Effects of Equivalence Classification on the Production of Foreign Language Speech Sounds*

James Emil Flege

1. INTRODUCTION

Languages may differ according to the inventory of categories used to signal differences in meaning, and also according to how those categories are realized on the phonetic surface during speech production (Ladefoged 1980, 1983). The adult learner of a foreign language (L2) must therefore learn not only how to produce new sounds (i.e., phones or phonetic segments) in L2, but also how to modify previously established patterns of production. This preliminary study examined the ability of adult L2 learners to produce two kinds of L2 sounds: 'new' sounds (which have no direct equivalent in L1), and 'similar' sounds (which differ acoustically from their counterpart in L1). The production of new and similar L2 sounds was examined in three different speaking tasks to determine if variations in attention to speech influence the extent to which L2 learners approximate the phonetic norms of L2. Phonetic approximation was objectively assessed in this study by acoustically analyzing L2 sounds produced by native and non-native speakers.

ACKNOWLEDGMENTS

*This research was supported by NIH grants NS07107 (to the Department of Communicative Disorders, Northwestern University) and NS20963 and NS20572 (to the Department of Biocommunication, University of Alabama at Birmingham). The author would like to thank students, staff, and faculty of the Departments of French and Italian at Northwestern University and Indiana University who served as unpaid subjects for this study. The author would also like to thank Annemarie Hornsby for her help in preparing the manuscript. Special thanks are due to S. Fletcher, N. Spector, D. Picard, S. Romanowski, and A. Valdman for their support and interest in this work.

Sound Patterns in Second Language Acquisition (James/Leather, eds.)
© 1987 Foris Publications
1.1. New versus similar L2 sounds

The underlying assumption of many segmentally-oriented studies of L2 production (e.g., Lado 1957) is that L2 learners may substitute an L1 sound for an L2 sound if they judge the L2 sound to belong to an already familiar (i.e., L1) category. This raises the question of whether sounds not identified with a counterpart in L1 are realized more authentically (i.e. more closely approximate L2 phonetic norms) than sounds which are identified with an L1 sound. That is, will L2 sounds without a counterpart in L1 be produced more authentically than L2 sounds which are sufficiently similar to a sound in L1 to be transcribed with the same IPA symbol, yet differ acoustically from it?

The view of some speech-language pathologists is that it is more difficult to remediate children’s production of ‘distorted’ sounds than to aid them in producing a sound not yet found in their phonetic repertoire (S.G. Fletcher, personal communication). If this experience with children acquiring L1 relates to L2 learning, it may be true that adult L2 learners have more difficulty producing similar L2 sounds – which requires the modification of previously established patterns of segmental articulation – than new L2 sounds. Valdman (1976), in fact, concluded that adult L2 learners succeed better in producing new than similar L2 sounds because they can avoid using previously established patterns of articulation.

It might be argued, on the other hand, that adult L2 learners will have more difficulty producing new than similar L2 sounds. It is generally accepted that the production of a sound (or sound type) requires the establishment of a central phonetic representation which contains information concerning both the perceptual ‘target’ as well as a motor plan specifying how that target is to be achieved in speech production (Schmidt 1976; Keele and Summers 1976). There also seems to be widespread acceptance of the view (Lenneberg 1967; Scovel 1969) that adults lose the ability to learn new sounds (or at least have diminished ability to do so). It is therefore possible that adults, unlike children, will have more difficulty adding a new L2 sound to their phonetic repertoire than modifying an L1 sound for use in L2.

The hypothesis that new sounds are more difficult for L2 learners to produce authentically than similar sounds was supported by two previous studies. Brière (1966) trained native English speakers to produce French, Arabic, and Vietnamese sounds varying in similarity to sounds found on the phonetic surface of English. Sounds with counterparts in English were generally perceived to have been produced more authentically than those without a close equivalent in English. However, the native English-speaking subjects showed surprisingly large differences in their ability to produce...
Effects of Equivalence Classification

various new sounds. For example, the voiceless velar fricative /x/ was learned faster and more successfully than its voiced cogener /γ/, even though neither sound occurs regularly on the phonetic surface of English. Flege and Port (1981) examined the English spoken by Saudi Arabians to determine if they would produce an English stop which has no counterpart in Arabic (i.e., /p/) better than English stops which do have a counterpart in Arabic (i.e., /t/ and /k/). Temporal acoustic measurements indicated that /p/, /t/, and /k/ were produced with equal authenticity, but a listening test showed that the Arabs’ /p/ was correctly identified significantly less often than their /t/ and /k/. This also suggested that new sounds may be more difficult to produce than similar sounds (see also Politzer and Weiss 1969; Johansson 1973; Walz 1979).

Despite these findings, the evidence which now exists is insufficient to permit a firm conclusion regarding the relative difficulty of new and similar L2 sounds. Most previous research bearing on this issue has examined the speech production of talkers with little or no knowledge of the L2 from which the L2 sounds being examined were taken; most examined immediate or delayed imitation rather than spontaneous speech; and most relied solely on phonetic transcription, often by non-native speakers. Thus, one aim of this study was to objectively compare the production of a new and similar sound in L2. Specifically, the study focused on production by native English speakers of the French vowels /u/ and /y/. French /u/ is realized with variants that are similar, yet acoustically non-identical, to the realizations of the /u/ category of English. French /y/, on the other hand, does not correspond directly to an English vowel, and can therefore be regarded as a new sound for English native speakers who learn French as an L2.

1.2. Attention to speech

Another aim of this study was to explore the extent to which authenticity of L2 production varies according to speaking task. The performance of a new motor skill generally seems to require progressively less concentration and effort as it becomes ‘automaticized’ through practice (Klein 1976). It has been proposed (Van Riper and Irwin 1958; Siegel, Fehst, Garber and Pick 1980) that motor control shifts from a predominantly ‘closed loop’ to ‘open loop’ (i.e., centrally pre-programmed) mode as talkers gain experience.

Some existing evidence suggests that talkers rely less on auditory feedback as speech develops. Although disruption of auditory feedback has little effect on speech intelligibility (see, e.g., Borden 1979), adventitious deafening gradually results in a deterioration of speech that may be more severe for young children than adults (Binnie et al. 1982). One study (Manning, Keap-
pock, and Stick 1976) showed that children undergoing speech therapy who succeeded best in producing their ‘errored’ sound in noisy conditions also succeeded best in maintaining the gains made during therapy afterward. Another study (Manning and Hein 1981) showed that masking noise had a less distorting effect on the production of new and similar L2 sounds (French /y/ and /r/) by native English-speaking adults as their production of the L2 sounds improved across ten training sessions. However, counter to the hypothesis that talkers make progressively less use of auditory feedback, Pick et al. (1982) observed no difference between adults and children for speech produced in noise.

