

## Chapter • 8

---

# Second Language Speech Learning *Theory, Findings, and Problems*

---

*James Emil Flege*

The aim of our research is to understand how speech learning changes over the life span and to explain why “earlier is better” as far as learning to pronounce a second language (L2) is concerned. An assumption we make is that the phonetic systems used in the production and perception of vowels and consonants *remain adaptive over the life span*, and that phonetic systems reorganize in response to sounds encountered in an L2 through the addition of new phonetic categories, or through the modification of old ones. The chapter is organized in the following way. Several general hypotheses concerning the cause of foreign accent in L2 speech production are summarized in the introductory section. In the next section, a model of L2 speech learning that aims to account for age-related changes in L2 pronunciation is presented. The next three sections present summaries of empirical research dealing with the production and perception of L2 vowels, word-initial consonants, and word-final consonants. The final section discusses questions of general theoretical interest, with special attention to a featural (as opposed to a segmental) level of analysis. Although nonsegmental (i.e., prosodic) dimensions are an important source of foreign accent, the present chapter focuses on phoneme-sized units of speech. Although many different languages are learned as an L2, the focus is on the acquisition of English.

### INTRODUCTION

Foreign accents in English are common in the speech of non-native speakers. Listeners hear foreign accents when they detect divergences from English phonetic norms along a wide range of segmental and suprasegmental (i.e., prosodic) dimensions. Foreign accents may have a number of undesirable consequences for non-native speakers. They



may make non-natives difficult to understand, especially in non-ideal listening conditions (e.g., Lane 1963). They may cause listeners to misjudge a non-native speaker's affective state (e.g., Gumperz 1982; Fayer and Krasinski 1987; Holden and Hogan 1993), or provoke negative personal evaluations, either as the result of the extra effort a listener must expend in order to understand, or by evoking negative group stereotypes (Lambert et al. 1960; Giles 1970).

Researchers have proposed many different explanations for foreign accents. For example, neurological maturation might reduce neural plasticity (Penfield 1965; Lenneberg 1967), leading to a diminished ability to add or modify sensorimotor programs for producing sounds in an L2 (Sapon 1952; McLaughlin 1977). Others have suggested that foreign accents are caused, at least in part, by the inaccurate *perception* of sounds in an L2 (Flege 1992a,b; Rochet this volume). Still others have proposed that the primary cause of foreign accents is inadequate phonetic input, insufficient motivation, psychological reasons for wanting to retain a foreign accent, or the establishment of incorrect habits in early stages of L2 learning (see Flege 1988b, for review). The diversity of these hypotheses attests to the complexity of the phenomenon of foreign accent.

Neurological maturation has often been cited as the primary impetus for a critical period for speech learning. Many believe that new forms of speech cannot be learned perfectly once a critical period has been passed (e.g., Lenneberg 1967; Scovel 1988; Patkowski 1990). Although this may be true, such a conclusion fails to provide insight into how L2 learning differs from L1 acquisition, or what actually *causes* foreign accent (Flege 1987b; Long 1990). Nor does a critical period hypothesis seem consistent with the foreign accent ratings obtained by Flege, Munro, and MacKay (1995a). This study assessed degree of perceived foreign accent in the English spoken by native Italian (NI) subjects who differed according to their age of learning (AOL) English. The 240 NI subjects examined had begun learning English between the ages of 3 and 21 years in Canada, where they had lived for over 30 years on average. Native English-speaking listeners used a continuous scale to rate English sentences for degree of accent in English. Figure 1 shows a strong linear component ( $r = 0.71$ ) in the relation between the NI subjects' AOL and their degree of perceived foreign accent. The later in life the NI subjects began learning English, the more strongly foreign-accented their English sentences were judged to be. If a critical period exists, it apparently does not result in a sharp discontinuity in L2 pronunciation ability at around puberty.

In 1981, Flege noted the following paradox, which eventually led to the formulation of the model presented below. At an age when children's sensorimotor abilities are generally improving, they seem to



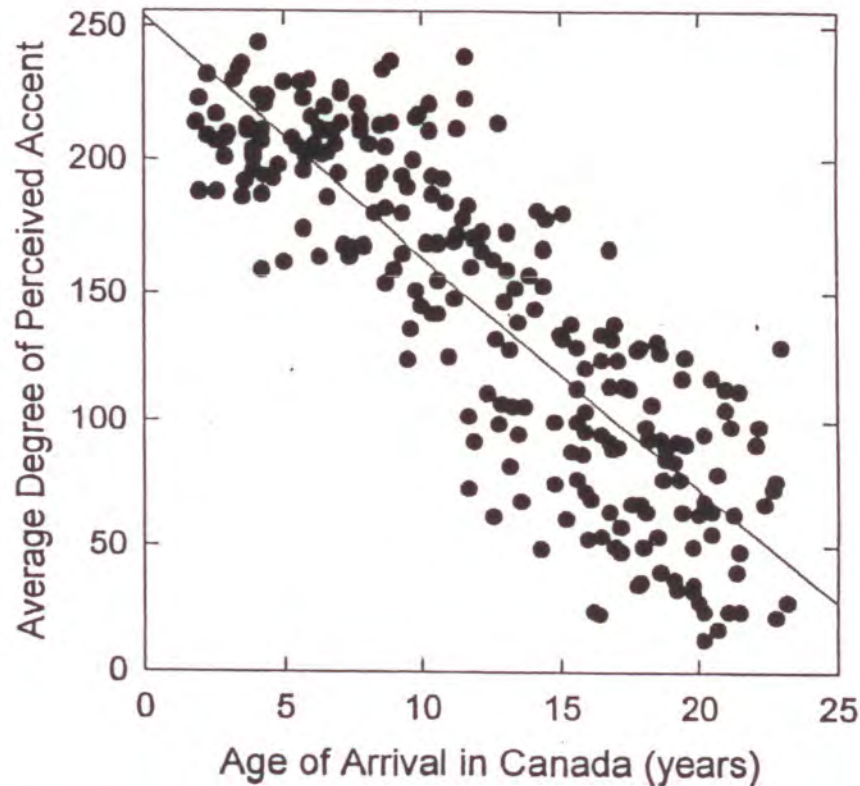


Figure 1. The mean foreign accent ratings accorded English sentences spoken by 240 native Italian speakers of Italian who differed according to age of learning (AOL) English, in years. Each mean is based on 150 judgments (5 sentences  $\times$  10 native English-speaking listeners) (data are from Flege, Munro, and MacKay 1995a).

lose the ability to learn the vowels and consonants of an L2. Segmental production research in the 1970s focused on the classroom learning of foreign languages, or on early stages of naturalistic L2 learning. Interference from the L1 was seen as the primary phonological cause of foreign accent. A common view was that 1) an L2 sound that is "identified" with a sound in the L1 will be replaced by the L1 sound, even if the L1 and L2 sounds differ phonetically; 2) contrasts between sounds in the L2 that do not exist in the L1 will not be honored; and 3) contrasts in the L1 that are not found in the L2 may nevertheless be produced in the L2 (e.g., Weinreich 1953; Lehiste 1988). Little attention was paid to the effect of AOL or variations in amount of L2 experience. Thus, it was common to encounter descriptions of L2 production errors that contained no reference to *when* in life L2 learning commenced, or *how long* the L2 had been spoken, or *with whom* (e.g., Koutsoudas and Koutsoudas 1983).

Second-language speech learning has been characterized as more "analytic" than L1 acquisition (e.g., Mack 1988), especially for those



whose early exposure to the target L2 comes primarily through the written word. Interference was usually described as occurring at a segmental level (but cf. Weinreich 1957, who emphasized featural-level analysis). This assumed that bilinguals decompose unfamiliar words of the L2 into phonemes and allophones. Given this conceptualization, accurate production of an L2 sound would require: (1) an accurate appraisal of the properties that differentiate the L2 sounds from one another, and from sounds in the L1; (2) the storing and structuring of this information in long-term memory; and (3) the learning of gestures with which to reliably reproduce the represented L2 sounds.

The view that foreign accent is caused by motoric difficulty (3, above) is inconsistent with the following observations. Although very common, foreign accents are apparently not *inevitable* (Hill 1970; Novoa, Fein, and Obler 1988). For example, Neufeld (1979) required his adult subjects to listen for a long time before talking. They were then apparently able to pronounce sentences in an unfamiliar foreign language without foreign accent. Adults may be as able as children to imitate foreign sounds. Finally, Snow and Hoefnagel-Höhle (1979) examined the naturalistic acquisition of Dutch by native English (NE) speakers differing in chronological age. Adults were at first better able than children to imitate and spontaneously produce Dutch sounds. By the end of the one-year study period, the spontaneous production of Dutch sounds by the NE adults and children was roughly equivalent.

Why do adults fail to exploit their motoric abilities fully? One possibility is that they fail to *perceive* L2 sounds accurately (1 and 2, above). Differences in segmental perception between native and non-native speakers have been shown to exist in many studies (e.g., Miyawaki et al. 1975; Flege and Eefting 1986; Flege and Hillenbrand 1986, 1987). In gating experiments, non-natives must hear a larger portion of a word than native speakers to recognize the word (Nooteboom and Truin 1980; Koster 1987). Perceptual differences are indirectly evident in the difficulty non-natives may have in comprehending English, especially distorted or synthetic speech versions of English or anomalous sentences (Oyama 1978; Nabelek and Donahue 1984; Greene Pisoni, and Gradman 1985; Mack 1988; Ozawa and Logan 1989; Mack et al. 1990; McAllister 1990).

The hypothesis that articulatory errors have a perceptual basis has been examined extensively in L1 acquisition research. Powers (1957) reviewed evidence that children's misarticulation of sounds cannot usually be traced to deficits in auditory acuity. A large number of studies that attempted to establish a link between the ability to discriminate a broad range of speech sounds, on the one hand, and speech sound production errors, on the other, yielded contradictory



results, at least for normally developing children (Weiner 1967). Other studies tested the hypothesis that the inaccurate production of an L1 sound could be attributed to an inadequate perceptual specification of it (Spriestersbach and Curtis 1951), or to difficulty discriminating a sound produced in error from its replacement. For example, Monnin and Huntington (1974) found that children who misarticulated /ɹ/ (usually as /w/) were less able to correctly identify /ɹ/ and /w/ in a picture pointing task than were children who produced /ɹ/ correctly. Other studies, however, failed to reveal a direct link between the perception and misarticulation of particular sounds (e.g., Haggard, Corrigan, and Legg 1971; Waldman, Singh, and Hayden 1978). Locke (1980) designed special tests for children aged 3–7 years, based on the specific consonants each child misarticulated. The children failed to perceptually distinguish a sound they misarticulated from its replacement sound in less than one-third of instances. Furthermore, many of the observed perceptual errors involved /f/ versus /θ/, and so might have had an auditory basis (see Sussman 1993). This led Locke (1980, p. 465) to conclude that the root cause of many L1 segmental production errors is to be found at a “motor” level rather than at a “mentalistic (level) of linguistic organization.”

The conclusion drawn for L1 acquisition may not apply equally to L2 learning, which differs from L1 acquisition in certain respects. Most importantly, bilinguals tend to interpret sounds encountered in an L2 through the “grid” of their L1 phonology (Trubetzkoy 1939; Wode 1978). This virtually ensures that non-native speakers will perceive at least some L2 vowels and consonants differently than do native speakers. For example, French /y/ is mispronounced as /i/ by Portuguese learners, but as /u/ by native English learners. Data reported by Rochet (this volume) suggests that native Portuguese learners of French may hear /y/ tokens as /i/, whereas English learners of French hear /y/ tokens as /u/. Japanese and Russian both have /s/ and /t/, but lack /θ/. Japanese learners mispronounce English /θ/ as /s/, Russians as /t/ (Weinberger 1990). This is not to say, however, that non-natives’ perception of L2 sounds remains constant. The results of feedback training experiments have suggested that language-specific perceptual patterns are modifiable to some extent (e.g., Logan, Lively, and Pisoni 1991; Strange 1992). This suggests that the perceived relation of L1 and L2 sounds may change during naturalistic L2 learning.

#### SPEECH LEARNING MODEL

Flege and his colleagues have developed a speech learning model (SLM) that aims to account for age-related limits on the ability to produce L2 vowels and consonants in a native-like fashion. The SLM is concerned



primarily with the ultimate attainment of L2 pronunciation, so work carried out within its framework focuses on bilinguals who have spoken their L2 for many years, not beginners. During L1 acquisition, speech perception becomes attuned to the contrastive phonic elements of the L1. Learners of an L2 may fail to discern the phonetic differences between pairs of sounds in the L2, or between L2 and L1 sounds, either because phonetically distinct sounds in the L2 are "assimilated" to a single category (see Best this volume), because the L1 phonology filters out features (or properties) of L2 sounds that are important phonetically but not phonologically, or both. The model claims that without accurate perceptual "targets" to guide the sensorimotor learning of L2 sounds, production of the L2 sounds will be inaccurate. The model does *not* claim, however, that all L2' production errors are perceptually motivated. For example, motoric output constraints based on permissible syllable types in the L1 may cause Spanish speakers to pronounce the word "school" as [eskul]. Still, a basic tenet of the model is that many L2 production errors have a perceptual basis.

The postulates and hypotheses that comprise the current version of the SLM are presented in Table I. The hypotheses derive from the postulates, and from empirical data presented in previous articles and chapters (see, e.g., Flege 1981, 1988b, 1991a, 1992a,b). It should be emphasized that this is a working model and is subject to further revision as new data are gathered. Regardless of whether the SLM is ultimately supported or disconfirmed, it serves as a useful heuristic for planning research. Also, it generates testable predictions and can be used to organize and interpret a wide range of empirical data. In the following section we discuss the model's hypotheses in general terms, and illustrate how they apply to important questions pertaining to students of L2 learning. Specific hypotheses are indicated by boldface (e.g., **H4**, **H7**, and so on).

The SLM differs from other approaches to L2 acquisition in important ways. According to the model's first hypothesis (**H1**), learners perceptually relate positional allophones in the L2 to the closest positionally defined allophone (or "sound") in the L1. Weinreich (1957) referred to L1 and L2 sounds that have been linked perceptually in this way as "diaphones." The phonetic level of analysis envisaged by **H1** is less abstract than the phonemic level of Contrastive Analysis (e.g., Lado 1957), but is, nevertheless, an *abstract* level of organization. Even within a single phonetic context, the production of a position-sensitive allophone is apt to vary considerably according to such factors as speaking rate, degree of stress, the talker's age and gender, and speaking style or clarity (Lindblom 1990a,b).

Support for **H1** comes from evidence that L2 learners are more successful at producing and perceiving certain allophones of English



Table 1. Postulates and hypotheses forming a speech learning model (SLM) of second language sound acquisition.

