

# Second Language Phonology and Speech

Ellen Broselow, State University of New York at Stony Brook

Yoonjung Kang, University of Toronto Scarborough

## 1. Introduction

The acquisition of a second language (L2) sound system poses significant challenges for learners, who must acquire a new system of sound contrasts, new restrictions on where sounds may occur, and a new prosodic system. The challenges facing researchers are to understand the characteristics of L2 speech and to explain how and why those characteristics arise. In this chapter, we focus on three major issues that guide research on the acquisition of L2 sound systems; (i) the influence of L1 and of linguistic universals in L2 speech sound patterns; (ii) the level of representation (phonological vs. phonetic) at which L2 acquisition occurs; and (iii) the relationship between the perception and the production of the second language.

### Transfer and Universals

Almost all studies on the acquisition of L2 sound systems address the question of how and to what extent L1 influence (“transfer”) contributes to shaping L2 learners’ sound patterns (cf. Major 2008). Lado’s (1957) *Contrastive Analysis Hypothesis* (CAH) predicted that those aspects of L2 that are similar to the L1 will be easily acquired, while those aspects that are different from the L1 will be difficult. While L1 transfer is undeniably a major factor in L2 acquisition, two recurring phenomena pose problems for the claim that L2 errors are based *solely* on the difference between the L1 and the L2. First, certain L2 structures not present in L1 appear to be more easily acquired than others, regardless of the particulars of the L1 and the L2 involved. Second, the speech of L2 learners in many cases exhibits patterns that are coherent and systematic, but are nonetheless distinct from those of both the target L2 and the native language. Phenomena of both types have frequently been attributed to universal principles such as “markedness,” where “unmarked” structures are generally considered to be those that are more basic, typologically common, and phonetically easier to perceive and produce than more marked structures (cf. Kager 1999; Rice 2007). Eckman (1977) proposed that the *Contrastive Analysis Hypothesis* must be supplemented by the *Markedness Differential Hypothesis*, which states that marked L2 structures are more difficult to acquire than equally novel but less marked structures (cf. Eckman 2008). Similarly, interlanguage patterns that appear to be rooted in neither the L1 nor the L2 can be analyzed as resulting from universal

preferences for unmarked structures or patterns.

Markedness has been encoded in theories of grammar (the implicit knowledge that makes it possible to hear and understand a language) in various ways. In a theory that views a grammar as a set of parameters with different possible settings (for example, a parameter allowing consonant clusters in syllable onsets, which may be set either on or off), the preference for less marked structures can be viewed as encoded as default parameter settings (e.g., Hayes 1995). More recently, within the framework of *Optimality Theory* (OT; Prince and Smolensky 2004), markedness is encoded in the form of universal constraints that favor less marked structures. These constraints are assumed to be present in the grammars of all languages, but may be rendered inactive by more highly ranked constraints; for example, a constraint banning consonant clusters in onsets may be outranked by a constraint demanding that lexical forms be realized faithfully. Exposure to L2 data may trigger resetting of parameters (e.g., Archibald 1993, 1994, 1995) or rerankings of constraints that may allow formerly dormant markedness constraints to become active (e.g., Hancin-Bhatt 2000, 2008; Broselow 2004; Eckman 2004).

As the cases discussed in this chapter illustrate, teasing apart L1 transfer effects from the effects of universal tendencies is not trivial. Complicating this issue is the fact that a number of researchers have argued that similarity to preexisting L1 structure may actually have a negative effect on the accurate perception and ultimate attainment of some aspects of L2 structure, particularly the acquisition of new phoneme contrasts. A number of models recognize a tendency for L2 sounds to be interpreted in terms of L1 sounds, giving rise to problems when, for example, two contrasting sounds in the L2 map onto a single L2 category. The tendency to map L2 sounds to similar but not necessarily identical L1 phoneme categories is recognized in the *Speech Learning Model* (SLM, Flege 1995), where it is termed “equivalence classification”; in the *Perceptual Assimilation Model* (PAM, Best 1995); and in the *Native Language Magnet Theory* (Kuhl and Iverson 1995), each of which outlines the ways in which this process may interfere with the accurate acquisition of L2 sounds, though the degree of L1 interference may be modulated by factors such as degree of experience with the L2, age of L2 exposure, speech rate, attention, and language mode in experimental conditions (Flege 1991; Piske et al. 2001; Ioup 2008; Antoniou et al. 2010, among others). Finally, Major and Kim (1996) propose that while L2 sounds that are dissimilar to L1 sounds are initially difficult, performance on these improves quickly, while performance on L2 sounds that are similar to L1 sounds stays the same or progressively worsens (the *Similarity Differential Rate Hypothesis*).

### **Phonology vs. Phonetics**

A second major issue in the acquisition of L2 sound systems concerns whether the characteristics of L2 sound patterns are best explained from the perspective of phonology (an internalized system of abstract rules or constraints defining the possible sound patterns of a

language) or of phonetics (the physical implementation of speech sounds in production and the interpretation of fine-grained acoustic cues in perception). Various accounts have assumed that the reason learners fail to produce L2 structures correctly is because their phonological grammar does not sanction them, causing L2 forms to be modified by the L1 grammar. In this approach, L2 acquisition involves moving from the starting point (the L1 grammar) toward an L2-like grammar in response to L2 input, which may trigger (depending on one's model of grammar) resetting of parameters; changes in phonological rules (Young-Scholten 2004, Eckman et al. 2003); or, in Optimality Theory, changes in the ranking of constraints (e.g., Broselow 2004; Escudero and Boersma 2004; Hancin-Bhatt 2008). However, evidence has emerged supporting the position that some if not all the modification of L2 forms may take place not in phonology but rather at the level of phonetics, where speakers may simply fall short in mastering the correct articulation (Colantoni and Steele 2007; Davidson 2010). (The question of whether modification may also reflect misperception is addressed in the next section.)

A related issue concerns the units of representation involved in L2 acquisition (Kang 2008)—that is, whether the L2 is perceived and produced in terms of abstract phonological concepts such as phonemes or phonological distinctive features (e.g., Hancin-Bhatt 1994; Weinberger 1997; Brown 2000; Larson-Hall 2004, among others); in terms of surface phonetic features (Brannen 2002; de Jong, Hao, and Park 2009; de Jong, Silbert, and Park 2009); or in terms of more phonetically-based notions such as articulatory gestures (Best and Halle 2010).

Researchers beginning from the perspective of phonology have proposed that the attainability of L2 structures is a function of the availability of corresponding phonological structures in the L1 (Brown 2000; Archibald 2005; Goad and White 2006). Distinctive features (which serve both to distinguish contrasting sounds of the language and to organize sounds into natural classes sharing a particular feature) have been appealed to for explanations of cross-linguistic differences in L2 perception as well as in phoneme substitution. For example, to explain the finding that speakers of Mandarin Chinese were better at discriminating English /l/ and /ɹ/ than speakers of Japanese, though both L1s lack the phonemic contrast of /l/ and /ɹ/, Brown (2000) argued that the difficulty of acquiring a new L2 phoneme contrast depends on whether the L1 employs a phonological feature to encode that contrast. She explained the Chinese speakers' relatively good perception by proposing that the Chinese contrast between alveolar and retroflex fricatives (e.g., /s/ vs. /ʂ/) requires a feature to encode subcoronal place contrasts, which can then be extended to the English contrast. The fact that no subcoronal contrast (and consequently no appropriate feature) exists in Japanese makes this contrast difficult for Japanese speakers. Similarly, Hancin-Bhatt (1994) argued that phonological features can explain the choice of L1 phoneme to replace an illegal L2 sound (for example, why English /θ/ is typically replaced with /t/ by speakers of Turkish but with /s/ by speakers of Japanese, even

though /t/ and /s/ are present in both L1 phoneme inventories). Investigating the differential substitution patterns for English interdental fricatives by speakers of German, Hindi, Japanese, and Turkish, she proposed that the likeliest L1 substitute is one that preserves those features of the L2 sound which carry the highest functional load in the L1 (that is, the features that encode the largest number of contrasts in the L1 phoneme inventory).

However, other studies have suggested that at least some L2 patterns must be described in terms of surface phonetic features rather than contrastive phonological features. According to Brannen (2002), the English interdental fricative /θ/ is replaced by /s/ in European French and by /t/ in Quebec French, although the two dialects have identical phoneme systems and consequently identical phonological feature inventories. The crucial difference between the two dialects therefore rests not in their phonological feature inventories but rather in their articulatory patterns: in European French, the coronal fricative /s/ has a dental place of articulation, while in Quebec French, /s/ is alveolar, a phonetic detail that is not considered to be contrastive in either dialect. Based on these findings, Brannen (2002) argues that L2 substitution may be sensitive to phonetically salient features, regardless of the contrastive status of the features.

Evidence suggesting that both phonological features and the mastery of articulatory routines play a role in L2 acquisition comes from comparing the rate at which contrasts defined by the same phonological feature are acquired. De Jong, Silbert, and Park (2009) and de Jong, Hao, and Park (2009) found that Korean learners' ability to perceive the English stop-fricative contrast both for labial /b-v/ and for coronal /d-ð/ proceeded similarly, suggesting that the learning of one contrast defined by the feature [continuant] generalized to continuant contrasts across all places of articulation. However, this effect in L2 perception was not accompanied by the same effect in the Korean speakers' production of English, where the accuracy of the continuancy contrasts in labial and coronal place did not correlate, presumably because implementation of /b-v/ involves articulatory gestures that differ from those involved in /d-ð/. There was, however, a correlation between the production accuracy of learners' stop-fricative contrasts in voiced labials /b-v/ and voiceless labials /p-f/, since in this case the gestures for producing the manner contrast are similar across the voiced and voiceless pairs.

Just as abstract features are realized differently in different phonemes, individual phonemes are realized differently in different contexts. Flege (1995) proposed that segmental acquisition takes place on the basis of contextual allophones, at a level more abstract than surface phones but less abstract than phonemes. This view is supported by numerous studies showing that a given L2 segment is not perceived or produced with uniform accuracy across different segmental contexts; rather, the same novel segment is acquired earlier in certain positions (e.g., onset, word-initial, etc.) than in others (e.g., coda, word-final, etc.). Some attribute such

positional variation to phonological markedness, while others seek explanations based on perceptual similarity, articulatory ease, and input frequency. For example, Trofimovich et al. (2007) showed that the production of English /ð/ by French speakers was more accurate in some positions (e.g., sentence-initial) than in others, and concluded that the target sound is more easily learned in positions where it was perceived as phonetically distinct from L1 sounds, as well as in positions where it more frequently occurred. Clearly, the problem of separating phonological from phonetic explanations is not trivial (see e.g., Archibald 2009), and a full account of L2 speech will require detailed analysis of both the phonology and the phonetics of the L1 and L2 systems.