Oro-sensory acuity has frequently been linked to articulatory development (Locke 1968; McNutt 1979). There is some evidence that the ordinarily minimal distortion of speech that results from the disruption of tactile and/or kinesthetic feedback (e.g., Ringel and Steer 1963) may be more severe for L2 than L1 sounds (Siegel et al. 1977; but cf. Borden 1980). However, a study by Borden, Harris, Fitch and Yoshioka (1981) raised doubts concerning the importance of sensory feedback for L2 learners. Three native English-speaking subjects showed somewhat greater distortion of L2 than L1 sounds when sensory feedback was minimized through nerve block injections, the presence of an oral prosthesis, or masking noise. However, four other talkers showed greater distortion of L1 than L2 sounds. Garber, Speidel, and Siegel (1980) also provided evidence that tactile/kinesthetic feedback may be no more important in developing than established speech. In that study, speech sounds were distorted with about equal frequency by five-year-olds and adults wearing experimental palatal prostheses.

The hypothesized decrease in reliance on peripheral sensory feedback has been linked to general attention or vigilance. Shelton and McReynolds (1979) suggested that articulatory motor control shifts from higher, more ‘conscious’ levels of control to lower, more ‘automatic’, levels as articulatory patterns are established in L1 acquisition. Abbs and Cole (1982, 174 ff.) posited that ‘consciously encoded ... motor subroutines’ which play a crucial role in the coordination of multi-articulator movements needed for speech production are eventually relegated to the level of ‘unconscious or automatic mediation’. Many researchers accept that motor skills shift from a predominantly closed to open loop (or central) mode of control. However, Kelso and Stelmach (1976) observed that this hypothesis lacks solid empirical support at present, noting that central control may remain prominent for the rapid ‘ballistic’ movements common in speech production.

Sociolinguistic research suggests that vigilance or ‘attention to speech’ may affect speech production. It is generally the case that individuals use
speech sound variants typical of the ‘prestige norm’ of their L1 more frequently when speaking carefully (or formally) than casually (or spontaneously). Researchers (e.g., Labov 1972) have used the ‘danger of death’ narrative to elicit casual varieties of speech. Individuals who are not native speakers of the prestige dialect of their L1 may use prestige norm variants less frequently in such a speaking task than in more formal speaking conditions (e.g., reading a paragraph). One hypothesis is that for such individuals, producing prestige norm variants requires special vigilance or attention to speech. The use of prestige norm variants (which are presumably learned after more basic variants have been established) may decrease because the demands of certain speaking tasks prevent speakers from effectively monitoring their speech output.

This hypothesis is consistent with the work of Krashen (1978), who distinguished ‘acquired’ from ‘learned’ aspects of L2 competence. His theory is that the learned modifications needed to produce L2 are accessed only through a relatively conscious ‘monitoring’ process, and that learners’ success in producing L2 may vary as a function of the extent to which L2 production can be monitored in various speaking tasks. If Krashen’s model were extended to the segmental articulation of L2 speech sounds, it would lead to the prediction that decreasing L2 learners’ opportunity to monitor their speech output will diminish the authenticity with which they realize L2 sounds. This assumes, of course, that L2 learners develop a rule for modifying the production of L1 sounds in order to authentically produce L2. Further, it assumes that talkers are capable of regulating the ongoing production of speech.3

Results reported by Oyama (1982), however, suggest that learners may not produce L2 more authentically in careful than casual speaking conditions. In Oyama’s study, listeners evaluated the degree of accent in paragraphs and excerpts of a ‘danger of death’ narrative produced by L2 learners. Contrary to expectation, the paragraphs (which had been read) were judged to be more accented than the speech produced spontaneously. Oyama (1982) noted the possibility that the listeners may have applied different criteria in judging the two speech samples, as well as the possibility of a series effect (the spontaneous speech was always evaluated after the paragraph). Another possibility is that specific sounds may have been produced more authentically in the spontaneous speech than in the paragraph, but these differences were overlooked because of the length of the speech samples (see Flege 1984 for a discussion of allocation of attention in an accent detection task). The second aim of this study was therefore to determine whether L2 learners realize L2 sounds more authentically in speaking conditions designed to reduce attention to speech.
1.3. Equivalence classification

The study’s third aim was to test a hypothesis concerning why phonetic interference so often persists in L2 speech production. A number of studies have shown that even highly experienced L2 learners often do not realize an L2 sound authentically if it differs acoustically from its counterparts in L1. This can be illustrated by considering the production of L2 stops that differ according to voice onset time (VOT) from corresponding L1 stops.

The VOT values of stops found in L1 and L2 may differ systematically according to VOT even though they are transcribed with the same IPA symbol. It usually takes children several years to learn to realize stops with the language-specific values of their L1 (see, e.g., Zlatin and Koenigsknecht 1976; Macken 1980). L2 learners typically manifest phonetic interference in producing L2 stops that differ from L1 stops according to VOT. For example, Flege and Port (1981) found that Arab L2 learners produced English /p,t,k/ with significantly shorter values than native English speakers because Arabic stops are realized with significantly shorter VOT values than English stops. Caramazza, Yeni-Komshian, Zurif, and Carbone (1973) found that native speakers of French produced English /p,t,k/ with substantially shorter VOT values than English monolinguals because the stops of French are realized with short-lag VOT values (as compared to the long-lag values of English).

The failure of L2 learners to match native speakers in producing L2 stops might be seen as support for the claim (Lenneberg 1967; Scovel 1969, 1981) that adults lose the ability to modify previously established patterns of segmental articulation. Selinker (1972) noted that the language abilities of L2 learners may become ‘fossilized’, but did not offer an explanation for why this might happen. An analogous phenomenon (known as ‘crystalization’) has been noted in birds. In certain avian species, song learning may be curtailed after a sensitive period has been passed (Marler and Mundinger 1971). However, unlike for humans, this phenomenon presumably results from hormonally-induced changes in neural functioning.

Flege (Flege 1981, 1985; Flege and Hillenbrand 1984) proposed an alternate hypothesis concerning why L2 learners may only partially modify segmental articulation in learning to produce an L2. He noted that it may be just as important to explain the phonetic approximation noted in L2 learning as it is to explain the persisting differences between native and non-native speakers of L2. For example, although the French Canadian subjects studied by Caramazza et al. (1973) did not match native English speakers in producing English /p,t,k/, they nonetheless produced English stops with longer VOT values than is typical for French stops produced by French monolinguals.
In Flege's view, partial approximation to L2 phonetic norms indicates that an L2 learner has noted an acoustic difference(s) between similar L1 and L2 sounds, and has demonstrated the ability to modify pre-existing patterns of segmental articulation. He hypothesized that learners' incomplete success in producing L2 sounds is due to a cognitive mechanism termed 'equivalence classification.' More specifically, he hypothesized that L2 learners may develop inaccurate perceptual targets for L2 sounds with a direct counterpart in L1 as the result of equivalence classification.