---

Postulates	
<b>P1</b>	The mechanisms and processes used in learning the L1 sound system, including category formation, remain intact over the life span, and can be applied to L2 learning.
<b>P2</b>	Language-specific aspects of speech sounds are specified in long-term memory representations called <i>phonetic categories</i> .
<b>P3</b>	Phonetic categories established in childhood for L1 sounds evolve over the life span to reflect the properties of all L1 or L2 phones identified as a realization of each category.
<b>P4</b>	Bilinguals strive to maintain contrast between L1 and L2 phonetic categories, which exist in a common phonological space.
Hypotheses	
<b>H1</b>	Sounds in the L1 and L2 are related perceptually to one another at a position-sensitive allophonic level, rather than at a more abstract phonemic level.
<b>H2</b>	A new phonetic category can be established for an L2 sound that differs phonetically from the closest L1 sound if bilinguals discern at least some of the phonetic differences between the L1 and L2 sounds.
<b>H3</b>	The greater the perceived phonetic dissimilarity between an L2 sound and the closest L1 sound, the more likely it is that phonetic differences between the sounds will be discerned.
<b>H4</b>	The likelihood of phonetic differences between L1 and L2 sounds, and between L2 sounds that are noncontrastive in the L1, being discerned decreases as AOL increases.
<b>H5</b>	Category formation for an L2 sound may be blocked by the mechanism of equivalence classification. When this happens, a single phonetic category will be used to process perceptually linked L1 and L2 sounds (diaphones). Eventually, the diaphones will resemble one another in production.
<b>H6</b>	The phonetic category established for L2 sounds by a bilingual may differ from a monolingual's if: 1) the bilingual's category is "deflected" away from an L1 category to maintain phonetic contrast between categories in a common L1-L2 phonological space; or 2) the bilingual's representation is based on different features, or feature weights, than a monolingual's.
<b>H7</b>	The production of a sound eventually corresponds to the properties represented in its phonetic category representation.

---

phonemes than others. For example, native Japanese (NJ) speakers have difficulty producing and perceiving English /l/ and /ɹ/. This is because Japanese has just one liquid, whereas English has two. However, allophones of English /ɹ/ and /l/ differ phonetically and are learned by non-native speakers at different rates and to differing extents. For example, NJ learners of English characteristically perceive and produce English liquids more accurately in word-final than word-initial position (Strange 1992), perhaps because the acoustic difference



between English /ɹ/ and /l/ is more robust in final than initial position (Sheldon and Strange 1982). Another possible explanation for the word-position effect is that final but not initial allophones of English /ɹ/ and /l/ are *categorized* differently by NJ learners of English. Takagi (1993) examined the perceived relation between Japanese /r/, English /ɹ/, and English /l/. Native Japanese subjects who had arrived recently in the United States used katakana symbols to write English nonsense words presented auditorily. As expected from loan-word phonology data, English /rV/ and /ɹV/ syllables were written with the symbols for Japanese /rV/. However, whereas tokens of word-final English /ɹ/ were written as "a" (78% of instances), word-final /l/ tokens were written as "ru" (53%) or "o" (30% of instances).

Rochet (this volume) suggested that all L2 vowels are identified as realizations of an existing L1 vowel category. Best, McRoberts, and Sithole (1988) suggested that foreign consonants are judged to be realizations of an L1 consonant, or else are heard as nonspeech (but cf. Best this volume). However, many L2 production studies have provided indirect evidence that, over time, L2 learners take note in some way of cross-language phonetic differences (e.g., Wenk 1979; Hammarberg 1988, 1990; Wieden 1990). According to the model's second hypothesis (H2), bilinguals sometimes establish a new phonetic category representation for sounds in the L2. And, according to H7, L2 sounds will eventually be produced as specified in phonetic category representations. If the new phonetic category established by a bilingual for an L2 sound matches native speakers', then the L2 sound will be produced accurately.

According to H3, the likelihood of cross-language phonetic differences being discerned increases with degree of perceived cross-language phonetic difference. There is evidence, for example, that Japanese /r/ (despite its symbolization) is closer perceptually to English /l/ than /ɹ/ (Sekiyama and Tohkura 1993; Takagi 1993). This leads to the expectation that a larger proportion of Japanese learners of English will discern some or all of the phonetic differences between Japanese /r/ and English /ɹ/ than between Japanese /r/ and English /l/. H4 states that the likelihood of cross-language phonetic differences being discerned decreases with AOL. So, for example, more German-speaking 10-year olds than adults should discern the phonetic difference between English /æ/ and the closest German vowel (probably /ɛ/). In fact, the perceived distance between /ɛ/ and /æ/ vowels is greater for German children than adults (Butcher 1976); and German adults but not children have difficulty discriminating /ɛ/ and /æ/ in an oddity task (Weiher 1975; see Figure 2 below).

As mentioned earlier, neurological maturation is often identified as the cause of foreign accentedness. Patkowski (1990, p. 87) suggested



that "the difference between child and adult learners (of speech and language) is of a fundamental, qualitative nature" and that preadolescent versus postadolescent L2 learners represent "different populations" of learners. We might suppose that if L2 production accuracy were limited by maturation, such a limitation would apply across the board to the full range of L2 sounds that differ phonetically from sounds in the L1. On the other hand, H3 and H4 predict that fewer sounds in the L2 will be produced accurately as AOL increases (both in terms of the range of sounds and the proportion of bilinguals). Degree of foreign accent is correlated with the number of segmental errors (e.g., Brennan, Ryan and Dawson 1975). Thus, the linear increase of foreign accent with AOL shown in Figure 1 seems to agree better with the SLM than with the prediction generated by a neurological maturation hypothesis.

A failure to discern cross-language phonetic differences may arise at one or more processing levels. In some circumstances, listeners can gain access to the sensory properties which distinguish pairs of unfamiliar L2 sounds (Miyawaki et al. 1975; Werker and Tees 1984; Mann 1986), or L1 and L2 sounds (Flege 1984; Flege and Munro 1994). However, they may fail to do so during the on-line processing of speech. Speech perception becomes automatic during L1 speech development (e.g., Linell 1982), which frees resources for other processing tasks (e.g., Mayberry and Eichen 1991). This may cause learners to attend less to phonetic detail when learning L2 than L1 sounds. Listeners may fail to gain access to sensory properties associated with certain L2 sounds as the result of pre-attentive processes. Neisser (1976, p. 20) spoke of "anticipatory schemata" that "prepare the perceiver to accept certain kinds of information rather than others." Jusczyk (1992, 1993) later posited the existence of automatic interpretative schemes that determine which properties of incoming words are attended to in early stages of processing. It is also possible that sensory information that has initially been processed is discarded at a later processing stage by non-native speakers as nondistinctive, or that non-natives weight features differently than do NE speakers (e.g., Weinreich 1953; Flege and Hillenbrand 1986, 1987; Flege 1988).

Traditionally, the term "interference" has applied only to the influence of the L1 on the production of an L2. According to H5 and H6, however, cross-language phonetic interference is bidirectional in nature. The model predicts two different effects of L2 learning on the production of sounds in an L1, depending upon whether or not a new category has been established for an L2 sound in the same portion of phonological space as an L1 sound. Previous studies have supported H5 by showing that L1 and L2 sounds that are perceptually linked to one another (diaphones) come to resemble one another in production.



For example, Flege (1987a) found that bilinguals produced stops in their L1 with VOT values resembling those typical for stops in the L2 (see also Peng 1993).

**H6** was added recently to the model. It is based on the observation that in the vowel system of languages, vowels tend to disperse so as to maintain sufficient auditory contrast (e.g., Liljencrants and Lindblom 1972; Lindblom 1990b). Recently obtained L2 production data (Bohn and Flege 1992) led us to accept that a bilingual's L1 and L2 categories *exist in a common phonological space*. By hypothesis, a bilingual's L1 and L2 vowels disperse so as to maintain contrast within that individual's phonological space. If so, then a category established by a bilingual for an L2 vowel may be "deflected" away from an L1 vowel, and so differ from a native speaker's category for the L2 vowel sound.

Analogies to **H6** can be found in historical sound change and dialect geography. As languages change, the raising of vowel A may precipitate the raising of B, which then causes C to rise. As the result of such push chains, the vowels A, B, and C may be produced differently, while the contrasts between them are preserved (Martinet 1955). Moulton (1945) observed that in Swiss German dialects that had an /æ/, the vowel /a/ was produced with backed variants. In /æ/-less dialects, on the other hand, /a/ was produced with central, or even fronted, variants. The results of a case study by Mack (1990) were consistent with **H6**. Acoustic measurements revealed that a 10-year-old boy who spoke French at home and English elsewhere produced /b d g/ with short-lag voice-onset time (VOT) values in both French and English. He produced /p t k/ with VOT values averaging 66 ms in French, but with values averaging 108 ms in English. Thus, this child managed to maintain *phonetic contrast* between three categories in his L1 and L2, but at the cost of producing /p t k/ inaccurately in *both* languages (i.e., with VOT values that were too long for both French and English).

The kinds of L1 production changes predicted by **H5** and **H6** are consistent with an incidental finding of a study by Flege, Munro, and MacKay (1995a). The 240 NI speakers who participated were asked to evaluate their own ability to pronounce Italian and English. There was a modest negative correlation between self-estimated ability to pronounce the L1 and L2. The NI subjects who began learning English before the age of 12 years said they pronounced English better than Italian; whereas, the reverse held true for those who began learning English later in life. Very few (6%) of the NI subjects pre- and post-adolescent learners indicated that they could speak both English *and* Italian without accent, including subjects who used English and Italian with equal frequency.



**H6** specifies a second circumstance in which a category established for an L2 sound by a non-native speaker may differ from a native speaker's category. This is predicted if the non-native speaker's phonetic category is based on different features than the native speaker's, which could arise if an L2 sound is distinguished from other L2 sounds by features not exploited in the L1. This important revision of the SLM leads us to expect that, even when categories are established for an L2 sound, *the L2 sound might not be produced exactly as it is produced by native speakers*. Thus, the model no longer predicts "mastery" of certain L2 sounds, and the model is now congruent with Grosjean's view of bilingualism. Grosjean (1989) suggested that the "mixing" of the L1 and L2 is inevitable, because a bilingual's two language systems are both constantly engaged. This view implies that bilinguals do not "switch" between two distinct phonetic systems, at least not as has been assumed traditionally.<sup>1</sup>

#### PRODUCTION AND PERCEPTION OF VOWELS

The SLM generates specific predictions concerning the production and perception of L2 vowels. First, even adult L2 learners are likely to discern the phonetic differences between certain L1 and L2 vowels, especially if the L1 has fewer vowels than the L2 (e.g., the 5-vowel system of Spanish in comparison to the 15-vowel English system). When this happens, new phonetic categories will be established for the L2 vowels (**H2**), and the L2 vowels will eventually be produced as specified in their phonetic category representations (**H7**). By **H3**, the greater the perceived distance of an L2 vowel from the closest L1 vowel, the greater is the likelihood that a new category will be established for the L2 vowel. So, for example, a native Spanish (NS) speaker should be more likely to establish a phonetic category for English /æ/ or /ɚ/ than for English /i/ (which differs only slightly from Spanish /i/). By **H4**, a native Spanish 8-year old should be more likely to establish a category for English /æ/ and /ɚ/ than a native Spanish 16-year old.

The model predicts different effects of L2 learning on the production of L2 vowels, depending upon whether or not a new category has been established for an L2 vowel. For example, using an orthographic classification task, Flege (1991c) showed that Spanish speakers with little or no experience in English tend to identify English /æ/ tokens as realizations of their Spanish /a/ category. If Spanish learners of English persist in doing so, and are, thereby, unable to establish a category for English /æ/, the model predicts that they will produce English /æ/

---

<sup>1</sup>Mack (1989) claimed, for example, that the "dominant" language may remain "impervious" to an influence by the nondominant language, at least for certain phonetic dimensions.



with Spanish /a/-like properties, and vice versa. (That is, their L2 /æ/ will have F2 (second formant) values that are too low, and their L1 /a/ will have F2 values that are too high.) However, if a category is established for English /æ/, then the model predicts accurate production unless the Spanish learner's new English /æ/ category is deflected away from Spanish /a/. An indirect consequence of a "deflection" would be a lowering of F2 values in Spanish /a/ (i.e., a backing of Spanish /a/ in phonological space away from English /æ/). Although testable, the data now available are insufficient to support or disconfirm such hypotheses. One gap in the literature is the virtual absence of work examining bilinguals' production of L1 vowels (but see Bohn and Flege 1992). The general pattern of available data reviewed below, however, are consistent with the framework just outlined.

### Categorical Discrimination of L1 and L2 Vowels

Best (1994) observed that properties defining category membership are likely to differ from those defining the systematic relationship among categories. Perceptual training data obtained by Morosan and Jamieson (1989) suggested that learning to distinguish a sound, X, from another sound, Y, will not necessarily generalize to an X-Z contrast because cues to an X-Y contrast may not suffice to distinguish X from Z. Thus, the notion "phonetic category" implies the perceptual ability to: 1) identify a wide range of different phones as being "the same," despite auditorily detectable differences between them along dimensions that are not phonetically relevant; and 2) ability to distinguish the multiple exemplars of a category from realizations of other categories, even in the face of noncriterial commonalities (Kluender, Diehl, and Killeen 1987).

A two-alternative forced-choice (2AFC) identification test is not well suited as a test of category formation. For example, Flege and Bohn (1989) examined NS subjects' identification of vowels in *beat-bit* (/i-ɪ/) and *bet-bat* (/ɛ-æ/) continua. In both, F1 (first formant) frequency and vowel duration were varied orthogonally. The NS subjects managed to partition both continua into two response categories, but in neither instance did the data provide compelling evidence for the establishment of categories for vowels not found in Spanish (namely, /æ/ and /ɪ/). As discussed by Bohn (this volume), many NS subjects identified members of the *beat-bit* continuum based primarily on vowel duration rather than on spectral quality, as was the case for NE subjects. These NS subjects may have based their identifications on a readily available auditory property (duration) rather than by referencing incoming stimuli to two distinct long-term memory representations. An over-reliance on duration was not evident for the *bet-bat* continuum. However, evidence obtained by Flege (1991c) suggested



that some of the NS subjects may have identified endpoints of this continuum (/ɛ/ and /æ/) in terms of two distinct *Spanish* categories (/e/ and /a/).

An oddity discrimination task has been used in L2 research to determine if learners can discriminate various L2 sounds (e.g., Weiher 1975; Lamminmäki 1979), and might be used to test for category formation. Gottfried (1983) administered an ABX (3 stimuli: A, B always different) "categorical" discrimination task (see Beddor and Gottfried this volume) to French native speakers and inexperienced NE speakers of French. Triads of stimuli contained potentially confusable French vowels. As expected, the NE subjects made more errors than the French subjects (25% vs. 17%), especially on triads containing the front rounded vowels /y œ ø/, which do not occur in English. The lack of a categorical representation for French front rounded vowels may have contributed to errors by the NE subjects.

