1

Generative Approaches to Bilingual Phonetics and Phonology

Ellen Broselow

1.1 Introduction

The goal of generative linguistics is to model the internalized grammars that allow members of a linguistic community to produce and understand utterances they have not previously encountered. Most generative models of the phonological component share the assumption that the grammar defines a mapping between two levels: *abstract mental representations*, which contain all nonpredictable information about the sound structure of morphemes, and *surface representations*, which encode predictable alternations. Opinions diverge, however, on the formal properties of the internalized grammar and the nature of the mapping from underlying to surface representations, as well as on the relationship between phonology and phonetics, on one hand, and usage-based factors such as frequency and variation, on the other.

Research in cross-language phonology has been informed by developments in generative phonology. At the same time, second language (L2) research has contributed insight to the understanding of the phonological grammar above and beyond what can be learned from the study of monolingual behavior. Speakers encountering a new language may be confronted with structures that do not occur in their first language (L1), and L2 production and perception of these novel structures may provide a valuable window into the nature of the L1 grammar by revealing how learners generalize beyond prior experience. Second language patterns may also provide insight into factors that are independent of experience with a particular language, such as the relative complexity, the articulatory difficulty, or the perceptual salience of particular structures, as well as the influence of universal grammatical principles and/or learning biases. And the study of bilingual speakers is especially crucial in shedding light on the interaction between distinct internalized grammars.

The goal of this chapter is to explore the strengths and limitations of different models of phonology in accounting for data from multilingual

speakers, including the realization of L2 phonemes, allophones, syllables, and suprasegmental structures (tone, stress, pitch accent, and intonation). While bilinguals fall on a continuum from early and consistently balanced users of two languages to adult learners of a new language, the majority of research in second language phonology has focused on data from speakers who have internalized an L1 before acquiring additional languages. Much of the work discussed here will therefore focus on the extent to which L2 patterns can be explained in terms of the L1 grammar. Issues to be addressed include the role of phonological versus phonetic structure in a grammar, the relationship between the production grammar and the perception grammar, and the role of language-independent factors such as putative innate learning biases, articulatory difficulty, and perceptual salience.

This chapter begins with a brief overview of competing models of the phonological grammar that have emerged in the last half-century of work in generative phonology. The focus then turns to how particular areas of L2 phonology have been approached within different models. We examine the limits of reliance on the L1 to explain L2 patterns, and then examine how different views of the nature of phonological representations and the nature of phonological derivations have intersected with differing explanations of L2 phonology.

1.2 Models of the L1 Phonological Grammar

Chomsky and Halle's 1968 foundational work *The Sound Pattern of English* (SPE) set forth a framework in which the job of the phonological grammar was to map abstract mental representations of stored, nonpredictable lexical structure into surface representations that display allophonic and allomorphic variation. Phonological representations at both the underlying and the surface levels were argued to consist of a linear string of segments, defined as a matrix of phonological feature specifications, along with indicators of morpheme and word boundaries. Surface representations were realized by a system of phonetic implementation which accounted for such aspects of phonetic detail as variation in the realization of a single surface representation.

SPE provided a comprehensive and explicit model of phonological analysis, inspiring researchers to extend this model to numerous languages. As the database expanded, so did the notion of phonological representations. Research on tonal processes, intonation, templatic morphology, and vowel harmony motivated abandonment of the principle of strict linear organization. In Autosegmental Phonology (e.g., Goldsmith, 1979), features were argued to occupy distinct tiers, with the possibility of one-to-many mappings between elements on different tiers; for example, a falling tone might be represented by two tones (High and Low) associated with a single vowel, while a sequence of high-toned syllables might be represented by a single High tone associated with multiple syllables. Research on word and phrasal stress and prominence

motivated the claims of Metrical Phonology (e.g., Hayes, 1985) that segments may be organized hierarchically into prosodic units such as syllables, feet, and prosodic words, and research on multiply-articulated segments motivated Feature Geometry (e.g., Clements, 1985; Sagey, 1986), which argued for hierarchical structure within segments, with features organized under place, manner, and laryngeal nodes. Perhaps the most radical departure from SPE phonology was associated with Articulatory/Gestural Phonology (e.g., Browman & Goldstein, 1989; Gafos, 2002), which included direct reference to the positions of articulators in phonological representations and to the timing of articulatory gestures in phonological derivations. This suggested a rethinking of some earlier analyses; for example, the loss of /t/ in connected pronunciation of the phrase *perfect memory* could be understood not as a result of a phonological rule deleting /t/ but rather as a result of the English pattern of overlap in time between the velar, the coronal, and the labial closures of the /ktm/ sequence which makes the coronal closure difficult to perceive (Browman & Goldstein, 1989).

The extension of the SPE model to a variety of languages also led to new ideas about the relationship between the grammars of individual languages and the role of typological generalizations pointing to cross-linguistic preferences for certain surface structures; for example, the finding that the presence of final voiced obstruent in a language implies the presence of final voiceless obstruents suggests that the latter are universally more natural (less marked) than the former. The assumption that particular surface structures may be considered universally more or less preferred has implications for the nature of phonological derivations. In SPE, the grammar makes no direct reference to surface structures, which simply emerge in the mapping from underlying to surface representations via ordered rewrite rules, where each rule effects a structural transformation at that level, without looking back to lexical representations or ahead to surface representations. The importance of surface structure restrictions was pointed out early on, however, by researchers who noted that individual languages often employ multiple rules that "conspire" to ensure that certain structures never appear in surface representations (Kisseberth, 1970) and furthermore that children often seem to apply rules that have the effect of simplifying children's productions in the direction of universally preferred surface structures, even where these rules are not supported by the data of the adult language, which may contain the dispreferred structures (Stampe, 1979).

The role of cross-linguistic surface naturalness in grammars was addressed in SPE in the form of the Evaluation/Simplicity Metric, which equates the naturalness of a grammar with the number of symbols contained in its rule set. While the Evaluation Metric rewards rules that refer to more natural classes of sounds, as these require fewer features to specify, it does not distinguish a cross-linguistically common rule such as devoicing of final obstruents from its rare and possibly unattested counterpart of final obstruent voicing, since both rules require the same number of symbols. The failure to distinguish rules

that create marked versus unmarked surface structures was acknowledged in the final chapter of SPE, where a system using default fill-in of the more natural value of a feature was proposed, though this system was not widely adopted. Alternative proposals for incorporating cross-linguistic naturalness into phonology included allowing constraints on preferred surface structures to either block or trigger the application of a rule, as well as extending the principles and parameters approach of syntax to include a universal set of phonological parameters to which learners have access, where the default setting defines the least marked (cross-linguistically most common and presumably most natural) option.