The VOT dimension observed in speech production is frequently related to the processing of VOT in speech perception (Abramson and Lisker 1970). For example, it generally takes longer VOT values to cue the perception of a voiceless stop for native speakers of English than French because /p,t,k/ is realized with longer VOT values in English than French. Despite the existence of speech perception differences between French and English monolinguals, and despite the fact that listeners can auditorily detect the acoustic differences distinguishing French from English stops (Flege 1984; see also Flege and Hammond 1982), it is likely that native speakers of French and English regard the /p,t,k/ of English and French as belonging to the same category. This assumption receives a priori support from the fact that the /p,t,k/ found in French and English manifest sufficient overall similarity to warrant being transcribed with the same IPA symbols.

Several studies have shown that the phoneme boundaries of native French and Spanish L2 learners occur at values which are intermediate to those observed for monolingual native speakers of L1 and L2 (Caramazza et al. 1973; Williams 1977, 1979, 1980; Obler 1983). This suggests that English learners whose L1 contains short-lag stops merge the phonetic properties of stops they have experienced in L1 and L2. If so, adult foreign language learners may produce L2 stops with VOT values intermediate to those produced by monolingual native speakers of L1 and L2 because they have developed a perceptual target specifying a VOT value for L2 stops which is intermediate in value to the VOT value specified in the perceptual target for stops by monolingual native speakers of L1 and L2.

The hypothesized role of equivalence classification receives a priori support from studies of infant speech perception. It has been shown (e.g., Kuhl and Miller 1982; Hillenbrand 1983) that infants are able to find perceptual constancy for categories by somehow identifying physically different sounds as belonging to the same category (Kuhl and Miller, 1982). This ability is a pre-requisite for speech learning, for children must learn to attend to common sensory properties in the many tokens of a phonetic category, while ignoring those properties which have no or little phonetic significance (Gibson 1969; Tversky and Gati 1978; Kuhl 1980). There is no
reason to suppose that adults lose the ability (or inclination) to judge acoustically different sounds as belonging to the same category. Therefore, an important difference between child L1 acquisition and L2 learning may be the existence of well-defined phonetic representations.

The child constructs perceptual targets for L1 stops based on the many physically different tokens of L1 stops he or she encounters. Children acquiring L1 develop accurate perceptual targets for L1 sounds because they hear only L1 sounds, and eventually acquire the ability to 'say what they hear'. Adult L2 learners have developed perceptual targets for L1 sounds by the time L2 learning commences. If they equate sounds in L1 and L2, the perceptual target they develop for L2 stops may be based on L1 stops as well as on the many tokens of L2 sounds encountered. Assuming that adult L2 learners eventually achieve their perceptual target for L2 stops in production, they will at best produce L2 stops with VOT values that are intermediate to the VOT values in stops produced by monolingual speakers of L1 and L2.

The third aim of this study was to test several predictions concerning the effect of equivalence classification on L1 and L2 speech production. If equivalence classification shapes the production of L2 sounds, we might expect to observe a number of specific phenomena in L2 speech production. First, not even highly experienced L2 learners should produce L2 sounds judged to be equivalent to an L1 sound with complete authenticity. Equivalence classification is thought to result in the development of inaccurate perceptual targets for L2 sounds through the merger of the phonetic properties of similar L1 and L2 sounds. Although highly experienced L2 learners may develop distinct perceptual targets for similar sounds in L1 and L2 (Elman, Diehl, and Buchwald 1977) we hypothesize that separate perceptual targets (if developed) will exert a kind of 'gravitational attraction' on one another (Schouten, 1975) because a single phonetic category representation underlies them.

A second prediction is that what we have been calling new L2 sounds will be produced more authentically than similar L2 sounds. This possibility exists because the L2 learner may not equate a new L2 sound with a sound(s) in L1 because of the greater dissimilarity between new L2 sounds and sounds found on the phonetic surface of L1. If so, L2 learners may establish a phonetic category representation for new L2 sounds, develop accurate perceptual targets for them, and eventually (although not necessarily) achieve the (accurate) perceptual target in speech production.

A third prediction is that L2 learning will affect the segmental articulation of sounds found in L1. If the perceptual target developed for a sound in L2 is influenced by the perceptual target previously developed for a similar (but acoustically non-identical) L1 sound, the reverse should also be true. And
if the perceptual target for a sound in L1 changes, the possibility exists that there will be a concomitant change in speech production. There is some evidence to support this prediction. Caramazza (Caramazza et al. 1973; Caramazza and Yeni-Komshian 1974) reported data indicating that native French speakers who learned English as an L2 produced /p, t, k/ with slightly longer VOT values in French words than French monolinguals. Similarly, Williams' (1979) data for native Spanish-speaking children who learned English indicated that they produced Spanish stops with VOT values that were longer than those expected for Spanish monolinguals.

2. METHODS

2.1. Subjects

Four subject groups differing principally in language background and experience were recruited. Each group consisted of seven women with self-reported normal hearing. The native French-speaking group (designated Group FR) was made up of French native speakers (mean age, 38 years) who had been living in the United States for an average of 12.2 years at the time of the study. Two of these subjects were from Belgium, four were from Paris, and one was from Annecy, France. Four of the native French subjects were married to native English-speakers, and all seven used English as their principal language at the time of the study. Although each spoke English fluently in the author's estimation, all but one rated her English-speaking ability lower than her ability to speak French (5.0 compared to 7.0 on a 7-point rating scale administered prior to the study).

There were three groups of native English speakers, most of whom had been born and raised in the Midwestern region of the U.S. The first group (designated Group Am0) consisted of women (mean age, 26 years) who spoke no foreign language. Subjects in the two other groups spoke French as an L2. The subjects in the group designated Group Am1 consisted of undergraduate students at Northwestern University (mean age, 22 years) who had spent the previous academic year in Paris. All seven rated their French-speaking ability lower than their ability to speak English (5.4 compared to 7.0). Subjects in the group designated Am2 consisted of women (mean age, 32 years) who held advanced degrees in French and taught French at the university level. Most speakers in this group had spent several periods of time in France, the total averaging 1.3 years. One was married to a native French speaker, and all but one rated her French-speaking ability lower than her ability in English (5.6 compared to 7.0).
The native English speakers of French began to study French in school sometime between the ages of 11–17 years, but none appears to have been able to converse effectively in French prior to the age of about 20 years. The subjects in both groups should therefore be regarded as adult L2 learners (as opposed to early childhood bilinguals). No attempt was made to objectively assess the French-speaking proficiency of subjects in Groups Am1 and Am2, for the intent was simply to constitute two extreme groups.

There was a clear difference between these two groups in terms of the length of time they had used French to communicate. This was less than a year for subjects in Group Am1, whereas the subjects in Group Am2 had used French as a functional second language for an average of about ten years. Moreover, the language background questionnaire indicated that, compared to subjects in Group Am1, the subjects in Group Am2 had substantially more formal instruction in French language and literature, rated their own production and comprehension of French somewhat higher, and used French somewhat more often on a daily basis in the period immediately preceding the experiment. Thus it seems reasonable to refer to subjects in Group Am2 as being more ‘experienced’ than those in Group Am1.