Flege, Munro, and Fox (1995) modified the categorical discrimination task to discourage within-category discrimination, and to encourage subjects to group sounds into *phonetically relevant* equivalence classes. As in the Gottfried (1983) study, the three stimuli in each triad were spoken by different talkers to encourage responses in a general rather than token-specific mode (e.g., Mullinex, Pisoni, and Martin 1989; Uchanski et al. 1991). The inter-stimulus interval between the members of each triad was somewhat longer than the one used by Gottfried (1983) to further reduce the possibility that correct responses could be based on information in auditory short-term memory (see, e.g., Ferrero et al. 1979; Werker and Logan 1985). Finally, the inclusion of "catch" triads, which consisted of realizations of a single category spoken by three different talkers (correct response: "no odd item out") encouraged subjects to respond only to phonetically relevant differences, not merely auditorily detectable ones.<sup>2</sup>

Table II summarizes the results obtained by Flege, Munro, and Fox (1995) for NE speakers and groups of native Spanish subjects who were relatively experienced (NS-1) or inexperienced (NS-2) in English. The pattern of between-group differences varied as a function of the acoustic and articulatory differences between the vowel contrasts tested. The subjects in all three groups readily discriminated Spanish /a/ tokens from realizations of English /e/, /i/ and /ɪ/. However, for triads testing the categorical discrimination of Spanish /a/ versus English /ɑ/, significantly lower percent correct scores were obtained

---

<sup>2</sup>The categorical discrimination task just described did not require training and obviated problems seen in identification tasks where listeners must choose from a large set of response alternatives, or in tasks that make use of written key words.



from the NS-1 and the NS-2 subjects than from NE subjects. For triads testing the /a/-/ε/, and /a/-/æ/ contrasts, the subjects in NS-2 but not in NS-1 responded correctly significantly less often than did the NE subjects ( $p < .05$ ).

These results were interpreted to mean that the NS subjects identified Spanish and English vowels that are distant from one another in phonological space (Moulton 1945) in terms of *two distinct* categories; whereas, they were less likely to do so for less distant pairs of vowels. The NE subjects were likely to have performed well on triads testing the contrast between Spanish /a/ and the English vowels /ε/ and /æ/ because they identified Spanish /a/ tokens as realizations of English /a/ (and, thus, distinct from the other two English vowels). The fact that inexperienced but not experienced NS subjects performed more poorly than did the NE subjects on /a/-/ε/ and a/-/æ/ contrasts suggests that some of the experienced NS subjects may have established categories for /ε/ and /æ/, neither of which occur phonemically in Spanish.

Flege, Munro, and Fox (1995, Experiment 1) provided additional evidence of categorical effects on the perception of L2 vowels. Native English and NS subjects rated pairs of Spanish and English vowels for degree of dissimilarity using a nine-point scale. The results for five different vowel pairs containing a Spanish /a/ token and a token of English /i/, /ɪ/, /ε/, /æ/, or /ʌ/ are shown in Figure 2. The NE subjects rated /a/-/ε/ and /a/-/æ/ pairs as significantly more dissimilar than did experienced (NS-1) or inexperienced (NS-2) native Spanish subjects.<sup>3</sup> It appears that the NE subjects were more likely than the NS subjects to identify vowels in /a/-/ε/ and /a/-/æ/ pairs as realizations of *two phonetically distinct* categories and that this augmented the degree of perceived vowel dissimilarity.

The NE and NS subjects responded to other vowel pairs in much the same manner, however. Both the NE and NS subjects rated the vowels in /a/-/ʌ/ pairs as very similar, probably because all subjects heard realizations of a single vowel category. The NE and NS subjects agreed in rating the vowels in /a/-/ɪ/ pairs as somewhat less dissimilar than vowels in /a/-/i/ pairs. Vowels in both of these pairs were likely to have been classified differently by all subjects, and so the common response by the NE and NS subjects may have reflected degree of *auditory difference*. Although not shown in Figure 2, the NS subjects rated English /i/-/ɪ/ pairs as significantly more similar than did the NE subjects. This finding, when taken together with converging evidence obtained using other paradigms (Flege and Bohn 1989; Flege

---

<sup>3</sup>Different subjects participated in the similarity rating experiment (Experiment 1) and in the categorial discrimination test (Experiment 2) mentioned earlier.



Table II. Percentage of correct responses obtained from native speakers of English and Spanish in a categorial discrimination task.<sup>a</sup>

Vowel Contrast	Subject group		
	NE <sup>b</sup>	NS-1 <sup>c</sup>	NS-2 <sup>d</sup>
/a/-/i/	98(4)	95(6)	95(5)
/a/-/ɪ/	91(15)	96(7)	89(17)
/a/-/eɪ/	98(5)	95(5)	97(6)
/a/-/ɛ/	80(29)	66(26)	51(27)
/a/-/æ/	77(31)	59(24)	38(25)
/a/-/ɑ/	57(36)	26(18)	18(5)

<sup>a</sup>Triads of stimuli tested the contrast between tokens of Spanish /a/ and tokens of English /i/, /ɪ/, /eɪ/, /ɛ/, /æ/ and /ɑ/ (data are from Flege, Munro, and Fox 1995).

<sup>b</sup>NE, monolingual native speakers of English.

<sup>c</sup>NS-1, relatively experienced Spanish speakers of English.

<sup>d</sup>NS-2, inexperienced Spanish speakers of English.

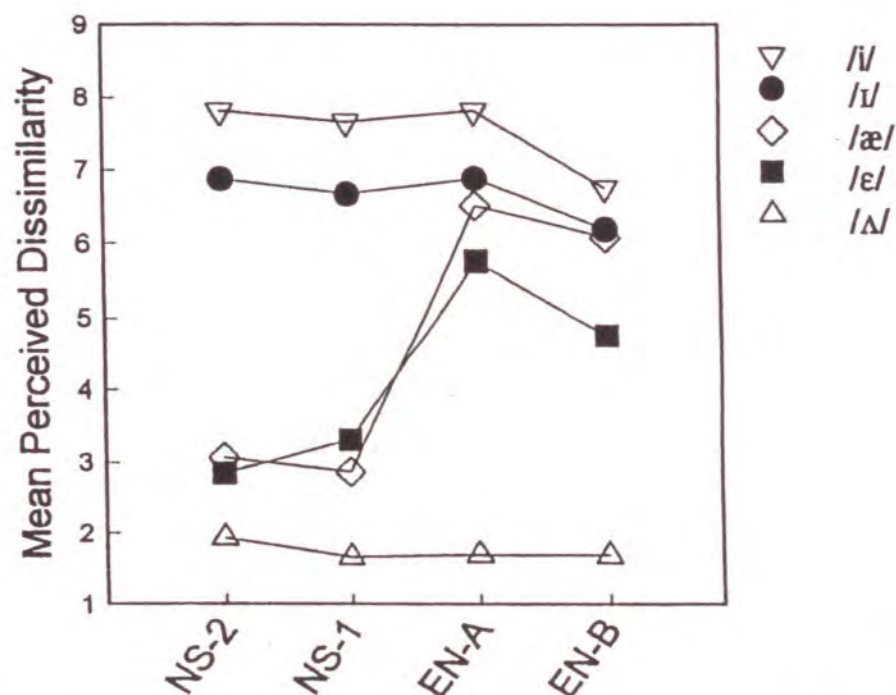


Figure 2. The mean perceived dissimilarity of vowel pairs by English-speaking listeners (EN-A, EN-B), relatively experienced native Spanish speakers of English (NS-1), and relatively inexperienced Spanish speakers of English (NS-2). Multiple tokens of Spanish /a/ were paired with tokens of English /i/, /ɪ/, /æ/, /ɛ/, and /ɑ/ (data are from Flege, Munro, and Fox 1994, Experiment 1).



1991c), suggests indirectly that some NS speakers of English detect the *auditory* difference between /i/ and /ɪ/ tokens, although they do not *categorize* these vowels as different. This is consistent with the model's claim that changes in L2 production can come about even when *phonetic* differences between L1 and L2 sounds are not discerned (i.e., when phonetic category formation is blocked by equivalence classification).

Flege (1995) used the categorial discrimination task described earlier to assess the perception of English vowels by non-native subjects differing in L1 background. The study's aims were to: 1) determine if the non-natives would fail to discriminate English vowels that could be discriminated by NE subjects; and 2) to determine if the non-natives' pattern of discriminative failures would vary according to the nature of their L1 vowel inventory. The /bVt/ words used as stimuli were spoken at a relatively rapid rate by five NE speakers.<sup>4</sup> Fourteen vowel contrasts (e.g., /i/ vs. /ɪ/) were tested by eight "change" triads, in which there was an odd item out, and by eight "catch" triads, which consisted of three different realizations of a single vowel category.

Figure 3 presents data obtained from 10 NE subjects and from 36 of the non-native subjects tested (6 L1 backgrounds x 6 talkers). The non-natives were slightly older than the NE subjects ( $M = 35$  vs. 31 years). They arrived in the United States at an average age of 27 years (range: 14–52 years) and had lived there for an average of 7 years (range: 0.6–38 years). The size of the non-native subjects' L1 vowel inventories appeared to influence whether they did, or did not, differentially classify certain English vowels. Of the six non-native groups shown in the figure, only the native Spanish and Portuguese subjects made significantly more errors on the catch triads (20% and 21%, respectively) than did the NE subjects (2%). For change triads, far more errors were made by the native Portuguese and Spanish subjects (26%, 27%) than by the NE subjects (4%). The German and Dutch subjects also made more errors (10%, 12%) than did the NE subjects. Given that German and Dutch have more vowels than either Portuguese or Spanish, speakers of the former languages were probably more likely to identify two different English vowels in terms of different L1 vowel categories than were speakers of the latter languages.

A failure to discriminate two English vowels was said to have occurred when the six subjects in an L1 group responded significantly less often to the change triads testing the contrast between those English vowels than in responding to triads testing at least three other English vowel contrasts. Just one such discriminative failure was noted

---

<sup>4</sup>Although this reduced the spectral and temporal contrasts seen in carefully articulated English vowels, all vowels were readily identifiable as intended by NE listeners.



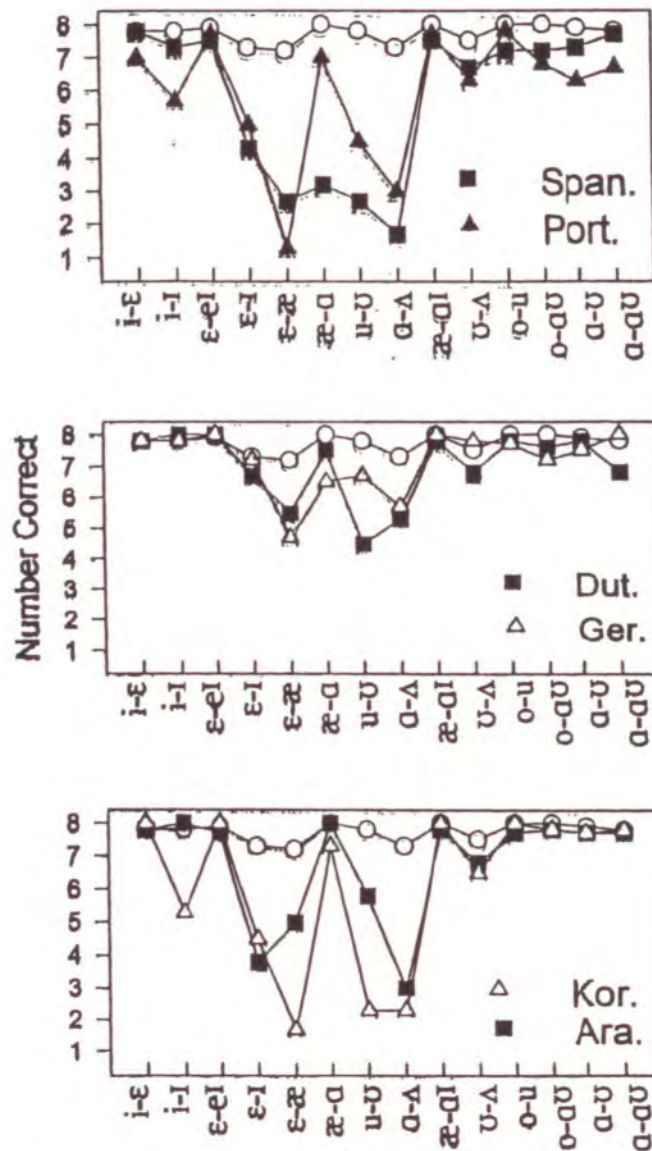


Figure 3. Data from a categorial discrimination study by Flege (1994) in which subjects attempted to choose the odd item out in triads testing 14 English vowel contrasts. The subjects were 10 native speakers of English (open circles), and 6 native speakers each of Spanish and Portuguese (top), German and Dutch (middle), and Korean and Arabic (bottom).

for the German subjects (/ɛ/-/æ/), and two for the Dutch subjects (/u/-/ʊ/, and /ɑ/-/ʌ/). Four such failures were noted for the Portuguese subjects (/ɛ/-/ɪ/, /ɛ/-/æ/, /ɑ/-/ʌ/, and /u/-/ʊ/), five for the Spanish subjects (/ɑ/-/ʌ/, /u/-/ʊ/, /ɛ/-/ɪ/, /ɛ/-/æ/, and /æ/-/ɑ/), and four for the native Arabic subjects (/ɛ/-/ɪ/, /ɛ/-/æ/, /u/-/ʊ/, and /ɑ/-/ʌ/).



The pattern of discriminative failures noted by Flege (1995) is consistent with the hypothesis that L2 learners will have difficulty discriminating a pair of English vowels if both vowels are identified in terms of a *single* L1 category (Best this volume). For example, the midvowels /e o/ of the five-vowel Spanish inventory (/i e a o u/) can be realized as [ɛ] and [ɔ]. The /ɛ/ of the seven-vowel Portuguese inventory (/i e ɛ a o ɔ u/) is often realized as [æ] (Major 1987a). The Portuguese subjects may have managed to discriminate /æ/-/ɑ/ by virtue of identifying English /æ/ tokens in terms of Portuguese /ɛ/, and English /ɑ/ tokens in terms of Portuguese /a/. The native Spanish subjects, on the other hand, may have failed to discriminate /æ/-/ɑ/ because they identified realizations of *both* English vowels in terms of Spanish /a/ (see Flege 1991c). Similarly, the native Spanish subjects may have failed to discriminate /ɛ/-/ɪ/ either because realizations of both English categories were identified as realizations of the Spanish /e/ category (Flege 1991c), or because *neither* vowel could be categorized in terms of an L1 vowel (Best 1994).

An unresolved question at present is how, or if, the kinds of discriminative failures just reported relate to the *production* of English vowels. The model predicts that if a bilingual is unable to discriminate categorially an L2 vowel from neighboring L2 vowels, as well as from neighboring L1 vowels that are distinct phonetically from the L2 vowel, then the L2 vowel will be produced inaccurately. To my knowledge, this has never been tested directly. However, the data now available are consistent with the hypothesis that certain L2 vowel production errors arise from discriminative failures.