The move to incorporate typological generalizations into the grammar ran up against the apparent paradox that even a principle that has clear cross-linguistic support is not necessarily obeyed in every language: some languages do allow surface structures that are considered marked. The emergence of Optimality Theory (OT; Kager, 1999; Prince & Smolensky, 2004) provided a way to make typological markedness of surface forms a driving force in phonological derivations while still allowing some languages to violate so-called universals. In this framework, rules were replaced by a set of constraints that define the optimal surface realization of an underlying representation. In classic OT, constraints are of two types: *markedness constraints* (or *well-formedness constraints*), which encode universal preferences by penalizing marked surface structures, and *faithfulness constraints* (or *correspondence constraints*), which penalize changes from underlying structure. While the set of constraints is assumed in classic OT to be universal, constraint rankings are considered language-specific, and any constraint may be violated under pressure from higher-ranking constraints. In accordance with the Subset Principle (Berwick, 1985), whereby the initial state grammar for L1 acquisition is the one defining the most restrictive grammar, children are argued to begin by ranking markedness constraints over faithfulness constraints (see papers in Kager, Pater, and Zonneveld, 2004). Reranking is motivated by positive linguistic evidence: hearing a marked structure motivates demotion of the violated constraint.

Alongside assumptions about the nature of phonological representations and derivations, a third area where assumptions diverge concerns the sorts of facts that the grammar must account for. While it has long been recognized that perception is affected by one's native language, explicit formal grammars of the mapping from acoustic signals to phonological representations (e.g., Boersma, 1998) were not proposed until the generative enterprise had been underway for several decades. And while sociolinguists tackled the question of variation in language early on, it is only relatively recently that mainstream phonologists have built grammars whose job is not to select a single grammatical realization of a given underlying representation, but rather to predict the probability of different surface representations of an underlying form (e.g., the

Gradual Learning Algorithm of Boersma and Hayes, 2001; the Maximum Entropy grammar of Hayes and Wilson, 2008).

Despite differing views on the nature of phonological representations and the mapping between underlying and surface levels, almost all generative phonologists would agree that acquiring a language involves internalizing a grammar. This makes data from second language acquisition (SLA) a fruitful area for research aimed at discovering the nature of the initial state grammar in language acquisition and the process of moving toward a target-like grammar based on the data to which the learner is exposed.

1.3 The Role of the L1 Grammar in L2 Acquisition

A major question in the study of multilingual speakers has been the extent to which the grammar of an already internalized language influences the acquisition of a new language. Much early work was driven by the Contrastive Analysis Hypothesis (Lado, 1957), which predicts *positive transfer* of features shared by the L1 and the L2 and *interference* or *negative transfer* where the two languages differ. In this view, learners' patterns that deviate from target language norms are classified as errors, indicative of incomplete learning. However, the generative view that learning a language involves constructing a grammar led to a shift in focus on learners' L2 productions as coherent systems (*interlanguages*; Selinker, 1972) that are worthy of study in their own right as reflections of the developing L2 grammar.

Much of the research inspired by the interlanguage approach maintains a crucial role for the L1 grammar, with most researchers assuming that for languages learned after early childhood, the initial state grammar is (a clone of) the L1 grammar, though developmental processes may also play some role (e.g., the Ontogeny Model; Major, 1987). In a rule-based framework, the learner must suppress rules of the L1 that do not apply in the new language and learn a new set of rules where warranted. In frameworks that assume universal parameters or constraints, the learner's task is to reset the parameters or rerank the constraints, and the assumption of universality predicts that any intermediate stages in the process will correspond to possible L1 grammars.

The study of interlanguage data as a coherent system led to the recognition of two sorts of patterns that are not readily explainable in terms of either transfer from the L1 or input from the L2: the fact that not all novel structures appear to be equally difficult for learners to master, and the fact that learners alter L2 structures in ways that may not be explainable as an effect of the L1 (which lacks the offending structure) or the L2 (which tolerates the structure). Such patterns have often been argued to provide evidence for the generative view that language acquisition is governed by innate, universal biases.

1.4 Representations in L2 Phonology: Phonemes, Features, Articulatory Gestures, and Acoustic Cues

A long tradition of work in the Speech Learning Model (Flege, 1995), the Perceptual Assimilation Model (Best & Tyler, 2007), and the Native Language Magnet Model (Kuhl & Iverson, 1995) focuses on how learners establish phoneme categories in L2 production and perception, a task that is complicated when the L1 and the L2 phoneme inventories do not match, or when a single phoneme category is realized differently in different positions (e.g., Flege, 1995; Trofimovich, Gatbonton, & Segalowitz, 2007). Yet in most generative models of the phonological grammar, phonemes per se are not primitives of mental representation; rather, segments are defined by a matrix of distinctive feature specifications. Furthermore, as Kang (2008) points out, generative models differ in their view of which features are present in lexical representations. SPE models assume that only contrastive features are specified, with rules filling in predictable features (such as aspiration for English voiceless stops in particular positions). In OT, on the other hand, the set of ranked constraints ensures the proper surface distribution of predictable features regardless of whether those features are specified in lexical representations or not, and the principle of Lexicon Optimization (Prince & Smolensky, 2004) assumes that learners will posit lexical representations that are as close as possible to surface representations. These differing views raise the question of whether L2 acquisition proceeds phoneme-by-phoneme, based on the phonetic similarity of L1 and L2 phonemes, or feature-by-feature – and if the latter, which types of features (contrastive versus predictable) are relevant.

1.4.1 Features versus Phonemes

The problem of phoneme substitution is one area where features have been argued to play a role in L2 phonology. Hancin-Bhatt (1994) laid out a feature-based approach to the long-standing puzzle of *differential substitution* illustrated by the observation that the preferred substitute for English /θ/ is /s/ for speakers of some languages (such as Turkish) but /t/ for speakers of others (such as Japanese), although each language has both /s/ and /t/ in its phoneme inventory. Hancin-Bhatt (1994) argues that the choice of substitute is a function of phoneme similarity, measured in terms of the number of specified features shared between two segments. She also argues that whether a feature value is specified depends on the functional load of that feature within a language; thus, the featural similarity of two phonemes can differ across languages. A conceptually similar appeal to the role played by a single feature within different languages is encoded in OT by Lombardi (2003) and by Wester, Gilbers, and Lowie (2007), who argue that speakers who choose /s/ as a substitute for /θ/ are those whose L1 grammar assigns a high rank to the

constraint requiring faithfulness to manner (the [+continuant] value shared by /s/ and /θ/); speakers whose L1 offers no evidence for such ranking choose the universally less marked substitute /t/. A challenge for such approaches, however, is posed by Brannen (2002), who points out that speakers of European and Canadian French, with essentially identical phoneme inventories and phonological systems, differ in their choice of replacement for English /θ/. Brannen attributes the differential choice of phoneme substitute to differences in the phonetic realization of /s/ in the two varieties of French.