2.2. Speech materials

Those subjects who spoke both French and English produced the following four sets of phrases:

<table>
<thead>
<tr>
<th>ENGLISH</th>
<th>FRENCH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set 1</td>
<td>Set 2</td>
</tr>
<tr>
<td>Two little boys</td>
<td>TV programs</td>
</tr>
<tr>
<td>Two little girls</td>
<td>TV schedules</td>
</tr>
<tr>
<td>Two little cats</td>
<td>TV ratings</td>
</tr>
<tr>
<td>Two little dogs</td>
<td>TV violence</td>
</tr>
<tr>
<td>Two little birds</td>
<td>TV reception</td>
</tr>
<tr>
<td>Two little mice</td>
<td>TV antennas</td>
</tr>
<tr>
<td>Two little men</td>
<td>TV commercials</td>
</tr>
</tbody>
</table>

The acoustic analyses focused on the words *tous*, *two*, and *tu*. Studies of first language acquisition (e.g., Barton 1980) have shown that word familiarity has an important effect on segmental articulation and phonetic perception, so it was considered important to choose test words that would be highly familiar to all subjects. *Tu and tous* occur frequently in conversa-
Effects of Equivalence Classification

tional French, and both are likely to be among the first words encountered by L2 learners (tu is the second person singular pronoun meaning 'you'; tous is an adjective meaning 'all'). The choice of test words was also partly dictated by the fact that French tous has a phonetic counterpart in English (two), making it possible to directly compare the syllable /tu/ across languages. Like tous, English two is also a frequently occurring word. The speech samples were designed in such a way that the initial syllable (/tu/ or /ty/) was always followed by the same sound (/l/). Subjects tended to produce the final word in each phrase with emphatic stress because these words — rather than the utterance-initial syllables being analyzed — varied.

The expected effect of cross-language phonetic interference is for native speakers of English to produce French /u/ with F2 values that are too high by French phonetic standards; and for native speakers of French to produce English /u/ with F2 values that are too low by English phonetic standards. Stevens (1983) concluded that an important cause of the relatively great F2 variation observed in past analyses of English /u/ stems from the fact that /u/ does not differ phonologically from a high front vowel category (/y/) in English as it does in other languages such as French. It is also possible that the lack of a phonological opposition between /u/ and /y/ in English may cause a pronunciation problem for native English learners of French. A contrastive analysis of English and French (e.g., LeBras 1981) leads to the prediction that English speakers will substitute the 'closest' English vowel (probably /u/) for French /y/.4

In fact, the substitution of an /u/-quality vowel for French /y/ and the reverse (i.e., the substitution of an /y/-quality vowel for French /u/) has been reported for relatively inexperienced native English learners of French (Walz 1979).

Phoneticians have described the /u/ of French as being more 'tense' (or less 'anterior') than its English counterpart (Delattre 1951, 1953; Adamczewski and Keen 1973; Valdman 1976, see also Lindblom and Sundberg 1971). If /u/ is articulated with a more anterior tongue position than French /u/, as suggested by auditory comparisons, F2 values should be higher for English /u/ than its French counterpart. F2 values published in several previous reports (Peterson and Barney 1952; Delattre 1951; Debrock and Forrez 1976; Riordan 1977) tend to undermine this interpretation, but the results of other acoustic studies support the traditional auditory description of the differences between French and English /u/. They indicate that /u/ may be produced by adult female talkers with F2 frequency values that are 400–500 Hz higher in French than English, at least in certain phonetic contexts (Stevens and House 1963) and speaking styles (Shockey 1974; Labov 1981). This finding is consistent with the observation that /u/ is
sometimes realized as an /y/-quality vowel in American speech, especially when preceded by a palatal glide (as in music).

As mentioned in the introduction, another expected effect of cross-language phonetic interference is for native English speakers to produce /t/ in French words with VOT values that are too long by French phonetic standards, and for native speakers of French to produce /t/ with VOT values that are too short in English words. I know of no comprehensive L2 production study which has examined the production of short-lag stops in an L2 (e.g., French) by learners whose native language (e.g., English) has long-lag voiceless stops. If short-lag stops are truly less difficult to produce for physiological reasons than long-lag stops (Kewley-Port and Preston 1974), native speakers of English might succeed better in producing the short-lag stops of French than native French speakers produce the long-lag voiceless stops of English (Caramazza and Yeni-Komshian 1974).

2.3. Data elicitation

The speech material was cued by the presentation of several written phrase lists. The order of language (French versus English) and phrase set (1 versus 2, 3 versus 4) was counterbalanced across talkers. The monolingual English speakers, of course, produced only the English phrases. The speech material was recorded on high-quality equipment (Sony model TCD5M) in a small, sound-treated room. An electret condenser microphone (Nakamichi model CM-300) was positioned about six inches from the talker’s mouth. The speech produced in the Story condition was highly natural and spontaneous in the author’s estimation.

The test phrases were produced in three conditions designed to influence the amount of attention subjects might allocate to monitoring the phonetic characteristics of their speech. In the first condition, the talkers simply read the test phrases in isolation. This task was expected to permit subjects to monitor their own speech production as fully as possible. In the second speaking task the subjects were required to generate original sentences by completing each of the phrases just produced in isolation. In the third speaking task, subjects were required to create a story based on the phrase sets. They were allowed to take as much time as necessary for silent rehearsal. The story they produced was to include a complete sentence initiated by each of the seven phrases in a set; additional sentences not initiated by one of the test phrases were also permitted. The talkers did not say a number before each utterance, as in the previous two speaking tasks. They were told instead to pause before each sentence (thereby ensuring that /tə/ and /tʃ/ occurred in absolute utterance-initial position). To help plan their story, the
Effects of Equivalence Classification

subjects were instructed to arrange a set of cards (each bearing one of the test phrases) as an outline.

Each subject heard the same set of recorded instructions modelling the speaking task in English. They were given an opportunity to practice before the experiment using a set of English practice phrases. All but one talker (who was replaced) performed satisfactorily. The subjects were naive as to the purpose of the study. They were told it examined an aspect of 'language creativity' in order to counteract for the natural propensity of talkers to speak as carefully as possible in the presence of a microphone (Labov 1972). Systematic debriefing after the experiment indicated that the subjects were not consciously aware the study focused on pronunciation.

It was expected that, if attentional factors influence the authenticity of L2 pronunciation, subjects would produce L2 sounds more authentically in the Isolated Phrase (and perhaps Sentence) condition than in the Story condition because the Story condition demanded more attention to previous and upcoming information. Moreover, the subjects might have gradually become relaxed across the three speaking tasks as they grew accustomed to being tape recorded. This might also have contributed to a less authentic production of L2 sounds, because as subjects relaxed they may have shown less tendency to realize a phonetic category with variants learned as the result of exposure to L2 than with an early-acquired (i.e., L1) variant.

