Assessing constraints on second-language segmental production and perception

James Emil Flege

1. Introduction

A great deal of research has examined the production and perception of phonetic segments in a second language (L2). The impetus for much of this work has been the desire to understand why individuals who learn an L2—especially those who began learning the L2 in late adolescence or adulthood—differ from monolingual native speakers of the target L2. A variety of proposals have been offered as to whether L2 speech learning is "constrained" in comparison to L1 speech learning, what is the basis for such constraints (should they exist), whether constraints differ for production and perception, and whether L2 learners must inevitably differ from L2 native speakers.

The purpose of this chapter is to review theory and evidence relating to the production and perception of L2 phonetic segments. Section 2 reviews theoretical issues, considering segmental production and perception separately. Section 3 summarizes the results of empirical research examining L2 vowel production and perception in light of the theoretical issues. The relation between segmental production and perception is considered in Section 4. Finally, Section 5 briefly sets goals for future research.

2. Theory

2.1. Production

Studies of L2 production have focused on the production of individual L2 vowels and consonants, consonant clusters, words, and whole

sentences (see Leather and James, 1996, for review). It is common to observe divergences from L2 phonetic norms in the speech of L2 learners. Controversy exists as to whether individuals who began learning the L2 in childhood ("early" learners) will differ from L2 native speakers. However, there is agreement that the magnitude of native versus non-native differences is generally greater for individuals who began learning the L2 in adolescence and adulthood ("late" learners) than for early learners (see Long 1990 for review). This has led to the proposal that L2 speech learning is constrained by a critical period arising from the loss of neural "plasticity" (McLaughlin 1977; Patkowski 1989).

Adherents of the critical period hypothesis suggest that the capacity for successful speech and language learning declines beyond the critical period. For example, DeKeyser (2000: 518-519) suggested that

> Somewhere between the ages of 6-7 and 16-17, everybody loses the mental equipment required for the abstract patterns underlying a human language, and the critical period really deserves its name … It may be that the severe decline of the ability to induce abstract patterns implicitly is an inevitable consequence of fairly general aspects of neurological maturation and that it simply shows up most clearly in language acquisition.

Production and perception are not usually differentiated in discussions of the critical period hypothesis. However, Scovel (1988: 62) observed that:

> Pronunciation is the only part of language which is directly "physical" and which demands neuromuscular programming. Only pronunciation requires an incredible talent for sensory feedback of where the articulators are and what they are doing. And only pronunciation forces us to time and sequence motor movements. All other aspects of language are entirely "cognitive" or "perceptual" in that they have no physical reality.

This suggests the possibility that if L2 acquisition is constrained by a critical period, it may affect segmental production and perception differently.

Bever (1981: 196) hypothesized that segmental production and

perception are aligned via a "psychogrammar" that is used in L1 acquisition to develop "conjoint" representations of perception and production. Bever hypothesized that the psychogrammar decays as L1 phonology acquisition reaches completion, which marks the end of a critical period for speech learning. Following the critical period, speech production and perception develop independently in such a way that individuals "often learn to discriminate sounds ... they cannot distinctively produce."

Other accounts of the relation between segmental production and perception have appeared in the literature. Pisoni (1995: 22-23) observed that the relation between production and perception is "complex" but that it nonetheless reflects the properties of a "unitary articulatory event." He concluded that talkers produce "precisely the same acoustic differences that are distinctive in perceptual analysis," and that the relation between speech production and perception is "unique" among category systems. However, a close relation between perception and action may be a general characteristic of brain functioning. Churchland (1986: 473) observed that "evolution [has] solved the problem of sensory processing and motor control simultaneously," so that "theories [must] mimic evolution and aim for simultaneous solutions as well." According to Edelman's theory of neuronal group selection (1989: 54-56), a "dynamic loop ... continually matches gestures and posture to several kinds of sensory signals," so that perception "depends upon and leads to action" and motor activity is seen as an "essential part of perceptual categorization."

Kuhl and Meltzoff (1996: 2425; see also Kuhl 2000: 11854) concluded that the information specifying auditory-articulatory relations must be "exquisitely detailed," and that even adults may have an "internalized auditory-articulatory 'map' that specifies the relations between mouth movements and sound." These authors noted, however, that an asymmetry exists in early stage of L1 speech learning. Specifically, the "formation of memory representations ... derives initially from perception of the ambient input and then acts as guides for motor output."

This last observation has been extended to L2 speech learning. Rochet (1995) examined the perception of a synthetic French /i/-/y/-

/u/ continuum by speakers of Portuguese and English. The native Portuguese participants tended to misidentify /y/-quality vowels as /i/ whereas native English participants tended to misidentify the same vowels as /u/. In a repetition task, the native Portuguese subjects realized /y/-quality vowels as /i/ whereas the native English participants tended to realize them as /u/. From this, Rochet concluded (1995: 404) that some L2 production errors are "the consequence of the target phones having been assigned to an L1 category." Flege (1995) suggested that L2 production accuracy is limited by perceptual accuracy. More specifically, he hypothesized that the production of an L2 phonetic segment will typically be no more native-like than its perceptual representation and might, in early stages of learning, be less native-like.

2.2. Perception

Several hypotheses regarding constraints on L2 perception have appeared in the literature, all of which assume that the perception of L2 phonetic segments is influenced by the L1 phonological system. Trubetzkoy (1939/1958) compared the L1 phonological system to a "sieve" through which L2 vowels and consonants must pass. Michaels (1974) noted that Russians tend to substitute /t/ for English /θ/ whereas Japanese learners substitute /s/ even though both Russian and Japanese have /t/ and /s/ (which are classified as non-strident and continuant sounds, respectively). Michaels hypothesized that Russians' perception of "non-stridency" in English /θ/ leads them to substitute the closest non-strident Russian sound, /t/, whereas Japanese speakers' perception of "continuancy" in English /θ/ leads them to substitute the closest continuant sound in Japanese, /s/. An implication of Michaels' (1974) hypothesis is that the relative importance of distinctive features may differ across languages, and that this influences L2 segmental perception.

The filtering hypothesis might be extended to continuously varying phonetic features or properties. Saudi Arabian Arabic has voiced stops (/b d g/) as well as voiceless stops (/t k/) in its inventory. If distinctive features were freely commutable, Saudi adults should have

no difficulty producing English /p/. However, Flege and Port (1981) found that Saudi adults who had lived in the United States for several years did have difficulty. They tended to produce English /p/ with the temporal properties of a bilabial stop, but with the closure voicing appropriate for /b/. These participants may have had difficulty integrating the glottal and supraglottal gestures needed for /p/. Alternatively, they may not have perceived the properties of English /p/ accurately. Sebastián-Gallés and Soto-Faraco (1999: 112) observed that children learn to weight acoustic features of speech in a way that is optimal for their L1 and that, later in life, L2 speech input will be "sieved" through L1-tuned feature weights. The extent to which the feature weights are "realigned" for the processing of L2 speech sounds may depend on the age of exposure to the L2.

The Perceptual Assimilation Model (PAM) developed by Catherine Best and colleagues (e.g., Best 1995; Best et al. 2001) proposes that the accuracy with which L2 speech sounds are discriminated will depend on how, or if, they are perceptually "assimilated" by L1 speech sounds. Instances of distinct L2 categories that are not perceptually assimilated by any L1 category will be discriminated well, even in the absence of prior experience. However, it appears that most L2 speech sounds are perceptually assimilated by an L1 category, at least initially. That being the case, L2 speech sounds will be discriminated more accurately if they are assimilated by two distinct L1 speech sounds than if they are assimilated by a single L1 speech sound category. The PAM predicts that discrimination accuracy may also be influenced by the degree of phonetic-articulatory similarity of L2 speech sounds to L1 speech sounds. Specifically, the PAM predicts that a pair of L2 speech sounds differing in perceived degree of goodness of fit to a single L1 category will be discriminated better than a pair of L2 speech sounds judged to have an equal goodness of fit to a single L1 category.