Appeals to the L1 feature system have also been used to explain why certain novel structures seem to pose greater difficulty in perception for some L1 groups than for others. Brown (2000) argued that the ability to perceive a new contrast in L2 correlates with whether the feature which encodes that contrast plays a role in defining some contrast in the L1 grammar. Brown argues that Mandarin speakers are more successful than Japanese speakers in perceiving the English /ɹ–l/ phoneme contrast because Mandarin contrasts two coronal fricatives (alveolar and retroflex), so Mandarin speakers can deploy the feature distinguishing these segments to perceive the English contrast. This approach is somewhat abstract, since the feature [lateral] which is assumed to distinguish the English liquids does not function in Mandarin. And even this more abstract view of featural correspondence cannot account for the fact that Japanese speakers distinguish Russian /r–l/ with relative ease, presumably because the Russian phonetic cues to the liquid contrast are more robust than those for the English contrast, as well as more distinct from the cues associated with the allophonic variants of the Japanese liquid phoneme (Archibald, 2009; Larson-Hall, 2004).

Another problem for the claim that the perceptibility of an L2 contrast depends on the availability of a relevant contrastive L1 feature is outlined by Martinez, Goad, and Dow (2023), who investigated the ability of native speakers of French, whose language contrasts oral and nasal mid vowels, to discriminate the oral-nasal contrast in Brazilian Portuguese mid and high vowels. The French speakers not only were significantly less able than native Portuguese speakers to discriminate the novel high vowel oral-nasal contrast, they in fact performed no better than English and Caribbean Spanish speakers, in whose native languages vowel nasality is not contrastive. Thus, the French listeners seemed unable to deploy their native language [nasal] feature to discriminate a new oral-nasal contrast. However, some evidence suggests that the ability to perceive a contrast in one phoneme may generalize across different places of articulation and different prosodic positions. De Jong, Silbert, and Park (2009) found that the ability of Korean learners of English to identify English (nonsibilant) obstruents contrasting in [voice] or [continuant] correlated across sounds sharing the same feature; both contrasts are new for Korean speakers, in whose L1 neither voicing nor continuancy is contrastive for nonsibilant sounds.

In production, however, the picture is more complex. Using the same set of English obstruents, De Jong, Hao, and Park (2009) found correlations between

the accurate production of voicing contrasts across different places of articulation, but no similar correlation for the production of a continuancy contrast for labial and coronal positions. They propose that while the gestures involved in the realization of [voice] values for labials and velars are similar enough to generalize across different places of articulation, the realization of [continuant] contrasts requires mastery of different gestures for different places of articulation. This is consistent with the findings of Olson (2019a), who trained English learners of Spanish on Spanish-like production of one of the three Spanish voiceless stop consonants (which differ in voice onset time [VOT] from their English counterparts). The move to more Spanish-like VOT values did generalize beyond the trained segment to other stops, with a comparable degree of improvement for all three places of articulation. De Jong, Hao, and Park (2009) suggest that feature-based acquisition plays a greater role in perception than in production, since production requires the acquisition of gestural routines associated with particular feature values that may not be uniform across all sounds sharing a feature.

The role of features in learning new sequences of segments in L2 (new phonotactics) is less well explored. Moreton (2002) argued that English listeners' perception of ambiguous consonant-consonant-vowel (CCV) sequences showed a stronger bias against /dlV/ than /bwV/ because, although both combinations of phonemes are illegal in English onsets, /dl/ violates a prohibition on onset sequences of [coronal] consonants. However, Davidson (2010) questions the hypothesis that L2 onset clusters that are featurally most similar to L1 onsets are most easily produced; she found that although English and Catalan differ in permitted onset clusters, both language groups exhibited the same hierarchy of difficulty in the production of unfamiliar Russian cluster types, suggesting that difficulty is a function of language-independent phonetic factors rather than featural similarity to L1 clusters (though L1 differences did affect the types of errors).[1]

1.4.2 Integrated Models of Sound Structure

Like the papers discussed so far, much work in the L2 acquisition of new phonemes and phoneme sequences ends with the conclusion that both phonological and phonetic factors play a role in L2 phonology. However, this conclusion generally assumes the traditional separation of an abstract phonology from a phonetic component dealing with details of articulation and acoustics. As discussed earlier, that separation no longer represents unanimous opinion.

One model of a perception grammar incorporating acoustic information is that of Escudero and Boersma (2004), which addresses the perception by L1 Spanish speakers of the /i–ɪ/ contrast in either Scottish English or Southern

[1] The indirect role of phonological features in determining the sonority profile of syllable margins is another area that has attracted considerable attention; this is discussed in Section 1.5.2.

British English. Because these two target dialects differ in the cues used to signal the /i–ɪ/ contrast (durational versus spectral), learners develop different constraint rankings based on the dialect to which they are exposed. Their perception grammars map the L2 acoustic signals to phonological feature-based representations by employing constraints referring not only to abstract phonological features but also to phonetic cues such as segment durations and formant values. This groundbreaking analysis makes explicit the way in which learners can move from the initial state L1 constraint rankings toward an L2-like perception grammar, on the basis of the L2 data to which they are exposed. More recently, van Leussen and Escudero (2015) present a revision of this approach couched in terms of the strength of the connections along the paths mapping from the acoustic level to the lexical level.

A model of a production grammar incorporating articulatory information is Davidson's (2006) analysis of L1 English speakers' pronunciations of Russian onset clusters such as /zg/. This analysis is based on data from Davidson's (2005) study which found that while many English speakers produce the Russian clusters with what is perceived as an inserted schwa between the two consonants, this interconsonantal portion is both acoustically and articulatorily distinct from lexical schwas. Davidson (2005) argues that the L2 productions involve gestural mistiming rather than vowel insertion, and Davidson (2006) models this mistiming with a grammar that employs constraints referring to consonant release and the alignment of consonantal gestures, with the reranking of constraints over time accounting for the range of production patterns in the data.[2]

While these approaches depart quite radically from classic generative notions of the phonological grammar as abstract, categorical, and distinct from details of phonetic realization, they provide possible paths for moving beyond the phonetics–phonology dichotomy that has dominated much of the discussion in L2 phonology.