Flege (1995) found that native Korean (NK) subjects failed to discriminate English /ɛ/-/æ/ and /i/-/ɪ/.<sup>5</sup> This is understandable in the light of vowel production and perception data obtained in an unpublished study by Flege and Yang, which examined the production of English /i ɪ e æ/ in a /b\_t/ context by NK subjects from Seoul. The NK subjects began learning English in the United States as adults. The 10 subjects in group NK-1 had lived in the United States between 4 and 18 years ( $M = 7.3$  years), the 10 subjects in NK-2 for just 1.5 years ( $M = 0.8$  years). Vowels spoken by NE subjects were identified correctly in nearly every instance by native English-speaking listeners, whereas vowels spoken by the subjects in groups NK-1 and NK-2 were often misidentified. Intended /i/ tokens were heard as /ɪ/ (33% of instances), /ɪ/ as /i/ (23%), /ɛ/ as /æ/ (19%), and /æ/ as /ɛ/ (70% of instances). The rate of misidentifications of vowels spoken by the two Korean groups did not differ significantly.

---

<sup>5</sup>The other discriminative failures noted for NK subjects were /u/-/ʊ/, /ɑ/-/ʌ/, /ɛ/-/ɪ/.



When plotted in an F1 x F2 acoustic space, the Koreans' productions of English /i/ and /ɪ/ showed substantially more overlap than did vowels spoken by the NE subjects. However, the NK subjects produced a larger temporal contrast between /i/ and /ɪ/ (56 ms, or 45%) than did the NE subjects (30 ms, 21%). The NK subjects' /ɛ/ and /æ/ productions also showed much spectral overlap. However, although the NE subjects made /æ/ longer than /ɛ/ (by 56 ms, or 31%), the NK subjects did not produce a temporal difference between these vowels. The NK subjects also identified members of the *beat-bit* (/i/-/ɪ/) and *bet-bat* (/ɛ/-/æ/) continua (see Flege and Bohn 1989). Changes in F1 frequency had a far larger effect on NE subjects' identification of vowels in both vowel continua than did changes in duration. The NK subjects showed a substantially greater effect of duration changes than did NE subjects when identifying members of the *beat-bit* continuum, but they did not respond systematically to changes in vowel duration in the *bet-bat* continuum.

Korean is often analyzed as having a phonemic length contrast between /i/ and /i:/, but this distinction is being merged in modern Seoul Korean (Magen and Blumstein 1993). Other evidence points to the merger of two Korean vowels in the portion of the phonological space occupied by English /ɛ/ and /æ/. We might infer that the NK subjects tested by Flege and Yang (unpublished) incorrectly judged the duration difference between /i/ and /ɪ/ to be more important than the spectral difference between these English vowels by analogy to the Korean /i/-/i:/ distinction produced by an older generation of speakers of Seoul Korean. The NK subjects may have been unable to make any "phonetic sense" of the English /ɛ/-/æ/ contrast, on the other hand, because neither temporal nor spectral cues are used systematically to distinguish vowels in this portion of the Korean vowel space. This could explain the NK subjects' failure to discriminate /ɛ/-/æ/ and /i/-/ɪ/ in the categorial discrimination test (Flege 1995). We might speculate that it is especially difficult for non-natives to discriminate two L2 vowels if phones resembling realizations of the L2 vowels occur in free variation in the L1.

The protocol used by Flege and Yang (unpublished) was administered to native German subjects by Bohn and Flege (1990, 1992) and to native Spanish subjects (Flege, Bohn, and Schmidt, unpublished). The NS subjects in the Flege et al. study had all begun learning English as adults. The subjects in NS-1 had lived in the United States for an average of 9.0 years, those in NS-2 for just 0.5 years on average. Vowels produced by NS and NE subjects were presented for forced-choice identification to native English-speaking listeners. As mentioned earlier, the NE subjects' vowels were identified at near-perfect rates. The correct identification rates obtained for the subjects in NS-1 were lower than those obtained for NE subjects (/i/-57%, /ɪ/-61%, /ɛ/-99%,



*/æ/*-73% correct), and those for NS-2 subjects were lower still (*/i/*-69%, */ɪ/*-51%, */ɛ/*-91%, and */æ/*-70%).

Some NS subjects showed bidirectional errors in producing English */i/* and */ɪ/* (i.e., their */i/* attempts were heard as */ɪ/*, and vice versa). Acoustic analyses revealed that the NS subjects' productions of */i/* and */ɪ/* showed far more spectral overlap in an F1-F2 space than did vowels spoken by the NE subjects. These production data suggested that at least some of the NS subjects failed to discern the phonetic contrast between English */i/* and */ɪ/*, however, the NS subjects tested perceptually by Flege (1995) managed to discriminate English */i/* versus */ɪ/* categorically. Two explanations for this apparent discrepancy are possible. First, different NS subjects participated in the two studies. It is possible that more subjects in the Flege (1995) study than in the Flege et al. (unpublished) study had established a phonetic category for English */ɪ/*. Alternatively, the subjects in the Flege et al. study may have established an */ɪ/* category in which duration figured more prominently than it does in NE speakers' categories. These subjects produced temporal contrasts between English */i/* and */ɪ/* resembling those of NE subjects, although Spanish does not use duration to contrast vowel phonemes. Bohn (this volume) suggested that adult L2 learners may be more attentive to temporal than spectral cues to a novel L2 phonetic contrast.

According to the model, the production of an L2 vowel will change if a category is established for it. The production of an L2 vowel may also change if a new category is *not* established. This is predicted to happen if the phonetic category used to process an L2 vowel that is linked perceptually to L1 vowel (diaphone) changes to reflect a two-language source of input. *Both* developments should lead to more native-like productions of L2 vowels, although the former might result in more rapid and dramatic changes in L2 production than the latter.

A number of studies have shown that late learners' productions of L2 vowels become more native-like as they gain experience in their L2, and that subjects who pronounce their L2 well overall produce L2 vowels better than less proficient speakers from the same L1 background. For example, acoustic measurements by Flege (1987b) revealed that experienced but not inexperienced NE speakers of French produced French */y/* accurately. Wang (1988) found that relatively experienced but not inexperienced Mandarin speakers of English produced a significant spectral difference between two English vowels not distinguished in the L1 (namely */i/* and */ɪ/*). Three studies showed significant improvements in the production of English */æ/* by native speakers of L1s in which */æ/* does not occur phonemically; namely, Portuguese (Major 1987a), German (Bohn and Flege 1992), and Dutch (Flege 1992b).

The literature provides indirect support for H5, which predicts a kind of "merger" of L1 and L2 vowels that have been equated perceptually, and H6, which predicts the deflection of a new L2 vowel category



away from the category for a neighboring L1 vowel(s). In the study by Major (1987a), Portuguese subjects' production of English /ɛ/ deteriorated as their production of /æ/ improved. In a cross-sectional study, Flege (1992b) found that Dutch subjects' production of English /u/, which is more fronted than Dutch /u/, deteriorated as they gained proficiency in English. Bohn and Flege (1992) found that inexperienced German subjects produced English /ɛ/ less accurately than did more experienced subjects, despite the fact that /ɛ/ has a close counterpart in German. H5 and H6 will need to be investigated more closely by examining longitudinal changes in the production of vowels in both the L2 and the L1, as well as concomitant changes in the discriminability of pairs of L2 vowels, and pairs of adjacent L1 and L2 vowels.

Most non-native subjects examined in the vowel production studies just cited had never lived in an English-speaking country, or else had done so for less than 8 years. The results of two recent studies suggest that certain vowel errors persist in the speech of highly experienced L2 learners. Munro (1993) found that many English vowels spoken by native Arabic subjects who had lived in the United States for over 15 years were judged to be foreign-accented. The Arabic subjects greatly exaggerated the duration differences between tense/lax English vowel pairs, as if they were Arabic-like long versus short contrasts. It is, therefore, possible that use of non-English feature specifications was responsible for foreign accent in the Arabic subjects' English vowels, as stated by H6.

Although early learners generally produce L2 vowels more accurately than late learners, their productions of L2 vowels do not always match perfectly those of native speakers. Data presented by Flege (1992a) showed that native Spanish speakers who learned English by the age of 7 years produced English vowels that were identified as often as vowels spoken by NE speakers. The early learners' vowels were not scaled for degree of perceived foreign accent, however, and thus, the study may have overestimated the early learners' success in producing English vowels. Munro, Flege, and MacKay (1995) provided the first comprehensive analysis of the effect of AOL on L2 vowel production. The subjects were 240 NI subjects who began learning English between the ages of 3 and 21 years. These subjects produced English /i i e e æ ɒ ʌ ə o u u/ in a consonant-vowel-consonant (CVC) context. These English vowels vary in acoustic distance from the closest of the seven vowels of Italian (/ i e e a ɔ o u/). English vowels produced by nearly all of the Italian subjects, even those who began learning English in adulthood, were identified correctly.<sup>6</sup> Assuming that no Italian vowel would be

---

<sup>6</sup>The one exception to this was /ʌ/, which was often identified incorrectly when produced by subjects who began learning English after the age of 10 years.



identified as English /ɪ/, /æ/, /ə/ or /ʊ/, we might take this to mean that these English vowels had been "mastered."

A very different pattern of results emerged, however, when the Italian subjects' vowels were rated for degree of foreign accent. Figure 4 presents foreign accent ratings taken from the study by Munro, Flege, and MacKay (1995). Shown here are the number of NI subjects in subgroups of 24 subjects each whose vowels received a rating that fell within a 95% confidence interval of the mean rating obtained for 24 NE subjects. As AOL increased, fewer NI subjects met this criterion, both for English vowels with a counterpart in Italian (e.g., /i/, /u/) and for English vowels without an Italian counterpart (e.g., /æ/, /ə/). Certain English vowels produced by subjects who began learning English as children were found to be foreign-accented. Given that the NI subjects had lived in Canada for over 30 years on average and spoke English more often than Italian, the NI subjects' foreign-accented productions could not be attributed to a lack of native speaker input.

The discrepancy between the intelligibility data and the foreign accent ratings obtained by Munro, Flege, and MacKay (1995) was especially evident for /ə/. Although the NI subjects' /ə/ productions were identified correctly in nearly every instance, few NI subjects with an AOL greater than 10 years managed to produce /ə/ without foreign accent. One possible explanation for the accentedness of /ə/ is that, beyond an AOL of about 10 years, NI subjects failed to attend to the retroflex feature (i.e., energy in the region of F3) which is used to distinguish /ə/ from all other English vowels (Terbeek 1977) but apparently is not used in Italian. Additional work will be needed to test this featural hypothesis and to learn if NI subjects' ability to categorically discriminate English vowels from other English and Italian vowels shows the same effect of AOL as the one seen in Figure 4.

#### PRODUCTION AND PERCEPTION OF INITIAL CONSONANTS

Age of learning exerts an effect on consonant production that is comparable to the one described earlier for vowels. Flege, Munro, and MacKay (1995b) examined the production of consonants by 240 NI subjects differing in AOL. Native English-speaking listeners identified consonants as having been produced "correctly," in a "distorted" fashion, or as having been replaced by some other consonant. Italian does not possess a /ð/ or /θ/. As shown in Figure 5, the NI subjects who began learning English in Canada by about the age of 10 years were judged to have produced the English fricatives correctly as often as did the NE subjects. Beyond that AOL, the number of NI subjects who produced them correctly declined precipitously. When the NI



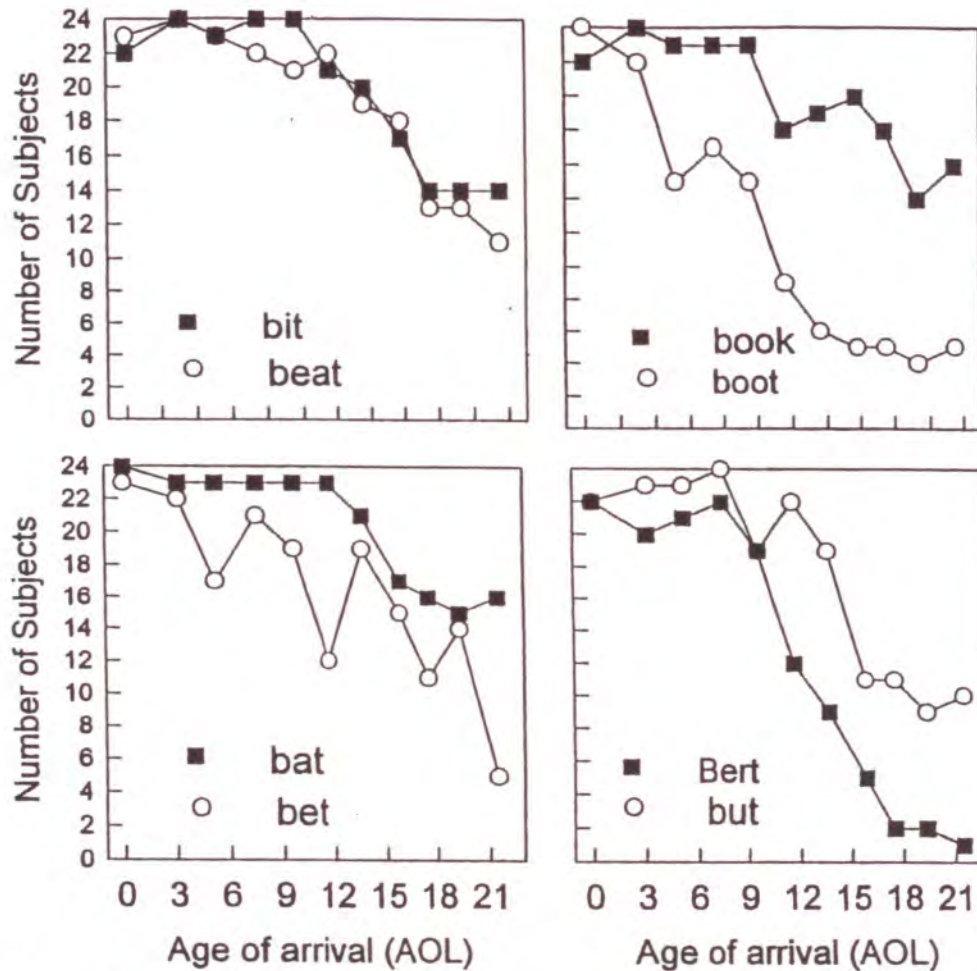


Figure 4. The number of subjects in subgroups of 24 whose production of English vowels received foreign accent ratings that fell within a 95% confidence interval of the mean ratings obtained for 24 native English (NE) subjects. Subgroups of native Italian subjects differed according to their age of learning (AOL) English. The NE subjects are designated as having an AOL of 0 years. The vowels shown here are those in *beat* (/i/), *bit* (/ɪ/), *book* (/u/), *boot* (/u/), *bat* (/æ/), *bet* (/e/), *Bert* (/ə/), and *but* (/ʌ/) (from Munro, Flege, and MacKay 1995).

subjects erred in producing /ð/ and /θ/, it was usually as /d/ and /t/, respectively (something never observed for the NE subjects). The Italian "r" is a trilled /r/ rather than a dorsal (or retroflexed) approximant /ɹ/, as in English. Age of learning exerted a comparable effect on the production of English /ɹ/.