2.4. Data analyses

The phrases were low-pass filtered (Khrone-Hite Model 3345) at 4 kHz before being digitized at a sampling rate of 10 kHz with 11-bit amplitude resolution. Values of the first two or three formants in a vowel correspond systematically to its auditory quality (Ladefoged 1967; Ladefoged, Harshman, Goldstein, and Rice 1978). However, the present study focused on just second formant frequency because F2 is the single most important acoustic dimension distinguishing front vowels (like /i/ or /y/) from back vowels such as (/u/). Assuming equal lip rounding, F2 should provide an good indication of the vocal tract differences distinguishing the French from English /u/ produced by bilingual subjects, as well as French /y/ from /u/.

The F2 frequency of /u/ and /y/ was estimated by linear predictive coding (LPC) analysis using a 12-coefficient model (Markel and Grey 1976). As illustrated in Figure 1, a full 25.6 ms Hamming window was positioned so that its left-hand tail coincided with the first high-amplitude waveform peak in the periodic portion ('vowel', for short) of the syllables analyzed. Also, as illustrated in Figure 1, voice-onset time (VOT) was measured to the nearest 0.1 ms from the display of a high resolution graphics terminal.
Figure 1. The VOT of /t/ was measured from the sharp increase in energy signalling stop release (a) to the onset of periodicity of the following vowel (b). The 25.6 msec Hamming window used in LPC analyses was positioned so that its left-hand margin was situated at the first positive high amplitude peak in the periodic portion of the waveform (c).

Figure 2. The mean VOT in ms of /t/ in the English word *two* produced by subjects in four groups in three speaking conditions (Isolated Phrases; Sentences; and Story). FR designates native speakers of French who had learned English as an L2; Am0 refers to native English speakers who spoke no L2; Am1 refers to relatively inexperienced native English speakers of French; and Am2 designates more experienced native English speakers of French.

(Tektroniks Model 4010). The cursor was set at the beginning of the noise burst (signalling stop release) and at the first upward-going zero crossing of the waveform (signalling vowel onset).

Mean VOT and F2 values were computed from the values measured in
five tokens of each syllable produced by each subject in each of the three speaking conditions. The mean values were submitted to mixed design analyses of variance in which Subject Group served as the between-subjects factor, and Speaking Condition and either Syllable or Language (depending on the analysis) served as with-subjects factors. When needed, post-hoc comparisons using the Scheffé procedure (alpha = .01) were performed.

3. RESULTS

3.1. Voice onset time

Figure 2 plots the mean VOT of /t/ in the English word two produced by talkers in the four subject groups in the three speaking conditions. The native and non-native speakers produced English /t/ with substantially different VOT values in all three conditions. The native French speakers of English (Group FR) produced /t/ with VOT values averaging 51 ms. Subjects in the three groups of native English speakers (AmO, AmI, Am2), on the other hand, produced /t/ with an average VOT value of about 75 ms. The Anova revealed that the effect of the factor Subject Group was significant \[F(3,24) = 5.58, p < .01\]. Post-hoc tests indicated that the native French speakers produced /t/ with significantly shorter VOT values than talkers in the three native English groups, and that none of the three native English groups differed significantly from one another. It is also apparent from Figure 2 that VOT varied little across the three speaking conditions. Not surprisingly, the effect of the Speaking Condition was non-significant (p < .01).

Figure 3 plots the mean VOT of /t/ in the French words tous and tu produced in the three speaking conditions. The relatively inexperienced English speakers of French (i.e., Group Am1) evidenced more phonetic interference from English than the more experienced English speakers of French (i.e., Group Am2). They produced French /t/ with substantially longer VOT values (69 ms) than the experienced English speakers of French (52 ms), who differed very little from the French native speakers (55 ms averaged across the three speaking conditions). The Anova indicated that the effect of Subject Group was significant \[F(2, 18) = 5.97, p < .01\]. Post-hoc tests revealed that, in the Isolated Phrase and Sentence conditions, subjects in Group Am1 produced French /t/ with longer VOT values than subjects in Group Am2 or Group FR, and that the experienced English speakers of French (Am2) did not differ significantly from the native speakers of French. There were no significant between-group differences, however, in
the Story Condition (p < .01). The effect of Speaking Condition was once again small and non-significant (p < .01).

Figure 4 directly compares the /t/ produced in French and English words (i.e., tous and two). The extent to which subjects alter VOT in /t/ when switching from English to French (or the reverse) might be regarded as an index of their bilingualism, since voiceless stops are realized with substantially longer VOT values in English than French. In fact, the inexperienced English speakers of French (Group Am1) showed a much smaller shift when switching from English to French (about 6 ms) than the more experienced English speakers of French (about 23 ms). The native speakers of French, who were highly proficient in English, produced French and English /t/ with VOT values that differed less than 1 ms (averaged across the three speaking conditions).

The Anova indicated that only the experienced English speakers of French (Group Am2) produced /t/ with significantly different VOT values in French and English. As a result, the interaction of Language × Subject Group reached significance [F(2, 18) = 9.33, p < .01]. Post-hoc tests indicated that the subjects in Group Am2 produced a significant VOT difference between French and English /t/ in the Isolated Phrase and Sentence conditions, but not the subjects in Groups Am1 and NF. Once again, the small effect of Speaking Condition did not reach significance (p < .01).
Effects of Equivalence Classification

Figure 4. The mean VOT in ms of /t/ in English two and French tous produced by subjects in three groups in three speaking conditions (see text for details). This data has been replotted from Figures 2 and 3 to afford an easy comparison of cross-language modifications in speech production.

3.2. Second formant frequency

Figure 5 plots the mean F2 frequency of the English /u/ produced by talkers in the four subject groups as a function of speaking condition. The native speakers of French were expected to produce English /u/ with lower F2 values than the native speakers of English because /u/ is realized with lower F2 values in French than English, and because phonetic interference is persistent in L2 speech production (Flege 1980; Flege and Port 1981). As expected, the French native speakers (Group FR) realized English /u/ with lower F2 values (averaging about 1450 Hz across the three speaking conditions) than the native English speaking subjects in Groups Am0, Am1 and Am2 (about 1650 Hz). However, the effect of Speaker Group did not quite reach significance [F(3, 24) = 2.22] because of variability among the seven subjects per group. Vowel formant frequency differences across the three speaking conditions were small and non-significant (p < .01).

Figure 6 plots mean F2 values for the French vowels /u/ and /y/. It is evident that the native French speakers produced a substantially larger F2 contrast between /u/ and /y/ than the native speakers of English. The magnitude of the contrast averaged 815 Hz for the native speakers of French (Group NF), 419 Hz for the experienced native English speakers of French.
Figure 5. The mean second formant frequency in Hz of the English vowel /u/ in two produced by subjects in four groups in three different speaking conditions (see text for details).