1.5 Derivations in L2 Phonology: Integrating Markedness and Frequency Effects

The shift from derivations involving rules to derivations involving constraints referring to universal principles of markedness has been arguably the widest departure from SPE phonology. This change was, however, predated by the recognition in the study of L2 phonology that equally novel L2 structures may pose different degrees of difficulty. Well before the emergence of OT in the 1990s, Eckman (1977) proposed the Markedness Differential Hypothesis, which supplements the Contrastive Analysis Hypothesis by correlating

[2] In a study of the production of sequences of consonants across word boundaries in the L2, Zsiga (2003) found that Russian learners of English transferred their native patterns of consonant release and gestural coordination to English, though only one L1 English speaker transferred the English timing pattern to Russian; the remaining speakers reverted to a default pattern that maximized recoverability of the word-final consonant.

difficulty of acquisition not only with novelty but with markedness. Eckman (1981) pointed out not only that marked structures were more difficult to acquire but that the apparently spontaneous emergence in interlanguage phonologies of rules repairing marked structures posed a challenge for frameworks which assumed that rules were learned from ambient data.

Another shift in generative phonology has been the move from grammars that derive a single surface form to grammars that predict the probability of different surface representations, taking into account the frequency of different structures in the input to the learner. These changes have pointed toward new ways of approaching SLA data and answering some of the questions raised in earlier accounts. In the rest of this section we examine evidence for markedness and frequency effects in three areas: the acquisition of laryngeal contrasts, of complex syllable margins, and of L2 stress and tone.

1.5.1 Acquisition of Final Laryngeal Contrasts

One of the best documented examples of the differential difficulty of equally novel structures involves the acquisition of obstruent voicing contrasts in syllable- or word-final position, with numerous studies showing greater accuracy in the production of English final voiceless than voiced obstruents among speakers of a wide range of L1s that lack any final obstruents (Mandarin, Tswana, Japanese, Angami; see Broselow, 2018; Eckman, 1981; Edge, 1991; Flege & Davidian, 1984; Wang, 1995; Weinberger, 1987; Wiltshire, 2006; Wissing & Zonneveld, 1996; Yavas, 1997). As Eckman (1981, 2004) points out, the greater difficulty of final voiced obstruents is consistent with the cross-linguistic generalization that while many languages (e.g., Dutch, Polish, Catalan) allow only voiceless obstruents in final position, the reverse pattern is unattested. But in a rule-based analysis, the repairs evidenced among L2 learners (deletion or devoicing of the final obstruent and insertion of a vowel after the obstruent) must reflect rules of the learners' grammar, and these rules are neither part of the L1 grammar nor motivated by the data of the L2.

As Broselow, Chen, and Wang (1998) point out, the existence of languages banning voiced obstruents in final position (but not the reverse) motivates a constraint against the marked structure, and the assumption that each constraint is universal means that even speakers whose L1 lacks any final obstruents have the constraint NoFinalVoicedObstruent as part of their grammar, although the effects of this constraint become visible only once learners have demoted the more general constraint banning all final obstruents. The principle that learners rank markedness constraints as high as possible predicts an L1 grammar that ranks NoFinalObstruent and NoFinalVoicedObstruent above faithfulness constraints, but, as Broselow (2004) points out, it is not clear why learners faced with both voiced and voiceless final obstruents in English should first demote only the more general constraint to arrive at an intermediate ranking NoFinalVoicedObstruent ≫ Faithfulness ≫ NoFinalObstruent. Broselow argues that this intermediate stage can be predicted once frequency is taken into

account; in the Gradual Learning Algorithm (Boersma & Hayes, 2001), the rate at which a constraint is demoted is a function of the frequency with which the constraint is violated. Any form that violates the more specific NoFinalVoicedObstruent will by definition also violate the more general NoFinalObstruent, causing the more general constraint to be demoted more quickly.

The problem of explaining why learners use rules that have no apparent L1 or L2 support, identified by Eckman (1981), is also avoided in the constraint-based approach, since in OT there are no rules to be learned; the grammar generates all possible surface outputs, and the task of the grammar is simply to determine the optimal surface correspondent of an underlying representation. In this approach, the choice of repair may (at least in some cases) be determined by markedness constraints. For example, Wang (1995) and Broselow et al. (1998) argue that Mandarin learners' choice of repair for forms with final voiced obstruents – deletion or devoicing of the obstruent versus insertion of a vowel after the obstruent – was governed by a preference for disyllabic words; learners were more likely to delete or devoice a final obstruent in disyllabic forms (thus preserving syllable count) than in monosyllables, where vowel insertion (addition of a syllable) was more likely. Though Mandarin has been argued to show a preference for disyllabic words, this preference has been argued to play a role in L2 phonology even in the absence of L1 support. In a study of Brazilian Portuguese speakers' productions of English coda stops, Cardoso (2007) found that learners who were sufficiently advanced in English to produce final stops were more likely to do so in polysyllabic words than in monosyllables, which were more likely to undergo insertion of a vowel after the final stop.

One unavoidable question is whether the greater difficulty of voiced obstruents is actually not a matter of the production grammar at all, but is rather based in perception – that is, whether speakers whose L1 lacks final laryngeal contrasts simply misperceive final voiced obstruents as voiceless, and then faithfully reproduce what they hear. However, several studies comparing L2 final obstruent perception and production cast doubt on this; for example, Wissing and Zonneveld's (1996) Tswana speakers performed better on perceiving than producing English final laryngeal contrasts, and L1 Mandarin speakers in Flege's (1988) study performed as well as native English speakers on identification of final voiceless and voiced obstruents in naturally produced English words. Nonetheless, difficulty in the perception of voicing most likely does play a role in Mandarin speakers' interlanguage; English obstruents in final position may be partially devoiced and also may be unreleased, and Flege's (1988) study found that when stimuli were edited to remove release bursts, the Mandarin speakers' accuracy in voicing identification dropped below that of English speakers. Thus, a full understanding of the Mandarin facts is incomplete without a model of the perception grammar alongside the production grammar.[3]

[3] We return to a related question – whether the failure to produce final voiced obstruents is properly regarded as a matter of the phonological grammar or of phonetic implementation – in Section 1.5.4.

Another example of differential difficulty that may be rooted in perceptual difficulty is the fact that the L1 Mandarin speakers in Broselow and Xu's (2004) study were more successful in producing final /m/ than final voiced obstruents. Since Mandarin lacks both final [m] and final voiced obstruents, learners who can produce the former but not the latter appear to have demoted NoFinalLabial more quickly than NoFinalVoicedObstruent, despite the fact that voiced obstruents are more frequent than labial consonants in English word-final position. Broselow and Xu (2004) suggest that the rate of constraint demotion depends not on absolute frequency but rather on perceived frequency: a structure must be perceived accurately to be identified as a violation, and they suggest that Mandarin listeners are simply better at perceiving final /m/ than final voiced obstruents, for two reasons. First, the Mandarin speakers' L1 contrast between coda /n/ and /ŋ/ requires them to attend to place cues in nasal codas. Second, English speakers may fail to produce the cues that Mandarin speakers rely on to identify voicing (Flege, 1988).