The primary aim of the Native Language Magnet (NLM) model developed by Patricia Kuhl and colleagues (e.g., Kuhl 2000) is to account for the transition from auditory to language-specific perceptual processing. The NLM proposes that perception of the acoustic properties of speech sounds is defined by early experience. Infants percep-

tually sort segment-sized units into categories based on the recurrence of features they have detected in speech input. This results in a language-specific "mapping" between the categories developed for L1 speech sounds and the phonetic input that drives this crucial aspect of language acquisition. For example, Kuhl et al. (1992) compared the perception of synthetic high front vowels by 6-month-old infants being raised in English- and Swedish-speaking environments in the United States and Sweden, respectively. The English-learning infants were found to generalize their conditioned response to a good instance of English /i/ to neighboring English /i/ tokens, but did not show the same kind of response generalization (i.e., failure to discriminate) when exposed to a similar array of Swedish /y/ vowels. Swedish-learning infants, on the other hand, showed response generalization to a good Swedish /y/ token but not to a good English /i/ token.

The NLM (Kuhl 2000: 11854) proposes that infants' perceptual mapping of ambient language speech sounds creates a "complex network, or filter, through which language is perceived." Perceptual attunement to L1 categories may later shape the perception of L2 speech sounds. Interference effects may arise because of the difficulty inherent in functionally separating L1 and L2 mappings (i.e., categories), and because a neural "commitment" to L1 category mappings will later influence the processing of L2 speech sounds (see also Flege 1992). Importantly, the NLM proposes that constraints on the perception of L2 speech sounds arise from prior experience, not from a loss of plasticity that arises from normal neural maturation.

Support for the NLM account of the basis for native versus nonnative perceptual differences was obtained in a study examining English /ɹ/ and /l/. Iverson et al. (2001) had native English adults and native Japanese adults living in Tokyo rate the acoustic similarity of the members of a grid of /ɹa/ and /la/ stimuli. The stimuli differed in terms of the frequencies of F2 and F3 transitions into the vowel. Multidimensional scaling analyses suggested that the perception of acoustic-phonetic dimensions was shaped by attunement to the L1 phonetic system in a way that might be conceptualized as a "warping" of the phonetic space. Specifically, the native English adults showed an

augmented sensitivity to F3 differences between stimuli perceivable as English /ɹ/ and /l/, and a reduced sensitivity to F3 differences between stimuli identifiable as instances of either the /ɹ/ or the /l/ category. Unlike the native English participants, the native Japanese participants did not show a heightened discrimination of stimuli straddling the English /ɹ/-/l/ boundary. Moreover, they did not show evidence of either a stretching or a shrinking of the F3 dimension. In fact, the native Japanese participants were more sensitive to variation in the F2 than the F3 dimension.

The Iverson et al. (2001) results suggested that native speakers of Japanese develop perceptual maps that, although well suited for Japanese, may impede acquisition of the English /ɹ/-/l/ contrast (see also Flege 1988). The authors suggested that as a result of L1 interference effects, Japanese adults who do manage to establish new categories for English liquids might develop "erroneous" long-term memory representations in which variation in F2 frequency is given too much prominence and F3 frequency is given too little prominence.

Iverson et al. (2001) suggested that L1 interference effects might become progressively stronger as the L1 develops. Influence of the L1 may be "self-reinforcing" for Japanese adults if, as the result of a warping of the phonetic space, they fail to experience the same "auditory distribution" of F3 differences in English /ɹ/ and /l/ tokens as do children who are learning English as an L1 (2001: 114-115). However, according to the NLM, perceptual learning by adults is not impossible. Kuhl suggested (2000: 11855) that the influence of prior experience may be minimal for children who learn two languages simultaneously in early childhood, at least if "two different mappings" are acquired for L1 and L2 speech sounds. The best way for adult learners of an L2 to circumvent L1 interference effects may be to recapitulate infants' experience of L1 speech, that is, to receive "exaggerated acoustic cues, multiple instances by many talkers, and massed listening experience."

To summarize so far, two broad proposals have been offered in the literature to account for why the speech of L2 learners is often foreign-accented. The ability to learn to produce speech accurately may be constrained by a critical period. Alternatively, the accuracy with

which L2 speech sounds are produced may be limited by the extent to which the perceptual representations developed for L2 phonetic segments resemble those of native speakers of the target L2. Several proposals have, in turn, been offered to account for inaccurate perception of L2 phonetic segments. These proposals converge on the notion that the features or properties needed to develop accurate perceptual representations may, in some instances, be inaccessible to L2 learners.

A question of central importance is whether the limitations described in the literature affect L2 perceptual learning permanently, regardless of the learner's age, the kind or amount of L2 input received, or the contexts in which the L2 has been learned or used. For example, Best and Strange (1992: 327) hypothesized that experience with an L2 may lead to the "reorganization of perceptual assimilation patterns" which may, in turn, affect discriminability. However, Best and Strange did not specify the conditions under which perceptual assimilation patterns might change, or whether limits exist on the extent of change that is possible.

2.3. The Speech Learning Model

The Speech Learning Model (SLM) developed by James Flege and colleagues (e.g., Flege 1988, 1992, 1995, 1999a, 2002) is the only extant theory that focuses explicitly on L2 speech acquisition. Its primary aim is to account for changes across the life span in the learning of segmental production and perception. The SLM starts with two broad assumptions. The first is that bilinguals cannot fully separate their L1 and L2 phonetic subsystems (see also Paradis 1993). The second, and more controversial of the two SLM starting assumptions is that the capacities underlying successful L1 speech acquisition remain intact across the life span. These capacities include the ability to accurately perceive featural patterns in speech input, to sort a wide range of segments possessing common properties into categories, and to relate vocal output to the properties perceived in speech sounds (see also Kuhl 2000). The second SLM assumption stands in contrast to the view that L2 speech learning is constrained by a critical period

(e.g., Scovel 1988; DeKeyser 2000).

The SLM does not discount the proposals described earlier regarding the filtering or warping of L2 speech input. Indeed, it seems reasonable to suppose that L2 learners may filter out phonetic features or properties that are used to distinguish L2 but not L1 speech sounds in early stages of L2 speech learning. In support of this, Munro (1993) found that native Arabic men who had lived in the United States for an average of 6 years learned to produce a native-like spectral difference between English /i/ and /ɪ/, which differ spectrally from the closest vowels of Arabic (viz., /i/ and /i:/). However, these participants exaggerated the temporal difference between English /i/ and /ɪ/, as if they were producing phonologically long and short Arabic vowels rather than a tense and a lax English vowel.

There is nonetheless evidence that cross-language phonetic differences are detectable by naïve, inexperienced listeners in certain conditions, and that adults are perceptually sensitive to small divergences from the phonetic norms of their L1 (e.g., Flege 1984). This led to the SLM proposal that the filtering of L2 speech input will not persist as learners acquire a dense network of L2 lexical items that need to be differentiated phonetically. In support of this, McAllister, Flege, and Piske (2002) found that some native speakers of English and Spanish who were long-time residents of Stockholm learned to distinguish Swedish words differing in phonological quantity even though vowel duration is not used as the primary cue to vowel contrasts in either English or Spanish.

The results of Gottfried and Beddor (1988; see also Francis and Nusbaum 2002) suggested that L2 learning might result in a change in feature weighting. A synthetic continuum ranging from French /o/-/ɔ/ was created through an orthogonal variation in the frequency of the first two vowel formants (F1, F2) and duration. Native English speakers responded to a greater extent to duration differences than native French speakers did owing to the greater overall prominence of duration as a cue to vowel identity in English than French. Native speakers of English who had studied French in school showed an English-like use of duration in classifying the French vowels. However, more advanced native English speakers of French showed less

use of duration, thereby resembling native speakers of French.

Best and Strange (1992) suggested that new categories are more likely to be established for L2 speech sounds that are perceived to be "discrepant" instances of an L1 category than for L2 sounds perceived to be distant from the closest L1 sound. Conversely, the SLM predicts that the greater is the perceived phonetic dissimilarity of an L2 speech sound from the closest L1 sound, the more likely it is that a new category will be created for the L2 sound. For example, Flege (1987) found that adult native English learners of French were more successful in learning to produce French /y/ than /u/. This was attributed to the greater perceived phonetic distance of French /y/ from the closest English vowel than of French /u/ from the closest English vowel. (Perceptual assimilation data was not collected to verify this, however.)