A principal components analysis (Flege, Munro, and MacKay 1995c) provided no evidence that attitudinal or motivational factors influenced the NI subjects' production of the English consonants. Perceptual factors may have been at work, however. Farnetani (1994,



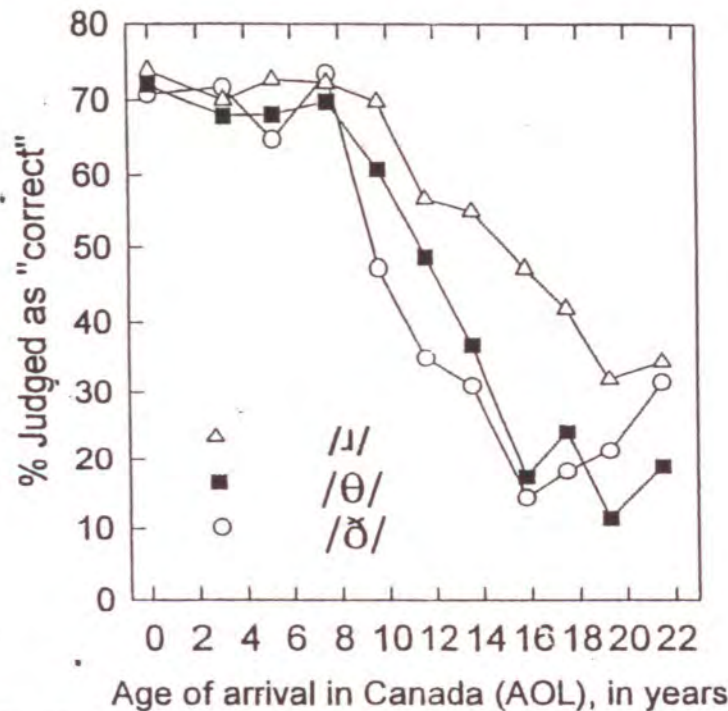


Figure 5. The mean percentage of subjects in subgroups of 24 whose productions of English word-initial consonants were judged by native English-speaking listeners to be "correct." Subgroups of native Italian subjects differed according to their age of learning (AOL) English. The NE subjects are designated as having an AOL of 0 years (data are from Munro, Flege, and MacKay 1995b).

personal communication) observed, for example, that Italians tend to hear word-initial English /ð/ tokens as /d/. It would, therefore, be useful to determine if inaccurate production of English /θ/ and /ð/ arises from the kind of "discriminative failure" discussed earlier for vowels and if so, whether their frequency increases with AOL, as predicted by H4. As for /ɹ/, it would be worthwhile to determine if AOL influenced the features used by NI subjects to distinguish /ɹ/ from other consonants in English and Italian.

Phones such as English [ɹ] and [l] do not occur systematically in Japanese. Native Japanese adults' difficulty producing and perceiving English /ɹ/ and /l/ is well known, but the effect of AOL is only partially understood (but see Yamada this volume). There is evidence, however, that Japanese children who learn English produce /ɹ/ and /l/ more accurately than adult learners (Cochrane 1977; see also Yeni-Komshian and Flege 1993). Yamada and Tohkura (1992b) examined the perception of a synthetic /ɹ/ to /l/ continuum by NJ subjects differing in AOL. Subjects exposed to English in the United States by the



age of 5 years performed in a native-like fashion, as did roughly two-thirds of subjects first exposed to English between the ages of 5 and 10 years. However, fewer than one-fourth of the NJ subjects who were first exposed to English after the age of 10 years identified the synthetic tokens in a native-like fashion (see also Nakauchi 1993). Thus, a question of interest is whether NJ subjects' ability to discern the phonetic difference between English /ɹ/ and /l/, and between these English consonants and the perceptually closest Japanese consonant(s) declines as AOL increases (H4) and whether a failure to discriminate L1 and L2 liquids (and glides) leads to production errors (H7).

Two recent studies provided evidence that, for NJ adults, Japanese /r/ is closer perceptually to English /l/ than /ɹ/ (Sekiyama and Tohkura 1993; Takagi 1993). Thus, the model (specifically: H2, H3, H7) predicts that NJ adults are more likely to establish a new category for /ɹ/ than /l/, and will, therefore, be more likely to produce English /ɹ/ accurately than to do so for English /l/. Results obtained by Flege, Takagi, and Mann (1995b) were consistent with the model's predictions. This study examined the identification of liquids in naturally produced English minimal pairs by three groups of listeners: NE subjects, experienced Japanese subjects (NJ-1) who had lived in the United States for more than 12 years ( $M = 21$ ), and inexperienced subjects (NJ-2) who had lived in the United States for less than 3 years ( $M = 1.6$ ). As expected, the NE subjects identified /l/ and /ɹ/ correctly in nearly every instance. As in previous studies with native Japanese subjects, /ɹ/ tokens were identified correctly at higher rates than /l/ by the subjects in NJ-1 (92% vs. 77% correct) and NJ-2 (76% vs. 63%). Before identifying word-initial liquids, the NJ subjects rated the words in which they occurred for subjective familiarity. Both NJ groups showed effects of subjective lexical familiarity on their identification of /ɹ/ and /l/. For example, they correctly identified the /ɹ/ in *room* (which is paired to a less familiar word, *loom*) more often than the /ɹ/ in *rook* (which is paired to a more familiar word, *look*). The effect of familiarity on the identification of /l/ was equally strong for the two NJ groups. For /ɹ/, on the other hand, it was significantly stronger for the inexperienced subjects in NJ-2 than for the subjects in NJ-1.

When unaffected by lexical familiarity, the experienced subjects in NJ-1 identified /ɹ/ tokens correctly at native-like rates. In minimal pairs consisting of equally familiar words, the correct identification rates for /ɹ/ and /l/ averaged 96% and 81% for the subjects in NJ-1, and 87% and 53% correct for the subjects in NJ-2. Lexical familiarity effects disappeared in a second experiment in which word-initial liquids were edited out of their original word contexts. When the data for one subject who reversed labels were excluded, the correct identification rate of the subjects in NJ-1 averaged 98% correct for /ɹ/, but just 83%



correct for /ɪ/.<sup>7</sup> The experienced NJ subjects' near-perfect identification rate for /ɪ/ is consistent with the hypothesis that they established a phonetic category for English /ɪ/.

In a study by Flege, Takagi, and Mann (1995a), the same Japanese subjects produced minimally paired English words beginning with /ɪ/ and /l/ in three speaking tasks. Native English listeners later attempted to identify the initial consonant in these words as /ɪ/ or /l/ and rated the degree of foreign accent. Liquids spoken by the experienced NJ-1 subjects were identified correctly at the same high rates as liquids spoken by NE subjects; whereas, liquids produced by the inexperienced NJ-2 subjects were often misidentified. Liquids produced by the NJ-1 subjects were identified as intended, and the /ɪ/ and /l/ tokens produced by 10 of the 12 NJ-1 subjects received ratings similar to those for liquids produced by the NE subjects. Inasmuch as the inexperienced NJ-2 subjects identified /ɪ/ tokens at somewhat higher rates than /l/ tokens in the earlier perception experiment, it came as a surprise (H3, H7) that their /l/ and /ɪ/ tokens were judged to have been produced with equal accuracy, at least insofar as the foreign accent ratings were concerned (see also Sheldon and Strange 1982).

A great deal of L2 research has examined the production and perception of English voiceless stops. In many languages (e.g., Dutch, French, Italian, Portuguese, Italian), /p t k/ are realized as unaspirated stops having short-lag VOT values. When adult native speakers of such languages learn English, they tend to produce /p t k/ with longer VOT values in English than in their L1, but with values that are nevertheless too short for English (e.g., Caramazza et al. 1973; Williams 1979; Flege and Port 1981; Major 1987b; Flege and Eefting 1987a). According to the model, category formation for English stops may be blocked by the continued perceptual linkage of L1 and L2 sounds (i.e., by equivalence classification). This limits the accuracy with which English /p t k/ can be produced (H5, H7). Learners nevertheless have access at an auditory level to cross-language phonetic differences (Flege and Hammond 1982; Flege and Munro 1994). When category formation is blocked, the phonetic norms of English may be approximated indirectly through a restructuring of the properties specified in a phonetic category used to process perceptually linked L1 and L2 diaphones. Several studies have provided evidence that, in such instances, the production of L1 stops begins to resemble that of corresponding L2 stops (H5, H7). Flege (1987b) found that experienced NE speakers of French produced English /p t k/ with shorter (French-like) VOT values than did English mono-

<sup>7</sup>The higher rate for /ɪ/ than /l/ probably cannot be attributed to an overall bias in favor of /ɪ/ responses, for NJ subjects identified /l/ correctly more often than /ɪ/ in word-initial clusters (in a case study by Sheldon and Strange 1982).



linguals. Conversely, experienced native French speakers of English produced French /p t k/ with longer (English-like) VOT values than did French monolinguals (see also Major 1992; Flege and Eefting 1987b).

According to the model, individuals who begin learning English as children are more likely than adults to discern the phonetic differences between /p t k/ in L1 and L2 and establish phonetic categories for English /p t k/. This leads to the prediction that early learners will produce these stops accurately more often than late learners. This prediction was confirmed by Flege (1991b) in a study that examined stops produced by native Spanish (NS) speakers who learned English as adults (late learners) or as children (early learners). The NS late learners produced English /p t k/ with the expected "compromise" VOT values; whereas, the early learners produced English stops with the same VOT values as did NE subjects. This was interpreted to mean that the early learners, unlike the late learners, had established phonetic categories for English /p t k/.

This conclusion is consistent with the results of an imitation study by Flege and Eefting (1988). Members of a consonant-vowel (CV) continuum in which VOT of the initial consonant varied were imitated by English monolinguals, Spanish monolinguals, and NS early learners. The monolinguals tended to produce stops with VOT values falling within the two modal VOT ranges used for stops in their L1 (Spanish: lead, short-lag; English: short-lag, long-lag). The NS early learners, on the other hand, produced stops with values in *all three* modal VOT ranges. Additional analyses revealed that the subjects covertly classified stops before imitating them. Thus, the results suggested that the early learners processed stops in the VOT continuum in terms of three phonetic categories: a prevoiced (Spanish or English) /d/ category, a short-lag (Spanish) /t<sup>-</sup>/ category, and a long-lag (English) /t<sup>h</sup>/ category.

Acoustic analysis by Flege, Munro, and MacKay (1995b) identified the AOL at which divergences from the phonetic norms of English first become apparent in NI subjects' productions of English stop consonants. Voice-onset time was measured in word-initial English stops produced by NI subjects who began learning English between the ages of 3 and 21 years. Their production of VOT in English /p t k/ varied inversely with AOL. A subgroup of 24 NI subjects who began learning English at an average age of 21 years produced English /p/, /t/ and /k/ with VOT values that were almost exactly intermediate to values measured for 24 NE subjects and values reported for Italian /p t k/. Native Italian subjects who arrived in Canada at earlier ages produced English stops with increasingly longer, and thus, more accurate, VOT values. The first NI subgroups to have produced /t/ and /p/ with significantly shorter VOT values than did the NE subjects consisted of



individuals with average AOLs of 11 and 17 years, respectively.<sup>8</sup>

It should be emphasized that the kind of age limit just mentioned does not necessarily apply to all individuals within a certain AOL range. For example, Flege and Schmidt (1995) recently used the speaking rate paradigm of Miller and Volaitis (1989) to test for category formation. Native English and NS late learners rated the members of a labial stop continuum for goodness. Voice-onset time ranged from short-lag values to long-lag values exceeding those typical for English /p/. The same VOT values occurred in short-duration syllables, which simulated a fast speaking rate, and in longer-duration syllables, which simulated a slower rate of speech. As expected, the NE subjects' goodness ratings increased as stimulus VOT values increased, then decreased as VOT increased beyond values typical for English. Also as expected, the NE subjects showed an effect of speaking rate on their goodness ratings. They gave higher ratings to stops with long-lag VOT values in the long-duration stimuli than to stops with the same VOT values in short-duration syllables. This matched how VOT varies according to speaking rate in the production of English.

According to Miller and Volaitis (1989), the speaking rate effects shown by NE subjects demonstrates "internal phonetic category structure." Flege and Schmidt (1995) reasoned that unless NS subjects had a category for long-lag voiceless stops, they would not show an English-like speaking rate effect on their goodness judgments of long-lag stops. When the NS subjects were subdivided into two groups based on length of residence in the United States, neither group showed a significant speaking rate effect. However, when subdivided according to how well they pronounced English sentences, proficient but not non-proficient NS subjects showed a significant speaking rate effect. This suggested that some of the late learners may have established a phonetic category for the long-lag /p/ of English. Schmidt and Flege (1995) examined the NS subjects' production of English /p/ at relatively slow and fast speaking rates. Roughly one-half of the proficient NS subjects produced /p/ with VOT values that were comparable to those observed for NE subjects. When speaking, these subjects adjusted VOT as a function of speaking rate (i.e., produced /p/ with longer VOT values at a slow than fast rate).

#### PRODUCTION AND PERCEPTION OF WORD-FINAL CONSONANTS

According to H1, position-sensitive allophones in the L2 and L1 are related perceptually to one another. This leads to the prediction that

<sup>8</sup>The VOT for Italian /k/ falls within the long-lag range. None of the 10 NI subgroups differed significantly from the NE subjects for /k/, apparently because the VOT difference between English and Italian /k/ is too small to permit the measurement of cross-language interference.



speakers of an L1 without word-final stops will not relate English word-final stops perceptually to word-medial or word-initial stops in their L1. If so, then we might expect them to eventually produce word-final stops in English accurately. This is because, if H1 is correct, L1 phonetic structures should not interfere with the establishment of new phonetic categories.

Contrary to H1 (and H7), inexperienced late learners of English have been shown to delete final stops, to devoice /b d g/, or to add an epenthetic vowel to CVC words (e.g., Eckman 1981; Flege 1988a, 1989; Flege and Davidian 1984; Weinberger 1987). Flege, Munro, and Skelton (1992) examined the production of final stops in English words such as *beat* and *bead* by groups of native Mandarin (NM) and native Spanish (NS) late learners differing in length of residence in the United States (NM-1 = 6 years, NM-2 = 1 year, NS-1 = 9 years, NS-2 = 2 years). Neither Mandarin nor Spanish permits word-final stops. Word-final stops produced by the NE subjects were almost always identified correctly by native English-speaking listeners. Stops produced by the non-native subjects were identified far less often (NM-1 = 62%, NM-2 = 65%, NS-1 = 73%, NS-2 = 71%). Surprisingly, the experienced non-native subjects' stops were identified correctly no more often than were stops spoken by the relatively *inexperienced* non-native speakers of English.