(Group Am2), and less than 200 Hz for the inexperienced native English speakers of French (Group Am1). As a result of these between-group differences, the interaction of Vowel x Subject Group reached significance [F(2, 18) = 14.3, p < .01]. Tests of simple main effects indicated that the interaction was due primarily to between-group differences in the production of /u/. The effect of Subject Group on the F2 of /u/ but not /y/ was significant in all three speaking conditions (p < .01). The inexperienced native English speakers of French (Group Am1) did not produce a significant F2 difference between French /y/ and /u/ in any condition. However, both the French native speakers (Group FR) and the experienced native English speakers of French (Group Am2) produced /y/ with significantly higher F2 values than /u/. As is evident from Figure 6, the effect of Speaking Condition on vowel formant frequency was small and non-significant.

Finally, Figure 7 juxtaposes the F2 values produced in French and English /u/. The native speakers of French (Group FR) and the experienced English speakers of French (Group Am2) produced /u/ with somewhat higher F2
Effects of Equivalence Classification

Figure 6. The mean second formant frequency in Hz of the /u/ in French *tu* produced by subjects in three groups in three speaking conditions (see text for details).

Figure 7. The mean second formant frequency in Hz of the /u/ vowels in French *tous* and English *two* produced by subjects in three groups in three speaking conditions (see text for details). This data is replotted from Figures 5 and 6 in order to afford an easy comparison of cross-language differences in speech production.
values in English than French. This means they *approximated* the phonetic norms of their L2, although it should be clear from Figure 7 that they did not actually match L2 native speakers. Somewhat surprisingly, the inexperienced native English speakers of French (Group Am1) produced /u/ with somewhat *higher* mean values in French than English. This means that, instead of approximating the phonetic norms of French for /u/, they produced /u/ with even higher (*less* French-like) F2 values than they did in English. Despite the different patterns evident for subjects in the three subject groups, however, the interaction of Subject Group x Language did not quite reach significance \[ F(1, 18) = 2.97, p = .078 \]. The effect of Subject Group was significant, however \[ F(2, 18) = 15.03, p < .01 \]. Tests of simple main effects revealed that, in all three speaking conditions, the native English speakers produced /u/ with higher F2 values than the native speakers of French, and the two native English groups (Am1 and Am2) did not differ significantly from one another. The Anova once again indicated that the effect of Speaking Condition was non-significant \( p < .01 \).

4. DISCUSSION

The results of this study demonstrated that adult learners of a foreign language (L2) make progress in learning to pronounce L2 sounds. Experienced learners of a foreign language appear to approximate the phonetic norms of L2 to a greater extent than less experienced L2 learners, but their progress may be limited by the mechanism of equivalence classification.

4.1. New versus phonetically similar sounds

Valdman (1976) and others have hypothesized that L2 learners tend to use native language (L1) sounds in place of L2 sounds that are similar yet not acoustically identical to the L1 sound. The maintenance in L2 speech production of articulatory patterns established during L1 acquisition is thought to be triggered by the *perceived* similarity of sounds occurring on the phonetic surface of L1 and L2 (see Flege, 1985 for a review). Flege (1981, 1985; Flege and Hillenbrand 1984) suggested that although even experienced L2 learners may only *partially* approximate the phonetic norms of L2 in producing 'similar' L2 sounds, they may succeed in authentically producing 'new' L2 sounds not identified with a sound occurring on the phonetic surface of L1.

Flege proposed that equivalence classification (also referred to as 'interlingual identification'; see Weinreich 1953) may cause the L2 learner to
Effects of Equivalence Classification

develop an incorrect perceptual target for L2 sounds by leading them to merge the phonetic properties of L1 and L2 sounds judged to be equivalent. A premise of this study was that experienced native English learners of French would identify the /u/, but not the /y/, of French with an English vowel category (namely /u/). It was predicted that, as a result of equivalence classification, even experienced native English speakers of French would produce /u/ with higher second formant frequency values than native speakers of French. (English /u/ is realized with higher F2 values than its French counterpart). It was also predicted that experienced native English speakers of French would produce French /y/ with about the same F2 values as French native speakers. (French /y/, which is produced with substantially higher F2 values than either French or English /u/, does not have a direct phonological counterpart in English).

The hypothesis concerning L2 learners' production of the new vowel /y/ and the similar vowel /u/ was supported by the data presented here. The subjects in two native English groups produced French /y/ with mean frequency values that did not differ significantly from the F2 values measured in the /y/ produced by native speakers of French. However, the subjects in both native English groups produced French /u/ with F2 values that were significantly higher (i.e., more English-like) than native speakers of French.

The subjects in the more experienced of the two native English groups produced French /u/ with F2 values that were only slightly lower (i.e., more French-like) than the F2 values measured in the English /u/ produced by English monolinguals. Thus this study provided little evidence that the experienced native English speakers of French approximated the phonetic norms of French for /u/ by lowering F2 values relative to their production of English /u/. Native English subjects who were less experienced in French produced French /u/ with F2 values that were actually higher than the F2 values they (and monolingual native English speakers) produced in English /u/. This suggests they may have confused the /y/ and /u/ categories of French. (Recall that French /y/ is realized with higher F2 values than French or English /u/).

Although the data presented here supports the hypothesis that L2 learners produce new L2 sounds better than similar ones, we cannot be certain that this finding has general validity. The study acoustically examined only one new and one similar L2 vowel (i.e., French /y/ and /u/). The differences in authenticity we observed in the production of these two vowels might conceivably have arisen from purely articulatory factors instead of from differences in the perceived similarity of vowels in French and English. Future studies examining the production of other pairs of new and similar L2 sounds (consonants as well as vowels) will be needed.
It would be useful for future studies to focus on 'new' L2 sounds that never occur on the phonetic surface of L1. One weakness of the present study is that the native English subjects examined here may have had some prior experience producing an /y/-quality vowel in American English. This might have aided them in producing the /y/ of French (see Brière 1966). Moreover, the vowels in *tous* and *tu* have been unrepresentative of the native English subjects' production of French /u/ and /y/. These two lexical items were chosen to ensure familiarity (see the 'Methods' section), but since they were likely to have been acquired early, they may represent non-optimal approximations to the phonetic norms of French that were superceded in the production of /u/ and /y/ in other, later-acquired French words.

### 4.2. Attention to speech

A second hypothesis that was tested in this study was that L2 learners produce L2 sounds with decreasing authenticity as the attentional demands of a speaking task increase (provided, of course, that the L2 sounds differ acoustically from their counterparts in L1). The results presented here did not support this hypothesis. The acoustic parameters measured (i.e., the F2 of /u/ and /y/, and the VOT of /t/) did not differ significantly in words produced in three different speaking tasks designed to manipulate attention to speech. It seems reasonable to think that the speaking tasks did influence attention to speech, for producing a spontaneous narrative based on a phrase list seems to require more attention than simply reading the same phrases from a list (see the 'Methods' section). However, the negative finding of this study does not serve to disconfirm the hypothesis (see Locke and Goldstein 1973) that changes in the level of general attention or vigilance affects the production of speech sounds. The present study provided NO DIRECT EVIDENCE that general vigilance or attention diminished across the three speaking tasks. It is conceivable (albeit unlikely, in the author's opinion) that the subjects were able to allocate as much attention to 'monitoring' their speech in the Story condition as when reading isolated phrases from a list.