Like the NLM, the SLM proposes that native versus non-native differences are more likely to arise as the result of interference from prior phonetic learning than from a loss of neural plasticity. The SLM proposes that even adults retain the capacities used by infants and children in successfully acquiring L1 speech, including the ability to establish new phonetic categories for the vowels and consonants encountered in an L2. However, the SLM proposes that phonetic category formation for L2 speech sounds becomes less likely with increasing age. According to the SLM, as L1 phonetic categories develop slowly through childhood and into early adolescence, they become more likely to perceptually assimilate L2 vowels and consonants. If instances of an L2 speech sound category persist in being identified as instances of an L1 speech sound, category formation for the L2 speech sound will be blocked. A limitation of the SLM is that it does not provide a metric for determining when cross-language phonetic differences will be too small to support category formation, and whether the triggering threshold varies as a function of age or L1 system development.

Baker et al. (2002) evaluated the SLM hypothesis that as L1 vowel categories develop, L2 vowels are more likely to be identified as instances of those categories. These authors carried out a perceptual assimilation experiment with Korean adults and children who had

lived in the United States for just 9 months. The results suggested that the perceptual assimilation of English vowels by Korean vowels was stronger for the Korean adults than children. Additional experiments examined Koreans who had arrived in the United States at average ages of 9 or 19 years (early and late learners) and had lived there for 9 years. One experiment examined the categorial discrimination of pairs of English vowels (/i/-/ɪ/, /ɛ/-/æ/, /u/-/ʊ/) that were perceptually assimilated by a single Korean vowel. The early learners discriminated the English vowels more accurately than the late learners did, and did not differ significantly from native English speakers. The same results were obtained in an experiment examining the production of English /i ɪ ɛ æ u ʊ/. The authors suggested that between-group differences in segmental production and perception may have been due to age-related differences in the strength of perceptual assimilation of English vowels by Korean vowels.

As mentioned earlier, language-specific attunement becomes evident in infancy (e.g., Kuhl et al. 1992). Children are generally credited with having acquired the phonemes of their L1 by the age of 8 years. However, the development of speech motor control and perceptual representations for L1 speech sounds appear to develop slowly through childhood and into adolescence (e.g., Hazan and Barrett 1999; Johnson 2000; Walley and Flege 2000). Unfortunately, the endpoint of L1 speech development has not yet been determined. If it coincided with the age thought to mark the end of a critical period for L2 acquisition, 12-15 years (Scovel 1988; Patkowski 1989), it would be difficult to differentiate a maturational account of age effects on L2 speech learning (the critical period hypothesis) from a developmental account (that of the SLM or the NLM).

The two accounts might be differentiated, however, by testing for effects of L2 learning on the production and perception of phonetic segments in the L1. The SLM proposes that the L1 and L2 phonetic subsystems of bilinguals necessarily interact because the phonic elements making up the L1 and L2 phonetic subsystems exist in a "common phonological space". According to the SLM, individual bilinguals strive to maintain contrasts between phonic elements in both the L1 and L2 phonetic subsystems in much the same way that

languages maintain contrasts between the phonic elements making up a single system (see de Boer 2000 for discussion).

The SLM proposes that phonetic categories interact through mechanisms called "phonetic category assimilation" and "phonetic category dissimilation." (See Flege 2000 for examples of how both mechanisms influence L2 speech learning.) When a new category is established for an L2 speech sound in a portion of phonetic space occupied by an L1 sound, the new L2 category and the pre-existing L1 category may dissimilate from one another. If this happens, neither the L1 category nor the new L2 category will be identical to the categories possessed by monolinguals. The modification of an L1 category as the result of category dissimilation is not predicted by either a critical period hypothesis or a filtering/warping hypothesis.

Category assimilation is predicted to occur when a new category has not been established for an L2 speech sound that differs audibly from the closest L1 speech sound. In such cases, an experienced L2 learner is predicted to develop a "composite" category that merges the properties of the L1 and L2 categories that have been perceptually equated, in proportion to the input received (perhaps with greater weight accorded recent input). As a result, productions of the L2 sound will remain L1-like and productions of the corresponding L1 sound will eventually become L2-like.

In support of this, Flege (1987) observed that native French adults who had learned English, and native English adults who had learned French, produced L2 stop consonants with voice onset time (VOT) values differing from the VOT values produced by native English and French speakers, respectively. The native French learners managed to increase VOT in English stops, but not sufficiently to match English monolinguals. Conversely, the native English learners decreased VOT in French stops, but not sufficiently to match French monolinguals. These native versus non-native differences might be attributed to the passing of a critical period or, indirectly, to the filtering/warping of L2 speech input. However, a critical period (or filtering/warping) hypothesis would not predict the observed changes in L1 production (viz., a lengthening of VOT in French stops produced by the native French participants, and a shortening of VOT in English

stops by the native English participants).

L2 speech learning, as envisaged by the SLM, takes place slowly and requires a large amount of native-speaker input to be successful. This consideration is based on the observation that L1 speech develops over a long period of time (see above). Evidence for constraints on L2 speech learning can only be considered persuasive, therefore, if obtained for individuals who have received as much L2 input as is needed by children learning that language as an L1 to fully acquire its phonetic segments.

The importance of input is illustrated by the divergent results obtained in two studies examining the production of English /p t k/ by groups of native Spanish adults who learned English in childhood. Participants who had learned English primarily from native speakers of English in the United States produced voiceless English stops with the long-lag VOT values typical of English (Flege 1991a). However, participants who learned English primarily from native speakers of Spanish in Puerto Rico (Flege and Eefting 1987) produced English stops with VOT values that were intermediate to the VOT values typical for /p t k/ in Spanish and English. It appeared that the early learners in Puerto Rico based their representation of English stops on the foreign-accented input they had received from other native Spanish speakers.

The results of Flege and Liu (2001) also illustrated the importance of input. These authors examined the identification of word-final English consonants by native Chinese adults who had lived in the United States for averages of 2 and 7 years. The participants differed in the nature of their daily activities. Half of the participants in each group were enrolled as full-time students, whereas the remaining participants held occupations (e.g., laboratory technician) thought likely to reduce the frequency of interactions with native English speakers. The long-residence students obtained significantly higher identification scores than the short-residence students did, but there was no difference between non-students who differed in length of residence. The students and non-students did not differ in terms of self-reported percentage use of English (roughly 50% for both groups). This led to the inference that what led to a difference in speech perception be-

tween the long-residence students and non-students was not how frequently they used English, but with whom.

3. L2 vowel acquisition research

The studies reviewed in this section examined the acquisition of English vowels by native speakers of Spanish (Section 3.1), the acquisition of Catalan vowels by native speakers of Spanish (Section 3.2), and the acquisition of English vowels by native speakers of Italian (Section 3.3). All of the studies reviewed here examined individuals who were highly experienced in the L2. The results therefore bear on the issue of how, or to what extent, L2 speech learning is constrained.

3.1. English /i/ *and* /ɪ/

Spanish has fewer vowels than English does (5 versus 14 in most dialects). This raises the issue of whether native speakers of Spanish will establish new vowel categories when they learn English, or whether they will simply try to adapt Spanish vowels when producing and perceiving English vowels. This general question can be illustrated by considering the acquisition of English /i/ and /ɪ/. Spanish has a single vowel, /i/, in the portion of vowel space occupied by English /i/ and /ɪ/. Physiological and acoustic measurements (Flege 1989; Bradlow 1995) indicate that Spanish /i/ is somewhat lower in vowel space than English /i/, but somewhat higher than English /ɪ/.