Flege, Munro, and Skelton (1992) carried out acoustic analyses to determine why the non-natives' production of both /t/ and /d/ were often misidentified. As with the NE speakers, the non-natives sustained closure significantly longer in /d/ than /t/, made vowels significantly longer before /d/ than /t/, and produced significantly lower F1 frequencies at the end of transitions leading into /d/ than /t/. However, the non-natives' acoustic distinctions were smaller than those of the NE speakers. The subjects in NM-1, NM-2, and NS-1 produced significantly smaller closure voicing differences between /t/ and /d/ than did the NE subjects, but the closure voicing difference produced by the experienced NS subjects did *not* differ significantly from the NE subjects'. This suggests that *closure voicing* may be learned more readily than other cues to the /t-/d/ distinction (see Kluender et al. 1988), but that a native-like production of closure voicing does not ensure correct identification by NE listeners.

Flege et al. (1995b) examined the production of word-final English stops by 240 NI subjects who had spoken English for over 30 years on average. The results of this study ruled out the possibility that the native Mandarin and Spanish subjects' stop production errors (Flege Munro, and Skelton 1992) were caused by a lack of sufficient native-speaker input. The NI subjects, who began learning English between the ages of 3 and 21 years, produced English /p t k/ accurately. Those who



began learning English by the age of 15 years also produced /b d g/ accurately. However, roughly 40% of NI subjects with an AOL of 15 to 21 years devoiced /b d g/.

Table III presents acoustic measurements of *tack* and *tag* tokens that were spoken by 24 NE subjects and 48 NI subjects with AOLs ranging from 12 to 21 years. Productions of /k/ and /g/ by the 24 subjects in NI-1 were identified correctly in every instance. Stops produced by the 24 subjects in NI-2 (matched in AOL to the NI-1 subjects) were identified correctly in only 64% of instances. The acoustic measurements make it apparent why this was so. The NI-2 subjects (unlike the NI-1 subjects) produced a significantly smaller vowel duration difference preceding /g/ versus /k/ than did the NE subjects, and a significantly smaller closure voicing difference ( $p < .01$ ). However, they produced a somewhat larger stop closure difference between /k/ and /g/ than did the NE subjects. Interestingly, the NI-1 subjects produced a significantly larger closure voicing contrast between /k/ and /g/ than did the NE subjects ( $p < .01$ ).

It is not clear why the NI-1 but not NI-2 subjects managed to produce a perceptually effective contrast between /k/ and /g/, which the model predicts for *all* highly experienced Italian learners of English. If certain NI learners of English treat word-final English stops as if they were word-*medial* Italian stops (contrary to H1), we would expect NI speakers to produce larger closure voicing differences between English /k/ and /g/ than NE speakers and to produce a larger stop closure duration difference. We would *not* expect an English-like vowel duration difference, however (Mack 1982; Vaggel et al. 1978; Magno Caldognetto et al. 1979; see also Crowther 1993). If, on the other hand, NI subjects treated the word-final stops in English CVC words as “new,” we might have expected a thoroughly English-like /k/ versus /g/ contrast, which was not observed for either the NI-1 or NI-2 subjects. Perhaps some NI subjects tend to rely on acoustic cues exploited in Italian more than others, and the use of L1 cues does not apply across the board to all relevant phonetic dimensions. This could explain the overuse of closure voicing and closure duration cues by certain subjects, and the underuse of vowel duration cues by other subjects (see also Flege 1989; Crowther 1993; Flege and Wang 1990).

#### THEORETICAL ISSUES

According to the contrastive analysis approach, the presence of phonemes in an L2 that do not occur in the L1 necessarily represents a learning problem (e.g., Lado 1957; Moulton 1962; Koutsoudas and Koutsoudas 1983). The learner's response may be to use the closest L1 phoneme as a “substitute” for the unfamiliar L2 phoneme (Lehiste 1988).



Table III. Temporal acoustic measurements of *tack* and *tag* as spoken by native English (NE) speakers and two groups of native Italian (NI) speakers who began learning English after the age of 12 years<sup>a</sup>

Variable	Subject group		
	NE	NI-1	NI-2
Number of subjects	24	24	24
Age at arrival in Canada	0(0)	18(2)	18(3)
% correct identification of /k g/	97(7)	100(0)	64(10)
Vowel duration— <i>tag</i>	246(48)	226(46)	210(53)
Vowel duration— <i>tack</i>	191(33)	170(38)	196(53)
Difference:	55	56	14
Stop closure duration— <i>tack</i>	102(28)	129(29)	149(44)
Stop closure duration— <i>tag</i>	75(24)	100(19)	113(33)
Difference:	27	29	36
Closure voicing duration— <i>tag</i>	51(22)	77(26)	41(28)
Closure voicing duration— <i>tack</i>	8(13)	13(13)	18(19)
Difference:	43	64	23

<sup>a</sup>The Italian subjects in NI-1 produced final stops that were always identified correctly. The stops produced by subjects in NI-2, who were matched according to AOL to the NI-1 subjects, were often misidentified by NE listeners. Values for the acoustic variables are in ms; standard deviations are in parentheses (data are from Flege, Munro, and MacKay 1995b).

An L1-for-L2 substitution implies that an L2 phoneme (or position-sensitive allophone) has been perceptually linked to a particular L1 phoneme (allophone) on some basis. It implies, further, that the learner has either failed to discern the phonetic difference between the L1 and L2 phonemes or allophones, is unable motorically to render a correctly perceived difference, or both.

First-language-for-second-language substitutions raise a number of theoretically important questions. According to some (e.g., Rochet this volume), most or all L2 sounds will be identified with an L1 sound. However, some analysts draw a distinction between L2 sounds that are relatively similar to sounds in the L1, as opposed to L2 sounds that are more dissimilar or even "new" (see Mueller and Niedzielski 1963; Pimsleur 1963; Brière 1966; Henning 1966; Delattre 1964, 1969; Flege 1981; Wode 1978). According to the SLM, the full range of L2 sounds may at first be identified in terms of a positionally defined allophone of the L1 but, as L2 learners gain experience in the L2, they may gradually discern the phonetic difference between certain L2 sounds and the closest L1 sound(s). When this happens, a phonetic category representation may be established for the new L2 sound that is independent of representations established previously for L1 sounds.



According to the model, the phonetic category established in childhood for an L1 sound may evolve gradually if it is linked perceptually to an L2 sound. For example, a French speaker's representation for word-initial /t/ may evolve to specify somewhat longer VOT values if English /t/ tokens are persistently identified as realizations of the French /t/ category. This might explain why highly experienced French learners of English tend to produce English /t/ with "compromise" VOT values and why their productions of /t/ in French may take on English phonetic characteristics (see above). According to the SLM, these developments are influenced by two important variables: AOL and perceived cross-language phonetic distance. The greater the perceived distance of an L2 sound from the closest L1 sound, the more likely it is that a separate category will be established for the L2 sound. Moreover, the earlier L2 learning commences, the smaller the perceived phonetic distance needed to trigger the process of category formation.

An obstacle to testing hypotheses such as these is the lack of an objective means for gauging degree of perceived cross-language phonetic distance. It is uncertain, also, what metric bilinguals use in doing so. Cross-language phonetic distance might be assessed in terms of the sensory (auditory, visual) properties associated with L1 and L2 sounds. However, Ladefoged (1990) concluded that although trained phoneticians may be able to agree as to whether or not sounds in two languages are "the same," their judgments may have no "principled basis," and their thresholds are "unknown." Another possibility is that cross-language distance is gauged in terms of differences in perceived gestures (Browman and Goldstein 1990). Best (this volume) suggested that L1 versus L2 distances can be gauged by the "spatial proximity of constriction locations and active articulators" and by "similarities in constriction degree and gestural phasing." Although appealing, this metric may also be difficult to apply. James (1984) suggested that *three different metrics* (gestural, acoustic phonetic, abstract phonological) may be applied depending on syllable position. His observation was motivated, in part, by the observation that native Dutch learners of English use Dutch /d/ for English /ð/ in word-initial position; whereas, they replace /ð/ with an alveolar fricative in word-final position.

The research reviewed in this chapter demonstrated that AOL exerts a powerful influence on the production of L2 sounds. (Similar tests of the effect of AOL on perception have yet to be undertaken, but see Oyama, 1978 and Yamada this volume.) Some studies provided evidence that L2 sounds not found in the L1 inventory may be produced more accurately than are L2 sounds with a counterpart in the L1 inventory. If we assume that the noninventory sounds are treated



as “new”; whereas, the within-inventory sounds are treated as “similar,” this might be taken as support for a new versus similar distinction (Flege 1988b). However, research reviewed in the chapter provided evidence that noninventory sounds may in some instances be produced inaccurately, even by highly experienced L2 learners.

It may be that, in certain instances, the positionally defined allophone is too coarse a unit of analysis to provide accurate predictions concerning L2 sound production. Trubetzkoy (1939) compared the mature L1 phonological system to a sieve. According to this conceptualization, the L1 phonology passes only that information about L2 sounds that is relevant to L1 phonemic contrasts (see also Koutsoudas and Koutsoudas 1983). Weinreich (1953, 1957) noted that feature mismatches between the L1 and L2 sounds may lead to overdifferentiation, reinterpretation, or underdifferentiation of feature contrasts (see also Lehiste 1988). The first two kinds of errors may not lead to detectable production errors, but the last kind of error may result in a sound substitution. For example, if NS speakers of English treat the [-continuant] feature of English word-initial /d/ tokens as a redundant feature (as it is in Spanish) rather than as a distinctive feature (as it is in English), they would not be expected to render word-initial /d/ tokens as stops, but as fricatives (Weinreich 1957).

Linguistically trained analysts tend to focus on “distinctive” features. Mismatches between L1 and L2 sounds may also be described in terms of differences in the *acoustic cues* used to contrast phones. So, for example, we might interpret the replacement of English word-initial /ð/ by /d/ as arising from lack of attention to, or inappropriate weighting of, the amplitude, frequency, and temporal differences between word-initial tokens of English /ð/ and /d/ (Morosan and Jamieson 1989; Crowther and Mann 1992; Crowther 1993). Some part or all of what is referred to as language’s “phonological filter” may consist of learned aspects of perceptual processing that are an outgrowth of “interpretative schemes” established in early childhood for word recognition (Jusczyk 1992). Such schemes are thought to focus attention automatically on auditory patterns important to meaning distinctions in the L1 and may become more abstract as children develop larger lexicons. Jusczyk (1985) suggested, for example, that allophonic variants may not be grouped into a single phonemic category by children until the age of 5 or 6 years, when they begin learning to read.

Alternatively, the feature differences between languages could be stated in terms of the “contrastive gestures” used to signal meaning differences (Browman and Goldstein 1990). According to Best’s direct realist account (this volume), perceivers have direct access through their senses to the elemental gestures used to form speech sounds.



During speech development, children learn to pick up information in speech more quickly and efficiently by virtue of learning what are the "critical" aspects of gestures used in the L1. This leads to the establishment of "lower-order invariants," which are generally relational in nature and which permit perceptual constancy in the face of acoustic variation. As the phonological system matures, the lower-order invariants may give way to higher-order, language-specific invariants that reduce the amount of lower-order phonetic detail that is detected (Best 1994).

One possible explanation for the ubiquitous age effects on L2 production discussed earlier in this chapter is that early learners of an L2 are more likely to "pick up" detailed information concerning the specification of L2 sounds than are individuals who begin learning an L2 later in life. Early learners may perceive L2 sounds in terms of "lower-order" rather than "higher-order" invariants. Changes in the level of auditory-acoustic analysis might also be important. Koutsoudas and Koutsoudas (1983) suggested that interlingual identification occurs at a phonemic level and is based on the articulatory phonetic "common denominator" of all allophones of a phoneme. Perhaps this statement is more true for individuals who begin learning the L2 after about the age of 10 years than it is for those who begin learning their L2 earlier in life.

Regardless of how the feature differences between L1 and L2 sounds are stated, it seems that non-natives often do not perceive L2 sounds in exactly the same way monolingual native speakers of the target L2 do. As noted earlier, they may fail to discern the phonetic differences between contrastive phones in the L2, or between L1 and L2 phones. Discriminative failures in L2 acquisition usually do not have an auditory basis. For example, Miyawaki et al. (1975) found that Japanese listeners who were unable to discriminate / $\mu$ / and / $\mu$ / syllables were, nonetheless, able to discriminate isolated third formants taken from those syllables. Werker and Tees (1984) found that NE subjects could discriminate short portions extracted from a foreign language contrast (Hindi dental vs. retroflex stops), but not the original CV syllables from which those portions had been edited. Mann (1986) noted an effect of preceding context (/ɹ/ versus /l/) on /d/-/g/ phoneme boundaries for native Japanese subjects who could not differentially identify /ɹ/ versus /l/ that was much the same as the effect noted for NE subjects. Finally, Best et al. (1988) observed surprisingly good discrimination of Zulu clicks by NE adults, apparently because the clicks were not identified as speech sounds.

What kind of features are used by L2 learners as they begin to analyze the phonic elements of their L2? What kind do they use once they have gained a thorough familiarity with the L2 sound system?



Answers to these questions are not yet possible, but several points can be made with some certainty. First, the features used to distinguish L1 sounds can probably not be freely recombined to produce new L2 sounds. Flege and Port (1981) examined native Arabic (NA) subjects who began learning English in adulthood, and had lived for several years in the United States. Arabic has the stop consonants /b t d k/, but no /p/ or /g/. Many of the NA subjects' English /p/ productions were heard as /b/ because they were produced with closure voicing (as are the /b/ and /d/ of Arabic). However, the NA subjects produced much the same temporal contrast between English /p/ and /b/ as they produced between /t/ and /d/. This suggested that they had recognized the phonological nature of the English /p/ versus /b/ contrast, but did not transfer the voiceless feature of their /t/ and /k/ to /p/.

Some production difficulties may arise because features used in the L2 are not used in the L1. In a recent multidimensional scaling (MDS) analysis, Fox, Flege, and Munro (1995) found that NE subjects used three dimensions in perceiving naturally produced vowels; whereas, NS subjects used just two, probably because Spanish has far fewer vowels than English. As mentioned earlier, Munro et al. (1995) found that NI subjects who began learning English after about the age of 10 years produced English /ɚ/ inaccurately. These NI subjects may have been unable to acquire sensitivity to the retroflex feature that NE listeners use to distinguish English /ɚ/ from all other English vowels (Terbeek 1977). Beyond a certain AOL, L2 learners may rely on features used in the L1 to interpret and represent sounds encountered in the L2, even those L2 sounds that are treated as "new" (i.e., as falling outside the L1 phonetic inventory). For example, neither Swedish nor Finnish has an /s/-/z/ contrast in final position. Flege and Hillenbrand (1986) found that Swedes and Finns used different acoustic cues (vowel duration, but not fricative duration) than did NE subjects (who used both) to identify final consonants in a *peace-peas* continuum. Work by Gottfried and Beddor (1988) and Munro (1992) suggested that the overall use made of duration in the L1 may influence the extent to which L2 learners exploit temporal cues to a novel L2 contrast.