1.5.2 Acquisition of Complex Syllable Margins

Second language syllable structure is another area in which universal markedness has been argued to correlate with difficulty of acquisition. The assumption that segments within a syllable tend to be organized in terms of sonority (with vowels having highest sonority, followed in order of decreasing sonority by glides, liquids, nasals, fricatives, and stops) has been a staple of investigations of L1 syllable typology. Selkirk (1984) articulates two principles defining the relative markedness of syllable types: (i) the Sonority Sequencing Generalization (SSG), which favors syllable onsets that increase in sonority and codas that decrease in sonority; and (ii) the Minimal Sonority Distance principle (MSD), which favors syllable margins whose consonants are more widely separated on the sonority scale. These principles together define a markedness scale along which (i) onset clusters such as English /sp, st, sk/, which violate the SSG, are more marked than clusters such as /sl, sn, pl/, which obey the SSG; and (ii) onset clusters with greater sonority distance, such as stop-liquid /pl/, are less marked than fricative-liquid /sl/, which is in turn less marked than onsets such as fricative-nasal /sn/. These principles predict a hierarchy of differential difficulty in the realization of novel syllable margins.

Studies investigating the L2 accuracy of sC onsets in terms of markedness as defined by the SSG have yielded mixed results. The Brazilian Portuguese speakers in Cardoso and Liakin's (2008) study produced SSG-conforming English /sn/, /sl/ more accurately than the marked /st/ onset, even though /st/ onsets were more frequent in the learners' input. Learners also produced /sl/ more accurately than /sn/, in accord with the MSD principle. Additional support for the role of markedness in production of sC was found by Cardoso, Collins, and Cardoso (2021), who trained Brazilian Portuguese speakers on

one of the onsets /st, sn, sl/. The participants trained on the most marked onset improved in their production of the untrained less marked onsets, but training on a less marked onset did not improve performance on more marked onsets. Markedness did not, however, account for the Portuguese speakers' perception of English onsets; Cardoso, John, and French (2008) found the highest perceptual accuracy rates for the most marked (but most frequent) /st/ onsets, suggesting that accuracy in perception is more affected by input frequency than by markedness. Furthermore, a number of studies of different L1–L2 combinations have found no clear correlation between the SSG-defined markedness of sC onsets and their rate of acquisition in L2; the reader is referred to Edwards (2014) for an overview.

Studies investigating the role of the MSD in a wider range of onsets and codas have also yielded mixed results. Broselow and Finer (1991) found that L1 speakers of Japanese and Korean, languages which allow only the least marked obstruent-glide onsets, produced stop-liquid onsets in English pseudowords more accurately than the more marked fricative-liquid onsets. Assuming a parameter-setting framework, they attribute this effect to the learners having moved to a parameter setting intermediate between the native and the target language settings. Hancin-Bhatt and Bhatt (1998) investigated the role of markedness in a constraint-based approach, examining the production of English complex onsets and codas by L1 speakers of Spanish and Japanese. They found that for L1 Spanish speakers, English stop-glide onsets were actually less accurately produced than the more marked stop-liquid onsets, though the Japanese speakers in their study did show accuracy rates consistent with MSD-defined onset markedness. A markedness effect in production of codas was found for the Spanish speakers, who were less accurate in producing the more marked liquid-nasal codas than the less marked liquid-obstruent codas, but not for the Japanese speakers. Hancin-Bhatt and Bhatt (1998) argue that the results that cannot be explained by markedness are attributable to effects of highly ranked L1 constraints determining the types of segments that are permitted in onset and coda in the native language, which also determine the preferred repair. Thus, the combination of constraints defining coda cluster markedness along with high-ranked L1 constraints defining acceptable coda single consonants together select the correct surface forms without the need for an L2-specific rule of consonant deletion. The OT approach therefore provides an account of phenomena that are unexplained in a rules-and-parameter-based approach.

Sonority-based markedness cannot, however, account for the full range of differential difficulty found in production of novel clusters. For example, Davidson's (2010) study of the production of novel onset types by L1 English and Catalan speakers revealed that for clusters involving a sonority plateau (either stop-stop or fricative-fricative), all of which are equally marked in terms of both SSG and MSD, participants were more accurate on onsets in which both consonants were voiceless (/fs/, /tp/) than on voiced clusters (/zv/, /db/). Davidson (2010) explains this pattern in terms of the articulatory

difficulty of sustaining voicing in obstruent clusters and the greater perceptual salience of fricatives versus stops in pre-consonantal position. As discussed in Section 1.4, Davidson (2006) provides a grammar incorporating constraints on gestural timing that accounts for patterns of non-native production as a failure to closely coordinate the articulation of the two onset consonants. Wilson and Davidson (2013) provide, in addition, a proposal for building frequency effects into the grammar with a Bayesian model of the English and Catalan productions that predicts the probability of different realizations of each type of onset, an approach that recognizes and accounts for the high degree of variability that is a feature of most L2 performances.

1.5.3 Acquisition of Suprasegmentals

The study of stress and tone systems has yielded several generally agreed upon universal preferences. Some research supports a role for markedness in L2 learners' productions even in the absence of direct evidence from the native or target languages.

In many languages, the position of stress within a word depends upon the weight of the syllables; for example, in Latin, the penultimate syllable of a word is stressed if and only if it is heavy (CVV or CVC, as opposed to light CV). In languages that take syllable weight into account in assigning stress, it is always heavier syllables that attract stress, and this is encoded in OT as the WSP (Weight-to-Stress Principle) constraint, which requires heavy syllables to be stressed (Kager, 1999). While English does exhibit some preference for stressing heavy syllables – particularly, syllables containing a long vowel – the English lexicon contains many violations of the WSP – for example, in *calendar*, where stress falls on the initial light syllable. Nonetheless, Guion, Harada, and Clark (2004) and Guion (2005) found a strong tendency among speakers of L1 Spanish and L1 Seoul Korean to place stress on heavy syllables in English pseudowords, a system that receives at best weak support in Spanish (in which that final CVC is somewhat more likely to be stressed than final CV) and no support in Seoul Korean, which is not a stress-based language. Similarly, Özçelik (2018) found that L2 learners of Turkish, where prominence falls on a word-final syllable, showed a tendency to accompany the move to final stress in Turkish with lengthening of vowels in word-final syllables.