### **Perception and Production**

A third area of longstanding interest in L2 phonological acquisition is the extent to which misproduction reflects misperception and the question of whether L2 perception and production develop in tandem. While it is generally assumed that in children's native language acquisition, accurate perception precedes accurate production (cf. Smolensky 1996), many L2 researchers, most notably Flege (1995), have suggested that many of the difficulties in L2 production stem from inaccurate perception of L2 targets. As in so many other areas, the literature provides conflicting evidence. For example, Hayes-Harb and Masuda (2008) found that some English speakers learning Japanese were successful at perceptually discriminating the singleton-geminate consonant contrast in Japanese, but were unable to implement this contrast in production. They suggest that this may not reflect a problem in production *per se*; rather, the English learners were able to notice the presence of novel contrasts as familiar vs. unfamiliar, but had not established an accurate lexical representation of the short-long contrast, which in turn led to failure to produce the contrast. Conversely, Weber and Cutler (2004) and Cutler et al. (2006) suggest that lexical encoding of a contrast does not necessarily lead to auditory discrimination. Using eye-tracking technology to examine the auditory processing of L2 speech, Cutler et al. (2006) found that when Japanese learners of L2 English presented with four different pictures heard English *rock*, they looked at both the picture of a rock and the picture of its minimal pair competitor, a lock. However, the confusion was asymmetrical; when the participants heard *lock*, they converged on the lock picture. This asymmetry suggests that the learners had encoded the /l/ vs. /ɹ/ contrast lexically but could not perceive the contrast in the on-line auditory word identification task.

In addition to cases showing production lagging perception, the L2 literature reveals cases showing the reverse; for example, both Goto (1971) and Sheldon and Strange (1982) found that some Japanese learners of English were more successful in producing the English /l/ vs. /ɹ/ contrast than in discriminating these sounds. Similarly, Kijak (2009) found, in a study discussed below, that learners of various L1 backgrounds generally performed better on production than

on perception of Polish stress.

In the remainder of this chapter, we will survey results of L2 acquisition research in the context of the questions outlined above. Section 2 will discuss L2 segmental acquisition—the perception and production of consonants and vowels; section 3 will discuss phonotactics—restriction on possible combination of segments; and section 4 will discuss the acquisition of prosody—stress, tone, pitch accent and intonation.

## 2. Segmental acquisition

The vastness of the literature on L2 segmental acquisition makes it impossible to provide a comprehensive review, so we will limit our discussion to the major research results on the production and perception of three groups of sounds—stops, vowels, and liquids.

### Voicing contrasts in stop consonants

Languages differ in their realization of the stop voicing contrast along the dimension of Voice Onset Time, or VOT (the time between the release of stop constriction and the onset of voicing on the following vowel). Even languages that share a two-way laryngeal contrast may implement this contrast differently. For example, English speakers produce /p t k/ with a *long lag* in VOT, resulting in aspiration, and /b d g/ with a *short lag*, so that voicing begins simultaneous with release of the stop into the following vowel. In Spanish, French, Dutch, Greek, and Portuguese, however, /p t k/ have a short lag (are unaspirated), while for /b d g/ voicing begins during the stop closure (i.e., a negative VOT). Many studies have found that when the L1 and L2 differ in phonetic realizations of voicing contrasts, L2 stops tend to show “compromise” VOT values intermediate between the L1 and the L2 stops (Flege 1991; Hazan and Boulakia 1993; Sancier and Fowler 1997; Kang and Guion 2006; Fowler et al. 2008; Antoniou et al. 2010, among others; see Zampini 2008 for a recent review). For example, Flege (1987) found that the mean VOT duration of French /t/ for less-experienced English-speaking learners of French was similar to the mean value for English /t/ produced by monolingual English speakers. For more experienced learners, however, the VOT was intermediate between the French norm and the English norm. These and similar findings are consistent with the claims discussed above that learners tend to map L2 sounds to similar L1 phoneme categories, even when the L1 and L2 categories may differ in phonetic detail (Flege 1995; Best 1995; Kuhl and Iverson 1995).

Another common finding is that L1 phonetic implementation patterns are not necessarily uniformly transferred. Learners whose L1 contrasts long lag unaspirated stops with prevoiced stops, such as Spanish, French and Dutch, tend to produce English voiceless aspirated stops as target-like, i.e., with a long lag, but to produce English voiced stops as prevoiced, showing

transfer from the L1 (Williams 1977; Hazan and Boulakia 1993; Simon 2009, among others). It is puzzling that the short lag stop category is not correctly produced in L2 even though it is already available in the L1, particularly given the fact that short lag stops are usually the first to be acquired in L1 acquisition (cf. Kager et al. 2007 cited in Simon 2009). Simon (2009) proposes possible explanations for this asymmetrical transfer of L1 structure based in both phonetics and in abstract phonological representations. First, the long VOT of aspirated stops is perceptually salient and therefore more easily heard by learners. Second, if the Dutch contrast between voiceless unaspirated and prevoiced stops is represented by a single phonological feature [voice] (which is specified for /b d g/ but absent in /p t k/), the stops specified for [voice] enjoy a privileged status and are therefore more likely to be transferred.

L1 transfer plays a role in the perception of L2 voicing contrasts as well as their production. Curtin et al. (1998) examined English speakers' perception of the Thai three-way contrast of voiced-voiceless-aspirated stops. English speakers who heard an auditory stimulus and were then asked to choose a matching picture over a minimal pair competitor were better at perceiving the Thai voicing contrast (b vs. p) than the aspiration contrast (p vs. p<sup>h</sup>), presumably due to transfer of the L1 voicing contrast. However, in an auditory discrimination task requiring participants to decide which of three stimuli were the same, the aspiration contrast was more accurately perceived than the voicing contrast. Curtin et al. (1998) suggest that the picture task, which associates sound differences with meaning differences, taps into a more abstract level of processing, while the discrimination task taps simple auditory processing, where discrimination is facilitated by the perceptual salience of aspiration (see Strange and Shafer 2008 for a recent review of task effects in L2 perception). A later study (Pater 2003) found significantly better performance on the more salient aspiration contrast than on the voicing contrast, regardless of task type, suggesting a stronger role for acoustic salience. However, the association of sound differences with lexical contrasts was found to facilitate acquisition of a new contrast by Hayes-Harb (2007), who found that in an artificial language learning experiment, English speakers performed better in distinguishing pseudo words that contain voiced [g] and voiceless unaspirated [k] (which do not normally serve a contrastive function in English) when stimuli were presented with contrastive meanings ([ga] 'pot' vs. [ka] 'mouse') than when they were presented with noncontrastive meanings ([ga], [ka] 'pot') or when no meaning was attached.

### **Vowel contrasts**

In an early influential study on cross-language vowel perception, Stevens et al. (1969:1) suggested that unlike consonant perception, vowel perception "is relatively independent of ...linguistic experience". Subsequent research, however, has revealed that both the perception and the production of L2 vowels is strongly influenced by listeners' L1, with abundant evidence that non-native speakers have difficulty perceiving certain L2 vowel contrasts not found in their

L1. For example, English speakers identified a (synthesized) French [y] vowel as English [u] while Portuguese speakers identified it as Portuguese [i] (Rochet 1995), presumably equating the L2 vowel with an L1 category in perception (Best 1995; Flege 1995; Kuhl and Iverson 1995). This perceptual distortion was also reflected in these speakers' imitative production of French [y]: English speakers' production of [y] was often judged to be [u] by French native speakers while Portuguese speakers' was judged to be [i], consistent with the claims of Flege et al. (1997) that learners' perceptual difficulties with novel L2 vowel contrasts are reflected in production difficulties.

In perceiving L2 vowel contrasts, speakers may rely on acoustic cues that are different from those utilized by native speakers of the L2; furthermore, when several cues to a contrast are present, L1 and L2 speakers may weight them differently. For example, many L2 learners of English who have difficulty perceiving the English [i]-[ɪ] contrast may rely on a durational difference rather than vowel quality differences, which are the primary cues for native speakers of most dialects of English to distinguish the vowels. Such a pattern is found not only for speakers whose L1 has durational contrasts in vowels—German (Bohn 1995), Finnish (Ylinen et al. 2009)—but also for speakers whose L1 has no durational contrasts—Spanish, Mandarin, Korean, Russian, and Catalan (Bohn 1995; Flege et al. 1997; Escudero and Boersma 2004; Cebrian 2006; Morrison 2009). Escudero et al. (2009) found that the “over”-reliance on duration for L2 vowel distinction is not specific to L2 English but is also found in the perception of the Dutch /a:/ - /ɑ/ contrast by Spanish speakers of L2 Dutch.

The reliance of L2 speakers on durational cues has been explained in terms of both language transfer and universal tendencies. Bohn (1995) proposes that when the vowel space containing two L2 vowels corresponds to a single vowel in L1, listeners will be “desensitized” to spectral distinctions in the region. He proposes furthermore that durational cues are universally salient, leading these to be the first cues used by inexperienced L2 learners. However, Kondaurava and Francis (2008) suggest that the use of duration for L2 vowel contrasts by Spanish and Russian speakers can be analyzed as transfer, since durational cues may be used in the L1 as a secondary cue for distinctions in stress and consonant voicing. Escudero and Boersma (2004) also challenge the claim that reliance on duration reflects a universal tendency, arguing that the sensitivity to duration evidenced in their studies followed from listeners' sensitivity to the statistical patterns of the input data, where duration did play a contrastive role, albeit a secondary one. Because duration was not used in the L1, the correlation of durational differences with different vowel categories in L2 was readily acquired, while the distribution of spectral differences was hindered by interference from the L1. Further evidence against the view that duration is a universally accessible cue for L2 contrasts comes from McAllister et al. (2002) and Ylinen et al. (2005), who found that quantity distinctions in L2 vowels (Swedish and Finnish) are better perceived by speakers whose L1 uses duration to signal vowel contrasts



(Estonian) than by speakers whose L1 (Spanish, Russian) does not contrast long and short vowels.

The acquisition of L2 vowel contrasts is complicated by the fact that individual vowel sounds are typically realized somewhat differently in different contexts. Levy and Strange (2008) found that the perception of the /u-y/ contrast in Parisian French was generally difficult for American English speakers, but that inexperienced L2 learners showed more errors in the context of alveolar consonants, where English /u/ is allophonically fronted and therefore more similar to /y/, than in bilabial contexts, where it is not. Although the /u-y/ contrast continued to be difficult for experienced speakers, the consonantal context effect disappeared, indicating that “learning L2 includes learning all of the language-specific variations that occur within its phonetic categories.” Similarly, in production, Oh (2008) examined the degree of coarticulation in alveolar stop + /u/ sequences in French and English speakers’ L1 and L2, and found that English had more extensive C-to-V coarticulation while French had more extensive V-to-C coarticulation. Although many learners acquired both target values and coarticulation patterns, some acquired only target values, suggesting that coarticulation patterns are language-specific and must be learned independent of the target vowel values.