Three studies examined native Spanish adults' perception of the relation between English /i ɪ/ and Spanish vowels. The results suggested that Spanish adults can detect differences between English /ɪ/ and Spanish /i/ whereas they can probably not reliably detect differences between English /i/ and Spanish /i/. Flege (1991b) found that native Spanish adults almost always classified English /i/ tokens as instances of Spanish /i/ and English /ɪ/ tokens as instance of either Spanish /i/ or /e/. Participants who spoke English as an L2 were more likely to identify English /ɪ/ tokens as "not a Spanish vowel" than Spanish monolinguals were. In a study by Flege, Munro, and Fox

(1994), native Spanish adults judged English /ɪ/ and Spanish /i/ tokens to be significantly more dissimilar from one another than English /i/ and Spanish /i/ tokens. In Wayland, Flege, and Imai (under review), 78 native Spanish adults who had lived in the United States for less than seven years consistently classified English and Spanish /i/ tokens as Spanish /i/, and English /ɪ/ tokens as either Spanish /i/ or /e/. The participants also rated English and Spanish vowels as instances of the selected Spanish category. The ratings for Spanish and English /i/ tokens that were classified as Spanish /i/ did not differ significantly (means = 4.1 on a 5-point scale for both). However, the English /ɪ/ tokens that were classified as Spanish /i/ received significantly lower ratings (mean = 2.1) than the English /i/ tokens classified as Spanish /i/ did ($p < .01$).

Based on the results just presented, the SLM predicts that some native Spanish adults—and even more native Spanish children who learn English—will eventually establish a phonetic category for English /ɪ/, and therefore produce and perceive this vowel accurately. The results of two studies are consistent with these predictions.

Flege, Bohn, and Jang (1997) examined 20 native Spanish adults who had lived in the United States for an average of 5 years. The participants identified the members of a synthetic "beat-bit" (/i/-/ɪ/) continuum in which F1 frequency values and vowel duration values varied orthogonally. Some participants showed no sensitivity to the spectral (F1) manipulation, basing their responses on duration alone. However, five participants showed as large an effect of F1 frequency variation as native English speakers did.

The late learners' production of English vowels was elicited in a word-reading task. Only a few participants were found to differentiate /i/ from /ɪ/ adequately. Native English-speaking listeners identified /ɪ/s produced by the native Spanish adults as intended somewhat less often (mean = 56%) than their /i/s (mean = 63%). Importantly, acoustic analyses revealed that two of the five participants who had shown perceptual sensitivity to the English /i/-/ɪ/ distinction produced spectral and temporal differences between /ɪ/ and /i/ that were as large, or larger, than differences produced by native English speakers. Not surprisingly, the /ɪ/s and /i/s produced by these native Spanish late

learners were identified correctly by native English listeners (see also Morrison 2002). Flege (1992) used the same procedures to evaluate vowels produced by native Spanish adults who had begun to learn English by school age. Acoustic analyses revealed that the early learners' /ɪ/s and /i/s differed in duration and showed no spectral overlap, and so were correctly identified by native English listeners.

In summary, the results just reviewed suggest that late L2 learners do not automatically filter out cross-language phonetic differences. It appears that most native Spanish early learners, and a few native Spanish late learners, manage to distinguish the English vowels /i/ and /ɪ/, which occur in a portion of vowel space occupied by a single Spanish vowel, /i/.

The results are consistent with predictions of the SLM, but are limited in several ways. Neither study reviewed here sought to determine if early learners use the same relative weighting of features as native English speakers do when perceiving /i/ and /ɪ/. Neither study tested the SLM prediction that native Spanish adults will be more likely to continue identifying English /ɪ/ tokens as instances of Spanish /i/ than native Spanish children will be. Finally, neither study linked native Spanish learners' perceptual assimilation patterns to their production and perception of English vowels. Late learners in the Flege, Bohn, and Yang (1997) study who perceived and produced a distinction between /i/-/ɪ/ may have done so for the reasons hypothesized by the SLM. (That is, they may have discerned the phonetic differences that distinguish English /i/ from /ɪ/, as well as the differences between English /ɪ/ and Spanish /i/.) However, they may have succeeded better than most other participants because they initially identified English /i/ and /ɪ/ as instances of two different Spanish categories (/i/ and /e/).

3.2. Catalan /e/ *and* /ɛ/

The results obtained in four studies carried out in Barcelona call into question an expectation generated by the SLM, viz. that most early bilinguals will establish phonetic categories for L2 vowels not found in the L1 inventory. This research examined native Spanish university students who had begun to learn Catalan by school age. The partici-

pants were said to be highly proficient in both Spanish and Catalan, and to use both of their languages frequently. Each study used a different technique to assess the perception of Catalan speech sounds. Pallier, Bosch, and Sebastián-Gallés (1997) examined the identification and discrimination of vowels in a synthetic continuum. Pallier, Colomé, and Sebastián-Gallés (2001) used the repetition priming paradigm. Sebastián-Gallés and Soto-Faraco (1999) used a version of the gating paradigm. Bosch, Costa, and Sebastián-Gallés (2000) employed the "perceptual magnet" paradigm used to evaluate the NLM.

All four studies examined Catalan /e/ and /ɛ/. These vowels occur in a portion of vowel space that is occupied by a single Spanish vowel, /e/.[1] Bosch et al. (2000) described Spanish /e/ as a vowel that occurs near the perceptual boundary between Catalan /e/ and /ɛ/, and has [e] and [ɛ] allophones. The four studies converged in showing differences between the native Spanish speakers of Catalan and native Catalan speakers of Spanish. Pallier et al. (1997: B14) concluded that even early and frequent exposure to an L2 may not enable the learning of "two new phonetic categories which overlap" a single L1 category. Sebastián-Gallés and Soto-Faraco (1999: 120) interpreted their findings to indicate a "lack of plasticity" in early bilinguals, and suggested that the malleability of the speech perception system may be limited "severely" by school age because exposure to the L1 exerts a "very strong constraint" on the "organization and acquisition of phonemic categories." Bosch et al. (2000) suggested that the early Spanish-Catalan bilinguals continued to represent Catalan vowels as "foreign" speech sounds for which "stable representations in long-term memory" were not established.

These conclusions are reminiscent of conclusions drawn from early studies examining the acquisition of English stops by native speakers of Romance languages. Caramazza et al. (1973) observed that the voiced-voiceless (e.g., /p/-/b/) boundaries of native French learners of English were intermediate to French and English monolinguals' boundaries. Caramazza et al. concluded that the early bilinguals they examined—young adults who began learning English as children—might show additional perceptual learning over time but would probably never match English monolinguals due to the contin-

ued influence of French stops. Similar findings were obtained for early Spanish-English bilinguals by Williams (1980), who concluded that early bilinguals may develop "compromise" categories reflecting the properties of phonetically different realizations of /p t k/ in the L1 and L2.

The Barcelona studies did not examine segmental production, nor assess the perceived relation between Spanish and Catalan vowels. The studies are nonetheless important because of their clear demonstration that even experienced early learners might differ from native speakers. As mentioned, Spanish does not possess an /e/-/ɛ/ contrast. The Barcelona results might, therefore, be taken as support for the hypothesis that L2 learners filter out features not needed to distinguish L1 speech sounds following attunement to the L1 phonological system in early childhood.

Additional research will be needed, however, to assess the generalizability of the Barcelona findings. Would different results have been obtained for early learners who did not frequently hear Spanish-accented Catalan (in which Catalan /e/ and /ɛ/ are sometimes not differentiated)? Would different results have been obtained for early learners who used Spanish infrequently, or who were dominant in Catalan? It would also be useful to determine if the status of [e] and [ɛ] as allophones of Spanish /e/ impeded learning through the mechanism of acquired equivalence (Goldstone 1994).

3.3. English /ʌ/ and /ɚ/

The studies reviewed in this section examined the production and perception of English vowels by native speakers of Italian who immigrated to English-speaking communities in Canada during the 1950s and 1960s. The results presented here are re-analyses of just three of the vowels that were examined originally.[2]

3.3.1. Perception of /ʌ/

The English vowel /ʌ/ is posterior in vowel space to Italian /i e ɛ/, anterior to Italian /ɔ o u/, and slightly higher than Italian /a/. When

measured acoustically, /ʌ/ occurs in a portion of the F1-F2 space that is not occupied by a standard Italian vowel. Flege and MacKay (under review) asked Italian university students who had recently arrived in Canada to classify English vowels as instances of one of the seven vowels of Italian, and then rate each vowel for goodness as an instance of the selected Italian vowel category. Multiple natural tokens of English /ʌ/ were usually identified as being instances of Italian /a/. Similar classifications and goodness ratings were obtained for English /ɒ/ tokens, suggesting that native Italian speakers will, at least initially, have difficulty distinguishing English /ʌ/ from /ɒ/.