In the area of tone, one well-established markedness principle is the OCP (Obligatory Contour Principle) which disfavors sequences of identical tones. Zhang (2016) found that native speakers of English, Korean, and Japanese studying Mandarin made more errors on words containing sequences of identical tones (whether contour or level tones) than nonidentical tones, consistent with the OCP. The exception was sequences of high tones, which showed high accuracy, a fact that Zhang (2016) ascribes to the universally least marked status of high tones. The learners' productions are analyzed as reflecting a grammar in which the constraint penalizing identical tone sequences is outranked by the constraint favoring high tone.

1.5.4 The Role of Markedness in Phonological Derivations

The introduction of markedness as the major determinant of the mapping from underlying to surface representations provided a way to use learners' grammars to account for why some novel structures seem to be more difficult than other equally novel structures as well as why learners choose particular repair strategies (for example, the preference for disyllabic words motivates insertion of a vowel in a monosyllable but not a disyllable).

However, structures that are considered more marked are arguably precisely those structures that are more difficult to produce and/or perceive. One position held by many is that the comparative rarity of marked structures is not within the province of grammatical explanation – the rarity of marked structures simply reflects the fact that marked structures are those most likely to be altered or lost in the process of transmitting a language from one generation to another (e.g., Blevins, 2004). And if marked structures are marked because of inherently greater difficulty, the fact that L2 speakers tend to be less successful in producing marked than unmarked structures does not in itself constitute an argument that markedness is built into phonological grammar. For example, voicing in final obstruents is not only difficult to perceive (particularly if cues like release burst are absent), it is also difficult to sustain during the narrow vocal tract constriction required for obstruent production. Even markedness principles that have no obvious explanation in terms of articulatory or perceptual difficulty may have a functional grounding; for example, Frisch (2009) argues that the OCP is based in processing difficulty, as the neurons that fire in the presence of one set of cues may not have sufficient time to recover if an identical stimulus follows too closely. The debate concerning the role of markedness in the grammar is far from being resolved, but it is clear that accounts of L2 phonology must take into consideration all the factors that may contribute to difficulty of acquisition in L2 (Archibald, 2021).

1.6 Grammar Interaction in Bilingual Speakers

The view that learning a language means inducing an internalized abstract grammar raises the question of whether bilinguals maintain distinct independent phonological grammars for each language, or whether these two systems may interact. Here factors such as age of acquisition and language dominance must be taken into account.

Flege's (1991) study of VOT values in the production of /t/ by early versus later Spanish-English bilinguals suggests that age of acquisition affects the degree of separation in systems of phoneme realization. Flege (1991) found that participants who had learned English at five to six years of age produced monolingual-appropriate VOT values in both Spanish and English words, suggesting that these speakers maintain not only independent phoneme categories but also independent systems of phonetic realization of those

categories. The productions of the early bilinguals contrasted with those of the L1 Spanish speakers who had learned English as adults, who showed VOT values for English /t/ that were significantly shorter (more Spanish-like) than those of monolingual English speakers. Similarly, Gonzales and Lotto (2013) found that when confronted with stop consonants with VOT values that fell within the voiced category in English but were voiceless in Spanish, early Spanish-English bilinguals (but not English monolinguals) shifted their identification of the stop's voicing depending on whether the pseudoword in which the stops were embedded contained Spanish-specific versus English-specific cues.

Yet even early bilinguals may show evidence of permeability across different linguistic systems. For example, Macleod and Stoel-Gammon (2005) found that French-English bilinguals who had acquired both languages before the age of three evidenced monolingual-like VOT values for voiceless stops in each language but not for English voiced stops, which were realized with VOT values more like those of French stops. It is difficult to determine whether such effects result from the interpenetration of grammatical systems or from external factors such as exposure to accented (nonmonolingual) varieties of one of the languages and/or to a range of variability in the realization of phoneme categories even among monolingual speakers. In any event, the findings that bilinguals sometimes exhibit realizations of a phoneme category that are somewhere between the monolingual norms of each language (e.g., Flege & Eefting, 1987) as well as the phenomenon of phonetic drift, whereby even a late-acquired L2 may cause changes in the phonetics of the L1 (Chang, 2019), suggest that, at least in terms of the pronunciation of segments, a speaker's coexisting internalized linguistic systems may interact.

Another area to look to for insight into the relationship between coexisting grammars is cases where a speaker's languages differ in the presence versus absence of a phonological process. For example, voiced stops are realized as spirants between vowels in Spanish but not in English, so bilinguals must map intervocalic voiced stops to spirants in Spanish but suppress this mapping in English productions. Amengual (2019) found that L1 English speakers who had learned Spanish after age eight showed less spirantization in read Spanish words than participants who had acquired Spanish earlier, suggesting incomplete mastery of the Spanish process by later learners. However, evidence that at least some Spanish-English bilinguals can successfully switch between phonological systems is presented by Olson (2019b), who analyzed the production of intervocalic voiced stops within all-Spanish, all-English, and code-switched sentences. Spirantization of word-initial voiced stops following a vowel-final word was more likely when the stop was in a Spanish word, regardless of whether the surrounding environment was English or Spanish, though the degree of difference in the stops in Spanish versus English words was more pronounced in participants who scored toward the balanced end of the language dominance scale. The difficulties of finding truly balanced

bilinguals and of controlling for factors such as the types of linguistic data to which bilinguals are exposed mean that the question of whether and how coexisting grammars interact remains elusive.

1.7 Future Directions

The past decades of research in generative phonology have seen a wealth of new proposals concerning the nature of phonological representations and phonological derivations, the role of markedness in the phonological grammar, the relationship between phonetics and phonology and between production and perception, and the role of usage factors such as frequency in language learning and grammar construction. New views of the phonological grammar have suggested new ways of approaching old problems in the study of L2 phonology: the inclusion of gestural targets in the elements making up phonological representations suggests ways to move beyond the phonetics–phonology dichotomy; the appeal to specific perception grammars clarifies the role of the L1 in L2 perception; the move to probabilistic rather than categorical grammar addresses the role of frequency and variation in L2 production; and the incorporation of markedness in phonological derivations sheds light on the problems of the differential difficulty of novel L2 structures and the differential repair strategies employed by speakers of different L1s or for different L2 structures. Models of L2 perception that draw on insights from connectionist approaches (e.g., van Leussen & Escudero, 2015; Bordag, Gor, & Opitz, 2021) provide avenues for insight into the nature of speakers' lexical representations, and new imaging techniques may help to reveal distinctions in mental representations that are not obvious from more traditional acoustic investigation of speakers' productions (e.g., Song & Eckman, 2021). Second language data have provided and will continue to provide a fruitful area in which to explore the predictions of different models of the internalized knowledge that constitutes a grammar.

References

Amengual, M. (2019). Type of early bilingualism and its effect on the acoustic realization of allophonic variants: Early sequential and simultaneous bilinguals. *International Journal of Bilingualism*, *23*(5), 954–970.