### **Liquid contrasts**

The majority of research on L2 liquids has focused on Japanese learners’ acquisition of the English /l/ vs. /ɹ/ contrast, which is notoriously difficult for Japanese speakers (e.g. Goto 1971; Logan et al. 1991; Bradlow et al. 1997; Bradlow et al. 1999; Takagi 2002; Aoyama et al. 2004, among others; see Bradlow 2008 for a recent review). Where English uses the /l/ vs. /ɹ/ contrast to distinguish words, Japanese has a single liquid whose distribution in the acoustic space straddles the two English categories. It has been shown that while native English speakers rely on the third formant (F3) for discrimination of English /ɹ-/l/, Japanese speakers rely on the second formant (F2) (Iverson et al. 2003). Aoyama et al. (2004) examines how the degree of perceived phonetic dissimilarity influences L2 learners’ success in acquiring L2 phonetic segments. The Japanese liquid /ɹ/ is considered more similar to English /l/ than /ɹ/, so in earlier stages of acquisition, /l/ is more accurately pronounced, showing the advantage of similarity to an L1 category. Later, however, /ɹ/ shows more improvement, presumably because its greater distance from L1 liquid prevents its assimilation to an L1 category (cf. the *Similarity Differential Rate Hypothesis*; Major and Kim 1996).

Research on acquisition of this contrast supports the view that the mapping of L2 sounds to L1 phoneme categories takes place not as a mapping from phoneme category to phoneme category; instead, learners map positional allophones (the variant pronunciations of a phoneme in different contexts) to L1 phoneme categories (Flege 1995). Ingram and Park (1998),

examining the perception of the English /l/-/ɫ/ contrast by Korean and Japanese speakers in various phonetic contexts, found contextual variation in the acquisition of this contrast. Korean speakers were better at perceiving the /l/-/ɫ/ contrast in intervocalic or cluster position, where the English /ɫ/-/l/ contrast can be equated with the native Korean contrast between native singleton /r/ vs. geminate liquid /ll/, than in initial position, where no singleton vs. geminate contrast is available in Korean. Japanese speakers, on the other hand, who have no comparable singleton vs. geminate liquid contrast in L1, showed generally poorer perception of the English /l/-/ɫ/ contrast than Korean speakers, but the Japanese speakers' perception was better in intervocalic and initial positions, where acoustic cues are generally more perceptible, than in clusters (a problem for the claim of Brown (2000) that Japanese speakers' difficulty in perceiving the English liquid contrast stems from their lack of an L1 phonological feature to encode the contrast, since the feature specification of the English liquids would be the same in all positions). Ingram and Park (1998) conclude that L1 background and general acoustic discriminability, but not universal markedness, affected the perception pattern.

Colantoni and Steele (2007), examining the production of French rhotic /ʀ/ by intermediate- and advanced-proficiency English-speaking learners, found additional evidence for the role of phonetic salience and articulatory ease in acquisition. The voiced uvular fricative realization of the French rhotic presents difficulties for English speakers not only because English lacks a dorsal fricative but also because the antagonistic articulatory requirements of voicing and frication make voiced fricatives inherently difficult to produce, particularly in dorsal place. The results show that the English learners produced frication at a level comparable to native speakers but failed to produce voicing correctly. Colantoni and Steele (2007) argue that because frication is a more salient acoustic property of the French rhotic than voicing, the L2 learners preserve the more easily perceived property of the novel sound at the expense of the less salient property.

### **3. Acquisition of restrictions on syllable structure and phonotactics**

In addition to differences in segment inventory, languages may differ in the number and types of segments that may be grouped into syllables and in restrictions on the types of segments that may occur in specific positions within a syllable. In general, while L2 segmental acquisition tends to show more obvious effects of L1 transfer than of universal tendencies, research on the acquisition of phonotactics and related phonological processes in the L2 sound system has been the source of a number of arguments for the role of universal markedness effects in L2 acquisition.

#### **Consonant clusters within and across syllables**

The perception and production of non-native consonant clusters has been extensively studied,

in large part because this area provides evidence for emergent hierarchies independent of the L1 or L2, as certain non-native cluster types seem to be acquired earlier than other equally novel clusters across various L1s and L2s (Broselow & Finer 1991; Eckman and Iverson 1993; Broselow et al. 1998; Hancin-Bhatt 2000; Hansen 2004; Berent et al. 2008; Berent et al. 2009; Davidson 2010, among others). Many of these studies address the question of whether ease of acquisition is related to sonority, assuming a universal sonority scale consisting of stops < fricatives < nasals < liquids < glides < vowels, organized from least to most sonorous. Two proposed universal principles govern the organization of segments within syllables: consonant clusters should increase in sonority approaching the vocalic nucleus, favoring rising sonority onset /pr/ over falling sonority onset /rp/; and consonants within a cluster should be distant in sonority, favoring onset /pr/ over onset /pn/, whose members are closer on the sonority scale.

Broselow and Finer (1991) examined the production of English pseudo words containing /Cj/ and /Cr/ onset clusters by speakers of Japanese and Korean, where the only onset clusters allowed are obstruent-*j*. Lower error rates were found for clusters with larger sonority distance such as /pr/ (stop – liquid) than for clusters with smaller sonority distance such as /fr/ (fricative-liquid). Similarly, studies of speakers' perception of unfamiliar cluster types have shown an effect of sonority: Berent et al. (2008) and Berent et al. (2009) argue that English and Korean speakers perceived non-native onset clusters of rising sonority in pseudo words more accurately than clusters of falling sonority. Berent et al. (2009) further argue that the sonority effect found in English speakers' perception of non-native clusters cannot be derived from the statistical generalizations of the English lexicon. However, Davidson (2010), in a study of English and Catalan native speakers' productions of pseudo words containing non-native word-initial clusters, argued that neither phonological markedness (as defined by sonority distance) nor analogy to L1 clusters was a good predictor of production accuracy, which was better predicted by general phonetic factors such as articulatory ease (voiceless obstruent clusters are easier to produce than voiced obstruent clusters) and perceptual salience (fricative-initial clusters contain more salient cues to the identity of the first consonant than stop-initial clusters).

While factors independent of L1 seem clearly to be a factor in the acquisition of L2 syllable structure and phonotactics, native language restrictions clearly affect not only the production but also the perception of L2. For example, Japanese speakers typically alter English words to fit the more restrictive syllable structure conditions of their native language; English 'pub' is borrowed as *pabu* with a vowel inserted after [b] because [b] cannot occur in Japanese syllable codas (except in geminates). Dupoux et al. (1999) investigated the perception of illegal structures such as the pseudo word *ebzo*, which constitutes a legal structure in French but not in Japanese. Presented with a series of stimuli ranging from e.g., *ebzo* to *ebuzo*, with a vocalic portion of varying length (from null to a full vowel), the majority of the Japanese speakers perceived an 'illusory vowel' in the illegal consonant sequence—that is, they perceived *ebzo* as

a possible native language structure *ebuzo* even when there was no vocalic signal in the stimulus. In contrast, the majority of French speakers, for whom *ebzo* is a possible structure, perceived a vowel between [b] and [z] only when the vocalic portion was at least 50 ms long. In a follow up study using ERP (Event-Related Potential) methodology, Dehaene-Lambertz et al. (2000) found that the illusory vowel effect held at early stages of speech processing. However, discrimination of CC-CvC can improve when the difference is associated with lexical contrasts; Davidson et al. (2007) found that American English speakers asked to distinguish pseudo words containing a non-native CC sequence vs. a CvC sequence in picture naming tasks performed better when the stimuli were presented with contrastive meanings.

Matthews and Brown (2004) and Idsardi and Kabak (2006) present additional evidence that illusory vowels arise in response to syllable structures that are illegal in the L1. Matthews and Brown's (2004) study investigating the ability of native speakers of Japanese and Thai to discriminate heterosyllabic [k.t] vs. [kVt] revealed that the Thai speakers were better able to discriminate [k.t] vs. [kVt] than Japanese speakers, even though the sequence [kt] is not possible in either language. This difference in performance arises, they argued, from the fact that [k] is a possible syllable coda in Thai (so long as it is not followed by [t]) but not in Japanese. Matthews and Brown (2004) conclude therefore that perception of an illusory vowel is triggered by native language constraints defining possible syllables, rather than simply by constraints on sequences of segments across syllable boundaries (see Idsardi and Kabak 2006 for similar arguments based on perception data from speakers of Korean). Further evidence for ascribing the illusory vowel effect (failure to distinguish CC vs. CvC) to the native language phonological system came from the fact that the failure to discriminate was found only when the interval between stimuli was long enough (1500ms) to force the listeners to access phonological representations stored in memory.

While vowel insertion in both production and perception is most frequently motivated by differences in L1 and L2 syllable structure, Korean speakers' production of English word-final stops presents a case of vowel insertion that is unmotivated by phonotactic restrictions of L1. Korean speakers frequently insert a vowel after a word-final stop of English (e.g., *tape* is pronounced as *t<sup>h</sup>eip<sup>h</sup>i ~ t<sup>h</sup>eip*) both in L2 speech and in loanword adaptation even though voiceless stops are allowed in coda position in Korean. Kang (2003) proposes that this seemingly gratuitous vowel insertion is due to the differences in the phonetics of final stops in English, where final stops are optionally released, and Korean, where no audible release is possible.

### **Restrictions on segments in syllable coda**

Many languages place restrictions on the types of consonant contrasts that can be realized in

syllable codas. One common pattern is a prohibition on voiced obstruents in syllable coda (as in German, Dutch, and Russian), resulting in a lack of obstruent voicing contrasts in coda position. Acquisition of an L2 that allows both voiceless and voiced obstruents in coda is expected to cause problems for learners of languages that disallow such contrasts, particularly since voiced obstruents in coda position are considered to be marked. In an investigation of German speakers' L2 production of the English voicing contrast in word-final stops, Smith et al. (2009) found that the German speakers produced more voicing contrasts for English word-final stops than for German words but that their English production still fell short of L1 English speakers' performance, showing a transfer of L1 final devoicing to L2. The fact that German speakers managed to partially acquire the voicing contrast in coda position may be attributed to the influence of English input or to the fact that many German speakers actually exhibited "incomplete neutralization" of the German voicing contrast—i.e., German speakers actually do make a subtle voicing distinction in their production of these coda consonants, although the difference is barely perceptible.

Final devoicing has attracted much attention in L2 phonology mainly because one often observes an emergent pattern that is not straightforwardly accounted for by either L1 transfer or L2 input—namely, patterns in which learners devoice final obstruents in their L2 production even though neither the L1 nor the target L2 manifests any evidence for such devoicing (Flege and Davidian 1984, Flege et al. 1987, Yavaş 1994, Wissing & Zonneveld 1996, Broselow et al. 1998, among others). Broselow (2004) proposed an OT analysis of such emergent patterns whereby a universal markedness constraint against voiced obstruents in coda position, which remains dormant in the L1, becomes visible as constraint rankings fluctuate in the process of L2 acquisition, though acknowledging that because voicing of obstruents in coda position is articulatorily difficult and perceptually not very salient, so it is likely that the L2 speakers may fail to produce final voiced obstruents not because their phonological grammar forbids them but because at least some of the final devoicing takes place during perception and/or production. Similarly, Simon (2010) examined final obstruent devoicing and cross-word voice assimilation in a corpus of L2 English conversations between native speakers of Dutch. The results show that while both L1 processes are frequently transferred into L2 English, a hierarchy emerges that is not motivated by the L1 data: final voiced fricatives are more frequently devoiced than final voiced stops. They attribute this asymmetry to the aerodynamic difficulty in producing voicing in final fricatives; while voicing requires adduction of glottis, frication requires a sufficient airflow through the glottis.