Two studies examined the categorial discrimination of English /ʌ/-/ɒ/. The participants examined by Flege, MacKay, and Meador (1999) and Flege and MacKay (under review) were highly experienced in English, having lived in Canada for averages of 35 and 36 years, respectively. Participants in both studies were selected on the basis of their age of arrival (AOA) in Canada. Research showing that degree of foreign accent is influenced by amount of continued L1 use (e.g., Piske, MacKay, and Flege 2001) motivated the decision to also use percentage self-reported Italian use as a selection criterion. In Flege, MacKay, and Meador (1999), participants in "early" and "late" groups (n = 18 each) had AOAs averaging 7 and 19 years, respectively. These participants reported using Italian more frequently (mean = 31%) than those in another group of early learners, "early-low" (mean = 8%). In Flege and MacKay (under review), the AOAs of early and late groups averaged 8 and 20 years, respectively. Half of both the early and late learners reported using Italian relatively seldom (mean = 8%) and half reported using Italian relatively often (mean = 48%).

A triadic test using multiple natural tokens of /ʌ/ and /ɒ/ was used to test /ʌ/-/ɒ/ discrimination. The three vowel tokens presented on each trial were always spoken by different talkers, and so differed physically. The participants' task was to choose the odd item out in change trials, and to indicate that there was no odd item out in no-change trials. The decision to include both change and no-change trials was motivated by the widely held view (e.g., Iverson et al. 2001) that phonetic category formation decreases sensitivity to within-

category differences and increases sensitivity to differences between the new category and adjacent categories. The discrimination (A') scores calculated for each vowel contrast were based on the proportion of hits (correct selections of the odd item out in change trials) and false alarms (incorrect selections of an odd item out in no-change trials).

Native English speakers in Flege, MacKay, and Meador (1999) obtained significantly higher discrimination scores than did early and late learners who used Italian often ($p < .05$), but not early learners who used Italian seldom. In Flege and MacKay (under review), the native English group obtained significantly higher /ɒ/-/ʌ/ discrimination scores than did late learners who used Italian often, late learners who used Italian seldom, and early learners who used Italian often ($p < .05$). However, the native English group did not differ significantly from early learners who used Italian seldom.

Flege and MacKay (under review) also examined native Italian students who had recently arrived in Canada. None of the Italian students obtained discrimination scores for /ɒ/-/ʌ/ that fell within 2 SDs of the mean value obtained for age-matched native English students. However, many of the 72 experienced Italian-English bilinguals obtained /ɒ/-/ʌ/ score that fell within 2 SDs of the mean scores obtained for age-matched native English speakers. Significantly more early than late learners (26 vs. 7) met the 2-SD criterion ($p < .05$). Importantly, some participants in all four native Italian groups met the 2-SD criterion (early-low 17, early-high 9, late-low 4, late-high 3).

In summary, the results reviewed here suggest that native Italian adults initially have great difficulty discriminating English /ɒ/-/ʌ/ because instances of both English vowels tend to be perceptually assimilated by a single Italian vowel (usually /a/). The problem created by single-category assimilation appears to persist for some late learners over many decades of English use. However, some late learners, and even more early learners, obtained /ɒ/-/ʌ/ discrimination scores that fell within the range of scores obtained for native English speakers. These participants may have established a new category for English /ʌ/. In both studies reviewed, early learners who used Italian of-

ten, but not early learners who used Italian seldom, discriminated /ɒ/-/ʌ/ less accurately than native English speakers did.

3.3.2. Production of /ʌ/

Flege, MacKay and Meador (1999) examined the production of ten English vowels including /ʌ/. Words containing the vowels of interest were repeated following an aural model, then presented to native English-speaking listeners for classification. The native Italian participants' /ʌ/s were identified as intended in 64% of instances. When misheard, the native Italian participants' /ʌ/s were usually classified as instances of Canadian English /ɒ/, suggesting replacement by the nearest Italian vowel, /a/. Intelligibility scores were significantly higher for /ʌ/s produced by the native English group than by the late learners ($p < .05$). However, the native English group did not differ significantly from either group of early learners (early-low, early).

Piske et al. (2002) examined productions of the same ten English vowels using a different evaluation technique. The English vowels were presented in separate blocks to native English-speaking listeners, who rated each token for goodness as an instance of its intended category. An analysis of the goodness ratings yielded the same results as the analysis of intelligibility scores reported earlier.

Flege, MacKay, and Schirru (in press) examined vowels produced by the participants whose vowel discrimination was examined by Flege and MacKay (under review). Words repeated following an aural model were rated by native English-speaking listeners for goodness as instances of the intended English vowel category. The native English speakers' /ʌ/s obtained significantly higher ratings than /ʌ/s spoken by both groups of late learners (late-low, late-high; $p < .05$), but not by either group of early learners (early-low, early-high). The ratings obtained for /ʌ/s spoken by 35 of the 72 native Italian participants fell within 2 SDs of the mean rating obtained for native English speakers' vowels. Significantly more early than late learners' vowels met the 2-SD criterion ($p < .05$). Importantly, vowels spoken by some participants in all four native Italian groups met the criterion (early-

low 15, early-high 11, late-low 3, late-high 6).

In summary, some highly experienced native Italian learners of English to mispronounce /ʌ/, a mid English vowel that does not occur contrastively in Italian. The nature of the mispronunciations suggested that the native Italian participants may have used Italian /a/ in producing English /ʌ/. Productions of /ʌ/ by late but not early learners, differed significantly from vowels spoken by native English speakers; however, some late learners produced English /ʌ/ accurately.

3.3.3 Perception of /ɚ/

English /ɚ/ is located in the middle of the F1-F2 vowel space. It is posterior to Italian /i e ɛ/, anterior to Italian /ɔ o u/, and higher than Italian /a/. The F2 value of English /ɚ/ is similar to the F2 values of back Italian vowels, but its F3 value is considerably lower than the F3 value of any Italian vowel. In Flege and MacKay (under review), Italian university students classified English /ɚ/ tokens as being an instance of a front Italian vowel (/e/ 63%, /ɛ/ 15%, or /i/ 18%). The students gave /ɚ/ tokens much lower ratings (mean = 1.5 on a 5-point scale) than any other English vowel, indicating they detected the phonetic difference between English /ɚ/ and Italian vowels. The SLM predicts, therefore, that many late learners, and most if not all early learners will eventually produce and perceive English /ɚ/ accurately.

Flege and MacKay (under review) examined the discrimination of /ɚ/-/ʌ/ by experienced native Italian speakers of English. Early and late bilinguals, regardless of amount of L1 use, obtained high scores. The Italians participants' excellent discrimination of /ɚ/-/ʌ/ was not necessarily due to the establishment of an /ɚ/ category, however. Italian students who had lived in Canada for just 3 months also obtained high discrimination scores for /ɚ/-/ʌ/, probably because they judged the /ɚ/ tokens to be poor instances of a front Italian vowel and the /ʌ/ tokens to be moderately good instances of Italian /a/. Additional research using a different testing procedure will be needed, therefore, to

assess the perception of /ɚ/ by experienced native Italian speakers of English.

3.3.4 Production of /ɚ/

Munro, Flege, and MacKay (1996) examined the production of English vowels by 240 native speakers of Italian who differed according to their AOA in Canada from Italy. The /ɚ/s spoken by most participants, even late learners, were usually heard as intended. However, AOA exerted a strong effect on the goodness ratings obtained for /ɚ/ productions. The /ɚ/s spoken by most (85%) participants with AOAs of 2-6 years obtained a rating that fell within 2 SDs of the mean rating obtained for 24 native English speakers. However, the /ɚ/s spoken by fewer participants with AOAs of 7-15 years (65%) and 15-23 years (10%) met the 2-SD criterion.

