12, 163–175.


Walsh, H. On certain practical inadequacies of distinctive feature systems. Journal of Speech and Hearing Disorders, 1974, 39, 32–43.


Revision received: 1 November 1984