The phenomenon of "differential substitution" shows that we need recourse to more than just a simple listing of features used in the L1 and L2 to explain certain L2 production errors. For example, although both Russian and Japanese have /s/ and /t/, Russians tend to substitute /t/ for English /θ/, whereas Japanese learners use /s/. Weinberger (1990) presented evidence that feature counting, feature weighting, and universal marking conventions cannot account for differential substitution patterns. Radical underspecification theory (which treats features, not segments, as primitives) on the other hand, may do so



given certain assumptions about the underlying feature matrix, redundancy rules, and feature pruning.

Certain features may enjoy an advantage over others because of the nature of their acoustic (or gestural) specification, or their reliability of occurrence. Polka (1991) found that Hindi dental versus retroflex place distinctions may be discriminated more accurately by native speakers of English in voiceless unaspirated stops than in prevoiced stops. This appeared to be the case because the place distinction is supported by greater formant transition differences in the former than the latter stops. Flege and Hillenbrand (1987) obtained data suggesting that native French speakers make greater use of release burst information in judging the voicing feature in word-final stops than do NE speakers, apparently because final stops are released consistently by speakers of French but not English.

Finally, features may be evaluated differently as a function of position in the syllable. Browman and Goldstein (1990) suggested that, in English, vowel gestures are coordinated (phased) with respect to the gestures used for a preceding consonant, and final consonant gestures are "phased with respect to preceding vocalic gestures." Samuel, Kat, and Tartter (1984) provided evidence that initial, medial, and final consonants are processed differently by NE listeners. Languages differ considerably in the range of permitted syllables, which leads to differences in syllable-processing strategies (Cutler et al. 1983, 1986; Flege 1989; Flege and Wang 1990). Not unexpectedly, then, studies have shown that the perceptual difficulty of a novel L2 phonemic contrast may vary according to syllable position (e.g., Sheldon and Strange 1982; James 1988; Weiden 1990; Pisoni and Lively this volume). For example, Major (1986) found that NE subjects' accuracy in producing Spanish /r/ varied considerably as a function of position in the syllable. Morosan and Jamieson (1989) found that effects of perceptual training on word-initial allophones did not transfer completely to medial or final positions, which suggested to them that listeners may learn "syllabically."

Given the ubiquity of foreign accents in L2 production, an important general question for future research is how, or if, the perception of L2 sounds changes as AOL increases. Hammarberg (1988, 1990) suggested that phonetic-level comparisons of L1 and L2 sounds may be more prevalent in early than later stages of L2 learning. Another important question is how, and to what extent, an individual's perception changes during L2 learning. Flege (1991c) found that as NS subjects gained familiarity with English, they became less likely to identify English /ɪ/ tokens as exemplars of their Spanish /i/ category. Thus, L2 learners may discover that certain sounds in an L2 are not "the same" phonetically as sounds they already know from the L1. The



observation that “schooled” native French (NF) speakers of French substitute /s/ for English /θ/, whereas “unschooled” NF subjects substitute /t/ suggests that the metric used to gauge cross-language distance might also change as a function of L2 experience or proficiency (Berger 1951; see also Wenk 1979). If the perception of L2 sounds does change, we need to know *how* it changes, and what impact perceptual changes have on L2 speech production. These, and many other questions, must be resolved in order to understand fully the contribution of perception to foreign accentedness.

#### REFERENCES

- Berger, M. 1951. The American English pronunciation of Russian immigrants. Ph.D. dissertation, Columbia University, New York.
- Best, C. 1993. Emergence of language-specific constraints in perception of non-native speech: A window on early phonological development. In *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*, ed. B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. MacNeilage, and J. Morton. Dordrecht, the Netherlands: Kluwer Academic Publishers.
- Best, C., McRoberts, G., and Sithole, N. 1988. The phonological basis of perceptual loss for non-native contrasts: Maintenance of discrimination among Zulu clicks by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance* 14:345–60.
- Bohn, O.-S., and Flege, J. E. 1990. Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics* 11:303–28.
- Bohn, O.-S., and Flege, J. E. 1992. The production of new and similar vowels by adult German learners of English. *Studies in Second Language Acquisition* 14:131–58.
- Brennan, E., Ryan, E., and Dawson, W. 1975. Scaling of apparent accentedness by magnitude estimation and sensory modality matching. *Journal of Psycholinguistic Research* 4:27–36.
- Brière, E. 1966. An investigation of phonological interference. *Language* 42:769–96.
- Browman, C., and Goldstein, L. 1990. Gestural specification using dynamically-defined articulatory structures. *Journal of Phonetics* 18:299–320.
- Butcher, A. 1976. The influence of the native language on the perception of vowel quality. M.Phil. thesis, Kiel University.
- Caramazza, A., Yeni-Komshian, G., Zurif, E., and Carbone, E. 1973. The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *Journal of the Acoustical Society of America* 54:421–28.
- Cochrane, R. 1977. The acquisition of /r/ and /l/ by Japanese children and adults Learning English as a second language. Unpublished Ph.D. dissertation, University of Connecticut.
- Crowther, C. 1993. The Operation of Vocalic Duration and F1 Offset Frequency as Voicing and Vowel Cues. Ph.D. dissertation, University of California–Irvine.
- Crowther, C., and Mann, V. 1992. Native language factors affecting use of vocalic cues to final consonant voicing in English. *Journal of the Acoustical Society of America* 92:711–22.



- Crystal, T., and House, A. 1988. The duration of American-English vowels: An overview. *Journal of Phonetics* 16:263–84.
- Cutler, A., Mehler, J., Norris, D., and Segui, J. 1983. A language specific comprehension strategy. *Nature* 304:159–60.
- Cutler, A., Mehler, J., Norris, D., and Segui, J. 1986. The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language* 25:385–400.
- Delattre, P. 1964. Comparing the vocalic features of English, German, Spanish, and French. *International Review of Applied Linguistics* 2:71–97
- Delattre, P. 1969. An acoustic and articulatory study of vowel reduction in four languages. *International Review of Applied Linguistics* 7:295–323.
- Eckman, F. 1981. On the naturalness of interlanguage phonological rules. *Language Learning* 31:195–216.
- Eilers, R., Wilson, W. and Moore, J. 1977. Developmental changes in speech discrimination in infants. *Journal of Speech and Hearing Research* 20:766–80.
- Fayer, J., and Krasinski, E. 1987. Native and non-native judgments of intelligibility and irritation. *Language Learning* 37:313–26.
- Ferrero, F., Mchelotti, E., Pelamatti, G., and Vaggas, K. 1978. Some acoustic characteristics of Italian vowels. *Journal of Italian Linguistics* 3:87–95.
- Ferrero, F., Michelotti, E., Pelamatti, G., Umilita, C., and Vaggas, K. 1979. Auditory and phonetic processing of Italian voiceless fricatives shortened in duration. *Italian Journal of Psychology* 6:237–49.
- Flege, J. E. 1981. The phonological basis of foreign accent. *TESOL Quarterly* 15:443–55.
- Flege, J. E. 1987a. A critical period for learning to pronounce foreign languages? *Applied Linguistics* 8:162–77.
- Flege, J. E. 1987b. The production of “new” and “similar” phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics* 15:47–65.
- Flege, J. E. 1988a. The development of skill in producing word-final English stops: Kinematic parameters. *Journal of the Acoustical Society of America* 84:1639–52.
- Flege, J. E. 1988b. The production and perception of foreign language speech sounds. In *Human Communication and Its Disorders, A Review—1988*, ed. H. Winitz. Norwood, NJ: Ablex.
- Flege, J. E. 1989. Chinese subjects' perception of the word-final English /t-/d/ contrast: Performance before and after training. *Journal of the Acoustical Society of America* 86:1684–97.
- Flege, J. E. 1991a. Perception and production: The relevance of phonetic input to L2 phonological learning. In *Crosscurrents in Second Language Acquisition and Linguistic Theory*, ed. T. Heubner, and C. Ferguson. Philadelphia: John Benjamins.
- Flege, J. E. 1991b. Age of learning affects the authenticity of voice onset time (VOT) in stop consonants produced in a second language. *Journal of the Acoustical Society of America* 89:395–411.
- Flege, J. E. 1991c. Orthographic evidence for the perceptual identification of vowels in Spanish and English. *Quarterly Journal of Experimental Psychology* 43:701–31.
- Flege, J. E. 1992a. Speech learning in a second language. In *Phonological Development: Models, Research, and Application*, ed. C. Ferguson, L. Menn, and C. Stoel-Gammon. Timonium, MD: York Press.
- Flege, J. E. 1992b. The intelligibility of English vowels spoken by British and



- Dutch talkers. In *Intelligibility in Speech Disorders*, ed. R. Kent. Amsterdam: John Benjamins.
- Flege, J. E. 1993. Production and perception of a novel, second-language phonetic contrast. *Journal of the Acoustical Society of America* 93:1589–1608.
- Flege, J. E. 1995. Oddity discrimination of English vowels by non-natives. Ms. in preparation. Department of Biocommunication, University of Alabama at Birmingham.
- Flege, J. E., and Bohn, O.-S. 1989. The perception of English vowels by native Spanish speakers. *Journal of the Acoustical Society of America* 85:S85(A).
- Flege, J. E., and Davidian, D. 1984. Transfer and developmental processes in adult foreign language speech production. *Applied Psycholinguistics* 5:323–47.
- Flege, J. E., and Eefting, W. 1986. Linguistic and developmental effects on the production and perception of stop consonants. *Phonetica* 43:155–71.
- Flege, J. E., and Eefting, W. 1987a. The production and perception of English stops by Spanish speakers of English. *Journal of Phonetics* 15:67–83.
- Flege, J. E., and Eefting, W. 1987b. Cross-language switching in stop consonant production and perception by Dutch speakers of English. *Speech Communication* 6:185–202.
- Flege, J. E., and Eefting, W. 1988. Imitation of a VOT continuum by native speakers of English and Spanish: Evidence for phonetic category formation. *Journal of the Acoustical Society of America* 83:729–40.
- Flege, J. E., and Hammond, R. 1982. Mimicry of non-distinctive phonetic differences between language varieties. *Studies in Second Language Acquisition* 5:1–18.
- Flege, J. E., and Hillenbrand, J. 1986. Differential use of temporal cues to the /s/-/z/ contrast by native and non-native speakers of English. *Journal of the Acoustical Society of America* 79:508–17.
- Flege, J. E., and Hillenbrand, J. 1987. Differential use of closure voicing and release burst as cue to stop voicing by native speakers of French and English. *Journal of Phonetics* 15:203–8.
- Flege, J. E., and Munro, M. 1994. The word unit in L2 speech production and perception. *Studies in Second Language Acquisition* 16:381–411.
- Flege, J. E., Munro, M., and Fox, R. 1994. Auditory and categorical effects on cross-language vowel perception. *Journal of the Acoustical Society of America* 95:3623–41.
- Flege, J. E., Munro, M., and MacKay, I. 1995a. Factors affecting degree of perceived foreign accent in a second language. *Journal of the Acoustical Society of America* 97:3125–34.
- Flege, J. E., Munro, M., and MacKay, I. 1995b. Effects of age of second-language learning on the production of English consonants. *Speech Communication* 16:1–26.
- Flege, J. E., Munro, M., and MacKay, I. 1995c. Factors affecting the production of word-initial consonants in a second language. In *Variation Linguistics and SLA*, ed. D. Preston. Amsterdam: John Benjamins. In press.
- Flege, J. E., Munro, M., and Skelton, L. 1992. Production of the word-final English /t/-/d/ contrast by native speakers of English, Mandarin, and Spanish. *Journal of the Acoustical Society of America* 92:128–43.
- Flege, J. E., and Port, R. 1981. Cross-language phonetic interference: Arabic to English. *Language and Speech* 24:125–46.
- Flege, J. E., and Schmidt, A. 1995. Native speakers of Spanish show rate-dependent processing of English stop consonants. *Phonetica* 52:90–111.



- Flege, J. E., and Takagi, N. 1994. Speaking task effects on Japanese speakers' production of English /ɹ/ and /l/. *Language and Speech*. Under review.
- Flege, J. E., Takagi, N., and Mann, V. 1995a. Japanese adults can learn to produce /ɹ/ and /l/ accurately. *Language and Speech* 38, Part I. In press.
- Flege, J. E., Takagi, N., and Mann, V. 1995b. Experienced Japanese speakers of English can identify English /ɹ/ accurately, but not /l/. *Journal of the Acoustical Society of America*. Under review.
- Flege, J. E., and Wang, C. 1990. Native-language phonotactic constraints affect how well Chinese subjects perceive the word-final English /t-/d/ contrast. *Journal of Phonetics* 17:299-315.
- Fox, R. A., Flege, J. E., and Munro, M. J. 1994. A multidimensional scaling analysis of Spanish and English vowels. *Journal of the Acoustical Society of America* 97:2540-51.
- Giles, H. 1970. Evaluational reactions to accents. *Educational Review* 22:211-27.
- Goldstein, L., and Browman, C. 1986. Representation of voicing contrasts using articulatory gestures. *Journal of Phonetics* 14:339-42.
- Gottfried, T. 1983. Effects of consonant context on the perception of French vowels. *Journal of Phonetics* 12:91-114.
- Gottfried, T., and Beddor, P. 1988. Perception of temporal and spectral information in French vowels. *Language and Speech* 31:57-75.
- Greene, B., Pisoni, D., and Gradman, H. 1985. Perception of synthetic speech by non-native speakers of English. *Research on Speech Perception* 11:419-28 (Department of Psychology, Indiana University).
- Grosjean, F. 1989. Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language* 36:3-15.
- Gumperz, J. 1982. *Discourse Strategies*. Cambridge: Cambridge University Press.
- Haggard, M., Corrigall, J., and Legg, A. 1971. Perceptual factors in articulatory defects. *Folia Phoniatrica* 23:33-40.
- Hammarberg, B. 1988. *Studien zur Phonologie des Zweitsprachhenerwerbs*. Stockholm: Almqvist and Wiksell.
- Hammarberg, B. 1990. Conditions on transfer in second language phonology acquisition. In *New Sounds 90, Proceedings of the 1990 Amsterdam Symposium on the Acquisition of Second-language Speech*, ed. J. Leather and A. James. Amsterdam: University of Amsterdam.
- Henning, W. 1966. Discrimination training and self-evaluation in the teaching of pronunciation. *International Review of Applied Linguistics* 4:7-17.
- Hill, J. 1970. Foreign accents, language acquisition, and cerebral dominance revisited. *Language Learning* 20:237-48.
- Holden, K., and Hogan, J. 1993. The emotive impact of foreign intonation: An experiment in switching English and Russian intonation. *Language and Speech* 36:67-88.
- James, A. 1984. Phonetic transfer: The structural bases of interlingual assessments. In *Proceedings of the Tenth International Congress of Phonetic Sciences*, ed. M. van den Broecke, and A. Cohen. Dordrecht, The Netherlands: Foris.
- James, A. 1988. *The Acquisition of Second Language Phonology*. Tübingen: Gunter Narr.
- Jamieson, D., and Morosan, D. 1986. Training non-native speech contrasts in adults: Acquisition of the English /ð/-/θ/ contrast by francophones. *Perception & Psychophysics* 40:205-15.
- Jusczyk, P. 1985. On characterizing the development of speech perception. *Neonate Cognition: Beyond the Blooming, Buzzing Confusion*, ed. J. Mehler, and R. Fox. Hillsdale, NJ: Lawrence Erlbaum.