Archibald, J. (2009). Phonological feature re-assembly and the importance of phonetic cues. *Second Language Research*, *25*(2), 231–233.

Archibald, J. (2021). Ease and difficulty in L2 phonology: A mini-review. *Frontiers in Communication*, 6. www.frontiersin.org/articles/10.3389/fcomm.2021.626529/full.

Berwick, R. (1985). *The Acquisition of Syntactic Knowledge*. Cambridge, MA: MIT Press.

Best, C. & Tyler, A. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn & M. J. Munro, eds., *Second*

Language Speech Learning: The Role of Language Experience in Speech Perception and Production. Amsterdam: John Benjamins, pp. 13–34.

Blevins, J. (2004). *Evolutionary Phonology: The Emergence of Sound Patterns*. Cambridge: Cambridge University Press.

Boersma, P. (1998). *Functional Phonology*. [Doctoral dissertation, University of Amsterdam]. The Hague: Holland Academic Graphics.

Boersma, P. & Hayes, B. (2001). Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry*, *32*, 45–86.

Bordag, D., Gor, K., & Opitz, A. (2021). Ontogenesis model of the L2 lexical representation. *Bilingualism: Language and Cognition*, 25, 185–201.

Brannen, K. (2002). The role of perception in differential substitution. *Canadian Journal of Linguistics*, *47*, 1–46.

Broselow, E. (2004). Unmarked structures and emergent rankings in second language acquisition. *International Journal of Bilingualism*, *8*(1), 51–66.

Broselow, E. (2018). Laryngeal contrasts in second language phonology. In L. Hyman & F. Plank, eds., *Linguistic Typology*. Berlin: Walter de Gruyter, pp. 312–340.

Broselow, E., Chen, S., & Wang, C. (1998). The emergence of the unmarked in second language acquisition. *Studies in Second Language Acquisition*, *20*, 261–280.

Broselow, E. & Finer, D. (1991). Parameter setting in second language phonology. *Second Language Research*, *7*(1), 35–59.

Broselow, E. & Xu, Z. (2004). Differential difficulty in the acquisition of second language phonology. *International Journal of English Studies*, *4*(2), 145–163.

Browman, C. & Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology*, *6*, 201–252.

Brown, C. (2000). Speech perception and phonological acquisition. In J. Archibald, ed., *Second Language Acquisition and Linguistic Theory*. Oxford: Blackwell, 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*, 219–248.

Cardoso, W., Collins, L., & Cardoso, W. (2021). Developmental sequences in second language phonology: Effects of instruction on the acquisition of foreign sC onsets. *Frontiers in Communication*, 6. www.frontiersin.org/articles/10.3389/fcomm.2021.662934/full.

Cardoso, W., John, P., & French, L. (2008). The variable perception of /s/ + coronal onset clusters in Brazilian Portuguese English. In M. A. Watkins, A. S. Rauber, & B. O. Baptista, eds., *Recent Research in Second Language Phonetics/Phonology: Perception and Production*. Newcastle upon Tyne: Cambridge Scholars Publishing, pp. 203–231.

Cardoso, W. & Liakin, D. (2008). When input frequency patterns fail to drive learning: The acquisition of sC onset clusters. In M. A. Watkins, A. S. Rauber, & B. O. Baptista, eds., *Recent Research in Second Language Phonetics/Phonology: Perception and Production*. Newcastle upon Tyne: Cambridge Scholars Publishing, pp. 174–202.

Chang, C. (2019). Phonetic drift. In M. S. Schmid & B. Köpke, eds., *The Oxford Handbook of Language Attrition*. Oxford: Oxford University Press, pp. 191–207.

Chomsky, N. & Halle, M. (1968). *The Sound Pattern of English*. New York: Harper & Row.

Clements, N. (1985). The geometry of phonological features. *Phonology Yearbook*, *2*, 225–252.

Davidson, L. (2005). Addressing phonological questions with ultrasound. *Clinical Linguistics and Phonetics*, 19, 619–633.
Davidson, L. (2006). Phonotactics and articulatory coordination interact in phonology: Evidence from nonnative production. *Cognitive Science*, *30*, 837–862.
Davidson, L. (2010). Phonetic bases of similarities in cross-language production: Evidence from English and Catalan. *Journal of Phonetics*, *38*(2), 272–288.
De Jong, K., 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*(4), 357–373.
De Jong, K., Silbert, N. H., & Park, H. (2009). Generalization across segments in second language identification. *Language Learning*, *59*(1), 1–31.
Eckman, F. R. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning*, *27*(2), 315–330.
Eckman, F. R. (1981). On the naturalness of interlanguage phonological rules. *Language Learning*, *31*, 195–216.
Eckman, F. (2004). From phonemic differences to constraint rankings: Research on second language phonology. *Studies in Second Language Acquisition*, *26*, 513–549.
Edge, B. (1991). The production of word-final voiced obstruents in English by L1 speakers of Japanese and Cantonese. *Studies in Second Language Acquisition*, *13*, 377–393.
Edwards, J. G. H. (2014). The role of input frequency, universals, and L1 transfer in the acquisition of English L2 onsets by native speakers of Cantonese, Mandarin Chinese, and Vietnamese. In M. Yavas, ed., *Unusual Productions in Phonology: Universals and Language-Specific Considerations*. New York: Psychology Press, pp. 206–225.
Escudero, P. & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition*, *26*, 551–585.
Flege, J. E. (1988). Chinese subjects' perception of the word-final English /t/-/d/ contrast: Performance before and after training. *Journal of the Acoustical Society of America*, *86*, 1684–1697.
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*(1), 395–411.
Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange, ed., *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Baltimore, MD: York Press, pp. 233–273.
Flege, J. E. & Davidian, R. D. (1984). Transfer and developmental processes in adult foreign language speech production. *Applied Psycholinguistics*, *5*, 323–347.
Flege, J. E. & Eefting, W. (1987). Production and perception of English stops by native Spanish speakers. *Journal of Phonetics*, *15*, 67–83.
Frisch, S. (2009). Language processing and segmental OCP effects. In B. Hayes, R. Kirchner, & D. Steriade, eds., *Phonetically-Based Phonology*. Cambridge: Cambridge University Press, pp. 346–371.
Gafos, A. (2002). A grammar of gestural coordination. *Natural Language and Linguistic Theory*, *20*, 269–337.
Goldsmith, J. (1979). *Autosegmental Phonology*. [Doctoral dissertation, Massachusetts Institute of Technology]. New York: Garland Press.
Gonzales, K. & Lotto, A. J. (2013). A *Bafri*, un *Pafri*: Bilinguals' pseudoword identifications support language-specific phonetic systems. *Psychological Science*, *24*(11), 2135–2142.