Another hierarchy emerges in Cardoso's (2007) examination of Brazilian Portuguese speakers' production of English word-final stops. Although Brazilian Portuguese allows only /N l r s/ and no stops at all in coda, coronal stops are more frequently attested than stops at other places of

articulation, in line with the observation that cross-linguistically, coronal place of articulation is the least marked of the three major places of articulation dorsal, coronal, and labial (Paradis and Prunet 1991).

The hierarchies of difficulty discussed above are horizontal, revealing different rates of acquisition for different members of a class. In addition, James (1987a, b) suggests a vertical hierarchy-- namely that the lexical level of L2 phonological representation (e.g., phonemes and word-level accents) is acquired before higher levels of representation (i.e., phrasal and rhythmic properties). We turn now to the acquisition of prosodic systems as they are realized on the word, phrase, and sentence levels.

#### **4. Prosodic Systems: Stress, Pitch Accent, Tone, and Intonation**

Like the acquisition of L2 segments and phonotactics, the acquisition of prosody shows evidence for the role of L1 as well as for universal principles not obviously grounded in either L1 or L2 input. We begin with some discussion of the typology of prosody. Languages are most often classified into three categories: tone languages, pitch accent languages, and stress languages. In a tone language such as Mandarin Chinese, at least some morphemes are distinguished solely by their pitch (e.g., *ma* with high level tone is 'mother' while *ma* with falling tone is 'horse'). Pitch accent languages such as Tokyo Japanese also use pitch to signal lexical contrasts, but the inventory of pitch patterns on words is generally restricted, with specific syllables within a word (accented syllables) associated with invariant tonal contours (Hayes 1995: 49–50). (Hyman (2006) argues that pitch accent should not be thought of as a discrete category; rather, pitch accent languages may combine features of stress-based and tone-based systems.) The defining characteristic of stress, according to Hyman (2006: 231), is that each lexical word contains exactly one syllable "marked for the highest degree of metrical prominence (primary stress)." Prominence is typically signaled by a combination of acoustic cues, including duration, intensity, and pitch changes, and is connected with the rhythmic organization of words and phrases. The function of pitch in a stress language like English is generally to signal sentence-level intonation. Intonational melodies convey "meanings that apply to phrases or sentences as a whole, such as sentence type or speech act, or focus and information structure" (Ladd 2008: 6), and stressed syllables may vary in pitch according to the intonational melody of the utterance in which they appear.

Languages characterized as stress-based may differ along a number of dimensions. A language may allow one or multiple stresses per word, and the position of stress may be lexically specified or phonologically predictable (or some combination of the two). In languages with predictable stress, the position of stress may depend on various factors. In a *fixed stress*

language, stress is determined purely by word position, as in Polish, where stress falls on the penultimate syllable of each word, or Hungarian, where the initial syllable is stressed. In *quantity-sensitive stress* languages, the position of stress may vary depending on syllable weight, as in Cairene Arabic, where stress falls on a penultimate syllable only if that syllable is heavy (containing a long vowel or a final consonant). Languages may also employ different foot types—with German favoring trochaic (strong-weak) feet and Choctaw (southeastern United States) favoring iambic (weak-strong) feet—and may impose a minimal size on feet (typically two syllables or moras, where the mora equals a weight unit within the syllable). Furthermore, languages may require all syllables to be contained within feet (giving rise to patterns of alternating stress), or may allow or even require syllables to be unfooted; for example, the English tendency to avoid stress on the final syllable of nouns (as in verb-noun pairs such as *reCORD/REcord*) is often analyzed as a requirement that the final syllable of a noun be extrametrical, i.e. outside a foot. Thus, even learners whose L1 is stress-based may encounter substantial differences between the L1 and L2 stress systems, which are generally characterized as differences in settings of parameters or different rankings of constraints.

Another dimension distinguishing languages is rhythm, with languages classified as stress-timed (English) vs. syllable-timed (Spanish) or mora-timed (Japanese). Impressionistically, English syllables tend to be of unequal durations and the intervals between syllable peaks tend to vary, while in Spanish or Japanese, syllables or moras (vocalic nuclei and coda consonants) tend to share similar durations and to be spaced at regular intervals. While cautioning that stress timing vs. syllable timing represent ends of a continuum rather than discrete categories, Dauer (1983) identifies two characteristics of stress-timed languages: (i) a tendency to reduce vowels in unstressed syllables and (ii) a larger inventory of syllable types than in syllable-timed languages. Both infants and adults have been demonstrated to be sensitive to rhythmic differences in both native and foreign languages, and such differences have been argued to play a role in L1 learners' ability to segment the speech stream (see Ramus et al. 1999 for a review).

### **5.1. Acquisition of Stress and Rhythm**

Since many studies have focused on L2 stress production or perception but not both, we will consider these areas separately.

#### **Production of L2 Stress**

Comparison of learners from various L1 backgrounds supports the claim that typological similarity confers an advantage in acquiring an L2 stress system. Learners whose L1 is a stress language, even when the placement of stress differs in the L1 and L2, appear to do better in producing L2 stress; for example, Altmann (2006) argued that native speakers of Arabic (a language with predictable stress, though following different parameter settings than English)

were more successful in producing English-like stress than were native speakers of Mandarin, a tone language. Learning also appears to be facilitated by overlap between L1 and L2 stress systems: in Kijak's (2009) study of the acquisition of L2 Polish, a language with fixed penultimate stress, learners whose L1 included penultimate stress as a possible option (German, Spanish, Italian, Russian, and English) showed an initial advantage over those whose L1 allowed only final stress (French) or initial stress (Czech), or was a tone language (Mandarin).

The types of errors found in L2 stress production also support transfer from the L1 grammar. The most common error in Archibald's (1993, 1994, 1995) studies of speakers of L1 fixed stress languages—Polish (penultimate) and Hungarian (initial)—involved placement of stress in the L1 position. Speakers of a quantity-sensitive L1 also showed a preference for maintaining L1 stress patterns; Youssef & Mazurkewich (1998) found that Cairene Arabic speakers produced near-perfect stress placement for English words whose stress was consistent with the Arabic pattern of stress on a superheavy (CVCC or CVVC) final syllable (*volunTEER*) or on a heavy (CVC or CVV) penultimate (*moMENTum*), while words whose stress deviated from the L1 pattern were produced correctly less than half the time, with the vast majority of errors involving putting stress in the appropriate L1 position (*bariTONE*, with final CVVC, and *cyLINDER*, with penultimate CVC). In Kijak's study of L2 Polish by speakers of both fixed stress and variable stress L1s, "the positions of the majority of non-L2 stresses produced were the possible positions for stress in the L1" (Kijak 2009: 322). Speakers of the tone language Mandarin, who presumably did not need to overcome a native language stress system, "were quite successful and consistent in their discovery of the correct word edge at which main stress in Polish is located" (Kijak 2009: 312-313), most frequently placing stress on the penultimate or (less frequently) the final syllable. In this case, as with segmental acquisition, the distance between the L1 and L2 actually seemed to facilitate learning by prevent equivalence classification.

Another area of transfer involves the choice of acoustic cues used to convey stress. Aoyama and Guion (2007) found that Japanese speakers used a wider pitch range on stressed syllables than did English speakers, suggesting that the Japanese speakers tended to rely on their native cue for pitch accent rather than the combination of acoustic cues for stress used by English native speakers.

L1 influence also appears in rhythm, associated in stress-timed languages such as English and German with the reduction of unstressed syllables. English vowel reduction has been shown to be problematic for speakers of Spanish (Flege and Bohn 1989), Japanese (Aoyama and Guion 2007), and Mandarin (Zhang et al. 2008). A study of the L2 German of Spanish and Korean speakers (Young-Scholten 1993) showed that Spanish and Korean speakers learning German had difficulty in producing L2 inflectional affixes containing reduced vowels.



While transfer is clearly a factor in L2 production, some patterns have appeared that appear independent of both the L1 and L2, but which nonetheless conform to patterns found in human language. Studies of the acquisition of English stress by speakers of Polish, Hungarian, and Spanish (Archibald 1993, 1994, 1995, 1997, 1998) and by speakers of Canadian French (Pater 1997) have been argued to present cases of this type, though Pater (1997) and van der Pas and Zonneveld (2004) suggest that some of the patterns Archibald describes can in fact be analyzed as a reflection of L1 transfer. One universal factor that receives support from several studies, both in production and perception, is a preference for assigning stress to heavy (CVV or CVC) as opposed to light (CV) syllables and in particular to syllables containing long vowels. Kijak (2009) found that learners of L2 Polish, in which stress falls on the penultimate regardless of syllable quantity, tended to prefer stress on closed syllables in Polish, even when their L1 was also quantity-insensitive, suggesting that “there may be a universal bias to perceive closed syllables as prominent” (Kijak 2009: 314). Guion et al. (2004), Guion (2005) found a similar preference for placing stress on long vowels in studies of English L2 produced by speakers of Spanish and of Seoul Korean. In the pronunciation of English-type pseudo words, learners were more likely to place initial stress on CVVCVCC than on CVCVCC pseudo words, and final stress on CVCVVC than on CVCVC. This pattern is not motivated in any direct way by either Spanish or Korean, but whether it can be said to be motivated by English depends on the degree of influence one is willing to grant to statistical tendencies rather than absolute requirements. While the grammar of English does not require stress to fall on long vowels, a corpus study of English examining the relationship between vowel length and stress reveals that long vowels are roughly twice as likely to be stressed as are short vowels (Guion et al. 2003).

A related finding from cross-linguistic studies of native language systems is a tendency for vowels of higher sonority (i.e., vowels that are lower in height) to attract stress (de Lacy 2006), presumably since higher sonority correlates with greater inherent duration. The existence of a natural bias toward this stress-sonority link was supported by an experiment in which both English and French speakers had more difficulty learning an artificial language in which lower sonority vowels attracted stress than one in which stress fell preferably on vowels of greater sonority (Carpenter 2010), even though neither English nor French appears to provide direct support for such a pattern (though a corpus study of frequency might conceivably reveal some tendency in this direction). The stress-sonority correlation is also attested in different varieties of English; Peng and Ann (2001) report a correlation between inherent vowel length and stress in Singapore English and Nigerian English, as well as the English of native Spanish speakers, with pronunciations like Chinese, where [aɪ] of the first syllable attracts stress in preference to the inherently shorter [i] of the second syllable. Loanword phonology also provides a number of cases in which prosodic prominence (i.e., a pitch accent or a contour tone) is placed on a heavy syllable in borrowed words, rather than on the originally stressed syllable of the input language (Kang 2010). The preferences for stressing heavier syllables and more sonorous vowels are

therefore strong candidates for universal biases that may play a role in both language acquisition and language contact.

To summarize, the body of research on the acquisition of L2 stress and rhythm suggests a correlation between the similarity of the L1 and L2 prosodic systems and success in the acquisition of L2 prosody, as well as a tendency for L2 stress errors to reflect the native language stress system. Furthermore, studies support a role for universal biases favoring the placement of stress on heavier over lighter syllables and on longer over shorter vowels, which appear to emerge even in the absence of direct evidence from the L1 or L2. What is striking, however, is the relative paucity of studies taking languages other than English as the target language, even though the fact that English stress is not fully predictable makes it less than ideal the investigation of a consistent L2 stress pattern.