Piske et al. (2002) also obtained goodness ratings for /ɚ/s produced by experienced Italian-English bilinguals in Canada. The /ɚ/s spoken by native English speakers obtained significantly higher ratings than those spoken by early and late learners who used Italian often ($p < .05$), but not /ɚ/s spoken by early learners who used Italian seldom. In Flege, MacKay, and Schirru (in press), native English speakers' /ɚ/s obtained significantly higher ratings than /ɚ/s produced by both groups of late learners (late-low, late-high; $p < .05$), but not the /ɚ/s spoken by either group of early learners (early-low, early-high). Once again, the vowels spoken by some participants in all four native Italian groups received ratings that fell within 2 SDs of the native English groups' mean rating (early-low 17, early-high 15, late-low 12, late-high 7).

Syrdal and Gopal (1986) observed that an F3-F2 difference of less than 3 bark distinguishes the rhotic English vowel /ɚ/ from other English vowels. One would not expect native Italian learners of English to develop accurate perceptual representations for English /ɚ/ if they filtered out the rhotic property of /ɚ/, which is not used to distinguish Italian vowels. Flege and MacKay (under review) did not obtain the similarity scaling data needed to evaluate native Italian

speakers' underlying perceptual space for /ɚ/ (see Iverson et al. 2001). Inferences concerning the participants' sensitivity to the rhotic property of English /ɚ/ was, therefore, drawn from acoustic analyses of their production of this vowel.

The /ɚ/s spoken by participants in Flege and MacKay (under review) were measured acoustically along with /ɚ/s spoken by Italian students (9 male, 9 female) living in Padua, Italy who did not speak English well or often. Vowels spoken by the 108 participants (6 groups x 18) were digitized at 22 kHz, then measured using the Kay Elemetrics Multi-Speech program. The frequencies of F2 and F3 were measured at the acoustic midpoint of each vowel token using linear predictive coding analysis (covariance method, 14 coefficients). The frequency values were converted from Hertz to bark values because there were slightly unequal numbers of males and females in each group. A F3-F2 bark difference score was then computed for each /ɚ/ token to estimate its degree of rhoticity.

The values obtained for /ɚ/s produced by the Italian students living in Italy exceeded 3 bark, so their vowels would not be classified as rhotic according to the Syrdal and Gopal (1986) criterion. However, all four groups of Italian-English bilinguals in Canada produced /ɚ/ with F3-F2 bark difference scores averaging less than 3 bark. An ANOVA revealed that the rhotic scores obtained for all four native Italian groups were significantly smaller than the scores obtained for the Italian students in Italy ($p < .05$). Another analysis revealed that the native English speakers' /ɚ/s had significantly smaller (and thus more rhotic) F2-F3 bark difference scores than the /ɚ/s produced by both groups of late learners (late-low, late-high; $p < .05$) but not by either group of early learners (early-low, early-high). More /ɚ/s produced by early than late learners had rhotic scores the fell within 2 SDs of the mean value obtained for native English speakers ($p < .05$). However, /ɚ/s produced by some participants in all four groups of Italian-English bilinguals met the 2-SD criterion (early-low 17, early-high 15, late-low 9, late-high 7).

In summary, English /ɚ/ is unlike any Italian vowel, both in terms of its position in an F1-F2 space and its low F3 frequency. A proposal offered by Best and Strange (1992) leads to the expectation that a

large phonetic distance between an L2 vowel and the closest L2 will impede establishment of a new phonetic category for the L2 vowel. However, the results reviewed here suggest that some native Italian learners of English establish a category for /ɚ/. The /ɚ/s spoken by most bilinguals were heard as intended. An analysis of goodness ratings revealed that /ɚ/s spoken by late but not early learner groups differed significantly from native English speakers' productions. However, some late bilinguals were found to have produced /ɚ/ accurately.

The results for /ɚ/ supported the SLM hypothesis that category formation remains possible across the life span, but will be less likely for late than early learners. Acoustic measurements of the rhotic property of English /ɚ/ were carried out to evaluate the hypothesis (Iverson et al. 2001) that Italian learners of English will develop erroneous categories for English /ɚ/. The results, although only preliminary, suggested that erroneous categories are not developed, and that L2 learners do not necessarily filter out properties of L2 speech sounds if those properties are not used to distinguish L1 speech sounds.

4. The relation between production and perception

This section will briefly review research relevant to the SLM hypothesis that L2 phonetic segments can be produced only as accurately as they are perceived. It will focus on the production and perception of English vowels by native speakers of Italian. A more thorough discussion is provided by Flege (1999b).

Flege, MacKay, and Meador (1999) tested the SLM hypothesis by examining the relation between vowel intelligibility and discrimination. Three scores were computed for each of 72 native Italian participants: the average discrimination of four pairs of English vowels, the average discrimination of four pairs of Italian vowels, and the average intelligiblity of ten English vowels. There was a significantly stronger correlation between English vowel production and English vowel discrimination than between English vowel production and the

discrimination of Italian vowels ($p < .01$). This established a link between the production and perception of language-specific phonetic segments, not just a general tendency to produce and perceive accurately or poorly.

As mentioned earlier, English /ʌ/ tokens tend to be identified as instances of the Italian /a/ category. A second analysis examined the discrimination of English /ʌ/ and Italian /a/. The native Italian participants were assigned to subgroups based on how accurately they had produced English /ʌ/. The participants with a relatively good pronunciation of /ʌ/ ($n = 41$) obtained significantly higher /ʌ/-/a/ discrimination scores than those ($n = 31$) with a poorer pronunciation of /ʌ/ ($p < .01$). Crucially, the two subgroups' discrimination of other vowels did not differ significantly.

Flege and MacKay (under review) also assessed the relation of vowel production and perception. The discrimination scores obtained by Italian-English bilinguals for /ɒ/-/ʌ/, /i/-/ɪ/, and /ɛ/-/æ/ were averaged, as were the ratings accorded each participant's production of /ɒ ʌ i ɪ ɛ æ/. The average discrimination scores and ratings of vowel production were then standardized. The two sets of z-scores showed a moderate positive correlation (r (70) = .68, $p < .01$). Significant correlations between segmental production and perception were also obtained when the early and late learners were considered separately (r (34) = .55, and .45, $p < .01$).

In summary, the results summarized here suggest that moderate positive correlations exist between the production and perception of L2 phonetic segments by experienced L2 learners. Bever (1981) hypothesized that a critical period for learning L2 speech arises from the loss of ability to align segmental production to segmental perception. The results do not support a strong version of this hypothesis. The finding that some L2 learners showed more accurate perception than production is consistent with the hypothesis that L2 segmental perception may "lead" (i.e., be more advanced than) segmental production. It also agrees with the results of laboratory training studies showing that gains derived from perceptual training may transfer to improved segmental production in the absence of production training (Rochet 1995; Bradlow et al. 1997, 1999).

5. Discussion

A number of hypotheses have been offered regarding possible constraints on L2 speech learning. Broadly speaking, two hypotheses have been offered regarding production. One is that the ability to learn to produce L2 phonetic segments not found in the L1 diminishes following a critical period. The other hypothesis is that L2 phonetic segments cannot be produced accurately unless they are perceived accurately.

Several proposals have been offered, in turn, for constraints on segmental perception. One is that features (properties) not needed to distinguish L1 speech sounds get filtered out, and so cannot be used in the development of new phonetic categories. Another hypothesis is that the auditory input associated with L2 speech sounds may be warped as the result of previous learning. Proponents of these perceptual hypotheses have suggested that the filtering/warping of L2 phonetic input may become stronger as the age of first exposure to an L2 increases. The SLM, on the other hand, proposes that the capacity to accurately perceive the phonetic properties of L2 speech sounds and to establish new categories based on those properties remains intact across the life span (although L2 category formation becomes less likely as L1 categories develop).

The evidence reviewed here provided support for each theoretical position. Iverson et al. (2001) showed that native Japanese adults weight acoustic properties that distinguish English /ɹ/ from /l/ differently than native English adults do. This supported the hypothesis that previous speech learning may lead to a warping of L2 speech input. The results obtained for early native Spanish learners of Catalan (e.g., Sebastián-Gallés and Soto-Faraco 1999) can be taken as support for the hypothesis that L2 learners filter out properties (or features) of L2 speech sounds that are not needed to distinguish L1 speech sounds. The results obtained for native Spanish and Italian learners of English supported SLM hypotheses. These studies revealed that although most late learners produced and perceived English vowels less accu-

rately than early learners did, some succeeded in learning English vowels not found in their L1. Also, acoustic analyses suggested that neither early nor late Italian-English bilinguals filtered out the rhotic property of English /ɚ/ even though this property is not used to distinguish Italian vowels.