Further, it is possible that variation in attention to speech does influence L2 speech production, but only in the earliest stages of L2 learning. Even the relatively inexperienced native English speakers of French examined here had lived for nearly a year in a French-speaking environment. Perhaps this enabled them to establish patterns of segmental articulation for French so firmly that they no longer relied on peripheral sensory feedback.

The finding that the phonetic parameters measured here in specific L2 sounds did not differ significantly across the three speaking tasks seems to
Effects of Equivalence Classification

differ from the finding of previous studies (e.g., Beebe 1980) which show that the perceived authenticity with which L2 sounds are produced varies as a function of speaking task. For example, Dickerson and Dickerson (1977) transcribed the English /r/ sounds produced by Japanese learners of English. The L2 learners were perceived to produce /r/ correctly about twice as often when reading word lists than speaking spontaneously. In the present study, authenticity of L2 speech sound production was assessed by acoustic analyses, without reference to the auditory effect of the dimensions being measured.

It is unlikely that the apparent discrepancy between previous studies and the present study was due simply to methodological differences. Flege and Hillenbrand (1984) provided detailed auditory evaluations of some of the speech material acoustically examined in the present study. In forced-choice identification and paired comparison tasks, listeners (both native English and native French) responded no differently to speech produced in isolated phrases read from a list and in a spontaneously generated story. We can only conclude that the authenticity of some L2 sounds, but not others, varies as a function of speaking task. Future studies should be aimed at determining which L2 sounds are most likely to be influenced by changes in speaking task.

4.3. Equivalence classification

We have already discussed the hypothesis that, as the result of equivalence classification, L2 learners will produce new L2 sounds more authentically than similar L2 sounds. A corollary of this hypothesis is that the authenticity with which L2 learners produce similar L2 sounds will be limited by equivalence classification. The production of a speech sound is thought to be guided by a perceptual target which develops as an individual hears many tokens of a sound. Monolingual native speakers of L2 develop an accurate perceptual target for L2 sounds because they hear only L2 sounds. The perceptual target developed by foreign language learners for L2 sounds which are acoustically non-identical to their counterpart in L1, on the other hand, may not be accurate as the result of equivalence classification.

It is hypothesized that when foreign language learners judge an L2 sound as belonging to a familiar (i.e., L1) phonetic category, previous phonetic experience in L1 is brought to bear on their production of the L2 sound. More specifically, it is hypothesized (Flege 1981, 1985) that the perceptual target a learner develops for a similar L2 sound will represent a merger of the phonetic properties of the L2 sound being learned and its counterpart in L1. Assuming that the L2 learner eventually achieves the perceptual tar-
get in speech production, he or she will (at best) only partially approximate the L2 norms for that sound. New L2 sounds not identified with an L1 sound, on the other hand, might evade the effect of equivalence classification and, as a result, be reproduced authentically.

The predicted limiting effect (Flege and Hillenbrand 1984) of equivalence classification was supported by vowel formant measurements. Highly experienced native French speakers of English produced English /u/ with significantly lower F2 values than monolingual native speakers of English. Highly experienced native English speakers of French were observed to produce French /u/ with substantially higher F2 values than has been reported previously for monolingual French-speaking women (987 Hz; Debrock and Forrez 1976).

The predicted limiting effect of equivalence classification was also supported by VOT measurements. Highly experienced native French speakers of English produced English /t/ with VOT values that were significantly shorter than those observed for English monolinguals. Highly experienced native English speakers of French produced French /t/ with VOT values that were substantially longer than has been previously reported for French monolinguals (ca. 20 ms; Benguerel, Hirose, Sawashima, and Ushijima 1978; Caramazza and Yeni-Komshian 1974). It is important to note, however, that the present study did not include data for French monolinguals, so the conclusion that the experienced English speakers of French did not achieve the phonetic norms of French in producing /t/ and /u/ must remain tentative.

Equivalence classification is predicted (Flege 1981, 1985) to have a second important effect. L2 learners who are observed to approximate L2 phonetic norms in producing an L2 sound are predicted to show an effect of L2 learning on L1 speech production. For example, a native English speaker who approximates French phonetic norms in producing French /t/ (i.e., by decreasing the VOT of /t/ relative to the long-lag VOT norm of English) is predicted to produce English /t/ with relatively shorter (i.e., French-like) VOT values than English monolinguals who have not learned French. This prediction was supported by an examination of the /t/ produced in French words by native French speakers of English. These subjects produced French /t/ with an average VOT value that was substantially longer than the short-lag VOT of about 20 ms reported in previous studies examining French monolinguals (Benguerel, Hirose, Sawashima, and Ushijima 1978; Caramazza and Yeni-Komshian 1974).

The predicted L2 effect on L1 production was not seen, however, in the production of English /u/ by native English subjects who had learned French as an L2. They did not produce English /u/ with lower (i.e., French-like) F2 values than English monolinguals who had never learned French.
It should be pointed out that this negative finding does not disconfirm the hypothesis, since even the experienced native English speakers of French failed to approximate French phonetic norms for /u/.

However, the negative findings for VOT represent a more serious challenge to the hypothesis. There was no L2 effect on the English /t/ produced by native English subjects who had learned French as an L2. The native English subjects who spoke French produced English /t/ with VOT values that did not differ significantly from the VOT values in /t/ produced by English monolinguals. This is not surprising, for the native English subjects who were relatively inexperienced in French did not approximate French phonetic norms in producing the /t/ in French words. However, the experienced native English speakers of French did approximate the phonetic norms of French for /t/, and should therefore have produced English /t/ with shorter VOT values than English monolinguals.

One possible explanation for why French did not influence the production of English /t/ by experienced native English speakers of French is that these subjects did not have sufficient experience speaking French. They had spent considerably less time in an L2-speaking country (less than 1.5 years) than the native French-speaking subjects who did show the predicted L2 effect on L1 stops (about 10 years). However, the experienced English speakers of French did have sufficient L2 experience to have noted the acoustic difference distinguishing French from English /t/, for they showed a clear cross-language switch in phonetic implementation. They produced /t/ with significantly shorter (by about 25 ms) VOT values in French than English words. The magnitude of this cross-language switch in VOT is very similar to that reported by Caramazza et al. (1973) for English and French words produced by French Canadians.