### **Perception of L2 Stress**

Some of the studies discussed above argue for transfer in the perception of L2 stress as well as in production. In fact, Kijak (2009) suggests that “the properties of an L1 stress system stand much more in the way of successful L2 perception than in the case of production” (Kijak 2009: 326). Two major sources of difficulty in perception have been identified: learners fail to attend to stress in the L2 input because L1 stress is fully predictable (*stress deafness*), and learners tend to misinterpret L2 acoustic cues in terms of the different functions that these cues serve in the L1.

Arguments for stress deafness come from a set of native language studies in which native speakers of languages with fixed, predictable stress (French, Finnish, and Hungarian) exhibited higher error rates and slower reaction times on tasks involving discrimination of stress differences than on tasks involving discrimination of phoneme differences. In this, the French, Finnish, and Hungarian speakers differed from speakers of Spanish, in whose native language stress is to some extent lexically determined (Dupoux et al. 1997, Peperkamp et al. 1999, Dupoux et al. 2001, Peperkamp and Dupoux 2002, Peperkamp 2004). The authors of these studies hypothesized that speakers of a language with fully predictable stress do not store stress in their phonological representations; thus, stress deafness should appear not at the level of acoustic processing, but rather in processing tasks that increase the load on working memory, requiring reference to a stored abstract representation of stress. Tasks that involved increased memory load (such as increasing the number of stimuli to be identified) did, as predicted, impair accuracy and reaction time for the speakers of predictable stress languages, though not for the Spanish speakers. However, although speakers of the predictably stressed languages French, Finnish, and Hungarian differed significantly from Spanish speakers, the Hungarians also differed from both the French and Spanish groups. Even more surprisingly, Polish speakers, whose native language has predictable stress, did not differ significantly from the Spanish

speakers in their ability to hear stress differences. In a follow up study, Peperkamp et al. (2010) confirmed that Polish speakers fell into an intermediate category between speakers of languages with nonpredictable stress and speakers of predictably stressed L1s French, Finnish, and Hungarian. She attributed this to the fact that Polish does have some lexical exceptions to the general pattern of penultimate stress, therefore defining stress deafness as gradient, correlating with the degree of predictability of stress in the speaker's language.

If stress deafness persists in an L2, we expect that speakers of languages with fully predictable stress should have difficulty accurately perceiving L2 stress. However, this is not necessarily the case. In Kijak's (2009) study, speakers of Czech, which has regular word-initial stress, performed very well in the perception of Polish stress, which is similarly predictable but falls in a different position. Kijak explains this result by pointing out that stress may serve functions other than signaling lexical contrasts: "stress fulfills an important 'demarcative' function which has been shown to be crucial in speech segmentation" (Kijak 2009: 320). Thus, Czech speakers would have good reason to attend to stress in their L1, as it can help them in segmenting the continuous acoustic stream into words (see Jusczyk 1997 on stress in child language acquisition). One might then ask why French speakers should not also attend to stress, where word-final stress could be expected to aid in segmentation. However, the analysis of final prominence in French as reflecting stress has been challenged by a number of researchers who argue that final prominence reflects an intonational accent assigned not to individual words but to the final syllable of a phrase (e.g. Jun and Fougeron 2002). In this case, the prosodic systems of French and Czech are crucially different.

Additional support for the claim that listeners do attend to stress which serves a demarcative function comes from loanword adaptation, where Languages with fixed word-edge stress often truncate stress-peripheral material. For example, the initially-stressed Spanish word *ígado* 'liver' is adapted into Huave (spoken in Mexico), a language with word-final stress, as *ik*. Broselow (2009) argues that the Huave speakers, for whom stress always signals the final syllable of a word, have assumed that material following the stress belongs to a different word (though see Kang 2010). Additional evidence that learners' segmentation of L2 strings into words is affected by L1 patterns is provided by Altenberg (2005), who showed that Spanish-speaking learners of English were significantly worse than native English speakers at identifying the position of word break in phrases such as *keep sparking* and *keeps parking*, finding it difficult to use cues to word position (such as aspiration) that are not relevant in their L1 (see also Broselow 1984/1987).

If a pattern of fully predictable word stress in L1 might interfere with learners' ability to attend to L2 stress, what of the absence of stress in the native language? Here, as with production, evidence suggests that the distance between the L1 and L2 plays a role. For example, Guion

(2005) found that Korean-English bilinguals were less like English L1 speakers in their perception of English stress than were Spanish-English bilinguals, presumably because Korean is not a stress-based language. Similarly, in Kijak's (2009) study of the perception of Polish L2, Mandarin speakers showed very poor perception, a result she attributes to the lack of overlap between cues for Polish stress and the acoustic cues that are important in the prosodic system of Mandarin. This hypothesis is bolstered by other studies suggesting that speakers are biased by their L1 to ignore acoustic cues that are crucial in the L2 but serve a different function (or no function) in their L1. In an investigation of Mandarin and English speakers' judgments of stress position in English words like DEsert/deSERT, Zhang and Francis (2010) found that the two groups differed in their weighting of four main cues to stress (vowel quality, pitch, duration, and intensity), with Mandarin speakers ignoring intensity cues when vowel quality cues were available.

Stress in English may function not only as a feature of words but as an indication of the relationships among words (*BLACKberry* vs. *black BERRY*). As with word stress, speakers of other languages may have difficulty interpreting the acoustic cues signaling these relationships. Nguyen et al. (2008) investigated the ability of speakers of Vietnamese (a tone language) to discriminate English compounds, which typically assign greater prominence to the initial constituent (*BLACKberry*), vs. phrases, which normally assign prominence to the phrase-final head (*black BERRY*). Pitch and duration were manipulated to determine which cues listeners used in their discrimination. Nguyen et al. found that beginning learners were not able to discriminate compounds from phrases using duration alone, because duration is not distinctive in Vietnamese. However, the Vietnamese speakers were more successful in discriminating phrases differing in the position of the focus, which is signaled in English by an intonational pitch accent on the focused element. For phrases with broad focus (*What is this? It's a black BERRY*) vs. phrases with narrow focus on the adjective (*It's not green, it's a BLACK berry*), the Vietnamese speakers "had no problem in manipulating the  $f_0$  and intensity contrastive levels on the accent-bearing syllables as a result of positive transfer from lexical tonal pitch."

Loanword adaptation offers additional examples in which the acoustic cues of a foreign language are interpreted through the filter of the L1. Words borrowed from stress languages into tone languages are most commonly assigned high tone to on the source word stressed syllables (e.g., Yoruba  $to^M MA^H to^L$ ,  $gua^M ran^M TEE^H$ , Kang 2010), suggesting that cues associated with stress in the lending language are interpreted as an indication of high tone in the borrowing language. A similar phenomenon was demonstrated by Chen (2007) for Mandarin-speaking learners of Spanish, who tended to interpret Spanish words with stress on the penultimate syllable as having a lexical rising tone on the stressed syllable and a falling tone on the following syllable.

In addition to differences related to L1 (Spanish, Korean, and Thai), a series of studies (Guion et al. 2004, Guion 2005, Wayland et al. 2006) found systematic differences between early and late learners of English. Early learners were similar to native speakers in recognizing the influence of syllable structure and noun-verb differences on stress, while “late learners of English may rely more heavily on word-by-word learning of stress patterns and are less likely to abstract generalities about stress placement by syllable structure and lexical class” (Wayland et al.: 298). However, Davis and Kelly (1997) found that speakers of 14 different L1 backgrounds did show awareness of the preferred noun-verb stress patterns in English (as in noun *REcord* vs. verb *reCORD*). Participants who were asked to create an English sentence using a pseudo word tended to use bisyllabic words with initial stress as nouns and with final stress as verbs, and were faster and more accurate in classifying actual words as nouns or verbs when the stress pattern conformed to the preferred structure for that category. They conclude that “Despite large individual differences in native language, age of arrival in the United States, and years of exposure to English, nonnative speakers have learned the English noun-verb stress difference” (Davis and Kelly 2006: 457).

The question of the level of attainment that learners can be expected to reach is still very much an open one. In a survey of L2 stress research, van der Pas and Zonneveld (2004: 125) claim that “despite more optimistic claims, the bulk of this literature fails to demonstrate that [parameter] resetting may occur.” While numerous studies demonstrate some degree of success in mastering L2 stress, it is difficult to determine whether these successes represent acquisition of the L2 grammar, particularly when the L2 is English, in which stress is partially lexically determined. Even in studies using pseudo words (Guion et al. 2004, Guion 2005, Wayland et al. 2006), which would seem to avoid the possibility that learners are simply relying on their knowledge of individual words, a major determinant of L2 stress position was the phonological similarity of the pseudo word to existing English words. In parameter-setting models of L2 stress acquisition, no consensus has emerged on whether particular parameters are more or less amenable to resetting, though Pater (1997) suggests that L2 restrictions on foot size and foot headedness may be learned before other parameter settings. Although the Optimality Theory literature offers explicit models of the interplay between linguistic data and constraint ranking (or reranking from the L1 grammar), we are not aware of attempts to test these models against a body of L2 stress data.

## **5.2 Acquisition of Tone, Pitch Accent, and Intonation**

Studies on the acquisition of L2 tone, pitch accent, and intonation are relatively rare. Wang, Jongman, and Sereno (2006) review several studies on the production of Mandarin tones by English-speaking learners, showing high rates of error in tone production by English speakers. Ioup and Tansomboon (1987) studied the production of both tone and segmental aspects of

Thai by four adult second language learners and four children, two of whom were learning Thai as a second language. They found that tone was one of the last aspects of Thai to be mastered by the adult learners but one of the earliest for the children learning Thai, either as L1 or L2.

Studies of cross-language tone perception involving participants who do not know the foreign language can be taken to offer some insight into the initial state for language learning. Such studies suggest that an L1 system in which pitch is used to realize lexical contrasts confers some advantage in tone perception; for example, Lee et al. (1996) found that native speakers of Cantonese performed better than English speakers in identifying pairs of Mandarin syllables as same or different in tone, even though the Cantonese tone inventory is different from that of Mandarin. Similarly, So and Best (2010) found that native speakers of Hong Kong Cantonese (a tone language) and Japanese (a pitch accent language) outperformed English speakers in identifying the tones of Mandarin syllables. However, the Cantonese speakers were actually disadvantaged in some respects by their L1 knowledge; the Mandarin tones that they found most difficult to discriminate were those that fell within the range of a single tone category in Cantonese, a pattern reminiscent of equivalence classification in the perception of phonemes. The error patterns of the Cantonese speakers were quite distinct from those of the Japanese and English speakers, whose error rate correlated more closely with the inherent acoustic similarity of the tonal pairs. Apparently, “the effect of linguistic experience is more related to the constraints of the phonological systems of listeners’ native languages than the degree of tonality use” (So and Best 2010: 290).