Additional research will be needed to determine whether persistent native versus non-native differences arise from neurological maturation, from the influence of previous phonetic learning, or both. An optimal study would be one that employed a longitudinal design and tested for changes over time in the perceived relation of L1 and L2 sounds, as well as changes in the production and perception of both L2 and L1 phonetic segments.

Such a study might be carried out to examine native English and Japanese adults' perception of the acoustic properties distinguishing English /ɹ/ from /l/. Iverson et al. (2001) observed difference in native and non-native participants' perception of properties in /ɹ/ and /l/. A question of theoretical and practical importance is whether such differences will persist indefinitely—and thus represents a true constraint on learning—or whether they will disappear following a certain amount of native-speaker input. Flege, Takagi, and Mann (1995, 1996) observed poor segmental production and perception of English /ɹ/ and /l/ by native Japanese adults who were inexperienced in English, but more accurate production and perception by those who were highly experienced in English. It would be useful to determine if Japanese adults who are highly experienced in English will show native-like perception at the feature (property) level and, if so, how much native-speaker input is needed for this kind of change in early stages of perceptual processing to occur.

It will be necessary to study a wide range of L1-L2 pairs and L2 speech sounds in order to draw general conclusions regarding the nature of constraints, if any, on L2 speech learning. Previous research suggests that the magnitude of differences between L2 learners and monolingual native speakers of the target L2 will depend, at least in part, on the degree of perceived phonetic dissimilarity of L2 sounds from the closest L1 sound. The magnitude of differences might also depend on allophonic status. For example, it might be more difficult

to learn the L2 phonemes /x/ and /y/ if they resemble the [x] and [y] allophones of a single L1 phoneme than if the L2 /x/ and /y/ phonemes resemble the primary allophones of two different L1 phonemes.

Grosjean (1999) noted that some published studies have provided little or no information about research participants, making cross-study comparison and replication difficult. He also suggested that the contexts in which languages are learned and used might influence performance in both the L1 and the L2. Care must be taken, therefore, in selecting participants for L2 research. For example, a number of factors relevant to L2 learning are typically confounded with age of L2 learning, at least in immigrant populations in North America. When the confounded variables are controlled, the effect of age of L2 learning may disappear for certain outcome measures (e.g., Flege, Yeni-Komshian, and Liu 1999).

It is especially important to consider L2 input and language use patterns when selecting participants for L2 research. It seems reasonable to think, for example, that learning an L2 /x/-/y/ contrast will be impeded if learners frequently hear fellow speakers of their L1 neutralize the L2 /x/-/y/ contrast. Some studies reviewed in Section 3 revealed differences between early learners and L2 native speakers. This might be taken as evidence that L2 speech learning is irreversibly constrained by previous learning, through a filtering or warping of L2 input, through a loss of neural plasticity, or both. However, other studies reviewed in Section 3 suggested that differences between native speakers and early learners may be confined to early learners who continue to use their L1 often.

Another factor to consider when selecting research participants is language dominance. Piske, MacKay, and Flege (2001) examined overall degree of foreign accent in English sentences produced by early learners. As in previous research, both of two groups of early learners were found to have detectable foreign accents, although those who used their L1 (Italian) often had significantly stronger accents than those who used their L1 seldom. Importantly, some members of both early learner groups received ratings that fell within the range of ratings obtained for native English speakers' sentences.

Flege, MacKay, and Piske (2002) re-examined the foreign accent ratings obtained for the participants examined by Piske et al. after assigning them to subgroups based on language dominance. Sentences produced by Italian-dominant early learners, but not by English-dominant early learners, received significantly lower ratings than the native English speakers' sentences did. This suggested that individuals who become dominant in their L2 might not show measurable L1 interference effects. If so, and if this finding can be shown to generalize to L2 segmental production and perception, one might conclude that L1 interference effects are not inevitable, and thus that L2 speech learning is not irreversibly constrained.

6. Acknowledgments

This chapter was supported by a grant from the National Institute of Deafness and Other Communicative Disorders (DC00257). The author thanks Katsura Aoyama, Satomi Imai, and Kimiko Tsukada for comments on an earlier version, and Ian R. A. MacKay for his long-term collaboration.

7. References

Baker, Wendy, Pavel Trofimovich, Molly Mack and James Flege
2002 The effect of perceived phonetic similarity on non-native sound learning by children and adults. In: Barbora Skarabela, Sarah Fish and Anna Do (eds.) *Proceedings of the 26th annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.

Best, Catherine
1995 A direct realistic perspective on cross-language speech perception. In: Winifred Strange (ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research*, 171-206. Timonium, MD: York Press.

Best, Catherine, Gerald McRoberts and Elizabeth Goodell
2001 Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America* 109: 775-794.

Best, Catherine and Winifred Strange
1992 Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics* 20: 305-330.

Bever, Thomas
1981 Normal acquisition processes explain the critical period for language learning. In: Karl Diller (ed.), *Individual Differences and Universal in Language Learning Aptitude*, 176-198. Rowley, MA: Newbury House.

Bosch, Laura, Albert Costa and Núria Sebastián-Gallés
2000 First and second language vowel perception in early bilinguals. *European Journal of Cognitive Psychology* 12: 189-221.

Bradlow, Ann
1995 A comparative acoustic study of English and Spanish vowels. *Journal of the Acoustical Society of America* 97: 1916-1924.

Bradlow, Ann, David Pisoni, Reiko Yamada and Yoh′ichi Tohkura
1997 Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America* 101: 2299-2310.

Bradlow, Ann, Reiko Yamada, David Pisoni and Yoh′ichi Tohkura
1999 Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production. *Perception and Psychophysics* 61: 977-985.

Caramazza, Alfonso, Grace Yeni-Komshian, Edgar Zurif and Evelyn Carbone
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: 421-428.

Churchland, Paul
1986 *Neurophilosophy*. Cambridge, MA: MIT Press.

de Boer, Bart
2000 Self-organization in vowel systems. *Journal of Phonetics* 28: 441-465

DeKeyser, Robert
2000 The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition* 22: 499-534.

Edelman, Gerald
1989 *The Remembered Present: A Biological Theory of Consciousness*. New York: Basic Books.

Flege, James
1984 The detection of French accent by American listeners. *Journal of the Acoustical Society of America* 76: 692-707.

Flege, James
1987 The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Jour-*

Journal of Phonetics 15: 47-65.

Flege, James
1988 The production and perception of speech sounds in a foreign languages. In: Harris Winitz (ed.) *Human Communication and Its Disorders, A Review 1988*, 224-401. Norwood, NJ: Ablex.

Flege, James
1989 Differences in inventory size affects the location but not the precision of tongue positioning in vowel production. *Language and Speech* 32: 123-147.

Flege, James
1991a 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, James
1991b Orthographic evidence for the perceptual identification of vowels in Spanish and English. *Quarterly Journal of Experimental Psychology* 43: 701-731.

Flege, James
1992 Speech learning in a second language. In: Charles Ferguson, Lise Menn and Carol Stoel-Gammon (eds.), *Phonological Development: Models, Research, and Implications*, 565-604. Timonium, MD: York Press.

Flege, James
1995 Second-language speech learning: Theory, findings, and problems. In: Winifred Strange (ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research*, 229-273. Timonium, MD: York Press.

Flege, James
1999a Age of learning and second-language speech. In: David Birdsong (ed.) *New Perspectives on the Critical Period Hypothesis for Second Language Acquisition*, 101-132. Hillsdale, NJ: Lawrence Erlbaum.

Flege, James
1999b The relation between L2 production and perception. In: John Ohala, Yoko Hasegawa, Manjari Ohala, Daniel Granville and Ashlee Bailey (eds.) *Proceedings of the XIVth International Congress of Phonetics Sciences*, 1273-1276. Berkeley, CA: Department of Linguistics.