A more likely explanation is that the English speakers of French were living in an English-speaking environment at the time of the study. They had not lived in a French-speaking environment for at least six months prior to the study, and a language background questionnaire indicated that many of them heard or spoke French infrequently in the period immediately preceding the experiment. Perhaps an L2 effect on L1 production will be diminished if a talker is using L2 infrequently when a speech sample is taken. It is also possible that the native English subjects accommodated their speech to that of the English-speaking experimenter. That is, the English-speaking subjects may not have shown a French influence on their production of English /t/ because this would have caused their speech to diverge from that of the English-speaking experimenter (see, e.g., Thakerar, Giles, and Cheshire 1982). Given these considerations, it would be useful to examine the English produced by native English speakers living in France, especially if the data were gathered by a native speaker of French.
4.4. Conclusions

The results of this study suggest that human speech does not become unmodifiable or 'crystallize' beyond some point in development or maturation, as do the vocal patterns seen in some songbirds (Marler and Mundinger 1971). Adults appear to remain capable of learning to pronounce new sounds in a foreign language, and to modify their production of L1 sounds differing from phonetic counterparts in the native language. We might conclude, based on the results of this study, that an important difference between children acquiring their native language and older individuals learning a foreign language pertains to the development of phonetic representations. Adults have developed central phonetic representations for L1 speech sounds by the time L2 learning commences. They appear to interpret L2 sounds in terms of sound categories established during L1 acquisition. That is, as a result of a mechanism termed 'equivalence classification' L2 sounds which are similar, but not identical to, sounds in L1 are identified with L1 sounds. A consequence of this is that the sounds of L1 influence sounds produced in L2 and, to a lesser extent, the sounds of L2 influence the sounds of L1.

Equivalence classification should be regarded as an adaptive mechanism during the process of L1 acquisition because it enables the child to find perceptual constancy in the wide range of acoustically different realizations of a phonetic category. It must also be regarded as highly useful for the perception of speech generally, for every phonetic category is realized with a wide range of variants. However, equivalence classification might be regarded as non-adaptive during the process of L2 learning, for it seems to place an upward limit on how authentically certain L2 sounds can be produced. In fact, equivalence classification may turn out to be an important cause for the persistent foreign accent of many adult L2 learners. Testing this hypothesis will require a great deal of future research that assesses not only L2 speech production, but also the perception of similarity between sounds occurring on the phonetic surface of L1 and L2.

NOTES

1. A small portion of the data presented here was previously reported in Flege and Hillenbrand (1984). Data for two additional subject groups (French monolinguals and a group of native English speakers who had resided in Paris for about 12 years) has been gathered since this chapter was accepted for publication. An expanded report including that data, along with more attention to individual subject data, is now being prepared.
2. The term 'phonetic norm' is used here in a purely statistical sense, without reference to the psychological or social reactions of listeners. For example, one aspect of the phonetic norm
for English /t/ is that this stop is realized with a VOT value of about 75 ms. Different individual native speakers of English may realize /t/ with somewhat different mean VOT values (see, e.g., Flege and Massey 1980), but there will be much less difference between individual English speakers than between English speakers (taken as a whole) and speakers of some other language such as French. The term 'phonetic approximation' refers to a change in speech sound production which diminishes the difference which exists between an L2 learner and native speakers of L2 along some specific acoustic dimension. For example, a native English speaker who produces /t/ with shorter VOT values when speaking French than English is said to have 'approximated' the phonetic norm of French because French monolinguals realize /t/ with substantially shorter VOT values than English monolinguals.

3. The existing literature does not provide strong support for the view that the production of individual phonetic segments is regulated, at least at cortical levels of processing, by the ongoing use of peripheral sensory feedback (see Flege, 1985 for a review). However, it is possible that L2 learners can modify broad production parameters (comparable to what have traditionally been referred to as 'bases of articulation') when sufficient attention can be allocated to pronunciation.

4. This raises the important unresolved issue of whether L2 learners relate sounds in a foreign language to L1 sounds on the basis of proprioceptive similarities (i.e., tongue shape, contact, position, and movement patterns), auditory similarities, or both. For example, I can think of no a priori theoretical reason why French /y/ should necessarily be identified with English /u/ rather than English /i/. Native speakers of certain West African languages are said to substitute an /i/ for French /y/ (Spector, personal communication), and French /y/ has been transformed into an /i/-quality vowel in Haitian Creole (Ferguson, personal communication). French /y/ may be identified with /i/ on the basis of similarities in tongue shape and position. The auditory similarity of /y/ and /i/ might also be greater than the auditory similarity of /y/ and /u/. One study (Eskenazi and Lienard 1983) showed that automatic matching algorithms based on the acoustic properties of French vowels resulted in more confusions between /y/ and /i/ than between /y/ and /u/. However, my own subjective impression (as a native English speaker of French), as well as that of other native English speakers I have queried, leads me to think that French /y/ is perceived to be closer in the phonetic vowel space to English /u/ than /i/. Perhaps cross-language similarity judgments are importantly influenced by the phonemic inventory of a listener's L1. Clearly, a great deal of work is needed to clarify which L1 and L2 sounds are identified with one another, and upon which criteria cross-language similarity judgments are based.

REFERENCES


French stop production: A fiberscopic, acoustic and electromyographic study’, *Folia 
Phoniatrica* 30, 175–198.

five-year-old following sudden loss of hearing’, *Journal of Speech and Hearing Disorders* 
47, 181–189.

Borden, G.J. (1979), ‘An interpretation of research on feedback interruption in speech’, *Brain 

Borden, G.J. (1980), ‘Use of feedback in established and developing speech’, in N. Lass (ed.), 
*Speech and Language: Advances in Basic Research and Practice*, Vol. 3, Academic Press, 

Borden, G.J., K.S. Harris, H. Fitch, and H. Yoshioka, (1981), ‘Producing relatively un-
familiar speech gestures: A synthesis of perceptual targets and production rules’, *Haskins 
Laboratories Status Reports on Speech Research* SR-66, 85–117.


et du français, Méthode et résultats’, *Revue de Phonétique Appliquée* 37, 27–73.


Dickerson, L., and W. Dickerson, (1977), ‘Interlanguage phonology: Current research and fu-
ture directions’, in S.P. Corder and E. Roulet (eds.), *The Notions of Simplification, Inter-


Flege, J.E. (1985), The production and perception of foreign language speech sounds, in H. 
Winitz (ed.), *Human Communication and Its Disorders*, Vol. 1, Ablex Publishing, Nor-
wood, in press.

Flege, J. and R. Hammond, (1982), ‘Mimicry of non-distinctive phonetic differences between 


Gibson, E. (1969), *Principles of Perceptual Learning and Development*, Appleton-Century-
Crofts, New York.


Schouten, M.E.H. (1975), Native-language interference in the perception of second-language vowels, Elinkwijk, Utrecht.


Effects of Equivalence Classification


Williams (Willis), L. (1979), ‘The modification of speech perception and production in second-language acquisition’, *Perception and Psychophysics* 26, 95–104.