One type of equivalence classification involves mapping pitch cues used for lexical contrasts in the L1 in terms of the intonational patterns of the L2. So and Best (2010) found evidence that English speakers were significantly less accurate in identifying Mandarin Tone 4 than were speakers of Cantonese (a tone language) and Japanese (a pitch accent language), and that Tone 4 was identified less accurately than the other tones. Since Mandarin Tone 4 is a falling contour similar to the intonation contour typically found at the end of English statements, it seems likely that English speakers assimilated this lexical tone to an L1 intonational contour. Consistent with this, Broselow, Hurtig, and Ringen (1987) found that for English speakers who received training in identifying the four Mandarin tones, identification of Tone 4 was significantly less accurate when Tone 4 occurred in final than in non-final positions, where it is less likely to be confused with a sentence-final declarative fall. Furthermore, the most frequent error involved (mis)identification of Tone 4 in final position as the high level Tone 1, presumably because English listeners take the falling pitch of Tone 4 to be associated with the sentence-final intonation and interpret only the high starting point of the syllable as signaling the lexical tone.

Even where the L1 and L2 have similar uses of pitch, interference may emerge when the meanings of intonation contours differ in the two languages. In tasks designed to determine how English-speaking learners of Portuguese and Portuguese-speaking learners of English interpreted L2 intonational contours, Cruz-Ferreira (1987) found that where the same meaning was conveyed by similar intonational contours in L1 and L2 the learners in both groups performed similarly to native speakers. Learners had difficulty, however, in cases where the match between meaning and intonation differed in L1 and L2.

A second source of interference may come from differences in the phonetic realization of pitch contours. Mennen (2004) studied two languages with relatively similar intonational systems, Dutch and Modern Greek. In both, declarative intonation includes a (non-final) rise (LH\*), but the languages differ in the alignment of the highest point of the rise: in Dutch, the peak occurs on the accented syllable, while in Greek, the peak occurs following the accented syllable. Comparison of Dutch and Greek sentences produced by L1 Dutch and L1 Greek speakers revealed that four out of five speakers produced patterns of peak alignment that differed from both L1 Greek and L1 Dutch, suggesting that learners had developed an interlanguage system intermediate between the L1 and the L2.

## **5. Summary**

Much research in L2 sound patterns has focused on whether L2 acquisition patterns should be understood as effects of L1 transfer or of universal preferences for particular linguistic structures; whether acquisition patterns should be explained at the level of abstract phonology or at a phonetic level; and whether patterns in L2 production correlate with patterns in L2 perception. Our goal in this chapter was to demonstrate that although these questions have provided a productive program for L2 research, the dichotomies they presuppose are overly simplistic. The literature has provided both plausible examples of transfer from the native language and emergent patterns that appear not to be motivated by input from either L1 or L2. Similarly, some L2 patterns seem to be well explained with reference to abstract phonological structures while others can only be explained as effects of phonetic salience and/or articulatory ease; furthermore, since much recent work in phonology assumes that the constraints of the phonological grammar are grounded in phonetics (e.g., Hayes et al. 2004), the boundary between the two components is increasingly blurred. Nor has a simple explanation of the relationship between production, perception, and lexical representations emerged, with accurate perception appearing to lag production in some areas and to precede it in others. Clearly, L2 speech represents a complex interplay of numerous factors, and despite many advances, no model has yet emerged to provide a fully comprehensive and predictive account of the patterns found in segmental and prosodic L2 speech. Furthermore, much work remains to be done. There is a regrettable dearth of studies charting longitudinal development, particularly the development of L2 prosody. Sorely needed are more studies of L2 stress that

take languages other than English as their L2, studies that investigate production and perception simultaneously, and studies that tap into both auditory and lexical levels.

However, there are grounds for considerable optimism, as methodological advances in the field of language study transform the way research into L2 sound acquisition is conducted (Schmid and Dusseldorp 2010). Behavioral probes such as eye-tracking methods and neurolinguistic tools, in particular ERP studies, provide fine-grained temporal information concerning the time course of L2 speech processing (Escudero et al. 2008; Strange and Shafer 2008), and the availability of ultrasound imaging and electropalatography allow direct observation of articulatory patterns (Gick et al. 2008; Mennen et al. 2010). And while the role of input in L2 acquisition has always been recognized (e.g., Young-Scholten 1994), increasingly large databases and increasingly sophisticated search tools may make possible the discovery of statistical tendencies in L2 input that might not otherwise be apparent (e.g., Guion et al. 2003), and developments in modeling statistically driven acquisition (Boersma and Hayes 2001; Hayes and Wilson 2008; Albright 2009) may shed light on the extent to which L2 patterns which are seemingly unmotivated by L1 or L2 data may be accounted for by the L2 learners' probabilistic knowledge of the learning data (Wilson and Davidson, to appear). Furthermore, artificial grammar learning experiments (e.g., Carpenter 2010), by allowing experimenters to strictly control access to learning input, provide a tool to determine whether typological tendencies reflect genuine learning biases or simply a function of linguistic change (e.g., Blevins 2004). Finally, within Optimality Theory, explicit models have been offered to describe the mapping from the acoustic signal to phonological representations (e.g., Escudero and Boersma 2004, Escudero 2009), an area long neglected in formal theories of phonology. Inevitably, new paradigms will emerge in this swiftly evolving field to offer fresh perspectives on the perennially intriguing problems posed by second language speech patterns.



## References

- Albright, A. 2009. Feature-based generalisation as a source of gradient acceptability. *Phonology* 26: 9-41.
- Altenberg, E. 2005. The perception of word boundaries in a second language. *Second Language Research* 21: 325-58.
- Altmann, H. 2006. *The perception and production of second language stress: a cross-linguistic experimental study*. PhD Thesis, University of Delaware.
- Antoniou, M., Best, C. T., Tyler, M. D., and Kroos, C. 2010. Language context elicits native-like stop voicing in early bilinguals' productions in both NL and L2. *Journal of Phonetics* 38: 640-53.
- Aoyama, K, Flege, J. E., Guion, S. G., Akahane-Yamada, R., and Yamada, T. 2004. Perceived phonetic dissimilarity and L2 speech learning: the case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics* 32: 233-50.
- Aoyama, K. and Guion, S. 2007. Prosody in second language acquisition: An acoustic analysis on duration and F0 range. In Bohn, O.-S. and Munro, M. J., pp. 281-97.
- Archibald, J. 1993. *Language learnability and L2 phonology: The acquisition of metrical parameters*. Dordrecht: Kluwer.
- Archibald, J. 1994. A formal model of learning L2 prosodic phonology. *Second Language Research* 10: 215-40.
- Archibald, J. 1995. The acquisition of stress. In Archibald J., 81-109.
- Archibald, J. (ed.) 1995. *Phonological acquisition and phonological theory*. New Jersey: L. Erlbaum Associates.
- Archibald, J. 1997. The acquisition of English stress by speakers of nonaccentual languages: lexical storage versus computation of stress. *Linguistics* 35: 167-81.
- Archibald, J. 1998. *Second language phonology*. Amsterdam/Philadelphia: John Benjamins.
- Archibald, J. (ed.) 2000. *Second language acquisition and linguistic theory*. Oxford: Blackwell.
- Archibald, J. 2005. Second language phonology as redeployment of NL phonological knowledge. *Canadian Journal of Linguistics* 50: 285-314.
- Archibald, J. 2009. Phonological feature re-assembly and the importance of phonetic cues. *Second Language Research* 25: 231-33.
- Bates, E., Tan, L., and Tzeng, O. (eds.) 2006. *Handbook of East Asian psycholinguistics, volume 1: Chinese*. Cambridge: Cambridge University Press.
- Berent, I., Lennertz, T., Jun, J., Moreno, M. A., and Smolensky, P. 2008. Language universals in human brains. *Proceedings of the National Academy of Science* 105: 5321-25.
- Berent, I., Lennertz, T., Smolensky, P., and Vaknin-Nusbaum, V. 2009. Listeners' knowledge of phonological universals: evidence from nasal clusters. *Phonology* 26: 75-108.
- Best, C. 1995. A direct realist view of cross-language speech perception. In Strange, W., pp. 171-204.
- Best, C. T. and Hall, P. A. 2010. Perception of initial obstruent voicing is influenced by gestural organization. *Journal of Phonetics* 38: 109-26.
- Blevins, J. 2004. *Evolutionary phonology: The emergence of sound patterns*. Cambridge: Cambridge University Press.

- Boersma, P. and Hamann, S. (eds.) 2009. *Phonology in perception*. Berlin and New York: Mouton de Gruyter.
- Boersma, P. and Hayes, B. 2001. Empirical Tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32: 45-86.
- Bohn, O.-S. 1995. Cross-language speech perception in adults: First language transfer doesn't tell it all. In Strange, W., pp.279-304.
- Bohn, O.-S. and Munro, M. J. (eds.) 2007. *The role of language experience in second-language speech learning: In honor of James Emil Flege*. Amsterdam: John Benjamins.
- Bradlow, A. R. 2008. Training non-native language sound patterns: Lessons from training Japanese adults on the English /ɹ/-/l/ contrast. In Edwards, J. G. H. and Zampini, M. L., pp. 287-308.
- Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., and Tohkura, Y. 1999. Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production. *Perception and Psychophysics* 61: 977-85.
- Bradlow, A. R., Pisoni, D. B., Yamada, R. A., and Tohkura, Y. 1997. Training Japanese listeners to identify English /r/ and /l/: IV. Some effect of perceptual learning on speech production. *Journal of the Acoustical Society of America* 101: 2299-310.
- Brannen, K. 2002. The role of perception in differential substitution. *Canadian Journal of Linguistics/Revue Canadienne de Linguistique* 47: 1-46.
- Broselow, E. 2004. Unmarked structures and emergent rankings in second language phonology. *International Journal of Bilingualism* 8: 51-65.
- Broselow, E. 1984. An investigation of transfer in second language phonology. *International Review of Applied Linguistics* 22, 253-269. Reprinted in Ioup, G. and Weinberger, S., pp. 261-278.
- Broselow, E. 2009. Stress adaptation in loanword phonology: perception and learnability. In Boersma, P. and Hamann and S., pp. 191-234.
- Broselow, E., Chen, S., and Wang, C. 1998. The emergence of the unmarked. *Studies in Second Language Acquisition* 20: 251-78.
- Broselow, E. and Finer, D. 1991. Parameter setting in second language phonology and syntax. *Second Language Research* 7: 35-59.
- Broselow, E., Hurtig, R., and Ringen, C. 1987. The perception of second language prosody. In Ioup, G. and Weinberger, S., pp. 350-62.
- Brown, C. 2000. Speech perception and phonological acquisition. In Archibald, J., pp. 4-63.
- Cardoso, W. 2007. The variable development of English word-final stops by Brazilian Portuguese speakers: A stochastic Optimality theoretic account. *Language Variation and Change* 19: 219-48.
- Carpenter, A. 2010. A naturalness bias in learning stress. *Phonology* 27: 345-92.
- Cebrian, J. 2006. Experience and the use of non-native duration in L2 vowel categorization. *Journal of Phonetics* 34: 372-87.
- Chen, Y. 2007. From tone to accent: the tonal transfer strategy for Chinese L2 learners of Spanish. *Proceedings of ICPHS XVI*: 1645-8.
- Colantoni, L. and Steele, J. 2007. Acquiring /ʋ/ in context. *Studies in Second Language Acquisition* 29: 381-406.