Flege, James
2002 Interactions between the Native and Second-language Phonetic Systems. In: Petra Burmeister, Thorsten Piske and Andreas Rohde (eds.), *An Integrated View of Language Development: Papers in Honor of Henning Wode*, 217-244. Trier: Wissen-

schaftlicher Verlag.
Flege, James, Ocke-Schwen Bohn and Sunyoung Jang
1997 The effect of experience on nonnative subjects' production and perception of English vowels. *Journal of Phonetics* 25: 437-470.
Flege, James and Wieke Eefting
1987 Production and perception of English stops by native Spanish speakers. *Journal of Phonetics* 15: 67-83.
Flege, James and Serena Liu
2001 The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition* 23: 527-552.
Flege, James and Ian MacKay
Under review Perceiving vowels in a second language.
Flege, James, Ian MacKay and Diane Meador
1999 Native Italian speakers' production and perception of English vowels. *Journal of the Acoustical Society of America* 106: 2973-2987.
Flege, James, Ian MacKay and Thorsten Piske
2002 Assessing bilingual dominance. *Applied Psycholinguistics* 23: 567-598.
Flege, James, Ian MacKay and Carlo Schirru
In press Interaction between the native and second language phonetic subsystems. *Speech Communication*.
Flege, James, Murray Munro and Robert Fox
1994 Auditory and categorical effects on cross-language vowel perception. *Journal of the Acoustical Society of America* 95: 3623-3641
Flege, James and Robert Port
1981 Cross-language phonetic interference: Arabic to English. *Language and Speech* 24: 125-146.
Flege, James, Naoyuki Takagi and Virginia Mann
1995 Japanese adults can learn to produce English /ɹ/ and /l/ accurately. *Language and Speech* 38: 25-55.
Flege, James, Naoyuki Takagi and Virginia Mann
1996 Lexical familiarity and English-language experience affect Japanese adults' perception of /ɹ/ and /l/. *Journal of the Acoustical Society of America* 99: 1161-1173.
Flege, James, Grace Yeni-Komshian and Serena Liu
1999 Age constraints on second language learning. *Journal of Memory and Language* 41: 78-104.
Francis, Alexander and Howard Nusbaum
2002 Selective attention and the acquisition of new phonetic categories. *Journal of Experimental Psychology: Human Perception and Performance* 28: 349-366.
Goldstone, Robert

1994 Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology: General* 123: 178-200.

Gottfied, Terry and Patrice Beddor

1988 Perception of temporal and spectral information in French vowels. *Language and Speech* 31: 57-75.

Grosjean, François

1999 Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition* 1: 117-130.

Hazan, Valerie and Sarah Barrett

1999 The development of phoneme categorisation in children aged 6 to 12 years. In: John Ohala, Yoko Hasegawa, Manjari Ohala, Daniel Granville and Ashlee Bailey (eds.), *Proceedings of the XIVth International Congress of Phonetics Sciences*, 2493-2496. Berkeley, CA: Department of Linguistics.

Iverson, Paul, Patricia Kuhl, Reiko Yamada, Eugen Diesch, Yoh′ichi Tohkura, Andreas Ketterman and Claudia Siebert.

2001 A perceptual interference account of acquisition difficulties for non-native phonemes. *Speech, Hearing and Language: Work in Progress* 13: 106-118.

Johnson, Carole

2000 Children's phoneme identification in reverberation and noise. *Journal of Speech, Language, and Hearing Research* 43: 129-143.

Kuhl, Patricia

2000 A new view of language acquisition. *Proceedings of the National Academy of Science* 97: 11850-11857.

Kuhl, Patricia and Andrew Meltzoff

1996 Infant vocalizations in response to speech: Vocal imitation and developmental change. *Journal of the Acoustical Society of America* 100: 2425-2438.

Kuhl, Patricia, Karen Williams, Francisco Lacerda, Kenneth Stevens and Bjorn Lindblom

1992 Linguistic experience alters phonetic perception in infants six months of age. *Science* 255: 606-608.

Leather, Jonathan and Allan James

1996 The acquisition of second language speech. In: William Ritchie and Tej Bhatia (eds.), *Handbook of Second Language Acquisition*, 269-316. San Diego: Academic Press.

Long, Michael

1990 Maturational constraints on language development. *Studies in Second Language Acquisition* 12: 251-285.

McAllister, Robert, James Flege and Thorsten Piske

2002 The influence of the L1 on the acquisition of Swedish vowel quantity by native speakers of Spanish, English and Estonian.

Journal of Phonetics 30: 229-258.

McLaughlin, Barry
1977 Second-language learning in children. *Psychological Bulletin* 84: 438-459.

Michaels, David
1974 Some replacements and phonological systems. *Linguistics* 126: 69-81.

Morrison, Geoffrey
2002 *Effects of L1 duration experience on Japanese and Spanish listeners' perception of English high front vowels*. Unpublished MA thesis, Simon Fraser University, Vancouver, BC.

Munro, Murray
1993 Productions of English vowels by native speakers of Arabic: Acoustic measurements and accentedness ratings. *Language and Speech* 36: 39-66.

Munro, Murray, James Flege and Ian MacKay
1996 The effect of age of second-language learning on the production of English vowels. *Applied Psycholinguistics* 17: 313-334.

Pallier, Christophe, Laura Bosch and Núria Sebastián-Gallés
1997 A limit on behavioral plasticity in speech perception. *Cognition* 64: B9-B17.

Pallier, Christophe, Àngels Colomé and Núria Sebastián-Gallés
2001 The influence of non-native phonology on lexical access. Exemplar-based versus abstract lexical entries. *Psychological Science* 12: 445-449.

Paradis, Michel
1993 Linguistic, psycholinguistic, and neurolinguistic aspects of the "interference" in bilingual speakers: The Activation Threshold Hypothesis. *International Journal of Psycholinguistics* 9: 133-145.

Patkowski, Mark
1989 Age and accent in a second language: A reply to James Emil Flege. *Applied Linguistics* 11: 73-89.

Piske, Thorsten, James Flege, Ian MacKay, and Diane Meador
2002 The production of English vowels by fluent early and late Italian-English bilinguals. *Phonetica* 59: 49-71.

Piske, Thorsten, Ian MacKay and James Flege
2001 Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics* 29: 191-215.

Pisoni, David
1995 Some thoughts on "normalization" in speech perception. *Research on Spoken Language Processing, Progress Report* 20: 3-30 (Department of Psychology, Indiana University).

Rochet, Bernard
1995 Perception and production of L2 speech sounds by adults. In: Winifred Strange (ed.) *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues*, 379-410. Timonium, MD: York Press.

Scovel, Thomas
1988 *A Time to Speak. A Psycholinguistic Inquiry into the Critical Period for Human Speech*. Cambridge, MA: Newbury House.

Sebastián-Gallés, Núria and Salvador Soto-Faraco
1999 On-line processing of native and non-native phonemic contrasts in early bilinguals. *Cognition* 72: 111-123.

Syrdal, Ann and H. S. Gopal
1986 A perceptual model of vowel recognition based on the auditory representation of American English vowels. *Journal of the Acoustical Society of America* 79: 1086-1100.

Trubetzkoy, Nikolaj.
1939/1958 Grundzüge der Phonologie. Göttingen: Vandenhoek & Ruprecht.

Walley, Amanda and James Flege
2000 Effects of lexical status on children's and adults' perception of native and non-native vowels. *Journal of Phonetics* 27: 307-332.

Wayland, Ratree, James Flege, and Satomi Imai
Under review The discrimination of English vowels by native speakers of Spanish.

Williams, Lee
1980 Phonetic variation as a function of second-language learning. In: Grace Yeni-Komshian, James Kavanagh and Charles Ferguson (eds.) *Child Phonology, Volume 2 Perception*, 185-216. New York: Academic Press.

---

[1] Sebastián-Gallés & Soto-Faraco (1999) also examined Catalan /o/-/ɔ/, /s/-/z/ and /ʃ/-/ʒ/, and Pallier et al. (2001) examined /o/-/ɔ/, /s/-/z/, and /m/-/n/.

[2] All *p*-values were adjusted to account for the number of between-group comparisons that were made. A full description of the analyses presented in this section is available upon request.