Guion, S. (2005). Knowledge of English word stress patterns in early and late Korean-English bilinguals. *Bilingualism: Language and Cognition*, *27*(4), 503–533.

Guion, S., Harada, T., & Clark, J. J. (2004). Early and late Spanish-English bilinguals' acquisition of English word stress patterns. *Bilingualism: Language and Cognition*, *7* (3), 207–226.

Hancin-Bhatt, B. (1994). Segmental transfer: A consequence of a dynamic system. *Second Language Research*, 10, 241–269.

Hancin-Bhatt, B. & Bhatt, R. (1998). Optimal L2 syllables: Interactions of transfer and developmental effects. *Studies in Second Language Acquisition*, *19*, 331–378.

Hayes, B. (1985). *A Metrical Theory of Stress Rules*. [Doctoral dissertation, Massachusetts Institute of Technology]. New York: Garland Press.

Hayes, B. & Wilson, C. (2008). A maximum entropy model of phonotactics and phonological learning. *Linguistic Inquiry*, *39*(3), 379–440.

Kager, R. (1999). *Optimality Theory*. Cambridge: Cambridge University Press.

Kager, R., Pater, J., & Zonneveld, W. (2004). *Constraints in Phonological Acquisition*. Cambridge: Cambridge University Press.

Kang, Y. (2008). Interlanguage segmental mapping as evidence for the nature of lexical representation. *Language and Linguistic Compass*, *2*, 103–118.

Kisseberth, C. (1970). On the functional unity of phonological rules. *Linguistic Inquiry*, *1*, 291–306.

Kuhl, P. K. & Iverson, P. (1995). Linguistic experience and the "perceptual magnet effect." In W. Strange, ed., *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Baltimore, MD: York Press, pp. 121–154.

Lado, R. (1957). *Linguistics Across Cultures: Applied Linguistics for Language Teachers*. 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*, 32–76.

Lombardi, L. (2003). Second language data and constraints on manner: Explaining substitutions for the English interdentals. *Second Language Research*, *19*, 225–250.

Macleod, A. N. & Stoel-Gammon, C. (2005). Are bilinguals different? What VOT tells us about simultaneous bilinguals. *Journal of Multilingual Communication Disorders*, *3* (2), 118–127.

Major, R. (1987). A model for interlanguage phonology. In G. Ioup & S. Weinberger, eds., *Interlanguage Phonology: The Acquisition of a Second Language Sound System*. Cambridge, MA: Newbury House, pp. 101–124.

Martinez, R. M., Goad, H., & Dow, M. (2023). L1 phonological effects on L2 (non-) naïve vowel perception: A cross-language investigation of the oral-nasal vowel contrast in Brazilian Portuguese. *Second Language Research*, *39*(2), 387–423.

Moreton, E. (2002). Structural constraints in the perception of English stop-sonorant clusters. *Cognition*, *84*, 55–71.

Olson, D. J. (2019a). Feature acquisition in L2 phonetic development: Evidence from phonetic training. *Language Learning*, *69*(2), 366–404.

Olson, D. J. (2019b). Phonological processes across word and language boundaries: Evidence from code-switching. *Journal of Phonetics*, *77*, 1–16.

Özçelik, Ö. (2018). Universal grammar and second language phonology: Full transfer/ prevalent access in the L2 acquisition of Turkish "stress" by English and French speakers. *Language Acquisition*, *25*(3), 231–267.

Prince, A. & Smolensky, P. (2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden, MA: Wiley-Blackwell.

Sagey, E. (1986). The representation of features and relations in non-linear phonology. [Doctoral dissertation, Massachusetts Institute of Technology].

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics*, *10*, 209–231.

Selkirk, E. (1984). On the major class features and syllable theory. In M. Aronoff & R. Oehrle, eds., *Language Sound Structure: Studies in Phonology*. Cambridge, MA: MIT Press, pp. 107–136.

Song, J. Y. & Eckman, F. (2021). Using ultrasound tongue imaging to study covert contrasts in second-language learners' acquisition of English vowels. *Language Acquisition*, *28*(4), 344–369.

Stampe, D. (1979). *How I Spent My Summer Vacation (A Dissertation on Natural Phonology)*. [Doctoral dissertation, University of Chicago]. New York: Garland Press.

Trofimovich, P., Gatbonton, E., & Segalowitz, N. (2007). A dynamic look at L2 phonological learning: Seeking processing explanations for implicational phenomena. *Studies in Second Language Acquisition*, *29*, 407–448.

van Leussen, J. W. & Escudero, P. (2015). Learning to perceive and recognize a second language: The L2LP model revised. *Frontiers in Psychology*, 6, 1000. https://doi.org/10.3389/fpsyg.2015.01000.

Wang, C. (1995). The acquisition of English word-final obstruents by Chinese speakers. [Doctoral dissertation, Stony Brook University].

Weinberger, S. H. (1987). The influence of linguistic context on syllable simplification. In G. Ioup & S. H. Weinberger, eds., *Interlanguage Phonology: The Acquisition of a Second Language Sound System*. Rowley, MA: Newbury House, pp. 401–417.

Wester, F., Gilbers, D., & Lowie, W. (2007). Substitution of dental fricatives in English by Dutch L2 speakers. *Language Sciences*, *29*, 477–491.

Wilson, C. & Davidson, L. (2013). Bayesian analysis of non-native cluster production. In S. Kan, C. Moore-Cantwell, & R. Staubs, eds, *Proceedings of the 40th Annual Meeting of the North East Linguistic Society*. Amherst, MA: Graduate Linguistics Student Association, pp. 265–278.

Wiltshire, C. (2006). Word-final consonant and cluster acquisition in Indian English(es). In D. Bamman, T. Magnitskaia, & C. Zaller, eds., *Boston University Conference on Language Development 30 Online Proceedings Supplement*. www.bu.edu/bucld/files/2011/05/30-WiltshireBUCLD2005.pdf.

Wissing, D. & Zonneveld, W. (1996). Final devoicing as a robust phenomenon in second language acquisition: Tswana, English, and Afrikaans. *South African Journal of Linguistics*, *14*, 3–23.

Yavas, M. (1997). The effects of vowel height and place of articulation in interlanguage final stop devoicing. *International Review of Applied Linguistics in Language Teaching*, *35*(2), 115–125.

Zhang, H. (2016). Dissimilation in the second language acquisition of Mandarin Chinese tones. *Second Language Research*, *32*(3), 427–451.

Zsiga, E. (2003). Articulatory timing in a second language: Evidence from Russian and English. *Studies in Second Language Acquisition*, *25*, 399–432.