- Cruz-Ferreira, M. 1987. Non-native interpretive strategies for intonational meaning: An experimental study. In James, A. and Leather, J. (eds.), pp. 102-20.
- Curtin, S., Goad, H., and Pater, J. 1998. Phonological transfer and levels of representation: the perceptual acquisition of Thai voice and aspiration by English and French speakers. *Second Language Research* 14: 389-405.
- Cutler, A., Weber, A., Otake, T. 2006. Asymmetric mapping from phonetic to lexical representations in second-language listening. *Journal of Phonetics* 34: 269-84.
- Dauer, R. M. 1983. Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics* 11: 51-62.
- Davidson, L. 2010. Phonetic bases of similarities in cross-language production: Evidence from English and Catalan. *Journal of Phonetics* 38: 272-88.
- Davidson, L., Shaw, J., and Adams, T. 2007. The effect of word learning on the perception of non-native consonant sequences. *Journal of the Acoustical Society of America* 122: 3697-709.
- Davis, S. M. and Kelly, M. H. 1997. Knowledge of the English noun-verb stress difference by native and nonnative speakers. *Journal of Memory and Language* 35: 445-60.
- de Jong, K. J., Hao, Y.-C., Park, H. 2009. Evidence for featural units in the acquisition of speech production skills: Linguistic structure in foreign accent. *Journal of Phonetics* 37: 357-73.
- de Jong, K. J., Silbert, N., and Park, H. 2009. Segmental generalization in second language segment identification. *Language Learning* 59: 1-31.
- Dehaene-Lambertz, G., Dupoux, E., and Gout, A. 2000. Electrophysiological correlates of phonological processing: A cross-linguistic study. *Journal of Cognitive Neuroscience* 12: 635-47.
- deLacy, P. 2006. *Markedness: Reduction and preservation in phonology*. Cambridge: Cambridge University Press.
- de Lacy, P. (ed.) 2007. *The Cambridge Handbook of Phonology*. Cambridge: Cambridge University Press.
- Dupoux, E., Pallier, C., Kakehi, K., and Mehler, J. 1999. Epenthetic vowels in Japanese: a perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance* 25: 1568-78.
- Dupoux, E., Pallier, C., Sebastián-Gallés, N., and Mehler, J. 1997. Destressing 'deafness' in French? *Journal of Memory and Language* 36: 405-21.
- Dupoux, E., Peperkamp, S., and Sebastian-Galles, N. 2001. A robust method to study stress 'deafness'. *Journal of the Acoustical Society of America* 110: 1606-18.
- Dupoux, E., Sebastian-Galles, N., Navarete, E., and Peperkamp, S. 2008. Persistent stress 'deafness': The case of French learners of Spanish. *Cognition* 106: 682-706.
- Eckman, F. R. 1977. Markedness and the Contrastive Analysis Hypothesis. *Language Learning* 27: 315-30.
- Eckman, F. R. 2004. From phonemic differences to constraint rankings: Research on second language phonology. *Studies in Second Language Acquisition* 26: 513-549.
- Eckman, F. R. 2008. Typological markedness and second language phonology. In Edwards, J. G. H. and Zampini, M. L., pp. 95-115.
- Eckman, F. R. and Iverson, G. K. 1993. Sonority and markedness among onset clusters in the interlanguage of ESL learners. *Second Language Research* 9: 234-52.

- Eckman, F. R., Iverson, G. K., Fox, R. A., Jacewicz, E. and Lee, S. 2003. Perception and production in the acquisition of L2 phonemic contrasts. *Second Language Research* 19: 169-208.
- Edwards, J. G. H. and Zampini, M. L. (eds.) 2008. *Phonology and second language acquisition*. John Benjamins: Amsterdam.
- Escudero, P. 2009. The linguistic perception of SIMILAR L2 sounds. In Boersma, P. and Hamann, S., pp. 151-190.
- Escudero, P., Benders, T., and Lipski, S. C. 2009. Native, non-native and L2 perceptual cue weighting for Dutch vowels: The case of Dutch, German, and Spanish listeners. *Journal of Phonetics* 37: 452-65.
- Escudero, P., and Boersma, P. 2004. Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition* 26: 551-85.
- Escudero, P., Hayes-Harb, R., and Mitterer, H. 2008. Novel second-language words and asymmetrical lexical access. *Journal of Phonetics* 36: 345-60.
- Flege, J. E. 1987. The production "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics* 15: 47-65.
- Flege, J. E. 1991. 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. 1995. Second language speech learning theory, findings, and problems. In Strange, W., pp. 233-77.
- Flege, J. and Bohn, O.-S. 1989. An instrumental study of vowel reduction and stress placement in Spanish-accented English. *Studies of Second Language Acquisition* 11: 35-62.
- Flege, J., Bohn, O.-S., & Jang, S. (1997). The effect of experience on nonnative subjects' production and perception of English vowels. *Journal of Phonetics*, 25, 437-70.
- Flege, J. and Davidian, R. 1984. Transfer and developmental processes in adult foreign language speech production. *Applied Psycholinguistics* 5: 323-47.
- Flege, J., McCutcheon, J. and Smith S. 1987. The development of skill in producing word final English stops. *Journal of the Acoustical Society of America* 82: 433-47.
- Flege, J., Bohn, O.-S., and Jang, S. 1997. The effect of experience on nonnative subjects' production and perception of English vowels. *Journal of Phonetics* 25: 437-70.
- Flynn, S., Martohardjono, G., and O'Neil, W. (eds.) 1998. *The Generative study of language acquisition*. Mahwah, NJ: Lawrence Erlbaum.
- Fowler, C. A., Sramko, V., Ostry, D. J., Rowland, S. A., Hallé, P. 2008. Cross language phonetic influences on the speech of French-English bilinguals. *Journal of Phonetics* 36: 649-63.
- Gick, B., Bernhardt, B., Bacsfalvi, P., and Wilson, I. 2008. Ultrasound imaging applications in second language acquisition. In Edwards, J. G. H. and Zampini, M. L., pp. 309-22.
- Goad, H. and White, L. 2006. Ultimate attainment in interlanguage grammars: A prosodic approach. *Second Language Research* 22: 243-68.
- Goto, H. 1971. Auditory perception by normal Japanese adults of the sounds "L" and "R". *Neuropsychologia* 9: 317-23.
- Guion, S. 2005. Knowledge of English word stress patterns in early and late Korean-English bilinguals. *Studies in Second Language Acquisition* 27: 503-33.
- Guion, S., Harada, T., and Clark, J.J. 2004. Early and late Spanish-English bilinguals' acquisition of English word stress patterns. *Bilingualism: Language and Cognition* 7: 207-26.

- Guion, S.G., Clark, J.J., Harada, T., and Wayland, R.P. 2003. Factors affecting stress placement for English words include syllable structure, lexical class, and stress patterns of phonologically similar words. *Language and Speech* 46: 403-27.
- Gussenhoven, C. and Warner, N. (eds.) 2002. *Laboratory Phonology 7*. Berlin; New York: Mouton de Gruyter.
- Hancin-Bhatt, B. 1994. Segmental transfer: A consequence of a dynamic system. *Second Language Research* 10: 241-69.
- Hancin-Bhatt, B. 2000. Optimality in second language phonology: codas in Thai ESL. *Second Language Research* 16: 201-32.
- Hancin-Bhatt, B. 2008. Second language phonology in Optimality Theory. In Edwards, J. G. H. and Zampini, M. L., pp. 117-51.
- Hannahs, S. J. and Young-Scholten, M. (eds.) 1997. *Focus on phonological acquisition*. Amsterdam/Philadelphia: John Benjamins.
- Hayes, B. 1995. *Metrical stress theory: Principles and case studies*. Chicago: The University of Chicago Press.
- Hayes, B., Steriade, D., and Kirchner, R. (eds.) 2004. *Phonetically based phonology*. Cambridge: Cambridge University Press.
- Hayes, B. and Wilson, C. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39: 379-440.
- Hayes-Harb, Rachel. 2007. Lexical and statistical evidence in the acquisition of second language phonemes. *Second Language Research* 23: 65-94.
- Hayes-Harb, R. and Masuda, K. 2008. Development of the ability to lexically encode novel second language phonemic contrasts. *Second Language Research* 24: 5-33.
- Hazan, V. and Boulakia, G. 1993. Perception and production of a voicing contrast by French-English bilinguals. *Language and Speech* 36: 17-38.
- Hyman, L. M. 2006. Word-prosodic typology. *Phonology* 23: 225-57.
- Ingram, J. C. and Park, S. G. 1998. Language, context, and speaker effects in the identification and discrimination of English /r/ and /l/ by Japanese and Korean listeners. *Journal of the Acoustical Society of America* 103: 1161-74.
- Ioup, G. and Tansomboom, A. 1987. The acquisition of tone: A maturational perspective. In Ioup, G. and Weinberger, S., pp. 333-49.
- Ioup, G. and Weinberger, S. (eds.) 1987. *Interlanguage Phonology*. Cambridge, MA: Newbury House.
- Ioup, G. 2008. Exploring the role of age in the acquisition of a second language phonology. In Edwards, J. G. H. and Zampini, M. L., pp. 41-62.
- Iverson, P., Kuhl, P.K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A., and Siebert, C. 2003. A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition* 87: B47-57.
- James, A. 1987a. The acquisition of phonological representation: a modular approach. In James, A. and Leather, J., pp. 225-49.
- James, A. 1987b. Prosodic structure in phonological acquisition. *Second Language Research* 3: 118-40