- Jusczyk, P. 1992. Developing phonological categories from the speech signal. In *Phonological Development: Models, Research, and Implications*, ed. C. Ferguson, L. Menn, and C. Stoel-Gammon. Timonium, MD: York Press.
- Kluender, K., Diehl, R., and Killeen, P. 1987. Japanese quail can learn phonetic categories. *Science* 237:1195-97.
- Kluender, K., Diehl, R., and Wright, B. 1988. Vowel-length differences before voiced and voiceless consonants: An auditory explanation. *Journal of Phonetics* 16:153-69.
- Koster, C. 1987. *Word Recognition in Foreign and Native Language*. Dordrecht The Netherlands: Foris.
- Koutsoudas, A., and Koutsoudas, O. 1983. A contrastive analysis of the segmental phonemes of Greek and English. In *Second Language Learning: Contrastive Analysis, Error Analysis, and Related Aspects*, ed. B. Robinett, and J. Schachter. Ann Arbor: University of Michigan Press.
- Ladefoged, P. 1990. Some reflections on the IPA. *Journal of Phonetics* 18:335-46.
- Lado, R. 1957. *Linguistics across Cultures*. Ann Arbor: University of Michigan Press.
- Lambert, W., Hodgson, R., Gardner, R., and Fillenbaum, S. 1960. Evaluation reactions to spoken languages. *Journal of Abnormal Social Psychology* 60:44-51.
- Lamminmäki, R. 1979. The discrimination and identification of English vowels, consonants, junctures, and sentence stress by Finnish comprehensive school pupils. *Papers in Contrastive Phonetics, Jyväskylä Cross-Language Studies* 7:165-231. (Department of English, Jyväskylä University.)
- Lane, H. 1963. Foreign accent and speech distortion. *Journal of the Acoustical Society of America* 35:451-53.
- Lehiste, I. 1988. *Lectures on Language Contact*. Cambridge, MA: MIT Press.
- Lenneberg, E. 1967. *Biological Foundations of Language*. New York: Wiley.
- Liljencrants, J., and Lindblom, B. 1972. Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language* 48:839-62.
- Lindblom, B. 1990a. On the notion of "possible speech sound." *Journal of Phonetics* 18:135-52.
- Lindblom, B. 1990b. Explaining phonetic variation: A sketch of the H and H theory. In *Speech Production and Speech Modeling*, ed. W. Hardcastle, and A. Marchal. Amsterdam: Kluwer.
- Linell, P. 1982. The concept of phonological form and the activities of speech production and speech perception. *Journal of Phonetics* 10:37-72.
- Locke, J. 1980. The inference of speech perception in the phonologically disordered child. Part II: Some clinically novel procedures, their use, some findings. *Journal of Speech and Hearing Disorders* 45:445-68.
- Logan, J., Lively, S., and Pisoni, D. 1991. Training Japanese listeners to identify /r/ and /l/: A first report. *Journal of the Acoustical Society of America* 89:874-86.
- Long, M. 1990. Maturation constraints on language development. *Studies in Second Language Acquisition* 12:251-85.
- Mack, M. 1982. Voicing-dependent vowel duration in English and French: Monolingual and bilingual production. *Journal of the Acoustical Society of America* 71:173-78.
- Mack, M. 1988. Sentence processing by non-native speakers of English: Evidence from the perception of natural and computer-generated anomalous L2 sentences. *Journal of Neurolinguistics* 3:293-316.
- Mack, M. 1989. Consonant and vowel perception and production: Early English-French bilinguals and English monolinguals. *Perception & Psychophysics* 46:187-200.



- Mack, M. 1990. Phonetic transfer in a French-English bilingual child. In *Language Attitudes and Language Conflict*, ed. P. H. Nelde. Bonn, Germany: Dümmler.
- Mack M., Tierney, J., and Boyle, M. 1990. The intelligibility of natural and LPC-vocoded words and sentences presented to native and non-native speakers of English. MIT Lincoln Laboratory, Technical Report 869 (32 pps).
- Magen, H., and Blumstein, S. 1993. Effects of speaking rate on the vowel length distinction in Korean. *Journal of Phonetics* 21:387-410.
- Magno Caldognetto, E., Ferrero, F., Vagges K., and Bagno, M. 1979. Indici acustici e indici percevetti nel riconoscimento dei suoni linguistici. *Acta Phoniatica Latina* 1:219-46.
- Major, R., 1986. The ontogeny model: Evidence from L2 acquisition of Spanish r. *Language Learning* 36:453-504.
- Major, R. 1987a. Phonological similarity, markedness, and rate of L2 acquisition. *Studies in Second Language Acquisition* 9:63-82.
- Major, R. 1987b. English voiceless stop production by speakers of Brazilian Portuguese. *Journal of Phonetics* 15:197-202.
- Major, R. 1992. Losing English as a first language. *The Modern Language Journal* 76:190-208.
- Mann, V. 1986. Distinguishing universal and language-dependent levels of speech perception: Evidence from Japanese listeners' perception of English "l" and "r". *Cognition* 24:169-96.
- Martinet, A. 1955. *Économie des Changements Phonétiques* Berne: A. Franke.
- Mayberry, R., and Eichen, E. 1991. The long-lasting advantage of learning sign language in childhood: Another look at the critical period for language acquisition. *Journal of Memory and Language* 30:486-512.
- McAllister, R. 1990. Perceptual foreign accent: L2 user's comprehension ability. In *New Sounds 90. Proceedings of the 1990 Amsterdam Symposium on the Acquisition of Second-language Speech*. ed. J. Leather, and A. James. Department of English, University of Amsterdam.
- Miller, J., and Volaitis, L. 1989. Effect of speaking rate on the perceptual structure of a phonetic category. *Perception & Psychophysics* 46:505-12.
- Miyawaki, K., Liberman, A., Jenkins, J., and Fujimura, O. 1975. An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception & Psychophysics* 18:331-40.
- Monnin, L., and Huntington, D. 1974. Relationship of articulatory defects to speech-sound identification. *Journal of Speech and Hearing Research* 17:352-66.
- Morosan, D., and Jamieson, D. 1989. Evaluation of a technique for training new speech contrasts: Generalization across voices, but not word-position or task. *Journal of Speech and Hearing Research* 32:501-11.
- Moulton, W. 1945. Dialect geography and the concept of phonological space. *Word* 18:23-32.
- Moulton, W. 1962. *The Sounds of English and German*. Chicago: University of Chicago Press.
- Mueller, T., and Niedzielski, H. 1963. The influence of discrimination training on pronunciation. *Modern Language Journal* 52:410-16.
- Mullennix, J., Pisoni, D., and Martin, C. 1989. Some effects of talker variability on spoken word recognition. *Journal of the Acoustical Society of America* 85:365-78.
- Munro, M. 1992. Perception and production of English vowels by native speakers of Arabic. Unpublished Ph.D. dissertation, University of Alberta.
- Munro, M. 1993. Non-native productions of English vowels: Acoustic measurements and accentedness ratings. *Applied Psycholinguistics* 36:39-66.



- Munro, M., Flege, J. E., and MacKay, I. 1995. Effects of age of second-language learning on the production of English vowels. *Applied Psycholinguistics*. To appear.
- Nabelek, A., and Donahue, A. 1984. Perception of consonants in reverberation by native and non-native listeners. *Journal of the Acoustical Society of America* 75:632-34.
- Nakauchi, J. 1993. The identification of /l/ and /r/ by Japanese early, intermediate and late bilinguals. *Speech, Hearing and Language, Work in Progress* 7:155-79 (University College, London, Dept. of Phonetics and Linguistics.)
- Neufeld, G. 1979. Towards a theory of language learning ability. *Language Learning* 29:227-41.
- Nooteboom, S., and Truin, P. 1980. Word recognition from fragments of spoken words by native and non-native listeners. *IPO Annual Progress Report* 15:42-47.
- Novoa, L., Fein, D., and Obler, L. 1988. Talent in foreign languages: A case study. In *The Exceptional Brain, Neuropsychology of Talent and Special Abilities*, ed. L. Obler, and D. Fein. New York: Guilford Press.
- Oyama, S. 1976. The sensitive period for the acquisition of a non-native phonological system. *Journal of Psycholinguistic Research* 5:261-85.
- Oyama, S. 1978. The sensitive period and comprehension of speech. *Working Papers on Bilingualism* 16:1-17.
- Patkowski, M. 1990. Age and accent in a second language: A reply to James Emil Flege. *Applied Linguistics* 11:73-89.
- Penfield, W. 1965. Conditioning the uncommitted cortex for language learning. *Brain* 88:787-98.
- Peng, S. H. 1993. Cross-language influence on the production of Mandarin /f/ and /x/ and Taiwanese /h/ by native speakers of Taiwanese Amoy. *Phonetica* 50:245-60.
- Pimsleur, P. 1963. Discrimination training in the teaching of French pronunciation. *Modern Language Journal* 47:199-203.
- Polka, L. 1991. Cross-language speech perception in adults: Phonemic, phonetic, and acoustic contributions. *Journal of the Acoustical Society of America* 89:2961-77.
- Powers, M. 1957. Functional disorders of articulation-symptomatology and etiology. In *Travis Handbook of Speech Pathology*. New York: Appleton Century Crofts.
- Samuel, A., Kat, D., and Tartter, V. 1984. Which syllable does an intervocalic stop belong to? A selective adaptation study. *Journal of the Acoustical Society of America* 6:1652-63.
- Schmidt, A. M., and Flege, J. E. 1995. Effects of speaking rate changes on native and non-native production. *Phonetica* 52:41-54.
- Scovel, T. 1988. *A Time to Speak, A Psycholinguistic Inquiry into the Critical Period for Human Speech*. Cambridge, MA: Newbury House.
- Sekiyama, K., and Tohkura, Y. 1993. Inter-language differences in the influence of visual cues in speech perception. *Journal of Phonetics* 21:427-44.
- Sheldon, A., and Strange, W. 1982. The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics* 3:243-61.
- Snow, C., and Hoefnagel-Höhle, M. 1979. Individual differences in second-language ability: A factor-analytic study. *Language and Speech* 22:151-62.
- Spiestersbach, R., and Curtis, J. 1951. Misarticulation and discrimination of speech sounds. *Quarterly Journal of Speech* 37:483-91.



- Strange, W. 1992. Learning non-native phoneme contrasts: Interactions among subject, stimulus, and task variables. In *Speech Perception, Production, and Linguistic Structure*, ed. E. Tohkura, E. Vatikiotis-Bateson, and Y. Sagisaka. Tokyo: Ohmsha.
- Sussman, J. 1993. Auditory processing in children's speech perception: Results of selective adaptation and discrimination tests. *Journal of Speech and Hearing Research* 36:380-95.
- Takagi, N. 1993. Perception of American English /r/ and /l/ by adult Japanese learners of English: A unified view. Unpublished Ph.D dissertation, University of California-Irvine.
- Terbeek, D. 1977. A cross-language multidimensional scaling study of vowel perception. *UCLA Working Papers in Phonetics*, 37.
- Trubetzkoy, N. 1939/1969. *Principles of Phonology*. Translated by C. A. Baltaxe, Berkeley, CA University of California Press.
- Uchanski, M., Miller, D., Ronan, C., Reed, M., and Braida, L. 1991. Effects of token variability on resolution for vowel sounds. *Journal of the Acoustical Society of America*, 90:2254(A).
- Vaggas, K., Ferrero, F., Magno-Caldognetto E., and Lavagnoli C. 1978. Some acoustic characteristics of Italian consonants. *Journal of Italian Linguistics* 3:69-85.
- Waldman, F., Singh, S., and Hayden, M. 1978. A comparison of speech-sound production and discrimination in children with functional articulation disorders. *Language and Speech* 21:205-20.
- Wang, C. 1988. The production and perception of English vowels by native speakers of Mandarin. Unpublished MA thesis, University of Alabama at Birmingham.
- Weiher, E. 1975. Lautwahrnehmung und Lautproduktion im Englischunterricht für Deutsche. Working Papers, 3, Institute of Phonetics, University of Kiel.
- Weinberger, S. 1987. The influence of linguistic context of syllable structure simplification. In *Interlanguage Phonology*, ed. G. Ioup, and S. Weinberber. Rowley, MA: Newbury House.
- Weinberger, S. 1990. Minimal segments in second language phonology. In *New Sounds 90, Proceedings of the 1990 Amsterdam Symposium on the Acquisition of Second-Language Speech*, ed. J. Leather, and A. James. Amsterdam: University of Amsterdam.
- Weiner, P. 1967. Auditory discrimination and articulation. *Journal of Speech & Hearing Disorders* 32:19-38.
- Weinreich, U. 1953. *Languages in Contact, Findings and Problems*. The Hague: Mouton.
- Weinreich, U. 1957. On the description of phonic interference. *Word* 13:1-11.
- Wenk, B. 1979. Articulatory setting and de-fossilization. *Interlanguage Studies Bulletin* 4:202-20.
- Werker, J., and Logan, J. 1985. Cross-language evidence for three factors in speech perception. *Perception & Psychophysics* 37:35-44.
- Werker, J., and Tees, D. 1983. Developmental changes across childhood in the perception of non-native speech sounds. *Canadian Journal of Psychology* 37:278-86.
- Werker, J., and Tees, D. 1984. Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America* 75:1866-78.
- Wieden, W. 1990. Some remarks on developing phonological representations. In *New Sounds 90, Proceedings of the 1990 Amsterdam Symposium on the*



- Acquisition of Second-language Speech*, ed. J. Leather, and A. James. Amsterdam: University of Amsterdam.
- Williams, L. 1979. The modification of speech perception and production in second-language learning. *Perception & Psychophysics* 26:95-104.
- Wode, H. 1978. The beginning of non-school room L2 phonological acquisition. *International Review of Applied Linguistics* 16:109-25.
- Yamada, R., and Tohkura, Y. 1992a. The effects of experimental variables on the perception of American English /r/ and /l/ by Japanese listeners. *Perception & Psychophysics* 52:376-92.
- Yamada, R., and Tohkura, Y. 1992b. Perception of American English /r/ and /l/ by native speakers of Japanese. In *Speech Perception, Production, and Linguistic Structure*, ed. E. Tohkura, E. Vatikiotis-Bateson, and Y. Sagisaka. Tokyo: Ohmsha.
- Yeni-Komshian, G., and Flege, J. E. 1993. The effect of age of L2 acquisition on accuracy of pronunciation of phonetic segments. Paper presented at the 6th International Congress for the Study of Child Language. Trieste, Italy, July 18-24, 1993.