- James, A. and Leather, J. (eds.) 1987. *Sound patterns in second language acquisition*. Dordrecht: Foris Publications.
- James, A. and Leather, J. (eds.) 1996. *Second-language speech: Structure and process*. Berlin: Mouton de Gruyter.
- Jun, S.-A. and Fougeron, C. 2002. Realizations of accentual phrase in French intonation. *Probus* 14: 147-72.
- Jusczyk, P. 1997. *The discovery of spoken language*. Cambridge, MA: MIT Press.
- Kabak, B. and Idsardi, W. 2007. Perceptual distortions in the adaptation of English consonant clusters: Syllable structure or consonantal contact constraints? *Language and Speech* 50: 23-52.
- Kager, R. 1999. *Optimality Theory*. Cambridge: Cambridge University Press.
- Kager, R., van der Feest, S., Fikkert, P., Kerkhoff, A., and Zamuner, T. 2007. Representations of [voice]: evidence from acquisition. In van der Torre, E. J. and van de Weijer, J., pp. 41-80.
- Kang, K.-H. and Guion, S. G. 2006. Phonological systems in bilinguals: Age of learning effects on the stop consonant systems of Korean-English bilinguals. *Journal of the Acoustical Society of America* 119: 1672-83.
- Kang, Y. 2003. Perceptual similarity in loanword adaptation: English postvocalic word-final stops in Korean. *Phonology* 20: 219-73.
- Kang, Y. 2008. Interlanguage segmental mapping as evidence for the nature of lexical representation. *Language and Linguistics Compass* 2: 103-18.
- Kang, Y. 2010. Tutorial overview: suprasegmental adaptation in loanwords. *Lingua???* : 2295-310.
- Kijak, A. 2009. *How stressful is L2 stress? A cross-linguistic study of L2 perception and production of metrical systems*. Utrecht: LOT (Netherlands Graduate School of Linguistics).
- Kondaurova, M. V. and Francis, A. L. 2008. The relationship between native allophonic experience with vowel duration and perception of the English tense/lax vowel contrast by Spanish and Russian listeners. *Journal of the Acoustical Society of America* 124: 3959-71.
- Kuhl, P. K. and Iverson, P. 1995. Linguistic experience and the “perceptual magnet effect”. In Strange, W., pp. 121-54.
- Ladd, R. 2008. *Intonational phonology, second edition*. Cambridge Studies in Linguistics 119. Cambridge: Cambridge University Press.
- Lado, R. 1957. *Linguistics across cultures*. Ann Arbor: University of Michigan Press.
- Larson-Hall, J. 2004. Predicting perceptual success with segments: A test of Japanese speakers of Russian. *Second Language Research* 20: 33-76.
- Lee, Y.-S, Vakoch, D.A., and Wurm, L. H. 1996. Tone perception in Cantonese and Mandarin: A cross-linguistic comparison. *Journal of Psycholinguistic Research* 25: 527-42.
- Levy, E. S. and Strange, W. 2008. Perception of French vowels by American English adults with and without French language experience. *Journal of Phonetics* 36: 141-57.
- Logan, J. S., Lively, S. E., and Pisoni, D. B. 1991. Training Japanese listeners to identify English /r/ and /l/: A first report. *Journal of the Acoustical Society of America* 89: 874-86.
- Major, R. C. 2008. Transfer in second language phonology: A review. In Edwards, J. G. H. and Zampini, M. L., pp. 63-94.

- Major, R. C., and Kim, E. 1996. The Similarity Differential Rate Hypothesis. *Language Learning* 46: 465-96.
- Matthews, J. and Brown, C. 2004. When intake exceeds input: Language specific perceptual illusions induced by NL prosodic constraints. *International Journal of Bilingualism* 8: 5-27.
- McAllister, R., Flege, J. E., and Piske, T. 2002. The influence of NL on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics* 30: 229-58.
- Mennen, I. 2004. Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics* 32: 543-63.
- Mennen, I., Scobbie, J. M., de Leeuw, E., Schaeffler, S., and Schaeffler, F. 2010. Measuring language-specific phonetic settings. *Second Language Research* 26: 13-41.
- Morrison, G. S. 2009. NL-Spanish speakers' acquisition of the English /i/-/ɪ/ contrast II: Perception of vowel inherent spectral change. *Language and Speech* 52: 437-62.
- Nguyen, T., Ingram, C., and Pensalfini, J. 2008. Prosodic transfer in Vietnamese acquisition of English contrastive stress patterns. *Journal of Phonetics* 36: 158-90.
- Oh, E. 2008. Coarticulation in non-native speakers of English and French: An acoustic study. *Journal of Phonetics* 36: 361-84.
- Paradis, C. and Prunet, J.-F. (eds.) 1991. *The special status of coronals: Internal and external evidence. (Phonetics and Phonology 2)*. San Diego: Academic Press.
- Pater, J. 1997. Metrical parameter missetting in second language acquisition. In Hannahs, S. J. and Young-Scholten, M., pp. 235-61.
- Pater, J. 2003. The perceptual acquisition of Thai phonology by English speakers: Task and stimulus effects. *Second Language Research* 19: 209-23.
- Peng, L. and Ann, J. 2001. Stress and duration in three varieties of English. *World Englishes* 20: 1-27.
- Peperkamp, S. 2004. Lexical exception in stress systems: Arguments from early language acquisition and adult speech perception. *Language* 80: 98-126.
- Peperkamp, S. and Dupoux, E. 2002. A typological study of stress 'deafness'. In Gussenhoven, C. and Warner, N., pp. 203-40.
- Peperkamp, S., Dupoux, E., and Sebastián-Gallés, N. 1999. Perception of stress by French, Spanish, and bilingual subjects. *Proceedings of EuroSpeech '99* 6: 2683-6.
- Peperkamp, S., Inga Vendelin, I., and Dupoux, E. 2010. Perception of predictable stress: A cross-linguistic investigation. *Journal of Phonetics* 38, 422-30.
- Piske, T., MacKay, I. R. A., and Flege, J. E. 2001. Factors affecting degree of foreign accent in an L2: a review. *Journal of Phonetics* 29: 191-215.
- Prince, A. and Smolensky, P. 2004. *Optimality Theory: Constraint interaction in Generative Grammar*. Malden, MA/Oxford: Blackwell Publishing.
- Ramus, F., Nespors, M., and Mehler, J. 1999. Correlates of linguistic rhythm in the speech signal. *Cognition* 73: 265-92.
- Rice, K. 2007. Markedness in phonology. In deLacy, P., pp. 79-97.
- Rochet, B. L. 1995. Perception and production of second-language speech sounds by adults. In Strange, W., pp. 379-410.
- Sancier, M. L. and Fowler, C. A. 1997. Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics* 25: 421-36.

- Schmid, M. S. and Dusseldorp, E. 2010. Innovative and quantitative methods for bilingualism research. *Second Language Research* 26: 5-11.
- 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.
- Simon, E. 2009. Acquiring a new second language contrast: an analysis of the English laryngeal system of native speakers of Dutch. *Second Language Research* 25: 377-408.
- Simon, E. 2010. Phonological transfer of voicing and devoicing rules: evidence from NL Dutch and L2 English conversational speech. *Language Sciences* 32: 63-86.
- Smith, B. L., Hayes-Harb, R., Bruss, M., and Harker A. 2009. Production and perception of voicing and devoicing in similar German and English word pairs by native speakers of German. *Journal of Phonetics* 37: 257-75.
- Smolensky, P. 1996. On the comprehension/production dilemma in child language. *Linguistic Inquiry* 27: 720-31.
- So, C. and Best, C. 2010. Cross-language perception of Non-native tonal contrasts: effects of native phonological and phonetic influence. *Language and Speech* 53: 273-93.
- Stevens, K. N., Libermann, A. M., Studdert-Kennedy, M., and Öhman, S. E. G. 1969. Crosslanguage study of vowel perception. *Language and Speech* 12: 1-23.
- Strange, W. 1992. Learning non-native phoneme contrasts: Interactions among subject, stimulus, and task variables. In Tohkura, Y., Sagisaka, Y., and Vatikiotis-Bateson, E., pp. 197-219.
- Strange, W. (ed.) 1995. *Speech perception and linguistic experience: Issues in cross-language speech research*. Timonium, MD: York Press.
- Strange, W., Akahane-Yamada, R., Kubo, R., Trent, S. A, Nishi, K., and Jenkins, J. J. 1998. Perceptual assimilation of American English vowels by Japanese listeners. *Journal of Phonetics* 26: 311-44.
- Strange, W. and Shafer, V. L. 2008. Speech perception in second language learners: the re-education of selective perception. In Edwards, J. G. H. and Zampini, M. L., pp. 153-91.
- Takagi, N. 2002. The limits of training Japanese listeners to identify English /r/ and /l/: Eight case studies. *Journal of the Acoustical Society of America* 111: 2887-96.
- Tohkura, Y., Sagisaka, Y., and Vatikiotis-Bateson, E. (eds.) 1992. *Speech perception, production and linguistic structure*. Tokyo, Japan: OHM.
- Trofimovich, P., Gatlinton, E., and Segalowitz, N. 2007. A dynamic look at L2 phonological learning: Seeking processing explanations for implicational phenomena. *Studies in Second Language Acquisition* 29: 407-48.
- van der Pas, B. and Zonneveld, W. 2004. L2 parameter setting for metrical systems. *The Linguistic Review* 21: 125-70.
- van der Torre, E. J. and van de Weijer, J. (eds.) 2007. *Voicing in Dutch*. Amsterdam: John Benjamins.
- Wang, Y., Jongman, A., and Sereno, J. 2006. Second language acquisition and processing of Mandarin tone. In Bates, E., Tan, L., and Tzeng, O., pp. 250-7.
- Wayland, R., Guion, S., Landfair, D., and Li, B. 2006. Native Thai speakers' acquisition of English word stress patterns. *Psycholinguistic Research* 35: 285-304.



- Weber, A. and Cutler, A. 2004. Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language* 50: 1–25.
- Weinberger, S. H. 1996. Minimal segments in second language phonology. In James, A. and Leather, J., pp. 263–312.
- Williams, L. 1977. The perception of stop consonant voicing by Spanish-English bilinguals. *Perception and Psychophysics* 21: 289-97.
- Wilson, C. and Davidson, L. to appear. Bayesian analysis of non-native cluster production. In Proceedings of NELS 40, MIT, Cambridge, MA.
- Wissing, D. and Zonneveld, W. 1996. Final devoicing as a robust phenomenon in second language acquisition: Tswana, English, and Afrikaans. *South African Journal of Linguistics* 34: 3-24.
- Yavaş, M. 1994. Final stop devoicing in interlanguage. In Yavaş, M., pp. 267-82.
- Yavaş, M. (ed.) 1994. *First and second language phonology*. San Diego: Singular.
- Ylinen, S., Shestakova, A., Alku, P., and Huottilainen, M. 2005. The perception of phonological quantity based on durational cues by native speakers, second-language users and nonspeakers of Finnish. *Language and Speech* 48: 313-38.
- Ylinen, S., Uther, M., Latvala, A., Vepsäläinen, S., Iverson, P., Akahane-Yamada, R., and Näätänen, R. 2009. Training the brain to weight speech cues differently: A study of Finnish Second-language users of English. *Journal of Cognitive Neuroscience* 22: 1319-32.
- Young-Scholten, M. 1993. *The acquisition of prosodic structure in a second language*. Tübingen: Max Niemeyer Verlag.
- Young-Scholten, M. 1994. On positive evidence and ultimate attainment in L2 phonology. *Second Language Research* 10: 193-214.
- Young-Scholten, M. 2004. Prosodic constraints on allophonic distribution in adult L2 acquisition. *International Journal of Bilingualism* 8: 67-77.
- Youssef, A. and Mazurkewich, I. 1998. The acquisition of English metrical parameters and syllable structure by adult native speakers of Egyptian Arabic (Cairene dialect). In Flynn, S., Martohardjono, G., and O'Neil, W., pp. 303-32.
- Zampini, M. L. 2008. L2 speech production research: Findings, issues, and advances. In Edwards, J. G. H. and Zampini, M. L., pp. 219-49.
- Zhang, Y. and Francis, A. 2010. The weighting of vowel quality in native and non-native listeners' perception of English lexical stress. *Journal of Phonetics* 38: 260-71.
- Zhang, Y., Nissen, S., and Francis, A.L. 2008. Acoustic characteristics of English lexical stress produced by native Mandarin speakers. *Journal of the Acoustical Society of America* 123: 4498-513.